SECTION 1: INTRODUCTION. Lagrange proved that any positive integer was the sum of four or fewer numbers of the form $x^2$ with $x$ a positive integer. Waring asked if given an $n \geq 2$, there is an $f = f(n)$ such that every positive integer is the sum of $f$ or fewer numbers of the form $x^n$ with $x$ a positive integer. Hilbert showed the answer was yes, via a very difficult and sophisticated proof. Subsequently, Y. V. Linnik discovered an elementary proof, reported in chapter 3 of the lovely little book Three pearls of Number Theory by A. Y. Khinchin, [K], (at this writing, available from Dover Press). We here present a rewriting of that chapter, and also carry Linnik’s ideas somewhat further. In particular, corollary 3 below will show that if $P(X)$ is a non-constant polynomial with integral coefficients and with positive leading coefficient, and if there is an integer $z$ with $P(z) = 1$, then there is an $f$ such that all positive integers are the sum of $f$ or fewer numbers of the form $P(x)$ with $P(x) > 0$. Waring’s problem concerns the special case $P(X) = X^n$, for which $P(1) = 1$.

Remark: Since $0^n = 0$, we could say that Hilbert proved there is an $f$ such that every non-negative integer is the sum of exactly $f$ numbers of the form $x^n$ with $x \geq 0$. However, for our $P(X)$, perhaps there is no integer $x$ with $P(x) = 0$. Thus, we need the ‘$f$ or fewer’ version of the statement. However, by that phrase we will mean at least 1. That is, we do not allow sums with 0 terms.
Notation: We will work in the integers. $P(X)$ will be a degree $n > 0$ polynomial having integral coefficients, with leading coefficient $c > 0$. For Waring’s problem, one considers integers $x \geq 1$.

We will consider integers $x \geq \alpha$ where $\alpha$ is either some fixed integer, or is minus infinity. (We will see that the choice of $\alpha$ is almost irrelevant.) Let $S = \{x \geq \alpha \mid P(x) > 0\}$.

Let $D = \text{GCD}\{P(x) \mid x \in S\}$. Obviously, there must be a finite set $\{x_1, \ldots, x_t\} \subseteq S$ such that $D = \text{GCD}\{P(x_i) \mid 1 \leq i \leq t\}$. Letting $d_i = P(x_i) > 0$, we have $D = \text{GCD}(d_1, \ldots, d_t)$.

Remark: We digress with an interesting comment about $D$. As defined, it appears to depend upon $S$, and so upon $\alpha$. Actually, we will now show that $D = \text{GCD}\{P(z) \mid z \text{ is an integer}\}$.

To see that, let $D' = \text{GCD}\{P(z) \mid z \text{ is an integer}\}$. Also select any integer $y$ with $P(y) \neq 0$, and let $D'' = \text{GCD}\{P(x) \mid y \leq x \leq y + |P(y)| - 1\}$. We claim $D' = D''$. Clearly $D'$ divides $D''$. To show $D''$ divides $D'$, it will suffice to show that $D''$ divides $P(z)$ for any integer $z$.

Since $D''$ divides $P(y)$, we have $D'' \leq |P(y)|$. Therefore, there is an $x$ with $y \leq x \leq y + D'' - 1 \leq y + |P(y)| - 1$, such that $z \equiv x \mod D''$. It follows that $P(z) = P(x) \mod D''$.

Since $D''$ divides $P(x)$, it must also divide $P(z)$. Thus $D' = D''$, as claimed.

We next note that because $c > 0$, $P(X)$ goes to infinity as $X$ does. Therefore, with $y \geq \alpha$ sufficiently large, we have $P(x) > 0$ for $x \geq y$. Thus $\{x \mid y \leq x \leq y + |P(y)| - 1\} \subseteq S$.

That tells us $D$ divides $D'' = D'$. As it is obvious that $D'$ divides $D$, we see that $D = D' = \text{GCD}\{P(z) \mid z \text{ is an integer}\}$, as desired.

We also note that the argument in the second paragraph of this remark gives a way of actually constructing $D$ for a given $P(X)$. 
Example: Let \( P(X) = X^2 - X \). We easily see that \( D = 2 \). However, the greatest common divisor of the coefficients of \( P(X) \) is 1. We therefore see that while the GCD of the coefficients of \( P(X) \) clearly is a divisor of \( D \), it might not equal \( D \).

Notation: For \( f > 0 \), let \( \mathcal{P}(f) = \{ k \mid k \text{ is the sum of } f \text{ or fewer numbers of the form } P(x) \text{ with } x \in S \} \).

Obviously every number in \( \mathcal{P}(f) \) is a multiple of \( D \). Equally obviously, \( \mathcal{P}(1) \subseteq \mathcal{P}(2) \subseteq \mathcal{P}(3) \subseteq \cdots \). Our goal is to show that sequence eventually stabilizes to a set we will call \( \mathcal{P} \), and that there is an integer \( H \) such that \( \{ mD \mid m \geq H \} \subseteq \mathcal{P} \). (The interested reader will be able to see that the only influence \( \alpha \) has concerns the size of \( H \) and how quickly the above sequence stabilizes.)

Suppose we can find an \( f \) such that there is an \( H \) with \( \{ mD \mid m \geq H \} \subseteq \mathcal{P}(f) \).

If \( f' > f \) and \( \mathcal{P}(f') \neq \mathcal{P}(f) \), then the numbers in \( \mathcal{P}(f') \) but not in \( \mathcal{P}(f) \) must all have the form \( mD \) with \( 1 \leq m < H \). Since there are only finitely many such \( mD \), we see that our sequence \( \mathcal{P}(1) \subseteq \mathcal{P}(2) \subseteq \mathcal{P}(3) \subseteq \cdots \subseteq \mathcal{P}(f) \subseteq \mathcal{P}(f+1) \subseteq \cdots \) will stabilize within a finite number of steps, showing \( \mathcal{P} \) exists, and completing the argument.

The rest of this work will be dedicated to showing there is an \( f \) and \( H \) with \( \{ mD \mid m \geq H \} \subseteq \mathcal{P}(f) \).

Recall that we have \( D = \text{GCD}(d_1, \ldots, d_i) \), with \( d_i = P(x_i) \) and with \( x_i \in S \).
The next lemma is rather well known.

Lemma 1: There is an \(H\) such that for all \(m \geq H\), \(mD\) has the form
\[m_1d_1 + \ldots + m_t d_t,\] with each \(m_i \geq 1\).

Proof: We will say that a linear combination \(m_1d_1 + \ldots + m_t d_t\) is ‘acceptable’ if each \(m_i\) is positive. We first do the case that \(D = 1\). There are integers \(u_1, \ldots, u_t\) with
\[u_1d_1 + \ldots + ud_t = 1.\] For \(1 \leq i \leq t\), let \(s_i\) and \(q_i\) be positive integers with \(s_i - q_i = u_i\).
We see that if \(k = q_id_1 + \ldots + q_id_t\), then \(s_id_1 + \ldots + s_id_t = k + 1\). Thus, \(k\) and \(k + 1\) have both been expressed as acceptable linear combinations. It is now clear that \(k + k, k + (k + 1), \) and \((k + 1) + (k + 1)\) can be expressed as acceptable linear combinations.
Thus \(2k, 2k + 1, \) and \(2k + 2\) have been expressed as acceptable linear combinations.
In the same manner, we see that \(2k + 2k = 4k, 4k + 1, 4k + 2, 4k + 3, \) and \(4k + 4 = (2k + 2) + (2k + 2)\) can all be expressed as acceptable linear combinations. Iterating, we eventually reach a list of \(d_1\) consecutive integers, each of which can be expressed as an acceptable linear combination. Call them \(H + j\) for \(0 \leq j \leq d_1 - 1\). If \(m \geq H\), then for some \(j\) \((0 \leq j \leq d_1 - 1)\), we have \(m = (H + j) + bd_1\) for some \(b \geq 0\). That form makes it clear that \(m = mD\) can be expressed as an acceptable linear combination.

In the general case, since \(\text{GCD}(d_1/D, \ldots, d_t/D) = 1\), we have just seen that for some \(H\), every \(m \geq H\) can be written as
\[m = m_1(d_1/D) + \ldots + m_t(d_t/D),\] with each \(m_i > 0\). Multiplying by \(D\) gives the desired result.
We reach a crucial point. We will now state a proposition, give a corollary to it, then use it to reach our desired goal, before finally turning to its elaborate proof.

Proposition 2: Let \( z \in S \). Then there is an \( f \) such that for all \( m \geq 1 \), \( mP(z) \) is the sum of \( f \) or fewer numbers of the form \( P(x) \) with \( x \in S \). (The proof will also show we can choose the \( P(x) \) to be multiples of \( P(z) \), a fact we do not need.)

Corollary 3: If there is an \( z \in S \) with \( P(z) = 1 \), then there is an \( f \) such that all positive integers are the sum of \( f \) or fewer numbers of the form \( P(x) \) with \( x \in S \) (so that \( P(x) > 0 \)).

Proof: Immediate from proposition 2.

Theorem 4: With notation as above, the sequence \( P(1) \subseteq P(2) \subseteq P(3) \subseteq \cdots \) eventually stabilizes to a set \( P \). Also, there is an integer \( H \) such that \( \{mD \mid m \geq H\} \subseteq P \).

Remark: We will use proposition 2 to prove the theorem 4. Conversely, if theorem 4 is true, proposition 2 must also be true. To see that, assume that \( P \) exists and equals \( P(f) \). Note that \( mP(y) \in P(m) \subseteq P = P(f) \), and so \( mp(y) \) is the sum of \( f \) or fewer numbers of the form \( P(x) \) with \( x \in S \).

Proof of theorem 4: We earlier pointed out that we only need to find an \( f \) and \( H \) such that \( \{mD \mid m \geq H\} \subseteq P(f) \). By lemma 1, there is an \( H \) such that for all \( m \geq H \), \( mD \) has the form
m_1d_1 + \ldots + m_td_t, with each m_i \geq 1. Recalling that d_i = P(x_i), we let z = x_i \in S in proposition 2, and learn that there is an f_i such that each m_id_i is the sum of f_i or fewer numbers of the form P(x) with x \in S. Letting f = f_1 + f_2 + \ldots + f_t, we see that for all m \geq H, mD is the sum of f or fewer numbers of the form P(x) with x \in S. Thus, \{mD \mid m \geq H\} \subseteq \mathcal{P}(f), and we are done.

Remark: Of course, the case D = 1 is of special interest, since it says there is an f such that any m \geq H is the sum of f or fewer numbers of the form P(x) with x \in S. Corollary 3 already covered the most special case, in which D clearly is 1.

SECTION 2: PROVING PROPOSITION 2.

In this section, we will prove proposition 2, modulo two facts. We will give a reference for the first of those facts, but the second fact will be proved in sections 3 through 7.

We now explain the two facts. First, we let B be an infinite subset of the non-negative integers, assuming 0 is in B. For N \geq 1 an integer, we let B(N) be the number of positive integers in B which are equal to or less than N. We define the Schnirelmann density of B to be GLB\{B(N)/N \mid N \geq 1\}. For an integer h \geq 1, we let hB = \{m \mid m \text{ is the sum of } h \text{ numbers in } B\}. (Notice that 0 \in B implies B \subseteq hB.)

Schnirelmann’s theorem: If the density of B is positive, then there is an h such that hB = \{m \mid m \geq 0\}. 
A proof of Schnirelmann’s theorem can be found in chapter 2 of [K]. The argument is simple and elegant. (That chapter also contains a result whose proof is elaborate, but which we do not need.)

The second fact we need is a fundamental lemma due to Linnik. Its proof appears in chapter 3 of [K]. However, despite the many virtues of that highly recommended little book, the presentation of the fundamental lemma is perhaps not quite as clear as it might be. In sections 3 through 7, we rewrite the proof of the fundamental lemma. In this section, we state and use it.

Notation: For integers $N \geq 1$, $g \geq 1$, and $m$, let $r_{PNg}(m)$ equal the number of $(x_1, \ldots, x_g)$ with each $x_i$ an integer with $|x_i| \leq N$, and such that $P(x_1) + \cdots + P(x_g) = m$.

Fundamental lemma: Given $P(X)$, there is a $g > n$ (depending solely on the degree $n$ of $P(X)$), and a constant $K$ (depending on the coefficients of $P(X)$) such that for any integers $m$ and $N \geq 1$, $r_{PNg}(m) \leq KN^{g-n}$.

We are ready to prove proposition 2 in section 1.

Proof of proposition 2: Suppose $z \in S$, and let $d = P(z) \geq 1$. Our goal is to show that for some $f$, for all $m \geq 1$, $md$ is a sum of $f$ or fewer numbers of the form $P(x)$ with $x \in S$.

Let $A = \{0\} \cup \{P(x)/d \mid x \in S \text{ and } d \text{ divides } P(x)\}$. Any $z'' \equiv z \mod d$ has $p(z'')$ a multiple of $d$, and so since the leading coefficient of $P(X)$ is positive (so that $P(z'')$ goes to infinity as $z''$ does), we see that $A$ is an infinite set of non-negative numbers that contains 0. Thus, it is the type of set
dealt with by Schnirelmann’s work. With g as in the fundamental lemma, we let $B = gA$, and will show that the Schnirelmann density of $B$ is positive. Therefore, by Schnirelmann’s theorem, there is an h such that $hgA = hB = \{m \mid m \geq 0\}$. Letting $f = hg$, we see that any $m \geq 1$ can be written as the sum of $f$ numbers from $A$. Now the nonzero numbers in $A$ have the form $P(x)/d$ with $x \in S$ and $d$ dividing $P(x)$. Thus, $m \geq 1$ is the sum of $f$ or fewer numbers of the form $P(x)/d$ with the $x \in S$ and with $d$ dividing $P(x)$. That is equivalent to the goal stated above. (We also see the unneeded fact that the $P(x)$ can be chosen to be multiples of $d = P(z)$.)

Let $B = gA$. We must show there is a positive lower bound to the set $B(N)/N$, where $N \geq 1$ is an integer and $B(N)$ is the number of positive integers in $B$ that are equal to or less than $N$.

We will now consider an integer $M \geq 1$, subject to two constraints concerning how large it must be. (There is will be no upper bound to its size.) Since the leading coefficient $c$ of $P(X)$ is positive, $P(X)$ eventually becomes strictly monotonically increasing, and goes to infinity as $X$ does. Therefore we can pick $M$ such that for any $M' \geq M$, we have $P(x) \leq P(M')$ for $0 \leq x \leq M'$. Also, since $P(X)$ asymptotically approaches $cX^n$ as $X$ goes to infinity, we may assume $M$ is large enough that for $M' \geq M$, $P(M') \leq 2cM'^n$. Taking these two constraints together, we see that for any $M' \geq M$ and any $x$ with $0 \leq x \leq M'$, we have $P(x) \leq 2cM'^n$. Notice that any integer larger than $M$ also satisfies this condition.

We next fix an integer $z' = z \mod d$. If the set $\{u \geq \alpha \mid u \not\in S\} = \{u \geq \alpha \mid P(u) < 0\}$ is empty, we insist that $z' \geq \max\{\alpha, 0\}$. However, if that set is non-empty, it clearly contains a maximal integer. In that case, we insist that both $z' \geq \max\{\alpha, 0\}$ and $z' > \max\{u \geq \alpha \mid u \not\in S\}$.
(We will write as if that set is non-empty. In the following, simply ignore any reference to it in the case that it is empty.)

Claim: With g and K as in the fundamental lemma, let $C = 2gc(z' + d)^a$, and $C' = \frac{1}{K(z' + d)^{t-n}}$. Then $B(CM^n) \geq C'M^n$.

Let $T = \{(x_1, ..., x_g) | \text{for } 1 \leq i \leq g, \text{ we have } x_i \in S, z' \leq x_i \leq z' + d(M - 1), \text{ and } d \text{ divides } P(x_i)\}$. Also let $T' = \{m | P(x_1)/d + ... + P(x_g)/d = m, \text{ for some } (x_1, ..., x_g) \text{ in } T\}$. Notice that the definitions of A, T and T' make it clear that $T' \subseteq gA = B$. Also notice that the definition of S implies that if $m \in T'$, then $m > 0$. Our plan is to show that every $m \in T'$ has $1 \leq m \leq CM^n$. That will show $B(CM^n) \geq |T'|$. We will also show $|T'| \geq C'M^n$. Together, those facts prove the claim.

We now turn to the details, beginning by showing $m \in T'$ implies $1 \leq m \leq CM^n$, the lower bound having already been noted. For $(x_1, ..., x_g)$ in T, and for $1 \leq i \leq g$, we have $0 \leq z' \leq x_i \leq z' + d(M - 1) \leq z'M + dM = (z' + d)M$. Since $d \geq 1$ and $z' \geq 0$, we have $(z' + d)M \geq M$. The choice of M shows that $P(x_i) \leq 2c((z' + d)M)^n$. Thus, for $(x_1, ..., x_g)$ in T, we have $P(x_1) + ... + P(x_g) \leq 2gc(z' + d)^aM^n = CM^n$. Therefore, if $m \in T'$, then $1 < m \leq CM^n$, as desired. We now know $B(CM^n) \geq |T'|$.

It remains to show that $|T'| \geq C'M^n$, which is a bit harder. We will do that by first finding upper and lower bounds for $|T|$, beginning with the lower bound. Let
T'' = \{(x_1, \ldots, x_g) | \text{for } 1 \leq i \leq g, \text{ we have } z' \leq x_i \leq z' + d(M - 1) \text{ and } x_i \equiv z' \mod d\}. \text{ We will show that } T'' \subseteq T. \text{ Consider some component } x_i \text{ of some } (x_1, \ldots, x_g) \text{ in } T''. \text{ We need to show that each } x_i \in S \text{ and that } d \text{ divides } P(x_i). \text{ Our first need is satisfied by the fact that } x_i \geq z' \geq \alpha \text{ and } x_i \geq z' > \max\{u \geq \alpha | u \notin S\}. \text{ Our second need is satisfied by the fact that } x_i \equiv z' \equiv z \mod d \text{ implies } P(x_i) = P(z) \mod d, \text{ and } P(z) = d. \text{ Thus } T'' \subseteq T. \text{ Now there are } M \text{ choices of } x_i \text{ with } z' \leq x_i \leq z' + d(M - 1) \text{ satisfying } x_i \equiv z' \mod d. \text{ Therefore } |T| \geq |T''| = M^g. \text{ That is our lower bound on } |T|.

For } m \text{ in } T', \text{ let } R(m) \text{ be the number of } (x_1, \ldots, x_g) \text{ in } T \text{ with } P(x_1)/d + \ldots + P(x_g)/d = m. \text{ Obviously } |T| = \sum_{m \in T'} R(m).

Let } (x_1, \ldots, x_g) \text{ be in } T. \text{ We previously saw that for } 1 \leq i \leq g, \text{ we have } 0 \leq x_i \leq (z' + d)M. \text{ Since } P(x_1)/d + \ldots + P(x_g)/d = m \in T \text{ implies } P(x_1) + \ldots + P(x_g) = md, \text{ the definition of } r_{PNg}(md) \text{ with } N = (z' + d)M \text{ shows that for } m \in T, R(m) \leq r_{P((z'+d)M)g}(md). \text{ By the fundamental lemma, we have } R(m) \leq K(z' + d)^g M^{g-n}. \text{ It follows from the conclusion of the previous paragraph that } |T| \leq |T'||K(z' + d)^g M^{g-n}. \text{ That is our upper bound for } |T|. \text{ Comparing our upper and lower bounds for } |T|, \text{ we see that } |T'| \geq \frac{M^g}{K(z' + d)^{g-n} M^{g-n}} = C'M^n, \text{ completing the proof of the claim.}

We now turn to showing that GLB\{B(N)/N | N \geq 1\} \text{ is positive. Consider the smallest integer } M_0 \geq 1 \text{ satisfying the constraints imposed on our integer } M. \text{ Suppose } N < CM_0^n. \text{ By hypothesis, we have } 1 = P(z)/d \in A \subseteq B. \text{ Thus } B(N)/N \geq 1/N > \frac{1}{CM_0^n}.\text{ \hfill \qed}
Now suppose $C M_0^n \leq N$. Any integer $M \geq M_0$ also satisfies those constraints, and so we may assume $M$ has been chosen with $C M^n \leq N < C(M + 1)^n$.

We have $B(N)/N \geq B(CM^n)/N > B(CM^n)/C(M + 1)^n$. By the claim, we get

$$B(N)/N > \frac{C'M^n}{C(M + 1)^n} = \left(\frac{C'}{C}\right) \left(\frac{M}{M + 1}\right)^n.$$ 

Since $M \geq 1$, we have $\left(\frac{M}{M + 1}\right)^n \geq (1/2)^n$, so that

$$B(N)/N > \frac{C'}{2^n C}.$$ 

Combining the two cases, we see that $B(N)/N > \min\left\{\frac{1}{CM_0^n}, \frac{C'}{2^n C}\right\} > 0$, and we are done.