CHAPTER 3

Primes in arithmetic progressions to large moduli

In this section we prove the celebrated theorem of Bombieri and Vinogradov

THEOREM 3.1 (Bombieri-Vinogradov). *For any $A \geq 1$, there exists $B = B(A) \geq 1$ such that for any $x \geq 1$*

$$\sum_{q \leq Q} \max_{(a,q)=1} \left|\psi(x; q, a) - \frac{1}{\varphi(q)}\psi^{(q)}(x)\right| \ll \frac{x}{\log^A x}$$

for $Q = x^{1/2}/\log^B x$. Here

$$\psi^{(q)}(x) = \sum_{\substack{n \leq x \\ (n,q)=1}} \Lambda(n).$$

REMARK 3.1. The term $\frac{1}{\varphi(q)}\psi^{(q)}(x)$ is the expected main term for the distribution of Λ in arithmetic progressions of modulus q and coprime to q; we can also replace this term by the seemingly more natural term $\frac{1}{\varphi(q)}\psi(x)$ at the cost of an error of size $O(\log q/\varphi(q))$. Observe that for Q a fixed positive power of x

$$\sum_{q \leq Q} \frac{1}{\varphi(q)}\psi(x) \simeq x \sum_{q \leq Q} \frac{1}{\varphi(q)} \geq x \log Q \gg x \log x.$$

Therefore the Bombieri-Vinogradov theorem states that the maximal error term on the distribution of primes in arithmetic progressions of modulus q

$$E(\Lambda, x; q) = \max_{(a,q)=1} E(\Lambda, x; q, a) = \max_{(a,q)=1} |\psi(x; q, a) - \frac{1}{\varphi(q)}\psi^{(q)}(x)|$$

is on average over $q \leq Q$ is $O(x/\log^A x)$ and is therefore negligible compared to the average of the main term; put in another way for any $A \geq 1$

$$E(\Lambda, x; q) \ll \frac{1}{\varphi(q)} \frac{\psi(x)}{\log^A x}$$

for almost all $q \leq Q = x^{1/2} \log^{-B} x$ for some $B = B(A)$.

Observe that the GRH would give that for for any $q \leq x$

$$\sum_{q \leq Q} E(\Lambda, x; q) \ll Q x^{1/2} \log^2 x = x/\log^{B-2} .x$$

therefore excepted for the dependency of B wrt to A the Bombieri-Vinogradov theorem does as good as the GRH for the distribution of primes in arithmetic progressions on average over the modulus.

1. Reduction to the large sieve inequality

We return to the special case of the von Mangolt function:

$$\begin{aligned}|\psi(x;q,a)-\frac{1}{\varphi(q)}\psi(x)| &= |\frac{1}{\varphi(q)}\sum_{1\neq\chi\,(\mathrm{mod}\,q)}\overline{\chi}(a)\sum_{n\leq x}\chi(n)\Lambda(n)|+O(\frac{\log q}{\varphi(q)})\\ &\leq \frac{1}{\varphi(q)}\sum_{1\neq\chi\,(\mathrm{mod}\,q)}|\sum_{n\leq x}\chi(n)\Lambda(n)|+O(\frac{\log q}{\varphi(q)}).\end{aligned}$$

the last term accounting for the contribution in the second term of the n not coprime with q. The total contribution of these lasts terms is bounded by

$$\ll\sum_{q\leq Q}\frac{\log q}{\varphi(q)}\leq\log Q\sum_{q\leq Q}\frac{q}{\varphi(q)}\frac{1}{q}\ll\log^2 Q.$$

here we have used the following

LEMMA 3.1. *For $Q\geq 1$, one has*

$$\sum_{q\leq Q}\frac{1}{\varphi(q)}\ll\log Q.$$

PROOF. This follows from the analytic properties of the L-function associated to the multiplicative function $q\mapsto q/\varphi(q)$: indeed for $\Re s>1$ one has

$$\sum_{q\geq 1}\frac{q}{\varphi(q)}\frac{1}{q^s}=\prod_p(1+(1-\frac{1}{p})^{-1}\frac{1}{p^s}(1-\frac{1}{p^s})^{-1})=\zeta(s)H(s)$$

with

$$H(s)=\prod_p(1+O(p^{-(s+1)}+p^{-2s}))$$

is holomorphic for $\Re s>1/2$. Therefore $\zeta(s)H(s)$ is meromorphic in $\Re s>1/2$ with a most a simple pole at $s=1$ (in fact this is a pole as more computation show that $H(1)\neq 0$). $\square$

We need therefore to evaluate

$$\sum_{q\leq Q}\frac{1}{\varphi(q)}\sum_{1\neq\chi\,(\mathrm{mod}\,q)}|\sum_{n\leq x}\chi(n)\Lambda(n)|.$$

We will also reduce this summation over primitive characters: given $\chi\,(\mathrm{mod}\,q)$ let $\chi^*\,(\mathrm{mod}\,q^*)$ be the primitive inducing χ, we have

$$|\sum_{n\leq x}\chi(n)\Lambda(n)|=|\sum_{\substack{n\leq x\\(n,q)=1}}\chi^*(n)\Lambda(n)|=|\sum_{n\leq x}\chi^*(n)\Lambda(n)|+O(\log q)$$

by bounding trivially the contribution of the n which are coprime to q^* but not coprime to q (and are therefore powers of primes dividing q). Writing $q = q^*q'$ we have

$$\sum_{q\le Q}\frac{1}{\varphi(q)}\sum_{1\neq\chi\,(\mathrm{mod}\,q)}|\sum_{n\le x}\chi(n)\Lambda(n)| =$$
$$\sum_{\substack{q^*q'\le Q\\ q^*>1}}\frac{1}{\varphi(q^*q')}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q^*)}|\sum_{n\le x}\chi^*(n)\Lambda(n)| + \sum_{q\le Q}O(\log q)$$

Here $\sum^\star$ mean that we average over primitive characters of modulus q^*. The second term is bounded by $O(Q\log Q)$ while for the first, we bound it using that $\varphi(q^*q')\ge\varphi(q^*)\varphi(q')$ so that this term is bounded by

$$\sum_{1<q^*\le Q}\frac{1}{\varphi(q^*)}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q^*)}|\sum_{n\le x}\chi^*(n)\Lambda(n)|(\sum_{q'\le Q/q^*}\frac{1}{\varphi(q')})$$
$$\ll \log Q\sum_{1<q^*\le Q}\frac{1}{\varphi(q^*)}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q^*)}|\sum_{n\le x}\chi^*(n)\Lambda(n)|.$$

by Lemma 3.1.

1.1. Applying Siegel's Theorem. We need to evaluate

$$\sum_{1<q\le Q}\frac{1}{\varphi(q)}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q)}|\sum_{n\le x}\chi(n)\Lambda(n)|$$

and for this we split the q-summation into two ranges: the small and the large moduli,

$$\sum_{1<q\le Q}\cdots = \sum_{1<q\le Q_1}\cdots + \sum_{Q_1<q\le Q}\cdots$$

where $Q_1 = \log^C x$ for some fixed $C\ge 1$ to be choosen later. For the small range we use the Siegel-Walfisz theorem: since $q > 1$ and each primitive $\chi\,(\mathrm{mod}\,q)$ being non-trivial, one has for any $A\ge 1$

$$\frac{1}{\varphi(q)}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q)}|\sum_{n\le x}\chi(n)\Lambda(n)| \ll_A \frac{x}{\log^A x}$$

and therefore

$$\sum_{1<q\le Q_1}\frac{1}{\varphi(q)}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q)}|\sum_{n\le x}\chi(n)\Lambda(n)| \lll_A \frac{x}{\log^{A-C} x} \tag{3.1}$$

which will be admissible as long as we take A sufficiently large compared to C.

It is to bound the large moduli range

$$\sum_{Q_1<q\le Q}\frac{1}{\varphi(q)}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q)}|\sum_{n\le x}\chi(n)\Lambda(n)| \tag{3.2}$$

that we need the so called multiplicative large sieve inequality.

2. Large Sieve inequalities

The above computations have reduce the proff of the Bombieri-Vinogradov theorem to the problem of evaluating on average of $q \leq Q$ and $\chi \,(\mathrm{mod}\, q)$ (primitive) the absolute values of linear forms

$$\ell(\Lambda, \chi; x) = \sum_{n \leq x} \Lambda(n)\chi(n).$$

The multiplicative large sieve inequality provide similar bounds for the average square of these linear forms for general arithmetic function (in place of just the van Mangolt function Λ):

2.1. The multiplicative large sieve inequality. For this additive version of the large sieve we deduce a multiplicative version

THEOREM 3.2. *For any $M \geq 1$ and $(\alpha_m)_{m \leq M}$ and any $Q \geq 1$ we have*

$$\sum_{q \leq Q} \frac{q}{\varphi(q)} \sideset{}{^\star}\sum_{\chi \,(\mathrm{mod}\, q)} |\sum_{m \leq M} \alpha_m \chi(m)|^2 \ll (Q^2 + M) \sum_{m \leq M} |\alpha_m|^2.$$

Here $\sum^\star$ mean that we average over primitive characters of modulus q.

Before embarking for the proof we deduce some corollaries

COROLLARY 3.1. *For any $(\alpha_n)_{n \geq 1}$ and any $Q_1, Q, N \geq 1$, we have*

$$\sum_{Q_1 \leq q \leq Q} \frac{1}{\varphi(q)} \sideset{}{^\star}\sum_{\chi \,(\mathrm{mod}\, q)} |\sum_{n \leq N} \alpha_n \chi(n)|^2 \ll \frac{\log Q}{Q_1}(Q^2 + N) \sum_{n \leq N} |\alpha_n|^2.$$

Here $\sum^\star$ mean that we average over primitive characters of modulus q.

PROOF. We decompose the sum into a sum of $O(\log Q)$ dyadic intervals

$$\sum_{Q_1 \leq q \leq Q} \frac{1}{\varphi(q)} \cdots = \sum_{Q'} \sum_{Q' < q \leq 2Q'} \frac{1}{\varphi(q)} \cdots ;$$

for each such sum we have

$$\sum_{Q' < q \leq 2Q'} \frac{1}{\varphi(q)} \cdots \leq \frac{1}{2Q'} \sum_{q \leq 2Q'} \frac{q}{\varphi(q)} \cdots$$

and we apply the multiplicative large sieve inequality. $\square$

2.2. Multiplicative large sieve inequalities for convolutions. We deduce from this result a bound for the average value of linear forms of non-trivial Dirichlet convolution:

COROLLARY 3.2. *For any sequences of complex numbers* $(\alpha_m)_{m\leq M}$, $(\beta_n)_{n\leq N}$ *and any* Q_1, Q *one has*

$$\sum_{Q_1\leq q\leq Q}\frac{1}{\varphi(q)}\sideset{}{^\star}\sum_{\chi\,(\mathrm{mod}\,q)}|\sum_{\substack{m\leq M\\ n\leq N}}\sum\alpha_m\beta_n\chi(mn)|$$

$$\ll \log Q(Q+M^{1/2}+N^{1/2}+\frac{(MN)^{1/2}}{Q_1})\|\alpha\|_2\|\beta\|_2.$$

REMARK 3.2. Observe that for $Q_1>1$ this bound is useless (with respect to the additional summation condition $q\geq Q_1$) if $N=1$ because then $M^{1/2}\geq (MN)^{1/2}/Q_1$. What we will show is that the von Mangolt function $(\Lambda(n))_{n\leq x}$ can be decomposed, up to admissible terms into a sum of functions of non-trivial convolution $(\alpha_m)_{m\leq M}\star(\beta_n)_{n\leq N}$ for $MN\sim x$ and $M,N>1$ so that one can apply Corollary 3.2.

PROOF. We decompose the q-sum into $O(\log Q)$ terms over dyadic intervals as above

$$\sum_{Q_1<q\leq Q}\frac{1}{\varphi(q)}\cdots\leq\sum_{Q_1\leq Q'\leq Q}\frac{1}{Q'}\sum_{q\sim Q'}\frac{q}{\varphi(q)}\cdots$$

and use the factorization

$$|\sum_{\substack{m\leq M\\ n\leq N}}\sum\alpha_m\beta_n\chi(mn)|=|\sum_{m\leq M}\alpha_m\chi(m)||\sum_{n\leq N}\beta_n\chi(n)|;$$

by Cauchy-Schwarz

$$\sum_{q\sim Q'}\frac{q}{\varphi(q)}\sum_{\chi\,(\mathrm{mod}\,q)}|\sum_m\cdots||\sum_n\cdots|\ll(M^{1/2}+Q')(N^{1/2}+Q')\|\alpha\|_2\|\beta\|_2$$

and we conclude with the bound

$$\sum_{Q_1\leq Q'\leq Q}\frac{1}{Q'}(M^{1/2}+Q')(N^{1/2}+Q')\ll\log Q(Q+M^{1/2}+N^{1/2}+\frac{(MN)^{1/2}}{Q_1}).$$

□

2.3. Proof of theorem 3.2. We will reduce the proof of this inequality involving multiplicative characters modulo q to an analoguous one involving additive character modulo q: for $\chi \mod q$ is primitive we have

$$\chi(n)=\frac{1}{\tau_{\overline{\chi}}}\sum_{a\,(\mathrm{mod}\,q)}\chi(a)e_q(na)$$

and therefore

$$\sum_{q\leq Q}\frac{q}{\varphi(q)}\sum_{\substack{\chi\,(\mathrm{mod}\,q)\\ \textit{primitive}}}|\sum_{n\leq N}\alpha_n\chi(n)|^2$$

$$= \sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\substack{\chi \,(\mathrm{mod}\, q)\\ primitive}} |\sum_{a \,(\mathrm{mod}\, q)} \sum_{n\leq N} \alpha_n \chi(a) e(\frac{an}{q})|^2$$

$$\leq \sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} |\sum_{a \,(\mathrm{mod}\, q)} \sum_{n\leq N} \alpha_n \chi(a) e(\frac{an}{q})|^2$$

$$= \sum_{q\leq Q} \frac{1}{\varphi(q)} \sum_{\chi \,(\mathrm{mod}\, q)} \sum_{a,a' \,(\mathrm{mod}\, q)} \chi(a)\overline{\chi}(a') \sum_{n,n'} \alpha_n \overline{\alpha_{n'}} e_q(an - a'n')$$

We have

$$\sum_{\chi \,(\mathrm{mod}\, q)} \sum_{a,a' \,(\mathrm{mod}\, q)} \chi(a)\overline{\chi}(a') = \varphi(q) \delta_{(aa',q)=1} \delta_{a=a'}$$

and therefore the above sum equals

$$= \sum_{q\leq Q} \sum_{\substack{a \,(\mathrm{mod}\, q)\\ (a,q)=1}} \sum_{n,n'} \alpha_n \overline{\alpha_{n'}} e_q(a(n-n'))$$

$$= \sum_{q\leq Q} \sum_{\substack{a \,(\mathrm{mod}\, q)\\ (a,q)=1}} |\sum_{n} \alpha_n e(\frac{an}{q})|^2$$

To conclude it will suffice to prove that

THEOREM 3.3. *We have for any* $(\alpha_n)_{n\leq N} \in \mathbb{C}^N$

$$\sum_{q\leq Q} \sideset{}{^\star}\sum_{a \,(\mathrm{mod}\, q)} |\sum_{n\leq N} \alpha_n e(\frac{an}{q})|^2 \ll (Q^2 + N) \sum_{n\leq N} |\alpha_n|^2 \tag{3.3}$$

Here $\sum^\star$ *mean that we average over the congruence classes* $a \,(\mathrm{mod}\, q)$ *which are coprime to* q.

□

2.4. The duality principle. Let $\mathcal{M}, \mathcal{N}$ be two finite sets and consider a matrix

$$\Phi := (\Phi(m,n))_{(m,n)\in\mathcal{M}\times\mathcal{N}} \in \mathbb{C}^{\mathcal{M}\times\mathcal{N}}.$$

this matrix defines a linear map

$$\Phi: \ \alpha = (\alpha_m)_{m\in\mathcal{M}} \in \mathbb{C}^{\mathcal{M}} \mapsto \beta = (\beta_n)_{n\in\mathcal{N}} = \Phi(\alpha) \in \mathbb{C}^{\mathcal{N}},$$

where

$$\beta_n = \sum_{m\in\mathcal{M}} \alpha_m \Phi(m,n).$$

Equipping $\mathbb{C}^{\mathcal{M}}$ and $\mathbb{C}^{\mathcal{N}}$ with their usual structure of Hilbert spaces

$$\|\alpha\|_2 = (\sum_{m\in\mathcal{M}} |\alpha_m|^2)^{1/2}, \ \|\beta\|_2 = (\sum_{n\in\mathcal{N}} |\beta_n|^2)^{1/2}$$

we have for any vector $(\alpha_m) \in \mathbb{C}^{\mathcal{M}}$

$$\|\Phi(\alpha)\|_2^2 \leq \|\Phi\|_2^2 \|\alpha\|_2^2$$

where $\|\Phi\|_2$ denote the operator norm of Φ: ie.

$$\|\Phi\|_2 = \sup_{\alpha\neq 0} \frac{\|\Phi(\alpha)\|_2}{\|\alpha\|_2} < \infty.$$

In other terms for any $\alpha \in \mathbb{C}^{\mathcal{M}}$ we have

$$\sum_{n\in\mathcal{N}} |\sum_m \alpha_m \Phi(m,n)|^2 \leq \|\Phi\|_2^2 \sum_{m\in\mathcal{M}} |\alpha_m|^2.$$

Let Φ^* be the transpose matrix

$$\Phi* := (\Phi(m,n))_{(n,m)\in\mathcal{N}\times\mathcal{M}} \in \mathbb{C}^{\mathcal{N}\times\mathcal{M}},$$

this matrix defines the transpose linear map

$$\Phi^* : \ \beta = (\beta_n)_{n\in\mathcal{N}} \in \mathbb{C}^{\mathcal{N}} \mapsto \alpha = (\alpha_m)_{m\in\mathcal{M}} = \Phi^*(\beta) \in \mathbb{C}^{\mathcal{M}},$$

where

$$\alpha_m = \sum_{n\in\mathcal{N}} \Phi(m,n)\beta_n.$$

The duality principle is the well known statement

THEOREM (Duality principle). *One has*

$$\|\Phi^*\|_2 = \|\Phi\|_2.$$

In other terms for any $\beta \in \mathbb{C}^{\mathcal{N}}$, one has

$$\sum_{m\in\mathcal{M}} |\sum_n \beta_n \Phi(m,n)|^2 \leq \|\Phi^*\|_2^2 \sum_{n\in\mathcal{N}} |\beta_n|^2 = \|\Phi\|_2^2 \sum_{n\in\mathcal{N}} |\beta_n|^2.$$

PROOF. We have

$$\|\Phi^*(\beta)\|_2^2 = \sum_{m\in\mathcal{M}} |\sum_n \beta_n \Phi(m,n)|^2 = \sum_m \sum_{n,n'} \beta_n \overline{\beta}_{n'} \Phi(m,n)\overline{\Phi(m,n')}$$

$$= \sum_n \beta_n \sum_m \alpha_m \Phi(m,n), \ \alpha_m = \sum_{n'} \overline{\beta}_{n'} \overline{\Phi(m,n')}.$$

By Cauchy-Schwarz this is bounded by

$$\|\beta\|_2 (\sum_n |\sum_m \alpha_m \Phi(m,n)|^2)^{1/2} = \|\beta\|_2 \|\Phi(\alpha)\|_2 \leq \|\beta\|_2 \|\Phi\|_2 \|\alpha\|_2$$

but

$$\|\alpha\|_2^2 = \sum_m |\sum_{n'} \overline{\beta}_{n'} \overline{\Phi(m,n')}|^2 = \sum_m |\sum_n \beta_n \Phi(m,n)|^2 = \|\Phi^*(\beta)\|_2^2$$

and therefore

$$\|\Phi^*(\beta)\|_2^2 \leq \|\Phi\|_2 \|\beta\|_2 \|\Phi^*(\beta)\|_2$$

and hence for any β,

$$\|\Phi^*(\beta)\|_2 \leq \|\Phi\|_2 \|\beta\|_2$$

or in other terms

$$\|\Phi^*\|_2 \leq \|\Phi\|_2;$$

the equality follows by symetry. $\square$

2.5. The additive large sieve inequality. To prove theorem 3.3, we apply the duality principle to the following situation:

$$\mathcal{M} = \mathcal{Q} = \{(a,q),\ q \leq Q,\ (a,q)=1\},\ \mathcal{N} = \{1,\cdots,N\}$$

and

$$\Phi((a,q),n) = e(\frac{an}{q}).$$

Theorem 3.3 states precisely that

$$\|\Phi^*\|_2^2 \ll N + Q^2.$$

By the duality principle this is equivalent to showing that

$$\|\Phi\|_2^2 \ll N + Q^2,$$

or in other terms, that for any $\alpha = (\alpha_{(a,q)})_{(a,q)\in\mathcal{Q}}$, one has

$$\sum_{n\leq N} |\sum_{q\leq Q} \sideset{}{^\star}\sum_{a \,(\mathrm{mod}\, q)} \alpha_{(a,q)} e(\frac{an}{q})|^2 \ll (N+Q^2)\|\alpha\|_2^2.$$

We will evaluate this last sum by computing the square and performing the n-summation; however before doing this we perform a *smoothing trick*: Let φ be a smooth, even, compactly supported function, and taking value 1 on $[-1,1]$. We have

$$\sum_{n\leq N} |\sum_{q\leq Q} \sideset{}{^\star}\sum_{a \,(\mathrm{mod}\, q)} \alpha_{(a,q)} e(\frac{an}{q})|^2 \leq \sum_{n\in\mathbb{Z}} \varphi(\frac{n}{N}) |\sum_{q\leq Q} \sideset{}{^\star}\sum_{a \,(\mathrm{mod}\, q)} \alpha_{(a,q)} e(\frac{an}{q})|^2$$

$$= \sum_{q,q'\leq Q} \sideset{}{^\star}\sum_{\substack{a \,(\mathrm{mod}\, q) \\ a' \,(\mathrm{mod}\, q)'}} \alpha_{(a,q)} \overline{\alpha_{(a',q')}} \sum_n \varphi(\frac{n}{N}) e((\frac{a}{q} - \frac{a'}{q'})n). \tag{3.4}$$

By Poisson's formula the n-sum equals

$$N \sum_{n\in\mathbb{Z}} \widehat{\varphi}(N(n + \frac{a}{q} - \frac{a'}{q'}))$$

Observe that by construction the function

$$x \mapsto \widehat{\varphi}_{N,\mathbb{Z}}(x) := \sum_{n\in\mathbb{Z}} \widehat{\varphi}(N(n+x))$$

is periodic of period 1 and therefor defines a smooth function on the additive group $\mathbb{R}/\mathbb{Z} \simeq S^1$. This implies that

$$\varphi_{N,\mathbb{Z}}(x) = \varphi_{N,\mathbb{Z}}(\pm\|x\|) = \varphi_{N,\mathbb{Z}}(\|x\|)$$

where $\|x\| = \inf_{n\in\mathbb{Z}} |x-n|$ denote the distance between x and the nearest integer: indeed either $+\|x\|$ or $-\|x\|$ is a representative of the class $x \,(\mathrm{mod}\, 1)$ in $\mathbb{R}/\mathbb{Z}$ and φ being even, $\widehat{\varphi}$ is also even. Moreover since φ is compactly supported and smooth, its Fourier transform is rapidly decreasing and in particular

$$\widehat{\varphi}(x) \ll \frac{1}{1+|x|^2}.$$

From we we deduce that

$$\varphi_{N,\mathbb{Z}}(x) \ll \frac{1}{1+(N\|x\|)^2}.$$

Using this bound and the trivial bound

$$\alpha_{(a,q)}\overline{\alpha_{(a',q')}} \leq |\alpha_{(a,q)}|^2 + |\alpha_{(a',q')}|^2$$

we obtain that (3.4) is bounded by

$$\ll \sum_{(a,q)} |\alpha_{(a,q)}|^2 \sum_{(a',q')} \frac{N}{1+(N\|\frac{a}{q}-\frac{a'}{q'}\|)^2}.$$

Observe that when $(a,q) \neq (a',q')$ the rational fractions a/q and a'/q' are distinct modulo 1 and we have for any $n \in \mathbb{Z}$

$$|\frac{a}{q}-\frac{a'}{q'}-n| = |\frac{(a-n)q'-a'q}{qq'}| \geq \frac{1}{qq'} \geq \frac{1}{Q^2}.$$

Therefore

$$\|\frac{a}{q}-\frac{a'}{q'}\| \geq \frac{1}{Q^2}$$

and for any other $(a'',q'') \neq (a',q')$ one hasv (the triangle inequality for the distance function $\|.\|$ on $\mathbb{R}/\mathbb{Z}$)

$$|\|\frac{a}{q}-\frac{a'}{q'}\| - \|\frac{a}{q}-\frac{a''}{q''} \geq \|\frac{a'}{q'}-\frac{a''}{q''}\| \geq \frac{1}{Q^2}.$$

Thereofore for any given (a,q) any interval in $\mathbb{R}$ of the shape $[kQ^{-2},(k+1)Q^{-2}[$, $k \in \mathbb{Z}$, contains at most one number of the shape $\|\frac{a}{q}-\frac{a'}{q'}\|$. It follows that

$$\sum_{(a',q')\neq(a,q)} \frac{N}{1+(N\|\frac{a}{q}-\frac{a'}{q'}\|)^2} \leq \sum_{k\geq 0} \frac{N}{1+(kNQ^2)^2} \ll N+Q^2.$$

Therefore we have proved that

$$\sum_{n\leq N} |\sum_{q\leq Q} \sum_{a \pmod q}^{\star} \alpha_{(a,q)} e(\frac{an}{q})|^2 \ll (N+Q^2)\|\alpha\|_2^2.$$

$\square$

3. Heath-Brown's identity

In order to apply Corollary 3.2, we need to show that the vonMangolt function Λ can be decomposed into a sum of airthmetic function which convolutions. We effectuate this using an identity due to Heath-Brown but there are many other possibilities (for instance Vaughan's identify).

THEOREM 3.4 (Heath-Brown's identity). *Let $J \geq 1$ an integer and $X > 1$, one has for any $n < 2X$*

$$\Lambda(n) = -\sum_{j=1}^{J}(-1)^j \binom{J}{j} \sum_{m_1,\cdots,m_j\leq Z} \mu(m_1)\cdots\mu(m_j) \sum_{m_1\cdots m_j n_1\cdots n_j=n} \log n_1,$$

where $Z = X^{1/J}$.

PROOF. This identity is an immediate consequence of the following identity for Dirichlet series: let

$$M_Z(s) = \sum_{n \leq Z} \frac{\mu(n)}{n^s}$$

be the truncation of the inverse of Riemann's zeta function

$$M(s) = \zeta(s)^{-1} = \sum_{n \geq 1} \frac{\mu(n)}{n^s}.$$

In particular (since $\zeta(s)\zeta(s)^{-1} = 1$ or equivalently

$$\sum_{d|n} \mu(d) = \delta_{n=1} \;)$$

one has

$$\zeta(s)M_Z(s) = 1 + \sum_{n>Z} \frac{a_Z(n)}{n^s};$$

in other terms the convolution of 1 with the function $\mu.1_{n \leq Z}$ takes value 0 between 2 and Z. It follows that for $J \geq 1$ the coefficients $b_{Z,J}(n)$ of the Dirichlet series $(1 - \zeta(s)M_Z(s))^J$ are zero for $n \leq Z^J = X$ and therefore given any Dirichlet series

$$L(s) = \sum_{n \geq 1} \frac{a(n)}{n^s}$$

associated to some arithmetic function $(a(n))_{n \geq 1}$ one has

$$L(s)(1 - \zeta(s)M_Z(s))^J = \sum_{n > Z^J} \frac{a * b_{Z,J}(n)}{n^s}.$$

We apply this observation to $L(s) = \frac{\zeta'(s)}{\zeta(s)}$. By the binomial law, we have

$$\frac{\zeta'(s)}{\zeta(s)}(1 - \zeta(s)M_Z(s))^J = \frac{\zeta'(s)}{\zeta(s)} + \sum_{j=1}^{J} (-1)^j \binom{J}{j} \zeta'(s)\zeta^{k-1}(s)M_Z^k(s).$$

this gives Heath-Brown's identity for $n < Z^J$ but we observe that since $\Lambda(1) = 1$, the coefficient of the Dirichlet series on the lefthand side are in fact zero for all $n < 2Z^J$. □

4. Proof of the Bombieri-Vinogradov theorem

The proof we present here is a bit of an overkill; for instance one can find in Kowalski-Iwaniec a very sleek and quite a bit shorter proof of the Bombieri-Vinogradov theorem. The purpose of this exposition is to propose alternative presentations which maybe useful in other contexts.

4.1. Exponent of distribution of arithmetic functions. Heath-Brown's identity states that on the interval $[1, 2x[$ the von Mangolt function $\Lambda(n)$ can be decomposed in a linear combination of functions of the shape

$$(1_{\leq Z}\mu)^{(\star j)} \star \log \star 1^{(\star j-1)},\ j = 1, \cdots, J, Z = x^{1/J}. \tag{3.5}$$

It is therefore sufficient to prove that any of the functions γ above one has

$$\sum_{q \leq Q} \max_{(a,q)=1} E(\gamma, x; q) \ll \frac{x}{\log^A x}$$

where

$$E(\gamma, x; q) = \max_{(a,q)=1} E(\gamma, x; q, a)$$

and

$$E(\gamma, x; q, a) = |\sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, q)}} \gamma(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leq x \\ (n,q)=1}} \gamma(n)|.$$

In is therefore worthwhile this problem (ie. estimating the quality of the distribution of γ in arithmetic progressions on average) for general arithmetic functions γ.

The case of arithmetic functions which are essentially bounded: functions γ for which there exists $K \geq 0$ such that for any $n \geq 1$

$$|\gamma(n)| \ll ((1 + \log n)d(n))^K \tag{3.6}$$

We have therefore the following trivial bounds: for $q \leq Q \leq x$

$$E(\gamma, x; q) \ll \frac{x(\log x)^{O(1)}}{\varphi(q)}$$

and

$$\sum_{q \leq Q} E(\gamma, x; q) \ll x(\log x)^{O(1)}.$$

Definition 3.1. *Given $\Delta \in [0, 1]$, an arithmetic function satisfying (3.6) has level of distribution $\geq \Delta$ if, for any $A \geq 0$, there exists $B = B(A)$ such that for $Q \leq x^{\Delta}/\log^B x$, one has*

$$\sum_{q \leq Q} E(\gamma, x; q) \ll_{K,A} \frac{x}{\log^A x}.$$

With this terminology we have

Theorem 3.1 (Bombieri-Vinogradov). *The von Mangolt function Λ has level of distribution $\geq 1/2$.*

The following simple result will be useful in the proof of the Bombieri-Vinogradov theorem:

Lemma 3.1. *Let P be a polynomial, the function $n \mapsto P(\log n)$ has level of distribution 1.*

PROOF. Il is sufficient to prove this for $n \mapsto \log^k n$ which is continuous monotone, therefore for $q \leq x$

$$\sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, q)}} \log^k(n) = \int_0^{\frac{x-a}{q}} \log^k(qt+a)dt + O(\log^k x) = \frac{1}{q}\int_a^x \log^k(t)dt + O(\log^k x)$$

and therefore for $Q \leq x$

$$\sum_{q \leq Q} E(\log^k, x; q) \ll \sum_{q \leq Q} \log^k x \ll Q \log^k x \ll \frac{x}{\log^A x}$$

as long as $Q \leq x^{1/2}/\log^B x$ with $B \geq k + A$. □

4.2. A Bombieri-Vinogradov theorem for factorable arithmetic functions. We will discuss now the problem of evaluating the exponent of distribution of essentially bounded arithmetic functions which admit factorizations $\gamma = \alpha \star \beta$ as a convolution of arithmetic functions (the idea then will be to use Corollary 3.2). By Heath-Brown identity this is essentially the case of the von Mangolt function which a linear combination of such functions.

For $k \geq 2$, let γ be an arithmetic function of the shape

$$\gamma(n) = \alpha_1 \star \cdots \star \alpha_k(n) = \sum_{n_1 \cdots n_k = n} \alpha_1(n_1) \cdots \alpha_k(n_k)$$

where α_i are arithmetic functions satisfying (3.6); therefore γ also satisfies (3.6).

We will give general sufficient conditions to insure that γ has level of distribution $\geq 1/2$.

4.2.1. *From hyperboloids to paralleloids.* For this we will need to make first a technical reduction: writing γ a a convolution, we need to evaluate sums of the shape

$$\sum_{\substack{n_1 . \cdots . n_k \leq x \\ \cdots}} \cdots$$

that is sums over integral point lying under the hyperboloid given by the equation

$$x_1 . \cdots . x_k = x.$$

In view of Corollary 3.2, we would rather evaluate sums of the shape

$$\sum_{\substack{n_1 \leq N_1, \cdots, n_k \leq N_k \\ \cdots}} \cdots, \text{ with } N_1 . \cdots . N_k \leq x$$

(this is sums over integral points contained in a paralleloid). We do this by approximating the region located under the hyperboloid by a union of sufficiently small paralleloids and evaluate the error made because of this approximation.

Given $\delta < 1$ such that $\delta^{-1} = \log^C x$ for some fixed constant $C \geq 1$ (to be choosen depending on A), we cover the cube $[1, x]^k$ by $O((\delta^{-1} \log x)^k) = \log^{O(C)} x$ paralleloids of the shape

$$\prod_{i=1}^{k}]N_i, N_i(1+\delta)]$$

where the N_i are powers of $1 + \delta$ and restricting to each paralleloid, we bound the sum

$$E(\gamma, x; q, a) = |\sum_{\substack{n \leq x \\ n \equiv a \,(\mathrm{mod}\, q)}} \gamma(n) - \frac{1}{\varphi(q)} \sum_{\substack{n \leq x \\ (n,q)=1}} \gamma(n)|$$

by a sum of $O((\delta^{-1} \log x)^k)$ sums of the shape

$$E(\gamma_{\mathbf{N}}, x; q) =$$
$$\max_{(a,q)=1} |\sum_{\substack{n_1, \cdots, n_k \\ n_1 \cdots n_k \equiv a \,(\mathrm{mod}\, q)}} \alpha_1(n_1) \cdots \alpha_k(n_k) - \frac{1}{\varphi(q)} \sum_{\substack{n_1, \cdots, n_k \\ (n_1 \cdots n_k, q)=1}} \alpha_1(n_1) \cdots \alpha_k(n_k)|$$

where $\mathbf{N}$ runs over $O((\delta^{-1} \log x)^k)$ k-tuples of the shape

$$\mathbf{N} = (N_1, \cdots, N_k), \ 1 \leq N_1 \cdots N_k \leq x$$

and the n_i are subject to the constraints

$$n_i \in]N_i, N_i(1+\delta)], \ i = 1 \cdots k$$

and the additional constraint

$$n_1 \cdots n_k \leq x. \tag{3.7}$$

Observe that in the sum $E(\gamma_{\mathbf{N}}, x; q)$ (3.7) is unnecessary if

$$(1+\delta)^k \prod_i N_i \leq x, \tag{3.8}$$

in such a case we will then write $E(\gamma_{\mathbf{N}}; q)$ instead of $E(\gamma_{\mathbf{N}}, x; q)$. For all the other terms, the corresponding $n = n_1 \cdots n_k$ satisfy

$$x(1+\delta)^{-k} \leq n \leq x$$

and the contribution of these terms to

$$\sum_{q \leq Q} E(\gamma_{\mathbf{N}}, x; q)$$

is bounded by

$$(\log x)^{O(1)} \sum_{q\leq Q} [\sum_{\substack{x(1+\delta)^{-k}\leq n\leq x\\ n\equiv a \,(\mathrm{mod}\, q)}} d(n)^K + \frac{1}{\varphi(q)} \sum_{\substack{x(1+\delta)^{-k}\leq n\leq x\\ (n,q)=1}} d(n)^K]$$

$$\ll \sum_{q\leq Q} \frac{1}{\varphi(q)} \delta x (\log x)^{O(1)} \ll \delta x \log^{O(1)} x = x \log^{O(1)-C} x.$$

Up to an error term of size $x\log^{O(1)-C} x$ (less that $x/\log^A x$ for C large enough) we are reducted to evaluate

$$\sum_{q\leq Q} E(\gamma_{\mathbf{N}}; q)$$

where

$$\gamma_{\mathbf{N}} = \alpha_1.1_{]N_1,N_1(1+\delta)]} \star \cdots \star \alpha_k.1_{]N_k,N_k(1+\delta)]}$$

and $\mathbf{N} = (N_1, \cdots, N_k)$ satisfying (3.8). We may in fact assume that

$$x^{1-1/100} \leq \prod_i N_i \ll_k x; \tag{3.9}$$

indeed the total contribution of terms with $\prod_i N_i \leq x^{99/100}$ is bounded trivially by

$$(\log x)^{O(1)} x^{99/100}.$$

We proceed as in the beginning of this chapter expressing the congruence condition $n \equiv a \,(\mathrm{mod}\, q)$ in terms of characters and then in terms of non-trivial primitive characters of moduli $q \leq Q$: writing $\chi^* \,(\mathrm{mod}\, q^*)$ for the primitive character underlying $\chi \,(\mathrm{mod}\, q)$ we have

$$\sum_{q\leq Q} E(\gamma_{\mathbf{N}}; q) \leq \sum_{q'\leq Q} \frac{1}{\varphi(q')} \sum_{1<q^*\leq Q/q'} \frac{1}{\varphi(q^*)} \sideset{}{^\star}\sum_{\chi^* \,(\mathrm{mod}\, q^*)} |\sum_{(n,q')=1} \gamma_{\mathbf{N}}(n)\chi^*(n)|$$

$$= \sum_{q'\leq Q} \frac{1}{\varphi(q')} \sum_{1<q\leq Q} \frac{1}{\varphi(q)} \sideset{}{^\star}\sum_{\chi \,(\mathrm{mod}\, q)} |\sum_{(n,q')=1} \gamma_{\mathbf{N}}(n)\chi(n)|,$$

upon changing notations. We separate the case of small and large moduli and are reduced to bound (here $Q_1 = (\log x)^D$ with D to be choosen large enough):

$$\sum_{1<q\leq Q_1} \frac{1}{\varphi(q)} \sideset{}{^\star}\sum_{\chi \,(\mathrm{mod}\, q)} |\sum_{(n,q')=1} \chi(n)\gamma_{\mathbf{N}}(n)|$$

and

$$\sum_{Q_1<q\leq Q} \frac{1}{\varphi(q)} \sideset{}{^\star}\sum_{\chi \,(\mathrm{mod}\, q)} |\sum_{(n,q')=1} \chi(n)\gamma_{\mathbf{N}}(n)|.$$

To each partition of the set $\{1, \cdots, k\} = I \sqcup I'$ into two subsets we associate a factorisation

$$\gamma_{\mathbf{N}} = \alpha \star \beta = \star_{i\in I}\alpha_i \star (\star_{i'\in I'}\alpha_{i'})$$

were α, β are supported on the intervals

$$[N_I, (1+O(\delta))N_I],\ [N_{I'}, (1+O(\delta))N_{I'}]$$

with

$$N_I = \prod_{i\in I} N_i,\ N_{I'} = \prod_{i'\in I'} N_{i'}$$

We have

$$\sum_{(n,q')=1} \chi(n)\gamma_{\mathbf{N}}(n) = \sum_{\substack{m\leq N_I\\ (m,q')=1}} \alpha(m)\chi(m) \sum_{\substack{n\leq N_{I'}\\ (n,q')=1}} \beta(n)\chi(n).$$

We may and will use different partitions depending on which method we use.

4.2.2. *Small moduli and the Siegel-Walfisz hypothesis.* To deal with small moduli, we make the following assumption on β:

Hypothesis 3.1 (Siegel-Walfisz type bound). There exist an absolute constant E such that given $q, q' \geq 1$ and $\chi \,(\mathrm{mod}\, q)$ non-trivial and primitive, one has for any $F \geq 1$, and any $y \geq 2$

$$\sum_{\substack{n\leq y\\ (n,q')=1}} \beta(n)\chi(n) \ll_A (d(q')q)^E \frac{y}{\log^F y}$$

Under this assumption we have for any $F \geq 1$

$$\sum_{1<q\leq Q_1} \frac{1}{\varphi(q)} \sideset{}{^\star}\sum_{\chi \,(\mathrm{mod}\, q)} |\sum_{n\leq x} \chi(n)\gamma_{\mathbf{N}}(n)| \ll_A (d(q')\log x)^{O(1)} \frac{x}{\log^F N_{I'}}.$$

Since

$$\sum_{q'\leq Q} \frac{d(q')^{O(1)}}{\varphi(q')} = (\log x)^{O(1)},$$

his bound is satisfactory as long as $N_{I'} \geq x^\eta$ for some fixed $\eta > 0$ and F is choosen so that

$$F \geq \eta^{-1}(A + O(1)).$$

In particular, suppose that all the $\alpha_i,\ 1 = 1, \cdots, k$ satisfy Hypothesis 3.1; taking $\beta = \alpha_i$ with N_i is maximal we have therefore $N_i \geq x^{\frac{99}{100}\frac{1}{k}}$ so that the above reasonning is valid with $\eta = \frac{99}{100}\frac{1}{k}$. Since we need to bound at most $O(\log^{O(1)})$ such sums, under this assumption, we obtain upon taking that the contribution of the moduli $q \leq Q_1$ is $\ll x/\log^A x$.

4.2.3. *Large moduli and the large sieve inequality.* For large moduli we apply Corollary 3.2 getting

$$\sum_{Q_1<q\leq Q} \frac{1}{\varphi(q)} \cdots \ll \log x(\frac{x^{1/2}}{Q_1} + N_I^{1/2} + N_{I'}^{1/2} + Q)\|\alpha\|_2\|\beta\|_2$$

with

$$\|\alpha\|_2 = (\sum_{n\ll N_I} |\alpha(n)|)^{1/2} \ll N_I^{1/2}\log^k x,$$

$$\|\beta\|_2 = (\sum_{n \ll N_{I'}} |\beta(n)|)^{1/2} \ll N_{I'}^{1/2} \log^k x$$

and therefore

$$\sum_{Q_1 < q \leq Q} \frac{1}{\varphi(q)} \cdots \ll x^{1/2}(\max(N_I, N_{I'})^{1/2} + \frac{x^{1/2}}{\log^D x} + \frac{x^{1/2}}{\log^B x}) \log^{1+2k} x$$

Let us recall that the number of $\mathbf{N}$ is bounded by $O(\log^{O(1)})$ therefore given $A \geq 1$ we will choose

$$B, D \geq A + O(1)$$

we declare a tuple $\mathbf{N}$ "good" if there is a factorisation $\gamma_{\mathbf{N}} = \alpha \star \beta$ such that

$$\max(N_I, N_{I'}) \leq \frac{x}{\log^{2(A+O(1))}}$$

and we obtain that

$$\sum_{\mathbf{N} \text{ good}} \sum_{Q_1 \leq q \leq Q} \frac{1}{\varphi(q)} \sum_{\chi \, (\mathrm{mod}\, q)}^{\star} |\sum_{n \leq x} \chi(n) \gamma_{\mathbf{N}}(n)| \ll \frac{x}{\log^A x}.$$

Suppose that $\mathbf{N}$ is not good, and let i be such that N_i is maximal in the set $\{N_j, \ j = 1, \cdots, k\}$ (to fix ideas, let us assume that $i = 1$); since $\mathbf{N}$ is not good and N_1 is maximal, one has necessarily

$$N_1 \geq \frac{x}{\log^{2(A+O(1))}}$$

and therefore

$$\prod_{i' \neq 1} N_{i'} \ll \log^{2(A+O(1))} x.$$

Setting $\beta = \star_{i' \neq 1} \alpha_i'$ and returning to the initial problem we have to bound

$$\begin{aligned}
\sum_{q \leq Q} E(\gamma_{\mathbf{N}}, x; q) &\leq \sum_{q \leq Q} \max_{(a,q)=1} \sum_{\substack{n' \ll \log^{O(1)} x \\ (n',1)=1}} |\beta(n')| \times E(\alpha_1, [N_1, (1+\delta)N_1]; q, a\overline{n'}) \\
&\ll (\log x)^{O(1)} \sum_{q \leq Q} E(\alpha_1, [N_1, (1+\delta)N_1]; q) \\
&\ll (\log x)^{O(1)} \sum_{q \leq Q} E(\alpha_1, (1+\delta)N_1; q) \\
&\quad + (\log x)^{O(1)} \sum_{q \leq Q} E(\alpha_1, N_1; q)
\end{aligned}$$

The latter sum is admissible if we show that the arithmetic function α_1 has level of distribution $\geq 1/2$.

5. Application to the Λ-function

We use Heath-Brown identity with $J = 2$ and are reduced to proving that the convolutions

$$\gamma = (1_{\leq Z}\mu)^{(\star j)} \star \log \star 1^{(\star j-1)},\ j = 1, 2,\ k = 2, 4,\ Z = x^{1/2}$$

have level of distribution $\geq 1/2$.

The following Proposition is left to the reader as an exercise:

Proposition 3.1. *The Moebius function satisfies a Siegel-Walfisz type bound: there exist $E \geq 0$ such that for any $y \geq 1$, $F \geq 1$, any $q, q' \geq 1$ and $\chi \,(\mathrm{mod}\, q)$ primitive non-trivial, one has*

$$\sum_{n \leq y} \beta(n)\chi(n) \ll_A (d(q')q)^E \frac{y}{\log^F y}.$$

It is also obvious that the constant function 1 and log satisfy a Siegel-Walfisz type bound, therefore we may apply the previous argument and it remains to bound the sum

$$\sum_{\mathbf{N} \text{ not good}} \sum_{q \leq Q} E(\gamma_{\mathbf{N}}, x; q).$$

With the notation of the previous section we have that

$$\alpha_1 \text{ is either } \alpha_1 = 1_{\leq Z}\mu \text{ or } \alpha_1 = \log \text{ or } \alpha_1 = 1$$

but since $Z = x^{1/2}$ and α_1 is supported around $x(\log x)^{O(1)}$ the first case is not possible and therefore $\alpha_1 = 1$ or log. Since these have level of distribution 1 we are done. More precisely, we have seen in the proof of Lemma 3.1 that for $Q \leq X$

$$\sum_{q \leq Q} E(\alpha_1, X; q) \ll \sum_{q \leq Q} \log X \ll Q \log X.$$

In particular for $X = x(\log x)^{O(1)}$ and $Q \leq x^{1/2} \log^{-B} x$ the later term is of size $x^{1/2}(\log x)^{O(1)} \ll x/\log^A x$ for any $A \geq 1$.