Math 571, Spring 2025, Waring's Problem: Simplest Upper Bound

Robert C. Vaughan

March 27, 2025

- We want to count $R_s(n)$, the number of solutions of
$$x_1^k + \cdots x_s^k = n$$
in positive integers $x_1, \ldots, x_s$.

- We want to count $R_s(n)$, the number of solutions of
$$x_1^k + \cdots x_s^k = n$$
in positive integers $x_1, \ldots, x_s$.
- Clearly we need only consider $x_j \le n^{1/k}$, and for brevity we write
$$N = \lfloor n^{1/k} \rfloor.$$

- We want to count $R_s(n)$, the number of solutions of

$$x_1^k + \cdots x_s^k = n$$

in positive integers $x_1, \ldots, x_s$.

- Clearly we need only consider $x_j \le n^{1/k}$, and for brevity we write

$$N = \lfloor n^{1/k} \rfloor.$$

- Then following the pattern established in studying the Goldbach problems we put

$$f(\alpha) = \sum_{x=1}^{N} e(\alpha x^k)$$

so that by the orthogonality of the additive characters $e(\alpha n)$ we have

$$R_s(n) = \int_0^1 f(\alpha)^s e(-\alpha n) d\alpha.$$

- We want to count $R_s(n)$, the number of solutions of
$$x_1^k + \cdots x_s^k = n$$
in positive integers $x_1, \ldots, x_s$.
- Clearly we need only consider $x_j \le n^{1/k}$, and for brevity we write
$$N = \lfloor n^{1/k} \rfloor.$$
- Then following the pattern established in studying the Goldbach problems we put
$$f(\alpha) = \sum_{x=1}^{N} e(\alpha x^k)$$
so that by the orthogonality of the additive characters $e(\alpha n)$ we have
$$R_s(n) = \int_0^1 f(\alpha)^s e(-\alpha n) d\alpha.$$
- Thus we would like to understand the behaviour of f when α is close to a rational number a/q.

- For smaller q there is a simple elementary result.
  Theorem 8.1 *Suppose that* $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a,q)=1$ *and* $\beta = \alpha - a/q$. *Then*

$$f(\alpha) = q^{-1}S(q,a)v(\beta) + O(q + qn|\beta|)$$

where

$$S(q,a) = \sum_{r=1}^{q} e(ar^k/q)$$

and

$$v(\beta) = \sum_{y=1}^{n} k^{-1}y^{\frac{1}{k}-1}e(\beta y)$$

- For smaller q there is a simple elementary result.
  Theorem 8.1 *Suppose that* $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, q) = 1$ *and* $\beta = \alpha - a/q$. *Then*

$$f(a) = q^{-1}S(q, a)v(\beta) + O(q + qn|\beta|)$$

where

$$S(q, a) = \sum_{r=1}^{q} e(ar^k/q)$$

and

$$v(\beta) = \sum_{y=1}^{n} k^{-1}y^{\frac{1}{k}-1}e(\beta y)$$

- There are a variety of possible choices for $v(\beta)$. An examination of the proof reveals that

$$v(\beta) = \int_0^n k^{-1}y^{\frac{1}{k}-1}e(\beta y)dy = \int_0^{n^{1/k}} e(\beta x^k)dx$$

are possible choices.

- Another is Estermann's $\displaystyle\sum_{h=0}^{n} \frac{\Gamma(h+1/k)}{h!k} e(\beta h)$.

- Another is Estermann's $\sum_{h=0}^{n} \frac{\Gamma(h+1/k)}{h!k} e(\beta h)$.
- I should also point out that by working a lot harder the error term can be replaced by $q^{1/2+\varepsilon}(1+n|\beta|)^{1/2}$

- Another is Estermann's $\sum_{h=0}^{n} \frac{\Gamma(h+1/k)}{h!k} e(\beta h)$.
- I should also point out that by working a lot harder the error term can be replaced by $q^{1/2+\varepsilon}(1+n|\beta|)^{1/2}$
- and if one supposes further that $|\beta| \le (2kq)n^{1/k-1}$ even by $q^{1/2+\varepsilon}$ and this can be very useful in some applications.

- Another is Estermann's $\displaystyle\sum_{h=0}^{n}\frac{\Gamma(h+1/k)}{h!k}e(\beta h)$.
- I should also point out that by working a lot harder the error term can be replaced by $q^{1/2+\varepsilon}(1+n|\beta|)^{1/2}$
- and if one supposes further that $|\beta|\le(2kq)n^{1/k-1}$ even by $q^{1/2+\varepsilon}$ and this can be very useful in some applications.
- Note that if we use Dirichlet's theorem to approximate a general α so that for some Q we have $q\le Q$, $|\beta|=|\alpha-a/q|\le q^{-1}Q^{-1}$, then

$$q^{1/2+\varepsilon}(1+n|\beta|)^{1/2}\ll q^{1/2+\varepsilon}+n^{1/2}q^{\varepsilon}Q^{-1/2}$$
$$\ll Q^{1/2+\varepsilon}+n^{1/2}Q^{\varepsilon-1/2}$$

- Another is Estermann's $\sum_{h=0}^{n} \frac{\Gamma(h+1/k)}{h!k} e(\beta h)$.
- I should also point out that by working a lot harder the error term can be replaced by $q^{1/2+\varepsilon}(1+n|\beta|)^{1/2}$
- and if one supposes further that $|\beta| \le (2kq)n^{1/k-1}$ even by $q^{1/2+\varepsilon}$ and this can be very useful in some applications.
- Note that if we use Dirichlet's theorem to approximate a general α so that for some Q we have $q \le Q$, $|\beta| = |\alpha - a/q| \le q^{-1}Q^{-1}$, then

$$q^{1/2+\varepsilon}(1+n|\beta|)^{1/2} \ll q^{1/2+\varepsilon} + n^{1/2}q^{\varepsilon}Q^{-1/2} \ll Q^{1/2+\varepsilon} + n^{1/2}Q^{\varepsilon-1/2}$$

- We want this to be smaller than $n^{1/k}$ and this will be so when $n^{1-2/k-\delta} < Q < n^{2/k-\delta}$ and this will work when $k < 4$.
- Thus we can actually give a major arcs only treatment to sums of cubes.

- Another is Estermann's $\sum_{h=0}^{n} \frac{\Gamma(h+1/k)}{h!k} e(\beta h)$.
- I should also point out that by working a lot harder the error term can be replaced by $q^{1/2+\varepsilon}(1+n|\beta|)^{1/2}$
- and if one supposes further that $|\beta| \le (2kq)n^{1/k-1}$ even by $q^{1/2+\varepsilon}$ and this can be very useful in some applications.
- Note that if we use Dirichlet's theorem to approximate a general α so that for some Q we have $q \le Q$, $|\beta| = |\alpha - a/q| \le q^{-1}Q^{-1}$, then

$$q^{1/2+\varepsilon}(1+n|\beta|)^{1/2} \ll q^{1/2+\varepsilon} + n^{1/2}q^{\varepsilon}Q^{-1/2}$$
$$\ll Q^{1/2+\varepsilon} + n^{1/2}Q^{\varepsilon-1/2}$$

- We want this to be smaller than $n^{1/k}$ and this will be so when $n^{1-2/k-\delta} < Q < n^{2/k-\delta}$ and this will work when $k < 4$.
- Thus we can actually give a major arcs only treatment to sums of cubes.
- This was first observed in RCV[1977] and lead to some substantial developments for cubic problems.

- Theorem 8.1 *Suppose that* $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, q) = 1$ *and* $\beta = \alpha - a/q$. *Then*

$$f(a) = q^{-1}S(q, a)v(\beta) + O(q + qn|\beta|)$$

where

$$S(q, a) = \sum_{r=1}^{q} e(ar^k/q)$$

and

$$v(\beta) = \sum_{y=1}^{n} k^{-1}y^{\frac{1}{k}-1}e(\beta y)$$

- Theorem 8.1 *Suppose that $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, q) = 1$ and $\beta = \alpha - a/q$. Then*

$$f(a) = q^{-1}S(q, a)v(\beta) + O(q + qn|\beta|)$$

where

$$S(q, a) = \sum_{r=1}^{q} e(ar^k/q)$$

and

$$v(\beta) = \sum_{y=1}^{n} k^{-1}y^{\frac{1}{k}-1}e(\beta y)$$

- As it stands for the error to be smaller than $n^{1/k}$ we need $q \le n^{1/k-\delta}$ and $|\beta| \le q^{-1}n^{1/k-\delta-1}$ and the total measure of all such $\alpha \in [0, 1]$ which satisfy this is

$$\ll \sum_{q \le n^{1/k-\delta}} \phi(q)q^{-1}n^{1/k-\delta-1} \ll n^{2/k-2\delta-1}$$

and this will certainly be small when $k \ge 3$.

- Theorem 8.1 *Suppose that $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, q) = 1$ and $\beta = \alpha - a/q$. Then*

$$f(a) = q^{-1}S(q, a)v(\beta) + O(q + qn|\beta|)$$

where $S(q, a) = \sum_{r=1}^{q} e(ar^k/q)$ *and*

$$v(\beta) = \sum_{y=1}^{n} k^{-1}y^{\frac{1}{k}-1}e(\beta y)$$

- Theorem 8.1 *Suppose that* $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a,q)=1$ *and* $\beta = \alpha - a/q$. *Then*

$$f(a) = q^{-1}S(q,a)v(\beta) + O(q + qn|\beta|)$$

where $S(q,a) = \sum_{r=1}^{q} e(ar^k/q)$ *and*

$$v(\beta) = \sum_{y=1}^{n} k^{-1} y^{\frac{1}{k}-1} e(\beta y)$$

- **Proof.** We start in the usual way by splitting the sum over x according to the residue class of x modulo q. Thus

$$f(\alpha) = \sum_{r=1}^{q} e(ar^k/q) \sum_{\substack{x=1 \\ x\equiv r \pmod q}}^{N} e(\beta x^k).$$

- Theorem 8.1 *Suppose that $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, q) = 1$ and $\beta = \alpha - a/q$. Then*

$$f(a) = q^{-1}S(q, a)v(\beta) + O(q + qn|\beta|)$$

where $S(q, a) = \sum_{r=1}^{q} e(ar^k/q)$ *and*

$$v(\beta) = \sum_{y=1}^{n} k^{-1}y^{\frac{1}{k}-1}e(\beta y)$$

- **Proof.** We start in the usual way by splitting the sum over x according to the residue class of x modulo q. Thus

$$f(\alpha) = \sum_{r=1}^{q} e(ar^k/q) \sum_{\substack{x=1 \\ x \equiv r \pmod{q}}}^{N} e(\beta x^k).$$

- When β is small we can expect that the sum would behave like the corresponding integral, so partial summation/integration is suggested.

- Thus

$$f(\alpha) = \sum_{r=1}^{q} e(ar^k/q) \sum_{\substack{x=1 \\ x\equiv r \pmod q}}^{N} e(\beta x^k).$$

- Thus

$$f(\alpha) = \sum_{r=1}^{q} e(ar^k/q) \sum_{\substack{x=1 \\ x\equiv r \pmod q}}^{N} e(\beta x^k).$$

- The inner sum here is

$$\sum_{\substack{x\le n^{1/k} \\ x\equiv r \bmod q}} \left(e(\beta n) - \int_x^{n^{1/k}} 2\pi i \beta k u^{\frac{1}{k}-1} e(\beta u^k) du \right)$$

$$= \sum_{\substack{x\le n^{1/k} \\ x\equiv r \bmod q}} e(\beta n) -$$

$$\int_0^{n^{1/k}} 2\pi i \beta k u^{k-1} \sum_{\substack{x\le u \\ x\equiv r \bmod q}} e(\beta u^k) du$$

- The inner sum here is

$$\sum_{\substack{x \le n^{1/k} \\ x \equiv r \bmod q}} e(\beta n) - \int_0^{n^{1/k}} 2\pi i \beta k u^{k-1} \sum_{\substack{x \le u \\ x \equiv r \bmod q}} e(\beta u^k) du$$

- The inner sum here is

$$\sum_{\substack{x\le n^{1/k}\\ x\equiv r \bmod q}} e(\beta n) - \int_0^{n^{1/k}} 2\pi i\beta k u^{k-1} \sum_{\substack{x\le u\\ x\equiv r \bmod q}} e(\beta u^k)du$$

- We have

$$\sum_{\substack{x\le u\\ x\equiv r \bmod q}} 1 = \frac{u}{q} + O(1).$$

- The inner sum here is

$$\sum_{\substack{x\le n^{1/k}\\ x\equiv r \bmod q}} e(\beta n) - \int_0^{n^{1/k}} 2\pi i\beta k u^{k-1} \sum_{\substack{x\le u\\ x\equiv r \bmod q}} e(\beta u^k)du$$

- We have

$$\sum_{\substack{x\le u\\ x\equiv r \bmod q}} 1 = \frac{u}{q} + O(1).$$

- Inserting this gives an error term $\ll 1+|\beta|n$ and a main term $n^{1/k}q^{-1}e(\beta n) - \int_0^{n^{1/k}} 2\pi i\beta k u^k q^{-1} e(\beta u^k)du$

- The inner sum here is
$$\sum_{\substack{x\le n^{1/k}\\ x\equiv r \bmod q}} e(\beta n) - \int_0^{n^{1/k}} 2\pi i\beta k u^{k-1} \sum_{\substack{x\le u\\ x\equiv r \bmod q}} e(\beta u^k)du$$
- We have
$$\sum_{\substack{x\le u\\ x\equiv r \bmod q}} 1 = \frac{u}{q} + O(1).$$
- Inserting this gives an error term $\ll 1 + |\beta|n$ and a main term $n^{1/k}q^{-1}e(\beta n) - \int_0^{n^{1/k}} 2\pi i\beta k u^k q^{-1} e(\beta u^k)du$
- and by integration by parts this is $q^{-1}\int_0^{n^{1/k}} e(\beta u^k)du$.

- The inner sum here is
$$\sum_{\substack{x\le n^{1/k}\\ x\equiv r \bmod q}} e(\beta n) - \int_0^{n^{1/k}} 2\pi i\beta k u^{k-1} \sum_{\substack{x\le u\\ x\equiv r \bmod q}} e(\beta u^k)du$$
- We have
$$\sum_{\substack{x\le u\\ x\equiv r \bmod q}} 1 = \frac{u}{q} + O(1).$$
- Inserting this gives an error term $\ll 1+|\beta|n$ and a main term $n^{1/k}q^{-1}e(\beta n) - \int_0^{n^{1/k}} 2\pi i\beta k u^k q^{-1} e(\beta u^k)du$
- and by integration by parts this is $q^{-1}\int_0^{n^{1/k}} e(\beta u^k)du$.
- Thus we find that
$$f(\alpha) = q^{-1}S(q,a)\int_0^{n^{1/k}} e(\beta u^k)du + O(q+qn|\beta|)$$
which gives the theorem with one of the alternative choices for v.

- Thus we find that

$$f(\alpha) = q^{-1}S(q,a)\int_0^{n^{1/k}} e(\beta u^k)du + O(q + qn|\beta|)$$

- Thus we find that

$$f(\alpha) = q^{-1}S(q,a)\int_0^{n^{1/k}} e(\beta u^k)du + O(q + qn|\beta|)$$

- The change of variable $t = u^k$ gives another form $\int_0^n k^{-1}t^{1/k-1}e(\beta t)dt$.

- Thus we find that

$$f(\alpha) = q^{-1}S(q,a)\int_0^{n^{1/k}} e(\beta u^k)du + O(q+qn|\beta|)$$

- The change of variable $t = u^k$ gives another form $\displaystyle\int_0^n k^{-1}t^{1/k-1}e(\beta t)dt.$
- The sum of a monotonic sequence equals the corresponding integral with an error largest. Thus

$$\sum_{y\le x} k^{-1}y^{1/k-1} = \int_1^x k^{-1}t^{1/k-1}dt + O(1) = x^{1/k} + O(1).$$

- Thus we find that

$$f(\alpha) = q^{-1}S(q,a)\int_0^{n^{1/k}} e(\beta u^k)du + O(q + qn|\beta|)$$

- The change of variable $t = u^k$ gives another form $\int_0^n k^{-1}t^{1/k-1}e(\beta t)dt$.
- The sum of a monotonic sequence equals the corresponding integral with an error largest. Thus

$$\sum_{y\le x} k^{-1}y^{1/k-1} = \int_1^x k^{-1}t^{1/k-1}dt + O(1) = x^{1/k} + O(1).$$

- Hence by the same process as before $\sum_{y\le n} k^{-1}y^{1/k-1}e(\beta y)$

$$= \sum_{y\le n} k^{-1}y^{1/k-1}e(\beta n) - \int_1^n 2\pi i\beta e(\beta t)\sum_{y\le t} k^{-1}y^{1/k-1}dv$$

$$= n^{1/k}e(\beta n) - \int_0^n 2\pi i\beta e(\beta t)t^{1/k}dt + O(1 + n|\beta|).$$

- We have

$$= n^{1/k}e(\beta n) - \int_0^n 2\pi i\beta e(\beta t)t^{1/k}dt + O(1 + n|\beta|)$$

- We have

$$= n^{1/k}e(\beta n) - \int_0^n 2\pi i\beta e(\beta t)t^{1/k}dt + O(1+n|\beta|)$$

- Therefore by integration by parts once more this is

$$\int_0^n k^{-1}t^{1/k-1}e(\beta t)dt.$$

- We have

$$= n^{1/k}e(\beta n) - \int_0^n 2\pi i\beta e(\beta t)t^{1/k}dt + O(1 + n|\beta|)$$

- Therefore by integration by parts once more this is

$$\int_0^n k^{-1}t^{1/k-1}e(\beta t)dt.$$

- Thus we can replace the integral version of $v(\beta)$ in our approximation by the sum version with a total error $\ll 1 + n|\beta|$.

- We have

$$= n^{1/k} e(\beta n) - \int_0^n 2\pi i \beta e(\beta t) t^{1/k} dt + O(1 + n|\beta|)$$

- Therefore by integration by parts once more this is

$$\int_0^n k^{-1} t^{1/k-1} e(\beta t) dt.$$

- Thus we can replace the integral version of $v(\beta)$ in our approximation by the sum version with a total error $\ll 1 + n|\beta|$.
- This completes the proof of the theorem

- We have already seen that the above theorem cannot be used to cover a unit interval when $k > 2$ so we are forced to divide into major and minor arcs.

- We have already seen that the above theorem cannot be used to cover a unit interval when $k > 2$ so we are forced to divide into major and minor arcs.
- We used the following in connection with primes.

- We have already seen that the above theorem cannot be used to cover a unit interval when $k>2$ so we are forced to divide into major and minor arcs.
- We used the following in connection with primes.
- **Lemma 8.2.** *Suppose that* X, Y, α *are real numbers with* $X\geq 1$, $Y\geq 1$ *and that* $q\in\mathbb{R}$, $a\in\mathbb{Z}$, $|\alpha-a/q|\leq q^{-2}$ *with* $(a,q)=1$. *Then*

$$\sum_{x\leq X}\min\left(XYx^{-1},\|\alpha x\|^{-1}\right)$$
$$\ll XY\left(\frac{1}{q}+\frac{1}{Y}+\frac{q}{XY}\right)\log(2XYq).$$

- Recall we basically used this to treat sums of the kind

$$\sum_{m,n}e(\alpha mn)$$

after performing the summation over n.

- We have already seen that the above theorem cannot be used to cover a unit interval when $k > 2$ so we are forced to divide into major and minor arcs.
- We used the following in connection with primes.
- **Lemma 8.2.** *Suppose that* X, Y, α *are real numbers with* $X \ge 1$, $Y \ge 1$ *and that* $q \in \mathbb{R}$, $a \in \mathbb{Z}$, $|\alpha - a/q| \le q^{-2}$ *with* $(a,q) = 1$. *Then*

$$\sum_{x \le X} \min\left(XYx^{-1}, \|\alpha x\|^{-1}\right) \ll XY\left(\frac{1}{q} + \frac{1}{Y} + \frac{q}{XY}\right)\log(2XYq).$$

- Recall we basically used this to treat sums of the kind

$$\sum_{m,n} e(\alpha mn)$$

after performing the summation over n.
- This we knew how to do since the exponent is linear in n.

- We have already seen that the above theorem cannot be used to cover a unit interval when $k > 2$ so we are forced to divide into major and minor arcs.
- We used the following in connection with primes.
- **Lemma 8.2.** *Suppose that* X, Y, α *are real numbers with* $X \geq 1$, $Y \geq 1$ *and that* $q \in \mathbb{R}$, $a \in \mathbb{Z}$, $|\alpha - a/q| \leq q^{-2}$ *with* $(a, q) = 1$. *Then*

$$\sum_{x \leq X} \min\left(XYx^{-1}, \|\alpha x\|^{-1}\right) \ll XY\left(\frac{1}{q} + \frac{1}{Y} + \frac{q}{XY}\right)\log(2XYq).$$

- Recall we basically used this to treat sums of the kind

$$\sum_{m,n} e(\alpha mn)$$

after performing the summation over n.
- This we knew how to do since the exponent is linear in n.
- But now our exponent is of higher degree.

- Herman Weyl makes the brilliant observation that if we have a polynomial Ψ of degree k, then $\Psi(x+h) - \Psi(x)$ is of degree $k-1$ in x and we can combine this with the Cauchy-Schwarz inequality to make progress.

- Herman Weyl makes the brilliant observation that if we have a polynomial Ψ of degree k, then $\Psi(x+h)-\Psi(x)$ is of degree $k-1$ in x and we can combine this with the Cauchy-Schwarz inequality to make progress.
- Thus central to his idea is the use of the forward difference operator which we define iteratively by

$$\Delta_1\big(\Psi(\alpha);\beta\big) = \Psi(\alpha+\beta)-\Psi(\alpha)$$

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

- Herman Weyl makes the brilliant observation that if we have a polynomial Ψ of degree k, then $\Psi(x+h)-\Psi(x)$ is of degree $k-1$ in x and we can combine this with the Cauchy-Schwarz inequality to make progress.
- Thus central to his idea is the use of the forward difference operator which we define iteratively by

$$\Delta_1\big(\Psi(\alpha);\beta\big) = \Psi(\alpha+\beta)-\Psi(\alpha)$$
$$\Delta_{j+1}\big(\Psi(\alpha);\beta_1,\ldots,\beta_{j+1}\big) = \Delta_1\big(\Delta_j\big(\Psi(\alpha);\beta_1,\ldots,\beta_j\big);\beta_{j+1}\big)$$

- For example

$$\Delta_1(\alpha^3;\beta_1) = (\alpha+\beta_1)^3-\alpha^3 = 3\alpha^2\beta_1+3\alpha\beta_1^2+\beta_1^3$$

- Herman Weyl makes the brilliant observation that if we have a polynomial Ψ of degree k, then $\Psi(x+h)-\Psi(x)$ is of degree $k-1$ in x and we can combine this with the Cauchy-Schwarz inequality to make progress.
- Thus central to his idea is the use of the forward difference operator which we define iteratively by

$$\Delta_1\big(\Psi(\alpha);\beta\big)=\Psi(\alpha+\beta)-\Psi(\alpha)$$
$$\Delta_{j+1}\big(\Psi(\alpha);\beta_1,\ldots,\beta_{j+1}\big)=\Delta_1\big(\Delta_j\big(\Psi(\alpha);\beta_1,\ldots,\beta_j\big);\beta_{j+1}\big)$$

- For example

$$\Delta_1(\alpha^3;\beta_1)=(\alpha+\beta_1)^3-\alpha^3=3\alpha^2\beta_1+3\alpha\beta_1^2+\beta_1^3$$

- and

$$\Delta_2(\alpha^3;\beta_1,\beta_2)=\beta_1\beta_2(6\alpha+3\beta_1+3\beta_2)$$

- Generally one has

$$\Delta_j(\alpha^k;\beta_1,\ldots,\beta_j)$$
$$=\sum_{\theta_1=0}^{1}\ldots\sum_{\theta_j=1}^{1}(-1)^{j-\theta_1-\cdots-\theta_j}(\alpha+\theta_1\beta_1+\cdots+\theta_j\beta_j)^k.$$

- Generally one has

$$\Delta_j(\alpha^k;\beta_1,\dots,\beta_j) = \sum_{\theta_1=0}^{1}\dots\sum_{\theta_j=1}^{1}(-1)^{j-\theta_1-\dots-\theta_j}(\alpha+\theta_1\beta_1+\dots+\theta_j\beta_j)^k.$$

- By the multinomial theorem here the k-th power is

$$\sum_{\substack{\ell_0+\ell_1+\dots+\ell_j=k\\ \ell_i\ge 0}} \frac{k!\alpha^{\ell_0}(\theta_1\beta_1)^{\ell_1}\dots(\theta_j\beta_j)^{\ell_j}}{\ell_0!\dots\ell_j!}$$

- Generally one has

$$\Delta_j(\alpha^k;\beta_1,\ldots,\beta_j) = \sum_{\theta_1=0}^{1}\ldots\sum_{\theta_j=1}^{1}(-1)^{j-\theta_1-\cdots-\theta_j}(\alpha+\theta_1\beta_1+\cdots+\theta_j\beta_j)^k.$$

- By the multinomial theorem here the k-th power is

$$\sum_{\substack{\ell_0+\ell_1+\cdots+\ell_j=k\\ \ell_i\ge 0}}\frac{k!\alpha^{\ell_0}(\theta_1\beta_1)^{\ell_1}\ldots(\theta_j\beta_j)^{\ell_j}}{\ell_0!\ldots\ell_j!}$$

- In formal power series there is a convention that $0^0=1$.

- Generally one has

$$\Delta_j(\alpha^k;\beta_1,\dots,\beta_j) = \sum_{\theta_1=0}^{1}\dots\sum_{\theta_j=1}^{1}(-1)^{j-\theta_1-\dots-\theta_j}(\alpha+\theta_1\beta_1+\dots+\theta_j\beta_j)^k.$$

- By the multinomial theorem here the k-th power is

$$\sum_{\substack{\ell_0+\ell_1+\dots+\ell_j=k\\ \ell_i\ge 0}} \frac{k!\alpha^{\ell_0}(\theta_1\beta_1)^{\ell_1}\dots(\theta_j\beta_j)^{\ell_j}}{\ell_0!\dots\ell_j!}$$

- In formal power series there is a convention that $0^0=1$.
- Thus if $\ell_i=0$ then $\displaystyle\sum_{\theta_i=0}^{1}(-1)^{1-\theta_i}\theta_i^{\ell_i}=1-1=0$

- Generally one has

$$\Delta_j(\alpha^k;\beta_1,\dots,\beta_j) = \sum_{\theta_1=0}^{1}\dots\sum_{\theta_j=1}^{1}(-1)^{j-\theta_1-\cdots-\theta_j}(\alpha+\theta_1\beta_1+\cdots+\theta_j\beta_j)^k.$$

- By the multinomial theorem here the k-th power is

$$\sum_{\substack{\ell_0+\ell_1+\cdots+\ell_j=k\\ \ell_i\ge 0}}\frac{k!\alpha^{\ell_0}(\theta_1\beta_1)^{\ell_1}\dots(\theta_j\beta_j)^{\ell_j}}{\ell_0!\dots\ell_j!}$$

- In formal power series there is a convention that $0^0=1$.
- Thus if $\ell_i=0$ then $\sum_{\theta_i=0}^{1}(-1)^{1-\theta_i}\theta_i^{\ell_i}=1-1=0$
- so we only get a non-zero term when each $\ell_i\ge 1$.

- Generally one has

$$\Delta_j(\alpha^k;\beta_1,\ldots,\beta_j) = \sum_{\theta_1=0}^{1}\cdots\sum_{\theta_j=1}^{1}(-1)^{j-\theta_1-\cdots-\theta_j}(\alpha+\theta_1\beta_1+\cdots+\theta_j\beta_j)^k.$$

- By the multinomial theorem here the k-th power is

$$\sum_{\substack{\ell_0+\ell_1+\cdots+\ell_j=k\\ \ell_i\ge 0}} \frac{k!\alpha^{\ell_0}(\theta_1\beta_1)^{\ell_1}\ldots(\theta_j\beta_j)^{\ell_j}}{\ell_0!\ldots\ell_j!}$$

- In formal power series there is a convention that $0^0=1$.
- Thus if $\ell_i=0$ then $\displaystyle\sum_{\theta_i=0}^{1}(-1)^{1-\theta_i}\theta_i^{\ell_i}=1-1=0$
- so we only get a non-zero term when each $\ell_i\ge 1$.
- Thus $\Delta_j(\alpha^k;\beta_1,\ldots,\beta_j)$ has $\beta_1\ldots\beta_j$ as a factor, is of degree $k-j$ and has leading coefficient $\dfrac{k!}{(k-j)!}\beta_1\ldots\beta_j$.

- The following theorem encapsulates *Weyl differencing*.

Theorem 8.3. *Let*

$$T(\Psi)=\sum_{x=1}^{Q}e\big(\Psi(x)\big)$$

where $\Psi(x)$ is an arbitrary arithmetical function. Then

$$|T(\Psi)|^{2^j}\le(2Q)^{2^j-j-1}\sum_{|h_1|<Q}\dots\sum_{|h_j|<Q}T_j$$

where

$$T_j=\sum_{x\in I_j}e\big(\Delta_j\big(\Psi(x);h_1,\dots h_j\big)\big)$$

and the intervals $I_j(h_1,\dots h_j)$ satisfy

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

- **Proof.** This is by induction on j.

- **Proof.** This is by induction on j.
- The case $j = 1$ follows by writing

$$|T(\Psi)|^2 = \sum_{x=1}^{Q} e\big(-\Psi(x)\big) \sum_{y=1}^{Q} e\big(\Psi(y)\big)$$

and then letting $y = x + h$.

- **Proof.** This is by induction on j.
- The case $j = 1$ follows by writing

$$|T(\Psi)|^2 = \sum_{x=1}^{Q} e\big(-\Psi(x)\big) \sum_{y=1}^{Q} e\big(\Psi(y)\big)$$

and then letting $y = x + h$.
- Then $-x < h \le Q - x$ and $1 \le x \le Q$, so that $1 - Q \le h \le Q - 1$ and $-h < x \le Q - h$.

- **Proof.** This is by induction on j.
- The case $j = 1$ follows by writing

$$|T(\Psi)|^2 = \sum_{x=1}^{Q} e\big(-\Psi(x)\big) \sum_{y=1}^{Q} e\big(\Psi(y)\big)$$

and then letting $y = x + h$.

- Then $-x < h \le Q - x$ and $1 \le x \le Q$, so that $1 - Q \le h \le Q - 1$ and $-h < x \le Q - h$.
- We now interchange the order of summation so that

$$|T(\Psi)|^2 = \sum_{1-Q \le h \le Q-1} \sum_{\substack{1 \le x \le Q \\ -h < x \le Q-h}} e\big(\Psi(x+h) - \Psi(x)\big)$$

- **Proof.** This is by induction on j.
- The case $j = 1$ follows by writing

$$|T(\Psi)|^2 = \sum_{x=1}^{Q} e\big(-\Psi(x)\big) \sum_{y=1}^{Q} e\big(\Psi(y)\big)$$

and then letting $y = x + h$.

- Then $-x < h \le Q - x$ and $1 \le x \le Q$, so that $1 - Q \le h \le Q - 1$ and $-h < x \le Q - h$.
- We now interchange the order of summation so that

$$|T(\Psi)|^2 = \sum_{1-Q\le h\le Q-1} \sum_{\substack{1\le x\le Q\\ -h<x\le Q-h}} e\big(\Psi(x+h) - \Psi(x)\big)$$

- Thus the x in the inner sum are precisely the x in the intersection of $[1, Q]$ and $[1 - h, Q - h]$ which is an interval $I_1(h)$ of the required kind.

- For the inductive step we begin by applying Cauchy. Thus

$$|T(\Psi)|^{2^{j+1}} \le (2Q)^{2^{j+1}-2j-2}\left(\sum_{|h_1|<Q}\dots\sum_{|h_j|<Q}T_j\right)^2$$
$$\le (2Q)^{2^{j+1}-j-2}\sum_{|h_1|<Q}\dots\sum_{|h_j|<Q}|T_j|^2.$$

- For the inductive step we begin by applying Cauchy. Thus

$$|T(\Psi)|^{2^{j+1}} \le (2Q)^{2^{j+1}-2j-2}\left(\sum_{|h_1|<Q}\dots\sum_{|h_j|<Q} T_j\right)^2$$
$$\le (2Q)^{2^{j+1}-j-2}\sum_{|h_1|<Q}\dots\sum_{|h_j|<Q} |T_j|^2.$$

- Now we treat $|T_j|^2$ as T in the initial case. Thus $|T_j|^2$

$$= \sum_{1-Q\le h\le Q-1}\sum_{\substack{x\in I_j\\ x+h\in I_j}} e\big(\Delta_1\big(\Delta_j\big(\Psi(x);h_1,\dots h_j\big);h\big)\big).$$

- For the inductive step we begin by applying Cauchy. Thus

$$|T(\Psi)|^{2^{j+1}} \le (2Q)^{2^{j+1}-2j-2}\left(\sum_{|h_1|<Q}\dots\sum_{|h_j|<Q} T_j\right)^2$$
$$\le (2Q)^{2^{j+1}-j-2}\sum_{|h_1|<Q}\dots\sum_{|h_j|<Q}|T_j|^2.$$

- Now we treat $|T_j|^2$ as T in the initial case. Thus $|T_j|^2$

$$= \sum_{1-Q\le h\le Q-1}\ \sum_{\substack{x\in I_j\\ x+h\in I_j}} e\big(\Delta_1\big(\Delta_j\big(\Psi(x);h_1,\dots h_j\big);h\big)\big).$$

- and the conclusion follows on taking $I_{j+1} = I_j \cap (h + I_j)$.

- We now come to one of the more famous results of analytic number theory.

- We now come to one of the more famous results of analytic number theory.
- **Theorem 8.4. [Weyl's inequality]** *Suppose that* $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, q) = 1$, $\alpha \in \mathbb{R}$, $|\alpha - a/q| \le q^{-2}$,
$$\Psi(x) = \alpha x^k + \alpha_1 x^{k-1} + \cdots \alpha_{k-1} x + \alpha_k$$
and $T(\Psi) = \sum_{x=1}^{Q} e\big(\Psi(x)\big)$. *Then*
$$T(\Psi) \ll Q^{1+\varepsilon}(q^{-1} + Q^{-1} + qQ^{-k})^{2^{1-k}}.$$

- We now come to one of the more famous results of analytic number theory.
- **Theorem 8.4. [Weyl's inequality]** *Suppose that* $q \in \mathbb{N}$, $a \in \mathbb{Z}$, $(a, q) = 1$, $\alpha \in \mathbb{R}$, $|\alpha - a/q| \le q^{-2}$,

$$\Psi(x) = \alpha x^k + \alpha_1 x^{k-1} + \cdots \alpha_{k-1} x + \alpha_k$$

and $T(\Psi) = \sum_{x=1}^{Q} e\big(\Psi(x)\big)$. *Then*

$$T(\Psi) \ll Q^{1+\varepsilon}(q^{-1} + Q^{-1} + qQ^{-k})^{2^{1-k}}.$$

- **Proof.** We use the case $j = k - 1$ of the previous theorem. From the comments surrounding Δ_j we have

$$\begin{aligned}\Delta_{k-1}\big(\Psi(x); h_1, \dots, h_{k-1}\big) \\ = \alpha k! h_1 \dots h_{k-1} x \\ + \alpha \frac{k!}{2} h_1 \dots h_{k-1}(h_1 + \cdots h_{k-1}) \\ + \alpha_1 h_1 \dots h_{k-1}.\end{aligned}$$

- We have

$$\Delta_{k-1}\big(\Psi(x); h_1, \dots, h_{k-1}\big) = \alpha k! h_1 \dots h_{k-1} x + \alpha \frac{k!}{2} h_1 \dots h_{k-1}(h_1 + \cdots h_{k-1}) + \alpha_1 h_1 \dots h_{k-1}.$$

- We have

$$\Delta_{k-1}\big(\Psi(x); h_1, \ldots, h_{k-1}\big) = \alpha k! h_1 \ldots h_{k-1} x + \alpha \frac{k!}{2} h_1 \ldots h_{k-1}(h_1 + \cdots h_{k-1}) + \alpha_1 h_1 \ldots h_{k-1}.$$

- Thus $T_{k-1} \ll \min(Q, \|\alpha k! h_1 \ldots h_{k-1}\|^{-1})$.

- We have

$$\Delta_{k-1}\big(\Psi(x);h_1,\dots,h_{k-1}\big) = \alpha k! h_1\dots h_{k-1}x + \alpha\frac{k!}{2}h_1\dots h_{k-1}(h_1+\cdots h_{k-1}) + \alpha_1 h_1\dots h_{k-1}.$$

- Thus $T_{k-1} \ll \min(Q, \|\alpha k! h_1\dots h_{k-1}\|^{-1})$.
- We apply this to the previous theorem and separate out the $\ll Q^{k-2}$ terms for which $h_1\dots h_{k-1}=0$.

- We have

$$\Delta_{k-1}\big(\Psi(x); h_1, \ldots, h_{k-1}\big) = \alpha k! h_1 \ldots h_{k-1} x + \alpha \frac{k!}{2} h_1 \ldots h_{k-1}(h_1 + \cdots h_{k-1}) + \alpha_1 h_1 \ldots h_{k-1}.$$

- Thus $T_{k-1} \ll \min(Q, \|\alpha k! h_1 \ldots h_{k-1}\|^{-1})$.
- We apply this to the previous theorem and separate out the $\ll Q^{k-2}$ terms for which $h_1 \ldots h_{k-1} = 0$.
- Thus $|T(\Psi)|^{2^{k-1}}$

$$\ll Q^{2^{k-1}-1} + Q^{2^{k-1}-k+\varepsilon} \sum_{1 \le h \le k! Q^{k-1}} \min(Q, \|\alpha h\|^{-1})$$

- We have

$$\Delta_{k-1}\big(\Psi(x); h_1, \ldots, h_{k-1}\big) = \alpha k! h_1 \ldots h_{k-1} x + \alpha \frac{k!}{2} h_1 \ldots h_{k-1}(h_1 + \cdots h_{k-1}) + \alpha_1 h_1 \ldots h_{k-1}.$$

- Thus $T_{k-1} \ll \min(Q, \|\alpha k! h_1 \ldots h_{k-1}\|^{-1})$.
- We apply this to the previous theorem and separate out the $\ll Q^{k-2}$ terms for which $h_1 \ldots h_{k-1} = 0$.
- Thus $|T(\Psi)|^{2^{k-1}}$

$$\ll Q^{2^{k-1}-1} + Q^{2^{k-1}-k+\varepsilon} \sum_{1 \le h \le k! Q^{k-1}} \min(Q, \|\alpha h\|^{-1})$$

- The result now follows from Lemma 8.2.

- OK, so we have a bound for the sup norm on the minor arcs, but the best that we can get is

$$\ll Q^{1-2^{1-k}+\varepsilon}.$$

- OK, so we have a bound for the sup norm on the minor arcs, but the best that we can get is

$$\ll Q^{1-2^{1-k}+\varepsilon}.$$

- and for k of any size this is not very good.

- OK, so we have a bound for the sup norm on the minor arcs, but the best that we can get is

$$\ll Q^{1-2^{1-k}+\varepsilon}.$$

- and for k of any size this is not very good.
- Typically we will have $Q = \lfloor n^{1/k} \rfloor$ and we will be trying to save $n = Q^k$.

- OK, so we have a bound for the sup norm on the minor arcs, but the best that we can get is

$$\ll Q^{1-2^{1-k}+\varepsilon}.$$

- and for k of any size this is not very good.
- Typically we will have $Q = \lfloor n^{1/k} \rfloor$ and we will be trying to save $n = Q^k$.
- With only this available one can see why Hardy and Littlewood could only obtain

$$G(k) \le (k-2)2^{k-1} + 5.$$

- OK, so we have a bound for the sup norm on the minor arcs, but the best that we can get is
$$\ll Q^{1-2^{1-k}+\varepsilon}.$$
- and for k of any size this is not very good.
- Typically we will have $Q = \lfloor n^{1/k} \rfloor$ and we will be trying to save $n = Q^k$.
- With only this available one can see why Hardy and Littlewood could only obtain
$$G(k) \le (k-2)2^{k-1} + 5.$$
- We need a strong mean value theorem.

- OK, so we have a bound for the sup norm on the minor arcs, but the best that we can get is

$$\ll Q^{1-2^{1-k}+\varepsilon}.$$

- and for k of any size this is not very good.
- Typically we will have $Q = \lfloor n^{1/k} \rfloor$ and we will be trying to save $n = Q^k$.
- With only this available one can see why Hardy and Littlewood could only obtain

$$G(k) \le (k-2)2^{k-1} + 5.$$

- We need a strong mean value theorem.
- Strangely it took more than 20 years before Hua came up with another application of Weyl differencing which is a bit better.

- OK, so we have a bound for the sup norm on the minor arcs, but the best that we can get is

$$\ll Q^{1-2^{1-k}+\varepsilon}.$$

- and for k of any size this is not very good.
- Typically we will have $Q = \lfloor n^{1/k} \rfloor$ and we will be trying to save $n = Q^k$.
- With only this available one can see why Hardy and Littlewood could only obtain

$$G(k) \le (k-2)2^{k-1} + 5.$$

- We need a strong mean value theorem.
- Strangely it took more than 20 years before Hua came up with another application of Weyl differencing which is a bit better.
- By then Vinogradov had come up with something better for large k, but for $k = 3$ or 4 it is still close to the best that we know.

- **Theorem 8.5 [Hua's Lemma, 1938]** *Suppose that* $1 \le j \le k$. *Then* $\int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.

- **Theorem 8.5 [Hua's Lemma, 1938]** *Suppose that* $1 \le j \le k$. *Then* $\int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- **Proof.** This is also by induction on j.

- **Theorem 8.5 [Hua's Lemma, 1938]** *Suppose that* $1 \le j \le k$. *Then* $\int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- **Proof.** This is also by induction on j.
- The case $j = 1$ is immediate from Parseval's identity.

- **Theorem 8.5 [Hua's Lemma, 1938]** *Suppose that* $1 \le j \le k$. *Then* $\displaystyle\int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- **Proof.** This is also by induction on j.
- The case $j = 1$ is immediate from Parseval's identity.
- Now assume the j-th case. By Theorem 8.3, $|f(\alpha)|^{2^j}$

$$\le (2N)^{2^j - j - 1} \sum_{\mathbf{h} \in [1-N, N-1]^j} \sum_{x \in I_j} e\big(\alpha h_1 \dots h_j p_j(x; \mathbf{h})\big)$$

where p_j is of degree $k - j$ with integer coefficients and leading coefficient $\frac{k!}{(k-j)!}$.

- **Theorem 8.5 [Hua's Lemma, 1938]** *Suppose that* $1 \le j \le k$. *Then* $\int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- **Proof.** This is also by induction on j.
- The case $j = 1$ is immediate from Parseval's identity.
- Now assume the j-th case. By Theorem 8.3, $|f(\alpha)|^{2^j}$

$$\le (2N)^{2^j - j - 1} \sum_{\mathbf{h} \in [1-N, N-1]^j} \sum_{x \in I_j} e\big(\alpha h_1 \dots h_j p_j(x; \mathbf{h})\big)$$

  where p_j is of degree $k - j$ with integer coefficients and leading coefficient $\frac{k!}{(k-j)!}$.
- We collect together the terms with $h_1 \dots h_j p_j(x; \mathbf{h}) = g$ and let $c(g)$ be the number of such $h_1, \dots, h_j, x$.

- **Theorem 8.5 [Hua's Lemma, 1938]** *Suppose that* $1 \le j \le k$. *Then* $\int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- **Proof.** This is also by induction on j.
- The case $j = 1$ is immediate from Parseval's identity.
- Now assume the j-th case. By Theorem 8.3, $|f(\alpha)|^{2^j}$

$$\le (2N)^{2^j - j - 1} \sum_{\mathbf{h} \in [1-N, N-1]^j} \sum_{x \in I_j} e\big(\alpha h_1 \dots h_j p_j(x; \mathbf{h})\big)$$

where p_j is of degree $k - j$ with integer coefficients and leading coefficient $\frac{k!}{(k-j)!}$.
- We collect together the terms with $h_1 \dots h_j p_j(x; \mathbf{h}) = g$ and let $c(g)$ be the number of such $h_1, \dots, h_j, x$.
- Then $|f(\alpha)|^{2^j} \le (2N)^{2^j - j - 1} \sum_{g \ll N^k} c(g) e(\alpha g)$ and

$$c(g) \ll N^{\varepsilon}\ (g \ne 0),\ c(0) \ll N^j.$$

- **Theorem 8.5 [Hua's Lemma, 1938]** *Suppose that* $1 \le j \le k$. *Then* $\int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- **Proof.** This is also by induction on j.
- The case $j = 1$ is immediate from Parseval's identity.
- Now assume the j-th case. By Theorem 8.3, $|f(\alpha)|^{2^j}$

$$\le (2N)^{2^j - j - 1} \sum_{\mathbf{h} \in [1-N, N-1]^j} \sum_{x \in I_j} e\big(\alpha h_1 \dots h_j p_j(x; \mathbf{h})\big)$$

where p_j is of degree $k - j$ with integer coefficients and leading coefficient $\frac{k!}{(k-j)!}$.
- We collect together the terms with $h_1 \dots h_j p_j(x; \mathbf{h}) = g$ and let $c(g)$ be the number of such $h_1, \dots, h_j, x$.
- Then $|f(\alpha)|^{2^j} \le (2N)^{2^j - j - 1} \sum_{g \ll N^k} c(g) e(\alpha g)$ and

$$c(g) \ll N^{\varepsilon}\ (g \ne 0),\ c(0) \ll N^j.$$

- We also have $\sum_g c(g) \ll N^{j+1}$.

- Also $|f(\alpha)|^{2^j} = f(\alpha)^{2^{j-1}} f(-\alpha)^{2^{j-1}} = \sum_g b(g)e(-\alpha g)$ where $b(g)$ is the number of solutions of
$$x_1^k + \cdots x_J^k = y_1^k + \cdots y_J^k \, (x_j, y_j \le N, J = 2^{j-1})$$

- Also $|f(\alpha)|^{2^j} = f(\alpha)^{2^{j-1}} f(-\alpha)^{2^{j-1}} = \sum_g b(g)e(-\alpha g)$ where $b(g)$ is the number of solutions of
$$x_1^k + \cdots x_J^k = y_1^k + \cdots y_J^k \,(x_j, y_j \le N, J = 2^{j-1})$$
- Then $\sum_g b(g) = f(O)^{2^j} = N^{2^j}$ and on the inductive hypothesis $b(0) = \int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.

- Also $|f(\alpha)|^{2^j} = f(\alpha)^{2^{j-1}} f(-\alpha)^{2^{j-1}} = \sum_g b(g)e(-\alpha g)$ where $b(g)$ is the number of solutions of
$$x_1^k + \cdots x_J^k = y_1^k + \cdots y_J^k \,(x_j, y_j \le N, J = 2^{j-1})$$
- Then $\sum_g b(g) = f(O)^{2^j} = N^{2^j}$ and on the inductive hypothesis $b(0) = \int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- We have
$$\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha \ll N^{2^j - j - 1} \int_0^1 |f(\alpha)|^{2^j} \sum_g c(g)e(\alpha g) d\alpha$$

- Also $|f(\alpha)|^{2^j} = f(\alpha)^{2^{j-1}} f(-\alpha)^{2^{j-1}} = \sum_g b(g)e(-\alpha g)$ where $b(g)$ is the number of solutions of
$$x_1^k + \cdots x_J^k = y_1^k + \cdots y_J^k \, (x_j, y_j \le N, J = 2^{j-1})$$
- Then $\sum_g b(g) = f(O)^{2^j} = N^{2^j}$ and on the inductive hypothesis $b(0) = \int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- We have
$$\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha \ll N^{2^j - j - 1} \int_0^1 |f(\alpha)|^{2^j} \sum_g c(g)e(\alpha g) d\alpha$$
- and so by Parseval
$$\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha \ll N^{2^j - j - 1} \sum_g b(g)c(g).$$

- Also $|f(\alpha)|^{2^j} = f(\alpha)^{2^{j-1}} f(-\alpha)^{2^{j-1}} = \sum_g b(g)e(-\alpha g)$ where $b(g)$ is the number of solutions of
$$x_1^k + \cdots x_J^k = y_1^k + \cdots y_J^k \; (x_j, y_j \le N, J = 2^{j-1})$$
- Then $\sum_g b(g) = f(O)^{2^j} = N^{2^j}$ and on the inductive hypothesis $b(0) = \int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- We have
$$\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha \ll N^{2^j - j - 1} \int_0^1 |f(\alpha)|^{2^j} \sum_g c(g)e(\alpha g) d\alpha$$
- and so by Parseval
$$\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha \ll N^{2^j - j - 1} \sum_g b(g)c(g).$$
- The term $g = 0$ is $\ll N^{2^j - j - 1} N^{2^j - j + \varepsilon} N^j = N^{2^{j+1} - j - 1 + \varepsilon}$

- Also $|f(\alpha)|^{2^j} = f(\alpha)^{2^{j-1}} f(-\alpha)^{2^{j-1}} = \sum_g b(g)e(-\alpha g)$ where $b(g)$ is the number of solutions of
$$x_1^k + \cdots x_J^k = y_1^k + \cdots y_J^k \, (x_j, y_j \le N, J = 2^{j-1})$$
- Then $\sum_g b(g) = f(O)^{2^j} = N^{2^j}$ and on the inductive hypothesis $b(0) = \int_0^1 |f(\alpha)|^{2^j} d\alpha \ll N^{2^j - j + \varepsilon}$.
- We have
$$\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha \ll N^{2^j - j - 1} \int_0^1 |f(\alpha)|^{2^j} \sum_g c(g) e(\alpha g) d\alpha$$
- and so by Parseval
$$\int_0^1 |f(\alpha)|^{2^{j+1}} d\alpha \ll N^{2^j - j - 1} \sum_g b(g) c(g).$$
- The term $g = 0$ is $\ll N^{2^j - j - 1} N^{2^j - j + \varepsilon} N^j = N^{2^{j+1} - j - 1 + \varepsilon}$
- and those $g \ne 0 \ll N^{2^j - j - 1} \sum_g b(g) N^\varepsilon \ll N^{2^{j+1} - j - 1 + \varepsilon}$ which completes the proof.

- Recall we are counting the number of solutions of

$$x_1^k + \cdots x_s^k = n.$$

- Recall we are counting the number of solutions of
$$x_1^k + \cdots x_s^k = n.$$
- For the major arcs we will need to have q somewhat smaller than $n^{1/k}$, so we define
$$P = n^{\nu/k}$$
for some smallish ν which could be
$$\nu = \frac{1}{100}.$$

- Recall we are counting the number of solutions of
$$x_1^k + \cdots x_s^k = n.$$
- For the major arcs we will need to have q somewhat smaller than $n^{1/k}$, so we define
$$P = n^{\nu/k}$$
for some smallish ν which could be
$$\nu = \frac{1}{100}.$$
- Now we define a typical major arc to be an interval
$$\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}.$$

- Recall we are counting the number of solutions of
$$x_1^k + \cdots x_s^k = n.$$
- For the major arcs we will need to have q somewhat smaller than $n^{1/k}$, so we define
$$P = n^{\nu/k}$$
for some smallish ν which could be
$$\nu = \frac{1}{100}.$$
- Now we define a typical major arc to be an interval
$$\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}.$$
- If $a/q \ne a'/q'$, then $|a/q - a'/q'| = |aq' - a'q|/(qq') \ge 1/(qq')$.

- Recall we are counting the number of solutions of

$$x_1^k + \cdots x_s^k = n.$$

- For the major arcs we will need to have q somewhat smaller than $n^{1/k}$, so we define

$$P = n^{\nu/k}$$

for some smallish ν which could be

$$\nu = \frac{1}{100}.$$

- Now we define a typical major arc to be an interval

$$\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}.$$

- If $a/q \ne a'/q'$, then $|a/q - a'/q'| = |aq' - a'q|/(qq') \ge 1/(qq')$.
- Thus certainly the major arcs will be disjoint when $q \le P$.

- Recall we are counting the number of solutions of
$$x_1^k + \cdots x_s^k = n.$$
- For the major arcs we will need to have q somewhat smaller than $n^{1/k}$, so we define
$$P = n^{\nu/k}$$
for some smallish ν which could be
$$\nu = \frac{1}{100}.$$
- Now we define a typical major arc to be an interval
$$\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}.$$
- If $a/q \ne a'/q'$, then $|a/q - a'/q'| = |aq' - a'q|/(qq') \ge 1/(qq')$.
- Thus certainly the major arcs will be disjoint when $q \le P$.
- Hence we define $\mathfrak{M}$ to be their union with $1 \le a \le q \le P$ and $(a, q) = 1$.

- Recall we are counting the number of solutions of
$$x_1^k + \cdots x_s^k = n.$$
- For the major arcs we will need to have q somewhat smaller than $n^{1/k}$, so we define
$$P = n^{\nu/k}$$
for some smallish ν which could be
$$\nu = \frac{1}{100}.$$
- Now we define a typical major arc to be an interval
$$\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}.$$
- If $a/q \ne a'/q'$, then $|a/q - a'/q'| = |aq' - a'q|/(qq') \ge 1/(qq')$.
- Thus certainly the major arcs will be disjoint when $q \le P$.
- Hence we define $\mathfrak{M}$ to be their union with $1 \le a \le q \le P$ and $(a, q) = 1$.
- Then $\mathfrak{M} \subset \mathfrak{U} = [\tau, 1 + \tau]$ where $\tau = P/N$

- Recall we are counting the number of solutions of
$$x_1^k + \cdots x_s^k = n.$$
- For the major arcs we will need to have q somewhat smaller than $n^{1/k}$, so we define
$$P = n^{\nu/k}$$
for some smallish ν which could be
$$\nu = \frac{1}{100}.$$
- Now we define a typical major arc to be an interval
$$\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}.$$
- If $a/q \ne a'/q'$, then $|a/q - a'/q'| = |aq' - a'q|/(qq') \ge 1/(qq')$.
- Thus certainly the major arcs will be disjoint when $q \le P$.
- Hence we define $\mathfrak{M}$ to be their union with $1 \le a \le q \le P$ and $(a, q) = 1$.
- Then $\mathfrak{M} \subset \mathfrak{U} = [\tau, 1+\tau]$ where $\tau = P/N$
- and we define the minor arcs by $\mathfrak{m} = \mathfrak{U} \setminus \mathfrak{M}$.

- **Theorem 8.6.** *There is a positive number δ such that if $s > 2^k$, then* $\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.

- **Theorem 8.6.** *There is a positive number δ such that if $s > 2^k$, then* $\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.
- **Proof.** By Dirichlet's theorem on diophantine approximation given $\alpha \in \mathfrak{m}$ there are a, q with $(a,q) = 1$, $q \le n/P$ and $|\alpha - a/q| \le \frac{1}{q(1+\lfloor n/P \rfloor}$ $< P/(qn)$.

- **Theorem 8.6.** *There is a positive number* δ *such that if* $s > 2^k$*, then* $\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.
- **Proof.** By Dirichlet's theorem on diophantine approximation given $\alpha \in \mathfrak{m}$ there are a, q with $(a,q)=1$, $q \le n/P$ and $|\alpha - a/q| \le \frac{1}{q(1+\lfloor n/P \rfloor)} < P/(qn)$.
- If $q \le P$, then we cannot have $a > q$, for then $\alpha > \frac{a}{q} - \frac{P}{qn} \ge 1 + \frac{1}{q} - \frac{P}{qn} > 1 + \frac{P}{n}$ contradicting $\alpha \in \mathfrak{U}$, so $a \le q$.

- **Theorem 8.6.** *There is a positive number* δ *such that if* $s > 2^k$*, then* $\displaystyle\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.
- **Proof.** By Dirichlet's theorem on diophantine approximation given $\alpha \in \mathfrak{m}$ there are a, q with $(a, q) = 1$, $q \le n/P$ and $|\alpha - a/q| \le \frac{1}{q(1+\lfloor n/P \rfloor}} < P/(qn)$.
- If $q \le P$, then we cannot have $a > q$, for then $\alpha > \frac{a}{q} - \frac{P}{qn} \ge 1 + \frac{1}{q} - \frac{P}{qn} > 1 + \frac{P}{n}$ contradicting $\alpha \in \mathfrak{U}$, so $a \le q$.
- Also $\alpha \in \mathfrak{U}$ implies that $P/n \le \alpha$ so $a \ge 1$.

- **Theorem 8.6.** *There is a positive number δ such that if $s > 2^k$, then* $\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.
- **Proof.** By Dirichlet's theorem on diophantine approximation given $\alpha \in \mathfrak{m}$ there are a, q with $(a, q) = 1$, $q \le n/P$ and $|\alpha - a/q| \le \frac{1}{q(1+\lfloor n/P \rfloor)} < P/(qn)$.
- If $q \le P$, then we cannot have $a > q$, for then $\alpha > \frac{a}{q} - \frac{P}{qn} \ge 1 + \frac{1}{q} - \frac{P}{qn} > 1 + \frac{P}{n}$ contradicting $\alpha \in \mathfrak{U}$, so $a \le q$.
- Also $\alpha \in \mathfrak{U}$ implies that $P/n \le \alpha$ so $a \ge 1$.
- Therefore $q \le P$ would imply that $\alpha \in \mathfrak{M}$.

- **Theorem 8.6.** *There is a positive number δ such that if $s > 2^k$, then* $\displaystyle\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.
- **Proof.** By Dirichlet's theorem on diophantine approximation given $\alpha \in \mathfrak{m}$ there are a, q with $(a,q)=1$, $q \le n/P$ and $|\alpha - a/q| \le \frac{1}{q(1+\lfloor n/P \rfloor)} < P/(qn)$.
- If $q \le P$, then we cannot have $a > q$, for then $\alpha > \frac{a}{q} - \frac{P}{qn} \ge 1 + \frac{1}{q} - \frac{P}{qn} > 1 + \frac{P}{n}$ contradicting $\alpha \in \mathfrak{U}$, so $a \le q$.
- Also $\alpha \in \mathfrak{U}$ implies that $P/n \le \alpha$ so $a \ge 1$.
- Therefore $q \le P$ would imply that $\alpha \in \mathfrak{M}$.
- Hence we have $P < q \le n/P$.

- **Theorem 8.6.** *There is a positive number δ such that if $s > 2^k$, then* $\displaystyle\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.
- **Proof.** By Dirichlet's theorem on diophantine approximation given $\alpha \in \mathfrak{m}$ there are a, q with $(a, q) = 1$, $q \le n/P$ and $|\alpha - a/q| \le \frac{1}{q(1+\lfloor n/P \rfloor}< P/(qn)$.
- If $q \le P$, then we cannot have $a > q$, for then $\alpha > \frac{a}{q} - \frac{P}{qn} \ge 1 + \frac{1}{q} - \frac{P}{qn} > 1 + \frac{P}{n}$ contradicting $\alpha \in \mathfrak{U}$, so $a \le q$.
- Also $\alpha \in \mathfrak{U}$ implies that $P/n \le \alpha$ so $a \ge 1$.
- Therefore $q \le P$ would imply that $\alpha \in \mathfrak{M}$.
- Hence we have $P < q \le n/P$.
- Thus, Weyl's inequality $f(\alpha) \ll n^{1/k-\delta}$ for a suitable small $\delta = \delta(k)$.

- **Theorem 8.6.** *There is a positive number δ such that if $s > 2^k$, then* $\displaystyle\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{\frac{s}{k}-1-\delta}$.
- **Proof.** By Dirichlet's theorem on diophantine approximation given $\alpha \in \mathfrak{m}$ there are a, q with $(a,q)=1$, $q \le n/P$ and $|\alpha - a/q| \le \frac{1}{q(1+\lfloor n/P \rfloor)} < P/(qn)$.
- If $q \le P$, then we cannot have $a > q$, for then $\alpha > \frac{a}{q} - \frac{P}{qn} \ge 1 + \frac{1}{q} - \frac{P}{qn} > 1 + \frac{P}{n}$ contradicting $\alpha \in \mathfrak{U}$, so $a \le q$.
- Also $\alpha \in \mathfrak{U}$ implies that $P/n \le \alpha$ so $a \ge 1$.
- Therefore $q \le P$ would imply that $\alpha \in \mathfrak{M}$.
- Hence we have $P < q \le n/P$.
- Thus, Weyl's inequality $f(\alpha) \ll n^{1/k-\delta}$ for a suitable small $\delta = \delta(k)$.
- Therefore, by Hua's Lemma with $j = k$ we have for any $s > k$

$$\int_{\mathfrak{m}} |f(\alpha)|^s d\alpha \ll n^{2^k/k-1+\varepsilon}(n^{1/k-\delta})^{s-2^k}.$$

- Now we are in the endgame.

- Now we are in the endgame.
- But as usual this will be the longest part.

- Now we are in the endgame.
- But as usual this will be the longest part.
- On each major arc we will use the approximation

$$f(\alpha) = q^{-1}S(q,a)v(\beta) + O(q + qn|\beta|)$$

- Now we are in the endgame.
- But as usual this will be the longest part.
- On each major arc we will use the approximation

$$f(\alpha) = q^{-1}S(q,a)v(\beta) + O(q + qn|\beta|)$$

- and so we need to understand the properties of $S(q,a)$ and $v(\beta)$.

- Now we are in the endgame.
- But as usual this will be the longest part.
- On each major arc we will use the approximation

$$f(\alpha) = q^{-1}S(q,a)v(\beta) + O(q + qn|\beta|)$$

- and so we need to understand the properties of $S(q,a)$ and $v(\beta)$.
- The latter of these is the easiest to deal with.

- Now we are in the endgame.
- But as usual this will be the longest part.
- On each major arc we will use the approximation

$$f(\alpha) = q^{-1}S(q,a)v(\beta) + O(q + qn|\beta|)$$

- and so we need to understand the properties of $S(q,a)$ and $v(\beta)$.
- The latter of these is the easiest to deal with.
- It is periodic with period 1, so we can concentrate on the interval $[-1/2, 1/2]$.

- **Lemma 8.7.** *Suppose that* $|\beta| \le \frac{1}{2}$. *Then*

$$v(\beta) \ll \min(n^{1/k}, |\beta|^{-1/k}).$$

- **Lemma 8.7.** *Suppose that* $|\beta| \le \frac{1}{2}$. *Then*

$$v(\beta) \ll \min(n^{1/k}, |\beta|^{-1/k}).$$

- **Proof.** We already saw in the proof of Theorem 8.1 that

$$\sum_{y \le x} k^{-1} y^{1/k-1} = x^{1/k} + O(1).$$

- We can suppose that $|\beta| > n^{-1}$.

- We can suppose that $|\beta| > n^{-1}$.
- Let $x = \lfloor |\beta|^{-1} \rfloor$. Then $\sum_{y \le x} k^{-1} y^{1/k-1} \ll |\beta|^{1/k}$

- We can suppose that $|\beta| > n^{-1}$.
- Let $x = \lfloor |\beta|^{-1} \rfloor$. Then $\displaystyle\sum_{y\le x} k^{-1}y^{1/k-1} \ll |\beta|^{1/k}$
- and $\displaystyle\sum_{x+1\le y\le n} k^{-1}y^{1/k-1}e(\beta y) = k^{-1}n^{\frac{1}{k}-1}\sum_{x+1\le y\le n} e(\beta y)$

$$-\int_{x+1}^{n}\left(\frac{1}{k^2}-\frac{1}{k}\right)u^{\frac{1}{k}-2}\sum_{x+1\le y\le u} e(\beta y)du$$

$$\ll k^{-1}n^{1/k-1}|\beta|^{-1} + \int_{x+1}^{n} k^{-1}(1-k^{-1})u^{1/k-2}|\beta|^{-1}du$$

$$\ll (x+1)^{1/k-1}|\beta|^{-1} \ll |\beta|^{-1/k}.$$

- The Gauss sum

$$S(q,a)=\sum_{x=1}^{q} e(ax^k/q)$$

is more interesting.

- The Gauss sum

$$S(q,a)=\sum_{x=1}^{q} e(ax^k/q)$$

is more interesting.

- There is a crude, but adequate for our purposes, bound for this.

- The Gauss sum

$$S(q,a)=\sum_{x=1}^{q} e(ax^k/q)$$

is more interesting.
- There is a crude, but adequate for our purposes, bound for this.
- **Lemma 8.8.** *Suppose that $a\in\mathbb{Z}, q\in\mathbb{N}$ and $(a,q)=1$. Then*

$$S(q,a)\ll q^{1-2^{1-k}+\varepsilon}$$

- The Gauss sum
$$S(q,a)=\sum_{x=1}^{q} e(ax^k/q)$$
is more interesting.
- There is a crude, but adequate for our purposes, bound for this.
- **Lemma 8.8.** *Suppose that* $a\in\mathbb{Z}, q\in\mathbb{N}$ *and* $(a,q)=1$. *Then*
$$S(q,a)\ll q^{1-2^{1-k}+\varepsilon}$$
- **Proof.** This follows immediately by Weyl's inequality, with $Q=q$, $\alpha=a/q$, $\Psi=aq^{-1}x^k$.

- The Gauss sum

$$S(q,a)=\sum_{x=1}^{q} e(ax^k/q)$$

  is more interesting.

- There is a crude, but adequate for our purposes, bound for this.

- **Lemma 8.8.** *Suppose that* $a\in\mathbb{Z}$,$q\in\mathbb{N}$ *and* $(a,q)=1$. *Then*

$$S(q,a)\ll q^{1-2^{1-k}+\varepsilon}$$

- **Proof.** This follows immediately by Weyl's inequality, with $Q=q$, $\alpha=a/q$, $\Psi=aq^{-1}x^k$.

- It is possible to do much better than this, and we may examine this later.

- Importantly $S(q, a)$ has a multiplicative property which we will use later.

- Importantly $S(q,a)$ has a multiplicative property which we will use later.
- **Lemma 8.9.** *Suppose that* $a,b\in\mathbb{Z}$, $q,r\in\mathbb{N}$ *and* $(a,q)=(bmr)=(q,r)=1$. *Then*

$$S(qr, ar+bq) = S(q,a)S(r,b).$$

- Importantly $S(q, a)$ has a multiplicative property which we will use later.
- **Lemma 8.9.** *Suppose that* $a, b \in \mathbb{Z}$, $q, r \in \mathbb{N}$ *and* $(a, q) = (bmr) = (q, r) = 1$. *Then*

$$S(qr, ar + bq) = S(q, a)S(r, b).$$

- **Proof.** $tr + uq$ runs over a complete set of residues modulo qr and t and u do modulo q and r respectively.

- Importantly $S(q,a)$ has a multiplicative property which we will use later.
- **Lemma 8.9.** *Suppose that* $a,b\in\mathbb{Z}$, $q,r\in\mathbb{N}$ *and* $(a,q)=(bmr)=(q,r)=1$. *Then*

$$S(qr,ar+bq)=S(q,a)S(r,b).$$

- **Proof.** $tr+uq$ runs over a complete set of residues modulo qr and t and u do modulo q and r respectively.
- Thus

$$S(qr,ar+bq)=\sum_{t=1}^{q}\sum_{u=1}^{r}e(ar^kt^k/q+bq^ku^k/r)$$

- Importantly $S(q,a)$ has a multiplicative property which we will use later.
- **Lemma 8.9.** *Suppose that $a, b \in \mathbb{Z}$, $q, r \in \mathbb{N}$ and $(a,q) = (bmr) = (q,r) = 1$. Then*

$$S(qr, ar + bq) = S(q,a)S(r,b).$$

- **Proof.** $tr + uq$ runs over a complete set of residues modulo qr and t and u do modulo q and r respectively.
- Thus

$$S(qr, ar + bq) = \sum_{t=1}^{q} \sum_{u=1}^{r} e(ar^k t^k/q + bq^k u^k/r)$$

- and tr and uq run over complete sets of residues modulo q and r as t and u do respectively.

- Importantly $S(q,a)$ has a multiplicative property which we will use later.
- **Lemma 8.9.** *Suppose that* $a,b\in\mathbb{Z}$, $q,r\in\mathbb{N}$ *and* $(a,q)=(bmr)=(q,r)=1$. *Then*

$$S(qr,ar+bq)=S(q,a)S(r,b).$$

- **Proof.** $tr+uq$ runs over a complete set of residues modulo qr and t and u do modulo q and r respectively.
- Thus

$$S(qr,ar+bq)=\sum_{t=1}^{q}\sum_{u=1}^{r}e(ar^kt^k/q+bq^ku^k/r)$$

- and tr and uq run over complete sets of residues modulo q and r as t and u do respectively.
- Thus it suffices to understand $S(q,a)$ when q is a power of a prime. It turns out that $S(q,a)\ll q^{1-1/k}$ and that sometimes the sum is this large, although often it is smaller, as we saw in homework 4 in the prime case.

- Recall that $P = n^{\nu/k}$ for some smallish ν and we defined the major arcs by $\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}$ and took $\mathfrak{M}$ to be their union with $1 \le a \le q \le P$ and $(a, q) = 1$.

- Recall that $P = n^{\nu/k}$ for some smallish ν and we defined the major arcs by $\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}$ and took $\mathfrak{M}$ to be their union with $1 \le a \le q \le P$ and $(a, q) = 1$.
- Now write $\beta = \alpha - a/q$ and $E = f(\alpha) - q^{-1}S(q, a)v(\beta)$ so that $E \ll q + qn|\beta| \ll P$.

- Recall that $P = n^{\nu/k}$ for some smallish ν and we defined the major arcs by $\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}$ and took $\mathfrak{M}$ to be their union with $1 \le a \le q \le P$ and $(a,q) = 1$.
- Now write $\beta = \alpha - a/q$ and $E = f(\alpha) - q^{-1}S(q,a)v(\beta)$ so that $E \ll q + qn|\beta| \ll P$.
- Then $f(\alpha)^s = \big(q^{-1}S(q,a)v(\beta) + E\big)^s$ so that, when $(a,q) = 1$,

$$\begin{aligned} f(\alpha)^s - q^{-s}S(q,a)^s v(\beta)^s & \\ &\ll (q^{-1}|S(q,a)v(\beta)|)^{s-1}|E| + |E|^s \\ &\ll n^{\frac{s-1}{k}}P + P^s \ll n^{\frac{s-1}{k}}P. \end{aligned}$$

- Recall that $P = n^{\nu/k}$ for some smallish ν and we defined the major arcs by $\mathfrak{M} = \{\alpha : |\alpha - a/q| \le P/n\}$ and took $\mathfrak{M}$ to be their union with $1 \le a \le q \le P$ and $(a,q) = 1$.
- Now write $\beta = \alpha - a/q$ and $E = f(\alpha) - q^{-1}S(q,a)v(\beta)$ so that $E \ll q + qn|\beta| \ll P$.
- Then $f(\alpha)^s = \big(q^{-1}S(q,a)v(\beta) + E\big)^s$ so that, when $(a,q) = 1$,

$$f(\alpha)^s - q^{-s}S(q,a)^s v(\beta)^s \ll (q^{-1}|S(q,a)v(\beta)|)^{s-1}|E| + |E|^s \ll n^{\frac{s-1}{k}}P + P^s \ll n^{\frac{s-1}{k}}P.$$

- Now integrating over $\mathfrak{M}(q,a)$ we obtain

$$\int_{\mathfrak{M}(q,a)} \big(f(\alpha)^s - q^{-s}S(q,a)^s v(\beta)^s\big)e(-\alpha n)d\alpha \ll q^{-1}n^{\frac{s-1}{k}-1}P^2.$$

- $\displaystyle\int_{\mathfrak{M}(q,a)} \left(f(\alpha)^s - q^{-s}S(q,a)^s v(\beta)^s\right) e(-\alpha n)d\alpha \ll n^{\frac{s-1}{k}-1}P^2/q$

- $\displaystyle\int_{\mathfrak{M}(q,a)} \big(f(\alpha)^s - q^{-s}S(q,a)^s v(\beta)^s\big) e(-\alpha n)d\alpha \ll n^{\frac{s-1}{k}-1}P^2/q$
- Let $K = 2^{1-k}$. Then

$$q^{-s}|S(q,a)|^s \int_{\frac{P}{qn}\le|\beta|\le\frac12} |v(\beta)|^s d\beta \ll q^{-sK+\varepsilon}\int_{\frac{P}{qn}}^{\infty} \beta^{-s/k}d\beta$$
$$\ll q^{-sK+\varepsilon}(qn/P)^{\frac{s}{k}-1}.$$

- $\displaystyle\int_{\mathfrak{M}(q,a)} \big(f(\alpha)^s - q^{-s}S(q,a)^s v(\beta)^s\big)e(-\alpha n)d\alpha \ll n^{\frac{s-1}{k}-1}P^2/q$
- Let $K = 2^{1-k}$. Then

$$q^{-s}|S(q,a)|^s \int_{\frac{P}{qn}\le|\beta|\le\frac12} |v(\beta)|^s d\beta \ll q^{-sK+\varepsilon}\int_{\frac{P}{qn}}^{\infty} \beta^{-s/k}d\beta \ll q^{-sK+\varepsilon}(qn/P)^{\frac{s}{k}-1}.$$

- Thus

$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)d\alpha - \mathfrak{S}(n;P)\int_{-1/2}^{1/2} v(\beta)^s e(-n\beta)d\beta \ll \Delta$$

where $\displaystyle\mathfrak{S}(n;Q) = \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s}S(q,a)^s e(-an/q)$ and

$$\Delta = \sum_{q\le P}\left(n^{\frac{s-1}{k}-1}P^2 + q^{\frac{s}{k}-sK+\varepsilon}(n/P)^{\frac{s}{k}-1}\right).$$

- Thus

$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)d\alpha - \mathfrak{S}(n;P)\int_{-1/2}^{1/2} v(\beta)^s e(-n\beta)d\beta \ll \Delta$$

where $\mathfrak{S}(n;Q) = \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s}S(q,a)^s e(-an/q)$ and

$$\Delta = \sum_{q\le P}\left(n^{\frac{s-1}{k}-1}P^2 + q^{\frac{s}{k}-sK+\varepsilon}(n/P)^{\frac{s}{k}-1}\right).$$

- Thus
$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)d\alpha - \mathfrak{S}(n;P)\int_{-1/2}^{1/2} v(\beta)^s e(-n\beta)d\beta \ll \Delta$$
where $\mathfrak{S}(n;Q) = \sum_{q\le Q}\sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s}S(q,a)^s e(-an/q)$ and
$$\Delta = \sum_{q\le P}\left(n^{\frac{s-1}{k}-1}P^2 + q^{\frac{s}{k}-sK+\varepsilon}(n/P)^{\frac{s}{k}-1}\right).$$
- Thus $\Delta \ll n^{\frac{s-1}{k}-1}P^3 + n^{\frac{s}{k}-1}P^{1-sK+\varepsilon} \ll n^{\frac{s}{k}-1}P^{-1}$ provided that $sK > 2$, i.e. $s > 2^k$.

- Thus
$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n) d\alpha - \mathfrak{S}(n;P) \int_{-1/2}^{1/2} v(\beta)^s e(-n\beta) d\beta \ll \Delta$$
where $\mathfrak{S}(n;Q) = \sum_{q\le Q} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q)$ and
$$\Delta = \sum_{q\le P} \left(n^{\frac{s-1}{k}-1} P^2 + q^{\frac{s}{k}-sK+\varepsilon} (n/P)^{\frac{s}{k}-1} \right).$$
- Thus $\Delta \ll n^{\frac{s-1}{k}-1} P^3 + n^{\frac{s}{k}-1} P^{1-sK+\varepsilon} \ll n^{\frac{s}{k}-1} P^{-1}$ provided that $sK > 2$, i.e. $s > 2^k$.
- We also have $\sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q) \ll q^{1-sK+\varepsilon}$ and
so $\sum_{q>P} \left| \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q) \right| \ll P^{2-sK+\varepsilon} \ll n^{-\delta}$
for some $\delta > 0$ provided that $s > 2^k$.

- Finally $\displaystyle\int_{-1/2}^{1/2} v(\beta)^s e(-\beta n)d\beta \ll$

$$\int_{-1/2}^{1/2}\left(\frac{n}{1+n|\beta|}\right)^{s/k} d\beta \ll n^{\frac{s}{k}-1}.$$

- Finally $\displaystyle\int_{-1/2}^{1/2} v(\beta)^s e(-\beta n)d\beta \ll$
$$\int_{-1/2}^{1/2} \left(\frac{n}{1+n|\beta|}\right)^{s/k} d\beta \ll n^{\frac{s}{k}-1}.$$
- Thus we have shown that for some $\delta > 0$
$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)d\alpha - \mathfrak{S}(n)J(n) \ll n^{\frac{s}{k}-1-\delta}$$

where $\displaystyle\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q)$ and

$$J(n) = \int_{-1/2}^{1/2} v(\beta)^s e(-n\beta)d\beta$$

- Finally $\displaystyle\int_{-1/2}^{1/2} v(\beta)^s e(-\beta n)d\beta \ll$
$$\int_{-1/2}^{1/2}\left(\frac{n}{1+n|\beta|}\right)^{s/k} d\beta \ll n^{\frac{s}{k}-1}.$$
- Thus we have shown that for some $\delta > 0$
$$\int_{\mathfrak{M}} f(\alpha)^s e(-\alpha n)d\alpha - \mathfrak{S}(n)J(n) \ll n^{\frac{s}{k}-1-\delta}$$
where $\displaystyle\mathfrak{S}(n) = \sum_{q=1}^{\infty}\sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s}S(q,a)^s e(-an/q)$ and
$$J(n) = \int_{-1/2}^{1/2} v(\beta)^s e(-n\beta)d\beta$$
- Combining this with Theorem 8.6 we have

Theorem 8.10. *Suppose that $s > 2^k$. Then there is a $\delta > 0$ such that for every large n we have*
$$r_s(n) = \mathfrak{S}(n)J(n) + O(n^{\frac{s}{k}-1-\delta}).$$

- There are now two further tasks to perform.

- There are now two further tasks to perform.
- What is the size of $J(n)$?

- There are now two further tasks to perform.
- What is the size of $J(n)$?
- Hopefully bigger than $n^{\frac{s}{k}-1}$.

- There are now two further tasks to perform.
- What is the size of $J(n)$?
- Hopefully bigger than $n^{\frac{s}{k}-1}$.
- And do we always have $\mathfrak{S}(n) \gg 1$?

- There are now two further tasks to perform.
- What is the size of $J(n)$?
- Hopefully bigger than $n^{\frac{s}{k}-1}$.
- And do we always have $\mathfrak{S}(n) \gg 1$?
- The first holds when $s > k$, but the second can fail for quite big s.

- There are now two further tasks to perform.
- What is the size of $J(n)$?
- Hopefully bigger than $n^{\frac{s}{k}-1}$.
- And do we always have $\mathfrak{S}(n) \gg 1$?
- The first holds when $s > k$, but the second can fail for quite big s.
- For example $k = 4$ and $s = 15$.

- There are now two further tasks to perform.
- What is the size of $J(n)$?
- Hopefully bigger than $n^{\frac{s}{k}-1}$.
- And do we always have $\mathfrak{S}(n) \gg 1$?
- The first holds when $s > k$, but the second can fail for quite big s.
- For example $k = 4$ and $s = 15$.
- Fortunately it does hold when $s > 2^k$.

- To see that there is a problem when $k=4$ and $s=15$, observe first that if x is odd, say $x=2y+1$, then $x^4=(2y+1)^4 \equiv 1+4(2y)+6(2y)^2 = 1+8x(1+3x) \pmod{16}$.

- To see that there is a problem when $k = 4$ and $s = 15$, observe first that if x is odd, say $x = 2y + 1$, then $x^4 = (2y + 1)^4 \equiv 1 + 4(2y) + 6(2y)^2 = 1 + 8x(1 + 3x) \pmod{16}$.
- Now $x(1 + 3x)$ is always even, so any fourth power is always $\equiv 0$ or $1 \pmod{16}$.

- To see that there is a problem when $k = 4$ and $s = 15$, observe first that if x is odd, say $x = 2y + 1$, then $x^4 = (2y + 1)^4 \equiv 1 + 4(2y) + 6(2y)^2 = 1 + 8x(1 + 3x) \pmod{16}$.
- Now $x(1 + 3x)$ is always even, so any fourth power is always $\equiv 0$ or $1 \pmod{16}$.
- Now consider $n = 2^{4k} \times 31$.

- To see that there is a problem when $k = 4$ and $s = 15$, observe first that if x is odd, say $x = 2y + 1$, then $x^4 = (2y+1)^4 \equiv 1 + 4(2y) + 6(2y)^2 = 1 + 8x(1+3x) \pmod{16}$.
- Now $x(1+3x)$ is always even, so any fourth power is always $\equiv 0$ or $1 \pmod{16}$.
- Now consider $n = 2^{4k} \times 31$.
- If n is the sum of 15 fourth powers, then they must all be even, so $n2^{-4}$ is also the sum of 15 fourth powers, and so on.

- To see that there is a problem when $k=4$ and $s=15$, observe first that if x is odd, say $x=2y+1$, then $x^4=(2y+1)^4 \equiv 1+4(2y)+6(2y)^2=1+8x(1+3x) \pmod{16}$.
- Now $x(1+3x)$ is always even, so any fourth power is always $\equiv 0$ or $1 \pmod{16}$.
- Now consider $n=2^{4k}\times 31$.
- If n is the sum of 15 fourth powers, then they must all be even, so $n2^{-4}$ is also the sum of 15 fourth powers, and so on.
- Hence 31 would have to be the sum of 15 fourth powers.

- To see that there is a problem when $k=4$ and $s=15$, observe first that if x is odd, say $x=2y+1$, then $x^4=(2y+1)^4\equiv 1+4(2y)+6(2y)^2=1+8x(1+3x)$ (mod 16).
- Now $x(1+3x)$ is always even, so any fourth power is always $\equiv 0$ or $1 \pmod{16}$.
- Now consider $n=2^{4k}\times 31$.
- If n is the sum of 15 fourth powers, then they must all be even, so $n2^{-4}$ is also the sum of 15 fourth powers, and so on.
- Hence 31 would have to be the sum of 15 fourth powers.
- But it isn't!. You have $31<3^4$ so you can only use 1^4 and 2^4, and then you can only use at most one 2^4 and there are not enough 1^4 to add up to 31.

- The function $J(n)$ is easy to bound. By orthogonality we have
$$J(n) = \sum_{\substack{x_1,\dots,x_s \\ x_1+\dots+x_s=n}} k^{-s}(x_1\dots x_s)^{1/k-1}$$

- The function $J(n)$ is easy to bound. By orthogonality we have
$$J(n) = \sum_{\substack{x_1,\ldots,x_s\\ x_1+\cdots+x_s=n}} k^{-s}(x_1\ldots x_s)^{1/k-1}$$
- Consider $x_1,\ldots x_{s-1}$ with $n(s^{-1}-\delta) \geq x_j \leq n(s^{-1}+\delta)$.

- The function $J(n)$ is easy to bound. By orthogonality we have
$$J(n) = \sum_{\substack{x_1,\dots,x_s \\ x_1+\dots+x_s=n}} k^{-s}(x_1\dots x_s)^{1/k-1}$$
- Consider $x_1,\dots x_{s-1}$ with $n(s^{-1}-\delta) \geq x_j \leq n(s^{-1}+\delta)$.
- When $\delta < \frac{1}{s(s-1)}$ we would have
$$n\big(s^{-1}-\delta(s-1)\big) \leq n-(x_1+\dots+x_{s-1}) \leq n\big(s^{-1}+\delta(s-1)\big)$$

- The function $J(n)$ is easy to bound. By orthogonality we have
$$J(n) = \sum_{\substack{x_1,\dots,x_s \\ x_1+\dots+x_s=n}} k^{-s}(x_1 \dots x_s)^{1/k-1}$$
- Consider $x_1, \dots x_{s-1}$ with $n(s^{-1} - \delta) \geq x_j \leq n(s^{-1} + \delta)$.
- When $\delta < \frac{1}{s(s-1)}$ we would have
$$n\big(s^{-1} - \delta(s-1)\big) \leq n - (x_1 + \dots + x_{s-1}) \leq n\big(s^{-1} + \delta(s-1)\big)$$
- and so there is an x_s with
$$n\big(s^{-1} - \delta(s-1)\big) \leq x_s \leq n\big(s^{-1} + \delta(s-1)\big)$$
and $x_1 + \dots + x_s = n$.

- The function $J(n)$ is easy to bound. By orthogonality we have

$$J(n) = \sum_{\substack{x_1,\ldots,x_s \\ x_1+\cdots+x_s=n}} k^{-s}(x_1\ldots x_s)^{1/k-1}$$

- Consider $x_1,\ldots x_{s-1}$ with $n(s^{-1}-\delta) \geq x_j \leq n(s^{-1}+\delta)$.
- When $\delta < \frac{1}{s(s-1)}$ we would have

$$n\big(s^{-1}-\delta(s-1)\big) \leq n-(x_1+\cdots+x_{s-1}) \leq n\big(s^{-1}+\delta(s-1)\big)$$

- and so there is an x_s with

$$n\big(s^{-1}-\delta(s-1)\big) \leq x_s \leq n\big(s^{-1}+\delta(s-1)\big)$$

and $x_1+\cdots+x_s=n$.
- Thus these x_j contribute

$$\gg n^{s-1}(n^s)^{1/k-1} = n^{s/k-1}.$$

- The function $J(n)$ is easy to bound. By orthogonality we have

$$J(n) = \sum_{\substack{x_1,\dots,x_s \\ x_1+\dots+x_s=n}} k^{-s}(x_1 \dots x_s)^{1/k-1}$$

- Consider $x_1, \dots x_{s-1}$ with $n(s^{-1} - \delta) \ge x_j \le n(s^{-1} + \delta)$.
- When $\delta < \frac{1}{s(s-1)}$ we would have

$$n\big(s^{-1}-\delta(s-1)\big) \le n-(x_1+\dots+x_{s-1}) \le n\big(s^{-1}+\delta(s-1)\big)$$

- and so there is an x_s with

$$n\big(s^{-1} - \delta(s-1)\big) \le x_s \le n\big(s^{-1} + \delta(s-1)\big)$$

and $x_1 + \dots + x_s = n$.
- Thus these x_j contribute

$$\gg n^{s-1}(n^s)^{1/k-1} = n^{s/k-1}.$$

- Hence

$$n^{s/k-1} \ll J(n) \ll n^{s/k-1},$$

the upper bound coming from our bound for the integral.

- The above is a method which works in most circumstances. Here we can do better.

- The above is a method which works in most circumstances. Here we can do better.
- **Theorem 8.11.** *Suppose that $s \geq 2$. Then*

$$J(n) = \frac{\Gamma\left(1+\frac{1}{k}\right)^s}{\Gamma\left(\frac{s}{k}\right)} n^{\frac{s}{k}-1} + O\left(n^{\frac{s}{k}-1-\frac{1}{k}}\right).$$

- The above is a method which works in most circumstances. Here we can do better.
- **Theorem 8.11.** *Suppose that $s \ge 2$. Then*

$$J(n) = \frac{\Gamma\left(1+\frac{1}{k}\right)^s}{\Gamma\left(\frac{s}{k}\right)} n^{\frac{s}{k}-1} + O\left(n^{\frac{s}{k}-1-\frac{1}{k}}\right).$$

- **Proof.** This is by induction on s. Both the initial case $s = 2$ and the inductive step depend on the following lemma.

- The above is a method which works in most circumstances. Here we can do better.
- **Theorem 8.11.** *Suppose that* $s \ge 2$. *Then*

$$J(n) = \frac{\Gamma\left(1+\frac{1}{k}\right)^s}{\Gamma\left(\frac{s}{k}\right)} n^{\frac{s}{k}-1} + O\left(n^{\frac{s}{k}-1-\frac{1}{k}}\right).$$

- **Proof.** This is by induction on s. Both the initial case $s = 2$ and the inductive step depend on the following lemma.
- **Lemma 8.12.** *Suppose that* α, β *are real numbers with* $\alpha \ge \beta > 0$ *and* $\beta \le 1$. *Then*

$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$

- **Lemma 8.12.** *Suppose that* α, β *are real numbers with* $\alpha \ge \beta > 0$ *and* $\beta \le 1$. *Then*

$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$

- **Lemma 8.12.** *Suppose that α, β are real numbers with $\alpha \ge \beta > 0$ and $\beta \le 1$. Then*

$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$

- **Proof.** If $\alpha + \beta = 2$ and $\alpha = 1$, then the sum in the lemma is trivial.

- **Lemma 8.12.** *Suppose that α, β are real numbers with $\alpha \ge \beta > 0$ and $\beta \le 1$. Then*
$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$
- **Proof.** If $\alpha + \beta = 2$ and $\alpha = 1$, then the sum in the lemma is trivial.
- When $\alpha + \beta \ne 2$, consider $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$.

- **Lemma 8.12.** *Suppose that α, β are real numbers with $\alpha \geq \beta > 0$ and $\beta \leq 1$. Then*

$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$

- **Proof.** If $\alpha + \beta = 2$ and $\alpha = 1$, then the sum in the lemma is trivial.
- When $\alpha + \beta \neq 2$, consider $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$.
- If $\alpha + \beta = 2$ and $\alpha \neq 1$, $g(x) = (n/x - 1)^{\alpha-1}$ which is monotonic.

- **Lemma 8.12.** *Suppose that* α, β *are real numbers with* $\alpha \geq \beta > 0$ *and* $\beta \leq 1$. *Then*

$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$

- **Proof.** If $\alpha+\beta=2$ and $\alpha=1$, then the sum in the lemma is trivial.
- When $\alpha+\beta \neq 2$, consider $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$.
- If $\alpha+\beta=2$ and $\alpha \neq 1$, $g(x) = (n/x-1)^{\alpha-1}$ which is monotonic.
- Then the interval $(0,n)$ this has at most one stationary point given by $(\alpha+\beta-2)x = (\beta-1)n$.

- **Lemma 8.12.** *Suppose that* α, β *are real numbers with* $\alpha \ge \beta > 0$ *and* $\beta \le 1$*. Then*

$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$

- **Proof.** If $\alpha+\beta=2$ and $\alpha=1$, then the sum in the lemma is trivial.
- When $\alpha+\beta\ne 2$, consider $g(x)=x^{\beta-1}(n-x)^{\alpha-1}$.
- If $\alpha+\beta=2$ and $\alpha\ne 1$, $g(x)=(n/x-1)^{\alpha-1}$ which is monotonic.
- Then the interval $(0,n)$ this has at most one stationary point given by $(\alpha+\beta-2)x=(\beta-1)n$.
- Thus this interval can be divided into two (or one if $X=0$ or n) intervals $(0,X)$, (X,n) such that g is monotonic on each interval.

- **Lemma 8.12.** *Suppose that α, β are real numbers with $\alpha \geq \beta > 0$ and $\beta \leq 1$. Then*

$$\sum_{m=1}^{n-1} m^{\beta-1}(n-m)^{\alpha-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} n^{\alpha+\beta-1} + O\left(n^{\alpha-1}\right).$$

- **Proof.** If $\alpha+\beta=2$ and $\alpha=1$, then the sum in the lemma is trivial.
- When $\alpha+\beta \neq 2$, consider $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$.
- If $\alpha+\beta=2$ and $\alpha \neq 1$, $g(x) = (n/x-1)^{\alpha-1}$ which is monotonic.
- Then the interval $(0, n)$ this has at most one stationary point given by $(\alpha+\beta-2)x = (\beta-1)n$.
- Thus this interval can be divided into two (or one if $X=0$ or n) intervals $(0, X)$, (X, n) such that g is monotonic on each interval.
- Thus our sum is

$$\int_1^{n-1} g(x)dx + O(n^{\alpha-1} + n^{\beta+\alpha-2}).$$

- $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$

$$\int_1^{n-1} g(x)dx + O(n^{\alpha-1} + n^{\beta+\alpha-2}).$$

- $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$

$$\int_1^{n-1} g(x)dx + O(n^{\alpha-1} + n^{\beta+\alpha-2}).$$

- Also

$$\int_0^1 g(x)dx \ll n^{\alpha-1}, \int_{n-1}^n g(x)dx \ll n^{\beta-1}.$$

- $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$

$$\int_1^{n-1} g(x)dx + O(n^{\alpha-1} + n^{\beta+\alpha-2}).$$

- Also

$$\int_0^1 g(x)dx \ll n^{\alpha-1}, \int_{n-1}^{n} g(x)dx \ll n^{\beta-1}.$$

- Hence the sum is

$$\int_0^n g(x)dx + O(n^{\alpha-1}).$$

- $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$

$$\int_1^{n-1} g(x)dx + O(n^{\alpha-1} + n^{\beta+\alpha-2}).$$

- Also

$$\int_0^1 g(x)dx \ll n^{\alpha-1}, \int_{n-1}^n g(x)dx \ll n^{\beta-1}.$$

- Hence the sum is

$$\int_0^n g(x)dx + O(n^{\alpha-1}).$$

- By a change of variable $x = yn$ the integral here is

$$n^{\beta+\alpha-1}\int_0^1 y^{\beta-1}(1-y)^{\alpha-1}dy$$

- $g(x) = x^{\beta-1}(n-x)^{\alpha-1}$

$$\int_1^{n-1} g(x)dx + O(n^{\alpha-1} + n^{\beta+\alpha-2}).$$

- Also

$$\int_0^1 g(x)dx \ll n^{\alpha-1}, \int_{n-1}^n g(x)dx \ll n^{\beta-1}.$$

- Hence the sum is

$$\int_0^n g(x)dx + O(n^{\alpha-1}).$$

- By a change of variable $x = yn$ the integral here is

$$n^{\beta+\alpha-1}\int_0^1 y^{\beta-1}(1-y)^{\alpha-1}dy$$

- and the new integral is the beta function which is well known to be

$$\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

- The lemma establishes the case $s = 2$ of the theorem.

- The lemma establishes the case $s = 2$ of the theorem.
- For $s \geq 2$ we have

$$J_{s+1}(n) = \sum_{m=1}^{n-1} k^{-1} m^{\frac{1}{k}-1} J_s(n-m)$$

- The lemma establishes the case $s = 2$ of the theorem.
- For $s \geq 2$ we have

$$J_{s+1}(n) = \sum_{m=1}^{n-1} k^{-1} m^{\frac{1}{k}-1} J_s(n-m)$$

- and so the general result follows by an easy induction.

- The lemma establishes the case $s = 2$ of the theorem.
- For $s \geq 2$ we have

$$J_{s+1}(n) = \sum_{m=1}^{n-1} k^{-1} m^{\frac{1}{k}-1} J_s(n-m)$$

- and so the general result follows by an easy induction.
- Thus we can summarize everything so far by a theorem.

- The lemma establishes the case $s = 2$ of the theorem.
- For $s \geq 2$ we have

$$J_{s+1}(n) = \sum_{m=1}^{n-1} k^{-1} m^{\frac{1}{k}-1} J_s(n-m)$$

- and so the general result follows by an easy induction.
- Thus we can summarize everything so far by a theorem.
- **Theorem 8.13.** *Let $r_s(n)$ denote the number of representations of n as the sum of s k-th powers of positive integers. Suppose that $s > 2^k$. Then there is a $\delta > 0$ such that for every large n we have*

$$r_s(n) = \frac{\Gamma\left(1+\frac{1}{k}\right)^s}{\Gamma\left(\frac{s}{k}\right)} \mathfrak{S}(n) n^{\frac{s}{k}-1} + O(n^{\frac{s}{k}-1-\delta}).$$

- If $s > 2^k$, then we have absolute convergence of

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q).$$

- If $s > 2^k$, then we have absolute convergence of

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q).$$

- For $a, b \in \mathbb{Z}$, $q, r \in \mathbb{N}$, $(a,q) = (bmr) = (q,r) = 1$, $S(qr, ar + bq) = S(q,a)S(r,b)$.

- If $s > 2^k$, then we have absolute convergence of
$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q).$$
- For $a, b \in \mathbb{Z}$, $q, r \in \mathbb{N}$, $(a,q) = (bmr) = (q,r) = 1$, $S(qr, ar+bq) = S(q,a)S(r,b)$.
- Let $\mathfrak{B}(q) = \sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q)$.

- If $s > 2^k$, then we have absolute convergence of
$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q).$$

- For $a, b \in \mathbb{Z}$, $q, r \in \mathbb{N}$, $(a,q) = (bmr) = (q,r) = 1$, $S(qr, ar+bq) = S(q,a)S(r,b)$.

- Let $\mathfrak{B}(q) = \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q)$.

- Then, when $(q,r) = 1$, we have $\mathfrak{B}(qr) =$
$$\sum_{\substack{a=1 \\ (a,q)=1}}^{q} \sum_{\substack{b=1 \\ (b,r)=1}}^{r} \frac{S(qr, ar+bq)^s}{(qr)^s} e\left(-\frac{an}{q} - \frac{bn}{r}\right) = \mathfrak{B}(q)\mathfrak{B}(r).$$

- If $s > 2^k$, then we have absolute convergence of

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q).$$

- For $a, b \in \mathbb{Z}$, $q, r \in \mathbb{N}$, $(a,q) = (bmr) = (q,r) = 1$, $S(qr, ar + bq) = S(q,a)S(r,b)$.
- Let $\mathfrak{B}(q) = \sum_{\substack{a=1 \\ (a,q)=1}}^{q} q^{-s} S(q,a)^s e(-an/q)$.
- Then, when $(q,r) = 1$, we have $\mathfrak{B}(qr) =$

$$\sum_{\substack{a=1 \\ (a,q)=1}}^{q} \sum_{\substack{b=1 \\ (b,r)=1}}^{r} \frac{S(qr, ar+bq)^s}{(qr)^s} e(-\tfrac{an}{q} - \tfrac{bn}{r}) = \mathfrak{B}(q)\mathfrak{B}(r).$$

- Hence the terms in the series $\mathfrak{S}(n)$ are multiplicative.

- **Theorem 8.14.** *Suppose that $s > 2^k$. Then for each prime p*

$$\mathfrak{T}(p) = 1 + \sum_{j=1}^{\infty} \mathfrak{B}(p^j)$$

converges absolutely and so does

$$\mathfrak{S}(n) = \prod_p \mathfrak{T}(p).$$

- **Theorem 8.14.** *Suppose that $s > 2^k$. Then for each prime p*

$$\mathfrak{T}(p) = 1 + \sum_{j=1}^{\infty} \mathfrak{B}(p^j)$$

converges absolutely and so does

$$\mathfrak{S}(n) = \prod_{p} \mathfrak{T}(p).$$

- The absolute convergence, and in particular the bound

$$\mathfrak{B}(q) \ll q^1 + \varepsilon - sK \ll q^{-1-\delta}$$

tells us that for some constant C

$$1 \ll \prod_{p>C} \mathfrak{T}(p) \ll 1.$$

- **Theorem 8.14.** *Suppose that $s > 2^k$. Then for each prime p*

$$\mathfrak{T}(p) = 1 + \sum_{j=1}^{\infty} \mathfrak{B}(p^j)$$

converges absolutely and so does

$$\mathfrak{S}(n) = \prod_p \mathfrak{T}(p).$$

- The absolute convergence, and in particular the bound

$$\mathfrak{B}(q) \ll q^1 + \varepsilon - sK \ll q^{-1-\delta}$$

tells us that for some constant C

$$1 \ll \prod_{p>C} \mathfrak{T}(p) \ll 1.$$

- Thus our main concern now is what happens with the small primes.

- **Theorem 8.14.** *Suppose that $s > 2^k$. Then for each prime p*

$$\mathfrak{T}(p) = 1 + \sum_{j=1}^{\infty} \mathfrak{B}(p^j)$$

converges absolutely and so does

$$\mathfrak{S}(n) = \prod_p \mathfrak{T}(p).$$

- The absolute convergence, and in particular the bound

$$\mathfrak{B}(q) \ll q^1 + \varepsilon - sK \ll q^{-1-\delta}$$

tells us that for some constant C

$$1 \ll \prod_{p>C} \mathfrak{T}(p) \ll 1.$$

- Thus our main concern now is what happens with the small primes.
- To this end we now begin to explore the local properties of $\mathfrak{S}(n)$.

- **Lemma 8.15.** *Let $M(q;n)$ denote the number of solutions of $x_1^k + \cdots x_s^k \equiv n \pmod q$. Then*

$$\sum_{d|q} \mathfrak{B}(d) = q^{1-s} M(q;n).$$

- **Lemma 8.15.** *Let $M(q;n)$ denote the number of solutions of $x_1^k + \cdots x_s^k \equiv n \pmod{q}$. Then*
$$\sum_{d|q} \mathfrak{B}(d) = q^{1-s} M(q;n).$$

- Thus we have
$$\mathfrak{T}(p) = \lim_{j\to\infty} p^{j(1-s)} M(p^j;n).$$

- **Lemma 8.15.** *Let $M(q;n)$ denote the number of solutions of $x_1^k + \cdots x_s^k \equiv n \pmod{q}$. Then*
$$\sum_{d|q} \mathfrak{B}(d) = q^{1-s} M(q;n).$$
- Thus we have
$$\mathfrak{T}(p) = \lim_{j\to\infty} p^{j(1-s)} M(p^j;n).$$
- This is usually interpreted as the density of the p-adic solutions of
$$x_1^k + \cdots + x_s^k = n.$$

- **Lemma 8.15.** *Let $M(q;n)$ denote the number of solutions of $x_1^k + \cdots x_s^k \equiv n \pmod q$. Then*

$$\sum_{d|q} \mathfrak{B}(d) = q^{1-s} M(q;n).$$

- Thus we have

$$\mathfrak{T}(p) = \lim_{j\to\infty} p^{j(1-s)} M(p^j;n).$$

- This is usually interpreted as the density of the p-adic solutions of

$$x_1^k + \cdots + x_s^k = n.$$

- **Proof.** We start from the orthogonality relation

$$\frac{1}{q}\sum_{r=1}^{q} e(hr/q) = \begin{cases} 1 & q|h, \\ 0 & q\nmid h. \end{cases}$$

- **Lemma 8.15.** *Let $M(q;n)$ denote the number of solutions of $x_1^k + \cdots x_s^k \equiv n \pmod{q}$. Then*
$$\sum_{d|q} \mathfrak{B}(d) = q^{1-s} M(q;n).$$
- Thus we have
$$\mathfrak{T}(p) = \lim_{j\to\infty} p^{j(1-s)} M(p^j;n).$$
- This is usually interpreted as the density of the p-adic solutions of
$$x_1^k + \cdots + x_s^k = n.$$
- **Proof.** We start from the orthogonality relation
$$\frac{1}{q}\sum_{r=1}^{q} e(hr/q) = \begin{cases} 1 & q|h, \\ 0 & q\nmid h. \end{cases}$$
- Thus
$$M(q;n) = \frac{1}{q}\sum_{r=1}^{q}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q} e\big(r(x_1^k + \cdots x_s^k - n)/q\big).$$

- Thus
$$M(q;n) = \frac{1}{q}\sum_{r=1}^{q}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q} e\big(r(x_1^k+\cdots x_s^k-n)/q\big).$$

- Thus

$$M(q;n)=\frac{1}{q}\sum_{r=1}^{q}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q}e\big(r(x_1^k+\cdots x_s^k-n)/q\big).$$

- We now sort the r according to the value $q/(r,q)=d$.

- Thus
$$M(q;n)=\frac{1}{q}\sum_{r=1}^{q}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q}e\big(r(x_1^k+\cdots x_s^k-n)/q\big).$$
- We now sort the r according to the value $q/(r,q)=d$.
- Thus $M(q;n)=$
$$\frac{1}{q}\sum_{d|q}\sum_{\substack{a=1\\(a,d)=1}}^{d}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q}e\big(\tfrac{a}{d}(x_1^k+\cdots x_s^k-n)\big)$$

- Thus

$$M(q;n)=\frac{1}{q}\sum_{r=1}^{q}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q}e\big(r(x_1^k+\cdots x_s^k-n)/q\big).$$

- We now sort the r according to the value $q/(r,q)=d$.
- Thus $M(q;n)=$

$$\frac{1}{q}\sum_{d|q}\sum_{\substack{a=1\\(a,d)=1}}^{d}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q}e\big(\tfrac{a}{d}(x_1^k+\cdots x_s^k-n)\big)$$

- Each x_j ranges q/d times over a set of residues modulo d.

- Thus

$$M(q;n)=\frac{1}{q}\sum_{r=1}^{q}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q}e\big(r(x_1^k+\cdots x_s^k-n)/q\big).$$

- We now sort the r according to the value $q/(r,q)=d$.
- Thus $M(q;n)=$

$$\frac{1}{q}\sum_{d|q}\sum_{\substack{a=1\\(a,d)=1}}^{d}\sum_{x_1}^{q}\dots\sum_{x_s=1}^{q}e\big(\tfrac{a}{d}(x_1^k+\cdots x_s^k-n)\big)$$

- Each x_j ranges q/d times over a set of residues modulo d.
- Therefor $M(q;n)=$

$$\frac{1}{q}\sum_{d|q}\sum_{\substack{a=1\\(a,d)=1}}^{d}(q/d)^s\sum_{x_1}^{d}\dots\sum_{x_s=1}^{d}e\big(a(x_1^k+\cdots x_s^k-n)/d\big)$$

$$=q^{s-1}\sum_{d|q}\mathfrak{B}(d)$$

as required.

- Given p choose $\tau = \tau(p)$ so that p^τ is the exact power of p dividing k, $p^\tau \| k$.

- Given p choose $\tau = \tau(p)$ so that p^τ is the exact power of p dividing k, $p^\tau \| k$.
- Now suppose that $(a, p) = 1$ and $t \geq \tau + 1$, and there is an x such that $x^k \equiv a \pmod{p^t}$.

- Given p choose $\tau = \tau(p)$ so that p^τ is the exact power of p dividing k, $p^\tau \| k$.
- Now suppose that $(a, p) = 1$ and $t \geq \tau + 1$, and there is an x such that $x^k \equiv a \pmod{p^t}$.
- For the time being suppose p is odd and g be a primitive root modulo p^{t+1}. Note that it is also one modulo p^t.

- Given p choose $\tau = \tau(p)$ so that p^τ is the exact power of p dividing k, $p^\tau \| k$.
- Now suppose that $(a, p) = 1$ and $t \ge \tau + 1$, and there is an x such that $x^k \equiv a \pmod{p^t}$.
- For the time being suppose p is odd and g be a primitive root modulo p^{t+1}. Note that it is also one modulo p^t.
- Choose u, v so $g^u \equiv a \pmod{p^{t+1}}$ & $x \equiv g^v \pmod{p^t}$.

- Given p choose $\tau = \tau(p)$ so that p^τ is the exact power of p dividing k, $p^\tau \| k$.
- Now suppose that $(a,p) = 1$ and $t \ge \tau + 1$, and there is an x such that $x^k \equiv a \pmod{p^t}$.
- For the time being suppose p is odd and g be a primitive root modulo p^{t+1}. Note that it is also one modulo p^t.
- Choose u, v so $g^u \equiv a \pmod{p^{t+1}}$ & $x \equiv g^v \pmod{p^t}$.
- Then $kv \equiv u \pmod{p^{t-1}(p-1)}$, & $p^\tau(k,p-1)|u$ and
$$\frac{k}{p^\tau(k,p-1)}v \equiv \frac{u}{p^\tau(k,p-1)} \pmod{p^{t-1-\tau}\frac{p-1}{(k,p-1)}}$$

- Given p choose $\tau = \tau(p)$ so that p^τ is the exact power of p dividing k, $p^\tau \| k$.
- Now suppose that $(a,p) = 1$ and $t \ge \tau + 1$, and there is an x such that $x^k \equiv a \pmod{p^t}$.
- For the time being suppose p is odd and g be a primitive root modulo p^{t+1}. Note that it is also one modulo p^t.
- Choose u, v so $g^u \equiv a \pmod{p^{t+1}}$ & $x \equiv g^v \pmod{p^t}$.
- Then $kv \equiv u \pmod{p^{t-1}(p-1)}$, & $p^\tau(k,p-1)|u$ and

$$\frac{k}{p^\tau(k,p-1)} v \equiv \frac{u}{p^\tau(k,p-1)} \pmod{p^{t-1-\tau}\frac{p-1}{(k,p-1)}}$$

- Hence $\left(\frac{k}{p^\tau(k,p-1)}, p^{t-\tau}\frac{p-1}{(k,p-1)}\right) = 1$ and so there is a v' such that

$$\frac{k}{p^\tau(k,p-1)} v' \equiv \frac{u}{p^\tau(k,p-1)} \pmod{p^{t-\tau}\frac{p-1}{(k,p-1)}}.$$

- Given p choose $\tau = \tau(p)$ so that p^τ is the exact power of p dividing k, $p^\tau \| k$.
- Now suppose that $(a,p) = 1$ and $t \geq \tau + 1$, and there is an x such that $x^k \equiv a \pmod{p^t}$.
- For the time being suppose p is odd and g be a primitive root modulo p^{t+1}. Note that it is also one modulo p^t.
- Choose u, v so $g^u \equiv a \pmod{p^{t+1}}$ & $x \equiv g^v \pmod{p^t}$.
- Then $kv \equiv u \pmod{p^{t-1}(p-1)}$, & $p^\tau(k,p-1)|u$ and
$$\frac{k}{p^\tau(k,p-1)}v \equiv \frac{u}{p^\tau(k,p-1)} \pmod{p^{t-1-\tau}\frac{p-1}{(k,p-1)}}$$
- Hence $\left(\frac{k}{p^\tau(k,p-1)}, p^{t-\tau}\frac{p-1}{(k,p-1)}\right) = 1$ and so there is a v' such that
$$\frac{k}{p^\tau(k,p-1)}v' \equiv \frac{u}{p^\tau(k,p-1)} \pmod{p^{t-\tau}\frac{p-1}{(k,p-1)}}.$$
- Thus $kv' \equiv u \pmod{\phi(p^{t+1})}$ is soluble, whence so is
$$x^k \equiv a \pmod{p^{t+1}}.$$

- When $p = 2$ things are more complicated since the multiplicative group of reduced residues modulo 2^t is no longer cyclic when $t \geq 3$.

- When $p = 2$ things are more complicated since the multiplicative group of reduced residues modulo 2^t is no longer cyclic when $t \geq 3$.
- However is easily shown that if k is odd, then x^k runs over a set of reduced residues as x does.

- When $p = 2$ things are more complicated since the multiplicative group of reduced residues modulo 2^t is no longer cyclic when $t \geq 3$.
- However is easily shown that if k is odd, then x^k runs over a set of reduced residues as x does.
- Thus that leaves the case when $\tau(2) > 0$.

- When $p = 2$ things are more complicated since the multiplicative group of reduced residues modulo 2^t is no longer cyclic when $t \geq 3$.
- However is easily shown that if k is odd, then x^k runs over a set of reduced residues as x does.
- Thus that leaves the case when $\tau(2) > 0$.
- In that case we work with $t \geq \tau + 2$ as above but with the generators -1 and 5.

- When $p = 2$ things are more complicated since the multiplicative group of reduced residues modulo 2^t is no longer cyclic when $t \geq 3$.
- However is easily shown that if k is odd, then x^k runs over a set of reduced residues as x does.
- Thus that leaves the case when $\tau(2) > 0$.
- In that case we work with $t \geq \tau + 2$ as above but with the generators -1 and 5.
- There are more complications of detail but the conclusion is the same.

- When $p = 2$ things are more complicated since the multiplicative group of reduced residues modulo 2^t is no longer cyclic when $t \geq 3$.
- However is easily shown that if k is odd, then x^k runs over a set of reduced residues as x does.
- Thus that leaves the case when $\tau(2) > 0$.
- In that case we work with $t \geq \tau + 2$ as above but with the generators -1 and 5.
- There are more complications of detail but the conclusion is the same.
- The argument above also shows that the number of k-th power residues modulo p^γ is

$$\frac{\phi(p^{\tau+1})}{(k, \phi(p^{\tau+1}))}.$$

- Let $M^*(q;n)$ denote the number of solutions of

$$x_1^k + \cdots x_s^k \equiv n \pmod{q}$$

with $(x_1, q) = 1$.

- Let $M^*(q;n)$ denote the number of solutions of

$$x_1^k + \cdots x_s^k \equiv n \pmod q$$

with $(x_1, q) = 1$.

- **Lemma 8.16.** *Suppose that τ is as above,*

$$\gamma = \begin{cases} \tau + 1 & \text{when } p > 2, \text{ or } p = 2 \text{ and } \tau = 0, \\ \tau + 2 & \text{when } p = 2 \text{ and } \tau > 0, \end{cases}$$

and $M^*(p^\gamma; n) > 0$ and $t \ge \gamma$. Then

$$M(p^t; n) \ge p^{(t-\gamma)(s-1)}.$$

- Let $M^*(q;n)$ denote the number of solutions of

$$x_1^k + \cdots x_s^k \equiv n \pmod{q}$$

with $(x_1, q) = 1$.

- **Lemma 8.16.** *Suppose that τ is as above,*

$$\gamma = \begin{cases} \tau + 1 & \textit{when } p > 2, \textit{ or } p = 2 \textit{ and } \tau = 0, \\ \tau + 2 & \textit{when } p = 2 \textit{ and } \tau > 0, \end{cases}$$

and $M^*(p^\gamma; n) > 0$ and $t \geq \gamma$. Then

$$M(p^t; n) \geq p^{(t-\gamma)(s-1)}.$$

- It follows immediately that

$$\mathfrak{T}(p) \geq p^{-\gamma(s-1)}.$$

- Let $M^*(q;n)$ denote the number of solutions of

$$x_1^k + \cdots x_s^k \equiv n \pmod q$$

with $(x_1, q) = 1$.

- **Lemma 8.16.** *Suppose that τ is as above,*

$$\gamma = \begin{cases} \tau + 1 & \text{when } p > 2, \text{ or } p = 2 \text{ and } \tau = 0, \\ \tau + 2 & \text{when } p = 2 \text{ and } \tau > 0, \end{cases}$$

and $M^*(p^\gamma; n) > 0$ and $t \ge \gamma$. Then

$$M(p^t; n) \ge p^{(t-\gamma)(s-1)}.$$

- It follows immediately that

$$\mathfrak{T}(p) \ge p^{-\gamma(s-1)}.$$

- Observe that this lower bound only depends on k and s.

- So now we only have to find a non-singular solution modulo p^{γ}.

- So now we only have to find a non-singular solution modulo p^γ.
- There are many ways of doing this.

- So now we only have to find a non-singular solution modulo p^γ.
- There are many ways of doing this.
- For example there is an *ad hoc* argument of Davenport in his notes on diophantine equations.

- So now we only have to find a non-singular solution modulo p^{γ}.
- There are many ways of doing this.
- For example there is an *ad hoc* argument of Davenport in his notes on diophantine equations.
- However I think it is instructive to proceed through a much more general result which I will prove in a while.

- So now we only have to find a non-singular solution modulo p^γ.
- There are many ways of doing this.
- For example there is an *ad hoc* argument of Davenport in his notes on diophantine equations.
- However I think it is instructive to proceed through a much more general result which I will prove in a while.
- Given a positive integer q and a collection $\mathcal{A}$ of residue classes modulo q, its local density $\rho = \rho(\mathcal{A})$ modulo q is defined by $\rho = q^{-1}\,\mathrm{card}(\mathcal{A})$.

- So now we only have to find a non-singular solution modulo p^γ.
- There are many ways of doing this.
- For example there is an *ad hoc* argument of Davenport in his notes on diophantine equations.
- However I think it is instructive to proceed through a much more general result which I will prove in a while.
- Given a positive integer q and a collection $\mathcal{A}$ of residue classes modulo q, its local density $\rho = \rho(\mathcal{A})$ modulo q is defined by $\rho = q^{-1}\,\mathrm{card}(\mathcal{A})$.
- **Theorem 8.17. [Cauchy–Davenport–Chowla]** *Suppose that q is a positive integer, that $\mathcal{A}$ and $\mathcal{B}$ are sets of residue classes modulo q of local density modulo q, α and β respectively, that $0 \in \mathcal{B}$ and that every non–zero residue class in $\mathcal{B}$ is a reduced residue class modulo q. Then*

$$\rho(\mathcal{A}+\mathcal{B}) \ge \min(1, \alpha+\beta-1/q).$$

- **Lemma 8.18.** *Suppose that*

$$s \geq \begin{cases} \frac{p}{p-1}(k, p^\tau(p-1)) & \text{when } \gamma = \tau + 1, \\ 2^{\tau+2} & \text{when } \gamma = \tau + 2 \text{ and } k > 2, \\ 5 & \text{when } p = k = 2. \end{cases}$$

Then $M^*(p^\gamma; n) > 0$.

- **Lemma 8.18.** *Suppose that*

$$s \ge \begin{cases} \frac{p}{p-1}(k, p^\tau(p-1)) & \text{when } \gamma = \tau + 1, \\ 2^{\tau+2} & \text{when } \gamma = \tau + 2 \text{ and } k > 2, \\ 5 & \text{when } p = k = 2. \end{cases}$$

Then $M^*(p^\gamma; n) > 0$.

- **Proof.** When $\gamma = \tau + 1$ the number of k-th power residues modulo p^γ is $N = \dfrac{\phi(p^{\tau+1})}{(k, \phi(p^\tau + 1))} = \dfrac{\phi(p^\gamma)}{(k, \phi(p^\gamma))}$.

- **Lemma 8.18.** *Suppose that*

$$s \ge \begin{cases} \frac{p}{p-1}(k, p^\tau(p-1)) & \text{when } \gamma = \tau+1, \\ 2^{\tau+2} & \text{when } \gamma = \tau+2 \text{ and } k > 2, \\ 5 & \text{when } p = k = 2. \end{cases}$$

Then $M^*(p^\gamma; n) > 0$.

- **Proof.** When $\gamma = \tau + 1$ the number of k-th power residues modulo p^γ is $N = \dfrac{\phi(p^{\tau+1})}{(k, \phi(p^\tau + 1))} = \dfrac{\phi(p^\gamma)}{(k, \phi(p^\gamma))}$.
- Let $\mathcal{B}^*$ be the set of k-th power (reduced) residues modulo p^γ and let $\mathcal{B} = \{0\} \cup \mathcal{B}^*$.

- **Lemma 8.18.** *Suppose that*

$$s \geq \begin{cases} \frac{p}{p-1}(k, p^\tau(p-1)) & \textit{when } \gamma = \tau + 1, \\ 2^{\tau+2} & \textit{when } \gamma = \tau + 2 \textit{ and } k > 2, \\ 5 & \textit{when } p = k = 2. \end{cases}$$

Then $M^*(p^\gamma; n) > 0$.

- **Proof.** When $\gamma = \tau + 1$ the number of k-th power residues modulo p^γ is $N = \dfrac{\phi(p^{\tau+1})}{(k, \phi(p^\tau + 1))} = \dfrac{\phi(p^\gamma)}{(k, \phi(p^\gamma))}$.
- Let $\mathcal{B}^*$ be the set of k-th power (reduced) residues modulo p^γ and let $\mathcal{B} = \{0\} \cup \mathcal{B}^*$.
- Then $\rho(\mathcal{B}^*) = Np^{-\gamma}$ and $\rho(\mathcal{B}) = (N+1)p^{-\gamma}$.

- **Lemma 8.18.** *Suppose that*

$$s \ge \begin{cases} \frac{p}{p-1}(k, p^\tau(p-1)) & \text{when } \gamma = \tau + 1, \\ 2^{\tau+2} & \text{when } \gamma = \tau + 2 \text{ and } k > 2, \\ 5 & \text{when } p = k = 2. \end{cases}$$

Then $M^*(p^\gamma; n) > 0$.

- **Proof.** When $\gamma = \tau + 1$ the number of k-th power residues modulo p^γ is $N = \dfrac{\phi(p^{\tau+1})}{(k, \phi(p^\tau + 1))} = \dfrac{\phi(p^\gamma)}{(k, \phi(p^\gamma))}$.
- Let $\mathcal{B}^*$ be the set of k-th power (reduced) residues modulo p^γ and let $\mathcal{B} = \{0\} \cup \mathcal{B}^*$.
- Then $\rho(\mathcal{B}^*) = Np^{-\gamma}$ and $\rho(\mathcal{B}) = (N+1)p^{-\gamma}$.
- Now consider $\mathcal{B}^* + (s-1)\mathcal{B}$.

- **Lemma 8.18.** *Suppose that*

$$s \geq \begin{cases} \frac{p}{p-1}(k, p^\tau(p-1)) & \text{when } \gamma = \tau + 1, \\ 2^{\tau+2} & \text{when } \gamma = \tau + 2 \text{ and } k > 2, \\ 5 & \text{when } p = k = 2. \end{cases}$$

Then $M^*(p^\gamma; n) > 0$.

- **Proof.** When $\gamma = \tau + 1$ the number of k-th power residues modulo p^γ is $N = \dfrac{\phi(p^{\tau+1})}{(k, \phi(p^\tau + 1))} = \dfrac{\phi(p^\gamma)}{(k, \phi(p^\gamma))}$.
- Let $\mathcal{B}^*$ be the set of k-th power (reduced) residues modulo p^γ and let $\mathcal{B} = \{0\} \cup \mathcal{B}^*$.
- Then $\rho(\mathcal{B}^*) = Np^{-\gamma}$ and $\rho(\mathcal{B}) = (N+1)p^{-\gamma}$.
- Now consider $\mathcal{B}^* + (s-1)\mathcal{B}$.
- By Cauchy-Davenport-Chowla and induction we have

$$\rho(\mathcal{B}^* + (s-1)\mathcal{B}) \geq \min(1, sNp^{-\gamma}).$$

- **Lemma 8.18.** *Suppose that*

$$s \ge \begin{cases} \frac{p}{p-1}(k, p^\tau(p-1)) & \text{when } \gamma = \tau+1, \\ 2^{\tau+2} & \text{when } \gamma = \tau+2 \text{ and } k > 2, \\ 5 & \text{when } p = k = 2. \end{cases}$$

  Then $M^*(p^\gamma; n) > 0$.
- **Proof.** When $\gamma = \tau + 1$ the number of k-th power residues modulo p^γ is $N = \dfrac{\phi(p^{\tau+1})}{(k, \phi(p^\tau + 1))} = \dfrac{\phi(p^\gamma)}{(k, \phi(p^\gamma))}$.
- Let $\mathcal{B}^*$ be the set of k-th power (reduced) residues modulo p^γ and let $\mathcal{B} = \{0\} \cup \mathcal{B}^*$.
- Then $\rho(\mathcal{B}^*) = Np^{-\gamma}$ and $\rho(\mathcal{B}) = (N+1)p^{-\gamma}$.
- Now consider $\mathcal{B}^* + (s-1)\mathcal{B}$.
- By Cauchy-Davenport-Chowla and induction we have

$$\rho(\mathcal{B}^* + (s-1)\mathcal{B}) \ge \min(1, sNp^{-\gamma}).$$

- We should note also that every element of $\mathcal{B}^* + (s-1)\mathcal{B}$ is a sum of s k-th powers modulo p^γ and that at least one of them is reduced.

- $\rho(\mathcal{B}^* + (s-1)\mathcal{B}) \geq \min(1, sNp^{-\gamma})$.

- $\rho(\mathcal{B}^* + (s-1)\mathcal{B}) \geq \min(1, sNp^{-\gamma})$.
- When $sNp^{-\gamma} \geq 1$ we are home and dry.

- $\rho(\mathcal{B}^* + (s-1)\mathcal{B}) \geq \min(1, sNp^{-\gamma})$.
- When $sNp^{-\gamma} \geq 1$ we are home and dry.
- In other words when s satisfies the hypothesis $s \geq \frac{p^{\gamma}\left(k, p^{\tau}(p-1)\right)}{\phi(p^{\gamma})}$.

- $\rho(\mathcal{B}^* + (s-1)\mathcal{B}) \geq \min(1, sNp^{-\gamma})$.
- When $sNp^{-\gamma} \geq 1$ we are home and dry.
- In other words when s satisfies the hypothesis $s \geq \frac{p^\gamma\left(k,p^\tau(p-1)\right)}{\phi(p^\gamma)}$.
- In the case $p=2$ and $k>2$ we are supposing $s \geq 2^\gamma$ and we can solve

$$x_1^k + \cdots x_s^k \equiv n \pmod{2^\gamma}$$

be taking each x_j to be 1 or 0 and at least one of them 1.

- $\rho(\mathcal{B}^* + (s-1)\mathcal{B}) \ge \min(1, sNp^{-\gamma})$.
- When $sNp^{-\gamma} \ge 1$ we are home and dry.
- In other words when s satisfies the hypothesis $s \ge \frac{p^\gamma\left(k, p^\tau(p-1)\right)}{\phi(p^\gamma)}$.
- In the case $p = 2$ and $k > 2$ we are supposing $s \ge 2^\gamma$ and we can solve

$$x_1^k + \cdots x_s^k \equiv n \pmod{2^\gamma}$$

be taking each x_j to be 1 or 0 and at least one of them 1.

- The third case has $k = 2$, so that $s \ge 5$ and $2^\gamma = 8$, and is likewise trivially soluble with $2 \nmid x_1$ since $x_j^2 \equiv 0$, 1 or 4 (mod 8).

- Bringing it all together we have
  Theorem 8.19. *Suppose that* $s > 2^k$. *Then* $\mathfrak{S}(n) \gg 1$.

- Bringing it all together we have
  Theorem 8.19. *Suppose that* $s > 2^k$. *Then* $\mathfrak{S}(n) \gg 1$.
- **Proof.** There are various cases to check.

- Bringing it all together we have
  Theorem 8.19. *Suppose that $s > 2^k$. Then $\mathfrak{S}(n) \gg 1$.*
- **Proof.** There are various cases to check.
- When $k \geq 2$ we have $2^k \geq 2k \geq \frac{pk}{p-1}$, so the first condition of the previous lemma will hold.

- Bringing it all together we have
  Theorem 8.19. *Suppose that* $s > 2^k$. *Then* $\mathfrak{S}(n) \gg 1$.
- **Proof.** There are various cases to check.
- When $k \geq 2$ we have $2^k \geq 2k \geq \frac{pk}{p-1}$, so the first condition of the previous lemma will hold.
- If $k = 2$, then $s \geq 5$ so the last condition also holds.

- Bringing it all together we have
  Theorem 8.19. *Suppose that $s > 2^k$. Then $\mathfrak{S}(n) \gg 1$.*
- **Proof.** There are various cases to check.
- When $k \geq 2$ we have $2^k \geq 2k \geq \frac{pk}{p-1}$, so the first condition of the previous lemma will hold.
- If $k = 2$, then $s \geq 5$ so the last condition also holds.
- When $\gamma = \tau + 2$, so $p = 2$, but $k > 2$ there are several possibilities.

- Bringing it all together we have
  Theorem 8.19. *Suppose that* $s > 2^k$. *Then* $\mathfrak{S}(n) \gg 1$.
- **Proof.** There are various cases to check.
- When $k \ge 2$ we have $2^k \ge 2k \ge \frac{pk}{p-1}$, so the first condition of the previous lemma will hold.
- If $k = 2$, then $s \ge 5$ so the last condition also holds.
- When $\gamma = \tau + 2$, so $p = 2$, but $k > 2$ there are several possibilities.
- If $\tau \ge 2$, then $k \ge 2^\tau \ge \tau + 2$, so that $2^k \ge 2^{\tau+2} = 2^\gamma$.

- Bringing it all together we have
  Theorem 8.19. *Suppose that $s > 2^k$. Then $\mathfrak{S}(n) \gg 1$.*
- **Proof.** There are various cases to check.
- When $k \geq 2$ we have $2^k \geq 2k \geq \frac{pk}{p-1}$, so the first condition of the previous lemma will hold.
- If $k = 2$, then $s \geq 5$ so the last condition also holds.
- When $\gamma = \tau + 2$, so $p = 2$, but $k > 2$ there are several possibilities.
- If $\tau \geq 2$, then $k \geq 2^\tau \geq \tau + 2$, so that $2^k \geq 2^{\tau+2} = 2^\gamma$.
- If $\tau = 1$, then $k = 2k_0$ with $k_0 \geq 3$, but $\gamma = \tau + 2 = 3$, so that $2^k \geq 2^6 > 2^\gamma$.

- Bringing it all together we have
  Theorem 8.19. *Suppose that $s > 2^k$. Then $\mathfrak{S}(n) \gg 1$.*
- **Proof.** There are various cases to check.
- When $k \geq 2$ we have $2^k \geq 2k \geq \frac{pk}{p-1}$, so the first condition of the previous lemma will hold.
- If $k = 2$, then $s \geq 5$ so the last condition also holds.
- When $\gamma = \tau + 2$, so $p = 2$, but $k > 2$ there are several possibilities.
- If $\tau \geq 2$, then $k \geq 2^\tau \geq \tau + 2$, so that $2^k \geq 2^{\tau+2} = 2^\gamma$.
- If $\tau = 1$, then $k = 2k_0$ with $k_0 \geq 3$, but $\gamma = \tau + 2 = 3$, so that $2^k \geq 2^6 > 2^\gamma$.
- The above argument can be refined to show that $\mathfrak{S}(n) \gg 1$ when $s \geq 2k$ and k is not power of 2 and $s \geq 4k$ when $k = 2^j$ with $j \geq 2$. However this requires better knowledge of the convergence of $\mathfrak{S}(n)$.