Problem

Problem Statement: Define $f(n,2)$ to be the maximal size of a subset $S \in K_n$, which $K_n = \{a|a \in \mathbb{N} \ \& \ a \leq n\}$ such that it follows the axiom that if $b \in S$, $2b$ is not in S. What is $f(n,2)$?

Cases:

1. $f(1,2) = 1$, only 1.
2. $f(2,2) = 1$, being 1 or 2.
3. $f(3,2) = 2$, either $1,3$ or $2,3$.
4. $f(4,2) = 3$, being $1,3,4$.

Observation 1.1: $f(n,2) \geq \lceil \frac{n}{2} \rceil$

Proof: Trivial because all odd positive integers less than or equal to n would work. □

Now, consider the power of 2 for each positive integers in the range, rewriting it in the form of $k = 2^n p_i$ for which $n = v_2(k)$, then when $n = 0$ would always work. Now, we define

$$G_k = \{2^k a \mid a \nmid 2 \ \& \ 2^k a < n\}$$

For n, the sets are defined as $G_0, G_1, ..., G_{\lfloor \log_2 n \rfloor}$, and define a set S_G that encompasses all of the above subsets.

Then clearly, we cannot take $2G_k$ if we take all the elements of G_k. Therefore, we either take $G_0, G_2, G_4, ..., G_{2s}$ or $G_1, G_3, ..., G_{2s+1}$. Now, it suffices to find the value of G_s.

Claim 1.2: $|G_s| = \left\lceil \frac{\lfloor \frac{n}{2^s} \rfloor}{2} \right\rceil$ which $\lfloor a \rfloor$ is defined as the largest integer less than or equal to a, and $\lceil a \rceil$ is defined as the smallest integer greater than or equal to a.

Proof: We divide this into two claims:

Subclaim 1: The number of positive integers k less than or equal to n such that $2^s \mid k$ is equal to $\lfloor \frac{n}{2^s} \rfloor$.

Proof: $2^s a \leq n$ for some positive integer a, and from here, $a \leq \frac{n}{2^s}$. If $\frac{n}{2^s}$ is not an integer, then the maximum pos.int a that is less than or equal to $\lfloor \frac{n}{2^s} \rfloor$ by definition. □

Subclaim 2: The number of odd positive integers less than n is $\lceil \frac{n}{2} \rceil$

Proof: Followed by the prove of observation 1.1. $\square$
The claim follows by combining these two subclaims.

Q.E.D.

To create a single expression of claim 1.2, we want to find a simpler form of G_s.

Conjecture 1.3: $|G_s|$ can be written as $\left\lceil \frac{m}{2^{s+1}} \right\rceil$ for some positive integer m.

$$\text{Now we define } p(x,y) = \left\lceil \frac{\left\lfloor \frac{x}{2^y} \right\rfloor}{2} \right\rceil \text{, then } p(a_1 2^m + b_1, m) = \left\lceil \frac{a_1}{2} \right\rceil \text{ such that } 0 \le b_1 < 2^m$$

From here, define n_s to be $n \pmod{2^s}$, then $f(n,s) = \left\lceil \frac{n-n_s}{2^{s+1}} \right\rceil$. Therefore, $\boxed{m = n - n_s}$. Then, we split into cases about the parity of a_1 :

1. If a_1 is even, then $a_1 2^m + b_1 \equiv b_1 \pmod{2^{m+1}}$
2. If a_1 is odd, then $a_1 2^m + b_1 = (a_1 - 1)2^m + 2^m + b_1 \equiv 2^m + b_1 \pmod{2^{m+1}}$

Assume that $n = 50$, then

1. $|G_0| = \left\lceil \frac{\left\lfloor \frac{50}{2^0} \right\rfloor}{2} \right\rceil = 25,\ 50 \equiv 0 \pmod 2$
2. $|G_1| = \left\lceil \frac{\left\lfloor \frac{50}{2^1} \right\rfloor}{2} \right\rceil = 13,\ 50 \equiv 2 \pmod 4$
3. $|G_2| = \left\lceil \frac{\left\lfloor \frac{50}{2^2} \right\rfloor}{2} \right\rceil = 6,\ 50 \equiv 2 \pmod 8$
4. $|G_3| = \left\lceil \frac{\left\lfloor \frac{50}{2^3} \right\rfloor}{2} \right\rceil = 3,\ 50 \equiv 2 \pmod{16}$
5. $|G_4| = \left\lceil \frac{\left\lfloor \frac{50}{2^4} \right\rfloor}{2} \right\rceil = 2,\ 50 \equiv 18 \pmod{32}$
6. $|G_5| = \left\lceil \frac{\left\lfloor \frac{50}{2^5} \right\rfloor}{2} \right\rceil = 1,\ 50 \equiv 50 \pmod{64}$

Conjecture 1.4:

1. $|G_s| = \left\lceil \frac{n}{2^{s+1}} \right\rceil$ if $n \equiv a \pmod{2^{s+1}}$ and $a \ge 2^s$
2.

Now, let $g(m,s) = \left\lceil \frac{m}{2^{s+1}} \right\rceil$, then $p(m,s) = g(m - 2^s, s)$

Now let $m = a_1 2^s + b_1$, which by euclidean algorithm, this will always be possible. Then, $f(m,s) = \lceil \frac{a_1}{2} \rceil$.

$$g(m-2^s, s) = \left\lceil \frac{a_1 2^s + b_1}{2^{s+1}} \right\rceil$$

Which because of the bound that $0 \leq b_1 < 2^s$, $\frac{b_1}{2^{s+1}} < \frac{2^s}{2^{s+1}} = \frac{1}{2}$. Therefore, b_1 will have no impact on the value of $g(m-2^s, s)$. $\square$

Therefore, $g(m-2^s, s) = \lceil \frac{a_1}{2} \rceil$, and the prove is complete.

In conclusion, $|G_m| = \lceil \frac{n-2^m}{2^{m+1}} \rceil$, and this allows us to massively simplify our expression. Before that, we need to prove that $|G_0| + |G_2| + ... > |G_1| + |G_3| + ...$ Consider the differences $|G_0| - |G_1|, |G_2| - |G_3|, ..., |G_{2n}| - |G_{2n+1}|$. For the remaining, consider $u = \lfloor \log_2 n \rfloor \equiv 1 \pmod 2$. Now we separate the cases by parity:

1. if $u \equiv 1 \pmod 2$, then we will be able to partition $G_0, ..., G_u$ into $\frac{u+1}{2}$ subsets of $\{G_{2n}, G_{2n+1}\}$ such that $0 \leq n \leq \frac{u-1}{2}$, and clearly $|G_{2n}| - |G_{2n+1}| > 0$, thus, the inequality

$$\sum_{m=0}^{\frac{u-1}{2}} |G_{2m}| - |G_{2m+1}| > 0$$

holds. $\square$

2. If $u \equiv 0 \pmod 2$, then consider the set $S \backslash \{G_u\}$. Then, $|S \backslash \{G_u\}| \equiv 1 \pmod 2$, which by subcase 1, the inequality holds. Now, this simply transforms to

$$\left(\sum_{m=0}^{\frac{u-2}{2}} |G_{2m}| - |G_{2m+1}| \right) + |G_u| > 0$$

Which is true since $|G_u|$ is always larger than 0 unless if $2^u > n$. $\square$

In conclusion, let $a = \lfloor \log_2 n \rfloor$, then $\boxed{f(n,2) = \sum_{m=0}^{\lfloor \frac{a}{2} \rfloor} \left\lceil \frac{n - 2^{2m}}{2^{2m+1}} \right\rceil}$

However, this will become much more complicated when k becomes large. Now, consider D_n to be the multiples of 2^n. Then, we can express all of these as $|D_{2n}| - |D_{2n-1}|$, and we can just combine as a single summation.

$$\boxed{f(n,2) = \sum_{i=0}^{\lfloor \log_2 n \rfloor} \left\lfloor \frac{n}{2^i} \right\rfloor}$$

And the idea is the same for all k other than 2 because we just want to take G_{2n} by subtracting all the multiples of k^{2n} by all the multiples of k^{2n+1}. Then,

$$\boxed{f(n,k) = \sum_{i=0}^{\lfloor \log_2 n \rfloor} \left\lfloor \frac{n}{k^i} \right\rfloor}$$

With simulation, it's evident that $\lim_{n\to\infty} f(n,k) = \frac{kn}{k+1}$

The Extended Problem Statement: Define $f(n, k_1, k_2)$ to be the maximal size of a subset $S \in K_n$, which $K_n = \{a | a \in \mathbb{N} \ \& \ a \leq n\}$ such that it follows the axiom that if $b \in S$, kb is not in S for all k_1, k_2. What is $f(n, k_1, k_2)$?
Intuition: What if I add another variable after k? What would that potentially change?

First of all, we can consider the multiples of k^0 with the previous strategy. Notice that for all $a \in K_n$, $a \equiv b \pmod{n}$ such that $b \in \frac{\mathbb{Z}}{k\mathbb{Z}}$, and I will first solve the problem when $\gcd(k_1, k_2) = 1$ and $k_1, k_2 \mid n$.

Observation 2.1: Assume that $\gcd(k_1, k_2) = 1$, then $f(n, k) \geq (k_1 - 1)(k_2 - 1) \left\lfloor \frac{n}{k_1 k_2} \right\rfloor$

Proof: This is a result of the Chinese Remainder Theorem: We can take every positive integers a that aren't divisible by k_1 or k_2, which by modular arithmetics,

1. $a \equiv 1, ..., k_1 - 1 \pmod{k_1}$
2. $a \equiv 1, ..., k_2 - 1 \pmod{k_2}$

Which since $\gcd(k_1, k_2) = 1$, then by Chinese Remainder Theorem, every pair of $a_1 \pmod{k_1}$ and $a_2 \pmod{k_2}$ will produce a unique residue modulo $k_1 k_2$. Therefore, there are a total $(k_1 - 1)(k_2 - 1)$ of residues in a cycle of length $k_1 k_2$ that is divisible by neither k_1 nor k_2. From here, we simply want to find the number of such cycles, which there are $\left\lfloor \frac{n}{k_1 k_2} \right\rfloor$ of them. $\square$.

However, this is not the end. We will also have $k_1^{2m}(k_1 x_1 + 1), k_1^{2m}(k_1 x_1 + 2)...$ and $k_2^{2m}(k_2 x_2 + 1), k_2^{2m}(k_2 x_2 + 2), ...$ Therefore, let $G_{k_1,m}$ be the set of all positive integers a such that $k_1^m a < n$ and $a \in \frac{\mathbb{Z}}{k_1 \mathbb{Z}} - \{0\}$.

Conjecture 2.2: Define $G_{k_1,m}$ be the set of positive integers p less than or equal to n such that $v_{k_1}(p) = m$. $G_{k_2,m}$ is defined analogously. then, the sets

$$S_{k_1,m}\{a \mid a \in k_1^{2m} G_{k_1 k_2,0} \text{ and } a \leq n\}$$

$$S_{k_2,m}\{b \mid b \in k_2^{2m} G_{k_1 k_2,0} \text{ and } b \leq n\}$$

will work.

Proof: Follow directly from the Chinese Remainder Theorem as a result of *Observation 2.1* and the extended version of *Conjecture 1.4*. $\square$.

Now, we simply want to find the set of quotients when a positive integer $a \in K_n$ is divided by powers of k_1 and k_2. There yields $p_1 = \left\lfloor \frac{n}{k_1^m} \right\rfloor$ such positive integers, and analogous defined for p_2, which the set will be in the form $\{1, ..., a_{p_1}\} \leftrightarrow \{k_1 m, ..., k_1^m a_{p_1}\}$ and the former set is simply K_{p_1}.

From here, use the Euclid's Algorithm by rewriting $p_1 = xk_1k_2 + r$, which there will simply be $x(k_1-1)(k_2-1)+r$ positive integers in the set $S_{k_1,m}$. From here, the answer is

$$x\left(\frac{p_1 - r}{xk_1k_2}\right)\varphi(k_1k_2) + r = \frac{(p_1 - r)\varphi(k_1k_2)}{k_1k_2} + r$$

To further simplify this, let's make clear of what $p_1 - r$ is. $p_1 - r$ is basically the largest multiple of k_1k_2 less than p_1, and p_1 is the largest quotient when a positive integer less than or equal to n divided by k_1^m, from our conjecture. $p_1 = \left\lfloor \frac{n}{k_1^m} \right\rfloor$

$p_1 - r$ is simply $k_1k_2 \left\lfloor \frac{\left\lfloor \frac{n}{k_1^m} \right\rfloor}{k_1k_2} \right\rfloor$. The conjecture so far is that this is equal to $\left\lfloor \frac{n}{k_1^{m+1}k_2} \right\rfloor$.

This conjecture is flawed because for $f(6,2,3)$, you can take $1, 4, 5, 6$ which will yield a result of 4 instead of 2. Conjecture 2.2 is disproved by this counterexample. $\blacksquare$.

Salvaged Conjecture 2.3: Assume that $k_1 < k_2$, then

$$f(n, k_1, k_2) = \sum_{\zeta=0}^{\infty} \left(\left\lfloor \frac{n}{(k_1 k_2)^{\zeta}} \right\rfloor - \left\lfloor \frac{n}{k_1 (k_1 k_2)^{\zeta}} \right\rfloor \right)$$

Proof: It's always possible to partition K_n into various strings of $\{p, ..., p k_1^a k_2^b\}$ such that $\gcd(k_1, k_2, p) = 1$. First, pick out all positive integers $a \in (\frac{n}{k_1}, n]$. It's easy to prove that this construction works because the function $\frac{n}{x}$ is monotonically decreasing, and the maximum value of of such multiple chain will be extended from n, which is $\frac{n}{k_1}$ and this is excluded from the set. $\square$

Claim: that the next stage is between the set $\left(\left\lfloor \frac{n}{k_1^2 k_2} \right\rfloor, \frac{n}{k_1 k_2} \right]$

Proof: For the sake of contradiction, assume that there exists an integer between $\left[\frac{n}{k_1 k_2}, \frac{n}{k_1} \right]$. This is equivalent to the statement that there exist a chain of multiples $a_1, k_1 a_1$ or $a_1, k_2 a_1$ for any integer $a_1 \in \left[\frac{n}{k_1 k_2} + 1, \frac{n}{k_1} - 1 \right]$.

Proceed with bounding: $\frac{n}{k_2} + k_1 \leq a_1 k_1 \leq n - k_1 < n$. $\square$

What if n is not divisible by k_1, nor k_2? In this case, we simply just take the floor value.

Example: Consider $f(1296, 2, 3)$, first we take all positive integers between $[649, 1296]$, and then we take $[109, 216], [19, 36], [4, 6], [1]$. This yields $648 + 108 + 18 + 3 + 1 = \boxed{778}$

For $f(1296, 3, 5)$ instead, we first take $[433, 1296]$. Then, the maximum value of $x \in [1, 432]$ is 86. Then, we take from 86 to $\left\lfloor \frac{86}{3} \right\rfloor + 1$, which is $[29, 86]$. Continue with this, we take $[2, 5]$. This gives $864 + 58 + 4 = \boxed{926}$

However, we see a counterexample that if $n = 20, k_1 = 2, k_2 = 5$, then we can choose $[11, 20]$ and $[2]$ by our algorithm, but notice that 5 can also be taken because $5 \cdot 5 > n$. Therefore, we have to also take positive integers $a \in \left[\left\lfloor \frac{n}{k_1 k_2} \right\rfloor, \left\lfloor \frac{n}{k_1} \right\rfloor \right]$, such that $\frac{a}{k_1} > \frac{n}{k_1^2 k_2}$ and $k_2 a > n$ and $k_1 a < \frac{n}{k_1}$.

Therefore, we will Salvage our current conjecture again:

Salvaged Conjecture 2.4: Assume that $k_1 < k_2$, then

$$f(n, k_1, k_2) = \sum_{\zeta=0}^{\infty} \left(\left\lfloor \frac{n}{(k_1 k_2)^{\zeta}} \right\rfloor - \left\lfloor \frac{n}{k_1 (k_1 k_2)^{\zeta}} \right\rfloor \right) + S_n$$

Such that S_n is a set that contains all positive integers a such that

$$\begin{cases} \frac{a}{k_1} > \frac{n}{k_1^2 k_2} \\ k_2 a > n \\ k_1 a < \frac{n}{k_1} \end{cases}$$

We bound the possible values a: $\frac{n}{k_2} < a < \frac{n}{k_1^2}$. We can ignore the first bound of $a > \frac{n}{k_1 k_2}$ because $\frac{n}{k_1 k_2} < \frac{n}{k_2}$. From here, we deduce that $k_1^2 < k_2$ must be true for $|S_n| > 0$, and this will produce about $n\frac{k_2 - k_1^2}{k_1^2 k_2}$ values.

Now, we continue to approximate the maximal number of elements.

$$f(n, k_1, k_2) = \sum_{\zeta=0}^{\infty} \left(\left\lfloor \frac{n}{(k_1 k_2)^{\zeta}} \right\rfloor - \left\lfloor \frac{n}{k_1 (k_1 k_2)^{\zeta}} \right\rfloor \right) + S_n$$

Which when $k_1^2 < k_2$, $|S_n| \approx \frac{(k_2 - k_1^2)n}{k_1^2 k_2}$. For the sake of approximation, assume that the floor values vanish. Then,

$$f(n, k_1, k_2) \approx \sum_{\zeta=0}^{\infty} \frac{n}{(k_1 k_2)^{\zeta}} \left(\frac{k_1 - 1}{k_1} \right) = \frac{(k_1 - 1)n}{k_1} \left(\frac{k_1 k_2}{k_1 k_2 - 1} \right) = n \left(\frac{k_1 k_2 - k_2}{k_1 k_2 - 1} \right)$$

Conjecture 2.5:

1. If $k_1^2 > k_2$, then
$$f(n, k_1, k_2) \approx n \left(\frac{k_1 k_2 - k_2}{k_1 k_2 - 1} \right)$$

2. If $k_1^2 < k_2$, then
$$f(n, k_1, k_2) \approx n \left(\frac{k_1 k_2 - k_2}{k_1 k_2 - 1} + \frac{k_2 - k_1^2}{k_1^2 k_2} \right)$$

Potential Ideas to prove or disprove this conjecture:

1. Pigeonhole Principle

2. Greedy Algorithm

Proof: When there is only 1 constraint, then we can partition K_n to be groups of size 1 or 2, which are $\{a, ka\}$ such that $v_k(a) \equiv 0 \pmod{2}$, which follows from conjecture 1.4.

Those positive integers s such that $ks > n$ will only contain 1 element, which by bounding, are between $\lceil \frac{n}{k} \rceil$ and n. Furthermore, for each $p \in K_n$ such that $\gcd(p, k) = 1$, let $g_k(p)$ be the number of elements in K_n it will contribute, then:

1. If $\left\lfloor \log_k \frac{n}{p} \right\rfloor = 2m + 1$ for some $m \in \mathbb{N}$, then $g_k(p) = m + 1$ by Pigeonhole Principle. □
2. If $\left\lfloor \log_k \frac{n}{p} \right\rfloor = 2m$ for some $m \in \mathbb{N}$, then $g_k(p) = m + 1$ by Pigeonhole Principle. □

Now, we simply want to split into the ranges of $\frac{n}{k^a}$ and $\frac{n}{k^{a+1}}$ and determine how many positive integers are there within that range that is relatively prime to k, which there are approximately $\frac{\varphi(k)}{k}\left(\frac{n}{k^m} - \frac{n}{k^{m+1}}\right)$.

Therefore, to sum up, we will have

$$n\left(1 - \frac{1}{k}\right)\frac{\varphi(k)}{k}\sum_{m=1}^{\infty}\left\lfloor\frac{m+1}{2}\right\rfloor k^{-m}$$

Which $\frac{\varphi(k)}{k}$ is the probability density function for the number of a within kp_1, kp_2 that is relatively prime to k.

For the extended problem statement of $f(n, k_1, k_2)$, consider the isomorphism

$$f : \mathbb{N}^2 \to \mathbb{R}$$

Then, for each value $a = k_1^{n_1} k_2^{n_2} r$ for some r such that $\gcd(r, k_1 k_2) = 1$, then f^{-1} maps every divisors of $\frac{a}{r}$ into lattice points $\mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2}$. Therefore, the number of lattice points is evidently the number of divisors of $\frac{a}{r}$, which is $(n_1 + 1)(n_2 + 1)$.

Subclaim 1. The maximum number of lattice points that can be selected such that no two are directly connected by an edge is $\left\lceil \frac{(n_1+1)(n_2+1)}{2} \right\rceil$

Proof: FTSOC assume there exists a number c larger than the bound, then it's equivalent to saying there is at least one pair of the points that are directly connected by the Pigeonhole Principle. □

This implies that if we define $(m_1, m_2) \in \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ to be the peak. Furthermore, we have

$$f(p, q) : \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \to \mathbb{Z}^+ \text{ defined by } f(p, q) = |p - m_1| + |q - m_2|$$

And if we call $f(p, q)$ to be the layer of point (p, q) in $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$, then from the proof of subclaim 1, we can only take $1, 3, 5, ...2y + 1$th layers, or $2, 4, 6, .., 2y$ layers, which the layers cannot be adjacent, and we can take all the numbers in that layer to maximize the size of the subset.

Now, notice that if a exists such that $ak_1^{m_1} k_2^{m_2}, ak_1^{n_1} k_2^{n_2} \in K_n$, then a would exist on the graph of both. However, this ensures our algorithm is correct when we select all the positive integers within the range of $(\frac{n}{k_1}, n]$, then every positive integers between will be the "peak" of the lattice points. Assume we take m as the peak, then we will need the 3rd layer, which is $\frac{m}{rk_1k_2}$, $\frac{m}{rk_2^2}$, $\frac{m}{rk_1^2}$ if they exist. This corresponds to $(p - 2, q), (p, q - 2)$ and $(p - 1, q - 1)$ if (p, q) is the peak.

This bounding gives

$$\frac{n}{k_1^2 k_2} < \frac{m}{k_2^2} < \frac{m}{k_1 k_2} \leq \frac{n}{k_1 k_2}$$

If $k_1^2 < k_2$, then $\frac{n}{k_2} < \frac{n}{k_1^2} < \frac{n}{k_1}$, which accounts for the second part of the conjecture and we are done. □

Final Final final problem statement: Let P_n be the set of prime divisors of n(for example, $P_{50} = \{2, 5\}$ and $P_{30} = \{2, 3, 5\}$). Define $K_n = \{a \in \mathbb{N} | a \leq n\}$ What is the maximal size of a subset $S \in K_n$ such that it follows the axiom that if $b \in S$, $kb \notin S$ for all $k \in P_n$.

To prove the strategy, we need to consider the similar mapping strategy:

$$\text{Consider the mapping } f : \mathbb{N}^{|P_n|} \to \mathbb{R}$$

Then, we can show that

$$f^{-1}\left(\prod_{m=1}^{|p_n|} p_m^{a_m}\right)$$

Will form an object with at most $|p_n|$ dimensional linear surface, which is trivial because the map connects a with $\frac{a}{p_i}$ for all $1 \leq i \leq |p_n|$. $\square$

Assume that we choose the peak to be an arbitrary value $a = p_1^{a_1} p_2^{a_2} ... p_{|p_n|}^{a_{|p_n|}}$. Ironically, the most number of vertices we can choose such that no two are directly connected is

$$\left\lceil \frac{1}{2} \prod_{m=1}^{|p_n|} (a_m + 1) \right\rceil$$

Proof. Assume contradiction that we can take more than that, then by the Pigeonhole Principle, we must have two nodes chosen that are consecutive. $\square$