A Detailed Exposition of a Proof of Hua's Lemma, following Bob Vaughan

Joe Clark

April 2016

1 Notation

I will be following Bob Vaughan's use of notation in this proof.

Let n be a sufficiently large integer, and let $N = \lfloor n^{1/k} \rfloor$.

Let k denote a natural number (usually $k \geq 2$). All statements with ϵ are true for every positive real ϵ.

The Vinogradov symbols $\ll$ and $\gg$ are used standardly: Given functions f and g (where g takes non-negative real fvalues), $f \ll g$ means $|f| \leq Cg$, where C is a constant. If f is also non-negative, then $f \gg g$ means $g \ll f$.The Vinogradov symbols may have implicit dependance on k and ϵ.

Given a function ϕ of a real variable α, iteratively define

$$\Delta_1(\phi(\alpha); \beta) = \phi(\alpha + \beta) - \phi(\alpha),$$

$$\Delta_{j+1}(\phi(\alpha); \beta_1, \ldots, \beta_{j+1}) = \Delta_1(\Delta_j(\phi(\alpha); \beta_1, \ldots, \beta_j); \beta_{j+1}).$$

Finally,

$$f(\alpha) := \sum_{m=1}^{N} e^{2\pi i m^k \alpha} \tag{1.1}$$

2 A Useful Fact from Number Theory

Let $d(n)$ denote the number of positive divisors of n for any natural number n. If n has prime factorization $n = p_1^{a_1} \dots p_k^{a_k}$, we know:

$$d(n) = \prod_{i=1}^{k} (a_i + 1)$$

It is worth noting first that the number of p^a satisfying $a + 1 > p^{\epsilon a}$ is finite.

Fix ϵ. Since the exponential function $p^{\epsilon a}$ eventually grows more quickly than the linear function $a + 1$, only finitely many powers of any p will satisfy the inequality. Specifically, as p gets sufficiently large, no power of p will satisfy the inequality. Since ϵ is fixed, $\exists p$ such that $p >$max$\{2^{1/\epsilon}, \epsilon^{1/\epsilon}\}$. By defining $f(x) = p^{\epsilon x}$ and $g(x) = x + 1$, we see that $f(1) = p^{\epsilon} > 2 = g(1)$ and $f'(x) = \epsilon(logp)p^{\epsilon x} > 1 = g'(x)$ for $x \geq 1$. Then, since $f(x) \geq g(x)$ for $x \geq 1$, no power of p satisfies the inequality.

This established, we now wish to prove that $d(n) \ll n^{\epsilon}$ *for every* $\epsilon > 0$.

Proof. Consider the prime factorization of $n = p_1^{a_1} \dots p_k^{a_k}$. Of these p_i, only finitely many satisfy $a + 1 > p_i^{\epsilon a}$. Rename these $q_1^{a_1}, \dots q_l^{a_l}$, and keep the remaining $p_{l+1}^{a_{l+1}}, \dots, p_k^{a_k}$. Now, since the product of the q_is is a constant dependent only on ϵ, say, Q, we have:

$$d(n) \leq Q \prod_{i=1}^{k} (a_i + 1) \leq Q \prod_{i=1}^{k} p^{\epsilon a_i} \leq Q n^{\epsilon}$$

which, as required, gives:

$$d(n) \ll n^{\epsilon} \tag{2.1}$$

3 A Comment on Δ-notation

We have previously defined, for a function ϕ of a real variable α:

$$\Delta_1(\phi(\alpha);\beta) = \phi(\alpha+\beta) - \phi(\alpha), \tag{3.1}$$
$$\Delta_{j+1}(\phi(\alpha);\beta_1,\dots,\beta_{j+1}) = \Delta_1(\Delta_j(\phi(\alpha);\beta_1,\dots,\beta_j);\beta_{j+1}).$$

Let $\phi(\alpha) = \alpha^k$. Then:

$$\begin{aligned}
\Delta_1(\alpha^k;\beta) &= (\alpha+\beta)^k - \alpha^k = \binom{k}{1}\alpha^{k-1}\beta + \dots + \binom{k}{k}\beta^k \\
\Delta_2(\alpha^k;\beta_1,\beta_2) &= \Delta_1(\binom{k}{1}\alpha^{k-1}\beta_1 + \dots + \binom{k}{k}\beta_1^k;\beta_2) \\
&= (\binom{k}{1}(\alpha+\beta_2)^{k-1}\beta_1 + \dots + \binom{k}{k}\beta_1^k) - (\binom{k}{1}\alpha^{k-1}\beta_1 + \dots + \binom{k}{k}\beta_1^k) \\
&= (\binom{k}{1}(\alpha^{k-2}\beta_1\beta_2 + \dots + \beta_1\beta_2^{k-1}) + \dots + \binom{k}{k-1}\beta_1^{k-1}\beta_2
\end{aligned}$$

Then, we can show that $\Delta_j(\alpha^k;\beta_1,\dots,\beta_j) = \beta_1\dots\beta_j p_j(\alpha;\beta_1,\dots,\beta_j)$, where $p_j(\alpha;\beta_1,\dots,\beta_j)$ is a polynomial in α of degree $k-j$, by induction:

Proof. The case j=1 has been demonstrated in (3.1). Suppose that

$$\Delta_{j-1}(\alpha^k;\beta_1,\dots,\beta_{j-1}) = \beta_1\dots\beta_{j-1}p_{j-1}(\alpha;\beta_1,\dots,\beta_{j-1})$$

Then, where c and d represent the appropriate binomial coefficients:

$$\begin{aligned}
\Delta_j(\alpha^k;\beta_1,\dots,\beta_j) =&\Delta_1(\beta_1\dots\beta_{j-1}p_{j-1}(\alpha;\beta_1,\dots,\beta_{j-1});\beta_j) \\
=&\Delta_1(c_{k-j+1}\beta_1\dots\beta_{j-1}\alpha^{k-j+1} + \dots + c_0\beta_1\dots\beta_{j-1};\beta_j) \\
=&(c_{k-j+1}\beta_1\dots\beta_{j-1}(\alpha+\beta_j)^{k-j+1} + \dots + c_0\beta_1\dots\beta_{j-1}) \\
&- (c_{k-j+1}\beta_1\dots\beta_{j-1}\alpha^{k-j+1} + \dots + c_0\beta_1\dots\beta_{j-1}) \\
=&d_{k-j}\beta_1\dots\beta_j\alpha^{k-j} + \dots + d_0\beta_1\dots\beta_j
\end{aligned}$$

which is exactly what was to be shown. Then:

$$\Delta_j(\alpha^k;\beta_1,\dots,\beta_j) = \beta_1\dots\beta_j p_j(\alpha;\beta_1,\dots,\beta_j) \tag{3.2}$$

where $p_j(\alpha;\beta_1,\dots,\beta_j)$ is a polynomial in α of degree $k-j$ with integer-valued coefficients.

4 Proof of Parseval's Identity

We will use a finite version of Parseval's Identity for the purposes of this proof.

Suppose $f : \mathbb{Z} \to \mathbb{C}$ has finite support - that is, $f(x) = 0$ for all x outside of some large interval, and define $\hat{f} : [0, 1) \to \mathbb{C}$ by :

$$\hat{f}(\alpha) = \sum_{x\in\mathbb{Z}} f(x)e^{2\pi i x\alpha}; \hat{g}(\alpha) = \sum_{x\in\mathbb{Z}} g(x)e^{2\pi i x\alpha}$$

Then

$$\int_0^1 \hat{f}(\alpha)\overline{\hat{g}(\alpha)}d\alpha = \sum_{x\in\mathbb{Z}} f(x)\overline{g(x)} \tag{4.1}$$

Proof.

$$\begin{aligned}
\int_0^1 \hat{f}(\alpha)\overline{\hat{g}(\alpha)}d\alpha &= \int_0^1 (\sum_{x\in\mathbb{Z}} f(x)e^{2\pi i x\alpha}\overline{\sum_{y\in\mathbb{Z}} g(y)e^{2\pi i y\alpha}})d\alpha \\
&= \int_0^1 (\sum_{x\in\mathbb{Z}} f(x)e^{2\pi i x\alpha}\sum_{y\in\mathbb{Z}} \overline{g(y)}e^{-2\pi i y\alpha})d\alpha \\
&= \int_0^1 (\sum_{x,y\in\mathbb{Z}} f(x)e^{2\pi i x\alpha}\overline{g(y)}e^{-2\pi i y\alpha})d\alpha \\
&= \int_0^1 (\sum_{x,y\in\mathbb{Z}} f(x)\overline{g(y)}e^{2\pi i(x-y)\alpha}d\alpha \\
&= \sum_{x,y\in\mathbb{Z}} (f(x)\overline{g(y)}) \underbrace{\int_0^1 e^{2\pi i(x-y)\alpha}d\alpha}_{=1 \text{ IFF } x=y, \text{ else } =0} \\
&= \sum_{x\in\mathbb{Z}} f(x)\overline{g(x)}
\end{aligned}$$

And, in particular, if $f(x) = g(x)$, then

$$\int_0^1 |\hat{f}(\alpha)|^2 d\alpha = \sum_{x\in\mathbb{Z}} |f(x)|^2$$

5 Proof of Weyl's Lemma

Let

$$T(\phi) = \sum_{x=1}^{Q} e^{2\pi i \phi(x)}$$

where ϕ *is an arbitrary arithmetical function: that is, a function* $f : \mathbb{N} \to \mathbb{C}$*. Then,*

$$|T(\phi)|^{2^j} \leq (2Q)^{2^j - j - 1} \sum_{|h_1|<Q} \cdots \sum_{|h_j|<Q} T_j \tag{5.1}$$

where

$$T_j = \sum_{x \in I_j} e^{2\pi i \Delta_j(\phi(x); h_1, \ldots, h_j)}$$

and the intervals $I_j = I_j(h_1, \ldots, h_j)$ *(possibly empty) satisfy*

$$I_1(h_1) \subset [1, Q], I_j(h_1, \ldots, h_j) \subset I_{j-1}(h1, \ldots, h_{j-1}).$$

Proof. We will use a proof by induction on j.

When $j = 1$, we wish to show that

$$|T(\phi)|^{2^1} \leq (2Q)^{2^1-1-1} \sum_{h_1 \leq Q} \sum_{x \in I_j} e^{2\pi i \Delta_1(\phi(x);h_1)}$$

That is, that

$$|\sum_{x=1}^{Q} e^{2\pi i \phi(x)}|^2 \leq \sum_{h_1 \leq Q} \sum_{x \in I_1} e^{2\pi i \Delta_1(\phi(x);h_1)}$$

Now, we know that:

$$\begin{aligned}
|\sum_{x=1}^{Q} e^{2\pi i \phi(x)}|^2 &= \sum_{y=1}^{Q} e^{2\pi i \phi(y)} \sum_{x=1}^{Q} e^{-2\pi i \phi(x)} \\
&= \sum_{x,y=1}^{Q} e^{2\pi i(\phi(y)-\phi(x))} \\
&\text{By substituting } y = x + h_1 \text{, we get:} \\
&= \sum_{x=1}^{Q} \sum_{y=1}^{Q} e^{2\pi i(\phi(x+h_1)-\phi(x))} \\
&= \sum_{x=1}^{Q} \sum_{x+h_1=1}^{Q} e^{2\pi i \Delta_1(\phi(x),h_1)} \\
&= \sum_{x=1}^{Q} \sum_{h_1=1-x}^{Q-x} e^{2\pi i \Delta_1(\phi(x),h_1)}
\end{aligned}$$

Since x ranges from 1 to Q, we know that h_1 ranges from $1-Q$ to $Q-1$. Since h_1 ranges from $1-x$ to $Q-x$ we know that x also ranges from $1-h_1$ to $Q-h_1$, so $x \in I_1 = [1, Q] \cap [1-h_1, Q-h_1]$.

Then,

$$|\sum_{x=1}^{Q} e^{2\pi i \phi(x)}|^2 = \sum_{h_1 \leq Q} \sum_{x \in I_1} e^{2\pi i \Delta_1(\phi(x);h_1)}$$

so

$$|\sum_{x=1}^{Q} e^{2\pi i \phi(x)}|^2 \leq \sum_{h_1 \leq Q} \sum_{x \in I_1} e^{2\pi i \Delta_1(\phi(x);h_1)}$$

The base case established, assume the conclusion (5.1) is true for j.

First, note that

$$\begin{aligned}
|T_j|^2 &= |\sum_{x\in I_j} e^{2\pi i \Delta_j(\phi(x);h_1,\ldots,h_j)}|^2 \\
&= \sum_{y\in I_j} e^{2\pi i \Delta_j(\phi(y);h_1,\ldots,h_j)} \sum_{x\in I_j} e^{-2\pi i \Delta_j(\phi(x);h_1,\ldots,h_j)}
\end{aligned}$$

By substituting $y = x + h_{j+1}, |h_{j+1}| < Q$, we get:

$$\begin{aligned}
&= \sum_{|h_{j+1}|<Q} \sum_{x+h_{j+1}\in I_j} \sum_{x\in I_j} e^{2\pi i(\Delta_j(\phi(x+h_{j+1});h_1,\ldots,h_j)-\Delta_j(\phi(x);h_1,\ldots,h_j))} \\
&= \sum_{|h_{j+1}|<Q} \sum_{x\in I_{j+1}} e^{2\pi i \Delta_{j+1}(\phi(x);h_1,\ldots,h_{j+1})} \\
&= T_{j+1}
\end{aligned}$$

where $I_{j+1} = I_j \cap \{x|x+h \in I_j\}$

Now, by squaring both sides of (5.1), we get

$$\begin{aligned}
|T(\phi)|^{2^{j+1}} &\leq ((2Q)^{2^j-j-1})^2 (\sum_{|h_1|<Q} \cdots \sum_{|h_j|<Q} T_j)^2 \\
&\leq (2Q)^{2^{j+1}-2j-2} \sum_{|h_1|<Q} \cdots \sum_{|h_j|<Q} |T_j|^2 (\text{Cauchy-Schwartz*}) \\
&\leq (2Q)^{2^{j+1}-2j-2}(2Q)^j \sum_{|h_1|<Q} \cdots \sum_{|h_j|<Q} |T_j|^2 \\
&= (2Q)^{2^{j+1}-(j+1)-1} \sum_{|h_1|<Q} \cdots \sum_{|h_j|<Q} |T_j|^2 \\
&= (2Q)^{2^{j+1}-(j+1)-1} \sum_{|h_1|<Q} \cdots \sum_{|h_j|<Q} T_{j+1}
\end{aligned}$$

**A well-known formulation of the Cauchy-Schwartz Inequality is:*

$$\sum a_i b_i \leq \sqrt{\sum a_i^2}\sqrt{\sum b_i^2}$$

When both sides are squared, this yields:

$$(\sum a_i b_i)^2 \leq \sum a_i \sum b_i$$

This is the form we use iteratively in this step, taking $a_i = T_j$ *and* $b_i = 1$.

The result is then proved.

6 Proof of Hua's Lemma

Suppose that $1 \leq j \leq k$. Then,

$$\int_0^1 |f(\alpha)|^{2j} d\alpha \ll N^{2^j - j + \epsilon} \tag{6.1}$$

Proof. We will use a proof by induction on j.

6.1 Base Case $j = 1$

First, suppose that $j = 1$. We know by the Fundamental Theorem of Calculus that

$$\int_0^1 e^{2\pi i x^k \alpha} = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases} \tag{6.2}$$

where $x \in \mathbb{Z}$.

The proof of Parseval's Lemma as given works just as well with $e^{2\pi i x^k \alpha}$ as it does with $e^{2\pi i x \alpha}$ (as shown in Section 4), since it is still true that $e^{2\pi i x^k \alpha} = 1$ IFF $x_m = x_n$, else $= 0$. So, Parseval's Identity holds, with $f(x) = 1$, so by definition of $f(\alpha)$,

$$\int_0^1 e^{2\pi i x^k \alpha} = \sum_{m=1}^{N} 1 = N \ll N^{2^1 - 1 + \epsilon} = N^{1+\epsilon}$$

This is clearly true. Done.

6.2 Inductive case

Now, let us suppose that (6.1) is true for $1 \leq j \leq k - 1$. By using $\phi(x) = \alpha x^k$ in Weyl's Lemma (5.1) along with (3.2), we obtain:

$$|f(\alpha)|^{2^j} \ll (2N)^{2^j - j - 1} \sum_{h_1} \underset{|\mathrm{h_i}| \leq \mathrm{N}}{\cdots} \sum_{h_j} \sum_{x \in I_j} e^{2\pi i \alpha h_1 \ldots h_j p_j(x; h_1, \ldots, h_j)}$$

By (3.2), we know that $p_j(x; h_1, \ldots, h_j)$ is a polynomial in x of degree $k - j$ with integer coefficients.

6.2.1 Defining and Bounding c_h

Since x and all h_i are integers, the value of the polynomial when evaluated must also be an integer. Reasoning thusly, we can simply rewrite the multiple sum as a single sum over the evaluated values of the polynomial - to wit, the integers, along with a constant c_h that is the number of solutions to $h_1 \dots h_j p_j(x; h_1, \dots, h_j) = h$.

Then, we have:

$$|f(\alpha)|^{2^j} \ll (2N)^{2^j-j-1} \sum_h c_h e^{2\pi i\alpha h} \tag{6.3}$$

Now, let us consider bounds on the c_h.

c_0 is the number of solutions to $h_1 \dots h_j p_j(x; h_1, \dots, h_j) = 0$. There are $(2N+1)^j$ distinct ways to fix the h_i such that $|h_i| \le N$, as specified by the bounded sums. Given fixed h_i, the polynomial can have at most $k-j$ roots, since it is of order $k-j$. Then, there are at most $(k-j)(2N+1)^j \ll N^j$ solutions. By the nature of the Vinogradov notation, we can then conclude that:

$$c_0 \ll N^j \tag{6.4}$$

Now, for $h \neq 0$, we make the key observation that p_j must be a factor of h. Since all the $h_i \le N$, we know that $|h| \le N^y$, where y is an arbitrary constant. By our useful fact from number theory (2.1), we know that:

$$d(h) \ll N^{y\epsilon}$$

Since p_j is a polynomial of degree $k-j$, only $k-j$ values of x can equal each divisor, so

$$c(h) \ll N^{y(k-j)\epsilon}$$

And if we substitute in $\dfrac{\epsilon}{y(k-j)}$ (for if it is true for this smaller value, it is surely true for the larger value that is ϵ), we get:

$$c_h \ll N^{\epsilon}(h \neq 0) \tag{6.5}$$

6.2.2 Defining and Bounding b_h

Consider again the expression $|f(\alpha)|^{2^j}$. By the definition of (1.1), we have

$$\overline{f(\alpha)} = \sum_{m=1}^{N} e^{-2\pi i m^k \alpha} = f(-\alpha)$$

Then,

$$\begin{aligned}
|f(\alpha)|^{2^j} &= \sqrt{f(\alpha)^{2^j}\overline{f(\alpha)^{2^j}}} \\
&= f(\alpha)^{2^{j-1}} f(-\alpha)^{2^{j-1}} \\
&= \sum_{\substack{|x_1|<N \\ 1\le i\le 2^{j-1}}} e^{2\pi i(x_1^k+\ldots+x_{2^j-1}^k)\alpha} \sum_{\substack{|y_1|<N \\ 1\le i\le 2^{j-1}}} e^{-2\pi i(y_1^k+\ldots+y_{2^j-1}^k)\alpha} \\
&= \sum_{\substack{|x_1|,|y_1|<N \\ 1\le i\le 2^{j-1}}} e^{2\pi i(x_1^k+\ldots+x_{2^j-1}^k-y_1^k-\ldots-y_{2^j-1}^k)\alpha} \\
&= \sum_h b_h e^{-2\pi i\alpha h}
\end{aligned}$$

Then,

$$|f(\alpha)|^{2^j} = \sum_h b_h e^{-2\pi i\alpha h} \tag{6.6}$$

where b_h is the number of solutions to $x_1^k+\ldots+x_{2^{j-1}}^k-y_1^k-\ldots-y_{2^{j-1}}^k=h$, $x_i, y_i \le N$.

If we let $\alpha = 0$, then we get:

$$\sum_h b_h(1) = f(0)^{2^j} = N^{2^j} \tag{6.7}$$

since

$$f(0) = \sum_{m=1}^{N} e^{2\pi i m^k 0} = \sum_{m=1}^{N} 1 = N$$

Now, by a similar argument presented in (6.2), we know that

$$\int_0^1 |f(\alpha)|^{2j} d\alpha$$

represents the number of times that

$$x_1^k+\ldots+x_{2^j}^k = 0, x_i \le N$$

which is equivalent to the definition of b_0, substituting $x=-y$ when applicable and re-labelling indices. By combining this insight with the inductive hypothesis (6.1), we have

$$b_0 = \int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j-j+\epsilon} \tag{6.8}$$

6.2.3 The Home Stretch

By substituting in (6.3) and (6.6), we can get:

$$\begin{aligned}\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha &= \int_0^1 |f(\alpha)|^{2^j} |f(\alpha)|^{2^j} d\alpha \\ &\ll \int_0^1 (2N)^{2^j-j-1} \sum_{h_1} c_{h_1} e^{2\pi i\alpha h_1} \sum_{h_2} b_{h_2} e^{-2\pi i\alpha h_2} d\alpha \\ &= (2N)^{2^j-j-1} \int_0^1 \sum_{h_1} c_{h_1} e^{2\pi i\alpha h_1} \sum_{h_2} b_{h_2} e^{-2\pi i\alpha h_2} d\alpha\end{aligned}$$

If we let $f(x) = c_h$ and $g(x) = \overline{g(x)} = b_h$ (since the b_h are all real-valued), then we can apply Parseval's Identity (4.1) to get:

$$\int_0^1 |f(\alpha)|^{2^j+1} d\alpha \ll (2N)^{2^j-j-1} \sum_h c_h b_h \tag{6.9}$$

But note, by substituting in results from (6.8), (6.4), (6.7), and (6.5), we get:

$$\sum_h c_h b_h = c_0 b_0 + \sum_{h\neq 0} c_h b_h \ll N^j N^{2^j-j+\epsilon} + N^\epsilon N^{2^j} = 2(N^{2^j+\epsilon}) \tag{6.10}$$

Then, by substituting (6.10) into (6.9), we achieve:

$$\begin{aligned}\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha &\ll (2N)^{2^j-j-1} \sum_h c_h b_h \\ &\ll (2N)^{2^j-j-1} 2(N^{2^j+\epsilon}) \\ &\ll (N)^{2^{j+1}-(j+1)+\epsilon}\end{aligned}$$

Q.E.D.

References

[1] Alex Rice, *MTH 391W Class Notes*, Unpublished, University of Rochester, 2016.

[2] R.C. Vaughan, *The Hardy-Littlewood method*, Cambridge Univeristy Press, Cambridge, 1981.