Naïve Sieve Theory – by Peter Braun

Introduction:
Naïve sieve theory is not strictly speaking sieve theory because the immediate goal is not about deducing results. It is about supporting the notion of explanation and looking at simple conditions on the sieve of Eratosthenes which provide possible ‘causal’ type explanations of theorems. In this sense it is different from results type number theory and also different from probabilistic number theory. But it is also about searching for unprovable statements which may be used as an underlying cause for things we would like to prove. In a way this may be viewed as a way out of some of the problem solving methods which have tried to deduce answers using manipulation backed by extraordinary ideas. These methods do seem to have limits. Chen’s theorem is tantalisingly close to the end result for Goldbach’s conjecture but the parity problem could be an obstacle which cannot be moved around. In this area of problem it seems more important that 2 is roughly 2 and 3 is roughly 3 and 5 is roughly 5 etc. This is not at all the case with the Riemann hypothesis where we only need a certain density of primes and it is not at all important how many of the early ones are left out. Thus we look for global statements which provide results and then wonder if there is some way we will be able to explain the statements as unprovable.

The mechanism is simply A (unprovable), A → I (true).
It is not suggested that the following A’s are unprovable (or even true) but we are dealing with I’s which look as if they should be true and it is believed that this approach cannot be ruled out as a method of proof. We should look for A’s which are unprovable and I’s which are interesting. Ideally, we would have logical equivalence between A and I.

The discussion below provides naïve explanations for some of the well-known questions of interest.

Naïve Theorems
Usage and notation:
Let B be any set of integers. By a+pB we mean the set \{a+pb: bє B\}.
An set of numbers A will be said to be covered by a set of distinct primes P if each element of A is divisible by at least one prime in P.
If \(p_i\) is the smallest prime dividing \(n\) we say \(p_i\) is a significant multiple of \(n\).
Let \(P_n\) denote the set of the first \(n\) prime numbers.

(a) Let \(T_n\) be the maximum number such that \(T_n\) consecutive numbers are covered by \(P_n\).
(b) Let \(S_n\) be the maximum number such that the sets \(a_1+\mathbb{Z}, a_2+\mathbb{Z}, \ldots, a_n+\mathbb{Z}\)
combined, include the numbers \(1,2,3,\ldots\) \(S_n\), where \(a_i\) is a residue modulo \(p_i\) for \(1 \leq i \leq n\), and \(\mathbb{Z}\) denotes the integers.

Lemma
\(S_n = T_n\)

Proof
To this end consider a maximal set of consecutive numbers \(M+1, M+2, \ldots, M+T_n\)
which are covered by \(P_n\).
For each prime number \(p_i\) select the smallest \(r_i \geq 1\) such that \(p_i \mid M+ r_i\).
We may assume \(r_i\) exists amongst the numbers \(1,2,\ldots, T_n\) since otherwise \(p_i\) has not been ‘used’ as a covering prime in arriving at the maximal sequence and we easily
construct a longer sequence $K+1, K+2... K+T_n, K+T_n+1$ in which the first $T_n+1$ numbers are covered by $P_n \sim \{p_i\}$ and $p_i | M+T_n+1$ which contradicts the maximal choice of $T_n$. Indeed for $j \neq i$ select $J \equiv M \mod p_i$ and $J \equiv M+T_n+1 \mod p_i$ using the Chinese remainder theorem. But then $J+1, J+2... J+T_n+1$ are covered by $P_n$.

Then the sets $r_1+p_1Z, r_2+p_2Z... r_n+p_nZ$ include the numbers 1, 2, 3... $T_n$

It follows that $S_n \geq T_n$.

To get the inequality in the other direction let $a_1+p_1Z, a_2+p_2Z... a_n+p_nZ$ be a choice which produces a maximal sequence 1, 2... $S_n$ of elements in these sets.

By the Chinese remainder theorem we may choose $J \equiv -a_i \mod p_i$ for $1 \leq i \leq n$.

But then $J+1, J+2... J+S_n$ are covered by $P_n$ which means $T_n \geq S_n$.

It is useful to hold the equivalence of (a) and (b) firmly in mind in considering certain problems in arithmetic.

We know that $T_n > p_{(n+1)}-2$ for $n > 5$ so that sieving using $M \equiv 0 \mod p_i$ ($1 \leq i \leq n$) does not immediately come across a maximum number of consecutive numbers covered by $P_n$. We need to look at larger numbers to find such a sequence.

Instead of starting $p_1$ on $p_1, p_2$ on $p_2... p_n$ on $p_n$ in the sieve of Eratosthenes we may modify the sieve and start $p_i$, for some specify choice of residues and immediately mark out $T_n$ consecutive numbers using the marking method of the sieve. Indeed, if $M+1, M+2... M+T_n$ is a maximal sequence and $M \equiv -a_i \mod p_i$ $(1 \leq i \leq n)$ we just mark $a_i$ and $a_i + \theta p_i$ and the maximum number of consecutive numbers will be marked. This provides an alternative method of determining the value of $T_n$.

We note that a maximum run of consecutive integers covered by $P_n$ includes significant multiples of each prime $p_1, p_2... p_n$.

**Goldbach's conjecture (GB)**

We describe here a small extension of the above thinking with an assumption which implies GB. The idea is not to wonder whether the assumption is true but just that it provides a simple explanation in terms of a cause.

Consider two standard number lines one below the other. 

On the top line mark out numbers according to the sieve for the prime $p_i$, for $p_1, p_2... p_n$.

On the bottom line choose an arbitrary selection of residues $a_i$ modulo $p_i$ ($1 \leq i \leq n$) to start the marking, the only condition being that $p_i$ starts in the same place on the bottom line as it does on the top line (i.e. consistent with the sieve of Eratosthenes).

What we are looking for is the maximum length away from unity before we get a miss on both lines, one above the other. Let $GB_n$ denote this number. It is easy to see that $GB_n$ is well defined.

**Proposition 1**

If $GB_n < (p_n)^2$ then every even number may be written as the sum of two primes.

**Proof**

It should be obvious that this is a much stronger condition than is necessary to find an equivalent statement to GB. Indeed the standard sieve of Eratosthenes is one of the 'runs' which needs to be considered to arrive at the maximum $GB_n$.

Hence if we apply the sieve of Eratosthenes to the two number lines one below the other, using the primes with $p_1, p_2... p_T$ denoting the primes less than or equal to $\sqrt{E}$, on

<table>
<thead>
<tr>
<th>1 2 3 4...</th>
<th>E-1 E</th>
</tr>
</thead>
<tbody>
<tr>
<td>2E-1 2E-2 2E-3 2E-4...</td>
<td>E+1 E</td>
</tr>
</tbody>
</table>

moving left to right on the top line and right to left on the bottom line avoiding the double count on $E$, we find 2 prime numbers adding up to $2E$.  

July 2011
As mentioned the idea here is not to be overly interested in the truth of the assumption but purely to appreciate a causal explanation in a simple setting.

If we wish to start from a statement logically equivalent to Goldbach’s theorem we also need to take into account the primes less than or equal to √E as these will be marked in the sieving process (Eratosthenes).

A similar assumption also implies unbounded twin primes. This may be framed in a number of ways but the original sieve provides a simple example.

We easily see that if m, m+2 co-prime to P occurs unboundedly. Let TP be the minimum number such that the set \{a_i+p_i: 1 \leq i \leq n\} necessarily contains two numbers m and m+2 which are both co prime to P, for every choice of residues \(a_1, a_2, \ldots, a_n\).

**Proposition 2**

If TP < \((p_n)^2\) then there are unbounded twin primes.

**Proof**

Since TP covers all possible starts for marking \(p_1, p_2, \ldots\) it includes the sieve of Eratosthenes.

Thus the numbers 1, 2, 3, …, TP contain numbers m and m+2 which are not covered by the sieving process if we sieve using the primes \(p_1, p_2, \ldots, p_T\) where \(p_T \leq \sqrt{TP} < p_{(T+1)}\).

Since these numbers are co-prime to \(p_1, p_2, \ldots, p_T\) they are necessarily prime.

**Dirichlet’s Theorem on primes in arithmetic progressions**

Dirichlet’s celebrated theorem has also been proven more recently by elementary methods at the same level of strength as the prime number theorem (Gelfond and Linnik [3]).

We may also find a possible explanation for ‘unbounded primes’ in terms of \(T_n\).

**Proposition 3**

Let \((a,b) = 1\) and suppose \(T_n + 1 \leq ((p_{n+1})^2 - b)/a\) for sufficiently large \(n\).

Then the progression \(an + b\) supports unbounded primes.

**Proof**

For sufficiently large \(n\) we assert the sequence

\(a+b, a^2+b, \ldots, a(T_n+1)+b\) contains a member co-prime to \(P_n\).

Indeed, if this is not the case then with \(a \equiv -b \mod p_i\) (1 \(\leq i \leq n\)) we see the set \(\{a_i+p_iZ: 1 \leq i \leq n\}\) contains \(T_n+1\) consecutive numbers. But \(S_n = T_n\).

So at least one of \(a+b, a^2+b, \ldots, a(T_n+1)+b\) is co-prime to \(P_n\).

If it were composite we would have \(a(T_n+1)+b \geq (p_{n+1})^2\) and this is against the assumption of the proposition. This result is also implicit in the next propositions.

**The growth rate of \(T_n\)**

We note there are \((p_1-1)(p_2-1)\ldots(p_n-1)\) numbers less than \(P_n\) which are co-prime to \(P_n\) and these numbers are the ones which finish runs of consecutive numbers covered by \(P_n\). The average distance between these numbers is \(K_n = P_n/(p_1-1)(p_2-1)\ldots(p_n-1)\).

Now \(\ln(P_n)/(p_1-1)(p_2-1)\ldots(p_n-1)) = \sum \ln(1/(1/p)) = \ln(\ln(x)+A-1 + O(1/\ln x))\) as \(x \to \infty\) using Merten’s theorem (Chandaresakaran [2]).

Consequently, \(K_n = e^A \ln x + O(1)\) as \(x \to \infty\). There are many solutions to \(m, m+2\) co-prime to \(P_n\) so the downward variation on this estimate is known. There is however no useful upper bound widely known if known at all.
The growth rate of $T_n$ as $n \to \infty$

If we start sieving according to Eratosthenes we find $p_{n+1} - 2$ consecutive numbers covered by $P_n$. How much better can we do by starting our 'sieving' at different start points?

$T_n = \partial(n) p_n$ where $\partial(n) = O(n^\epsilon)$ as $n \to \infty$ is probably the very best we could ever hope for but consequences of this assumption indicate how difficult this might be to prove by elementary methods.

In which case $GB_n = \partial(n) (p_n)^2$ where $\partial(n) = O(n^\epsilon)$ as $n \to \infty$ may be enough to prove the Goldbach conjecture but not using the simple discussion earlier.

In Braun [1], the approach was taken to see what sort of results followed from assumptions about $T_n$ and these results represent refined arithmetic which resulted from joint work with A. Zulauf.

The theorems are stated here without proof. The results cover 3 topics

1. The least prime in an arithmetic progression
2. The unbounded number of primes in a progression
3. The distance between consecutive prime numbers.

Let $J(x)$ be the maximum number of consecutive numbers each of which is covered by primes less than or equal to $x$.

**Proposition 4**

Let $J(x) \leq Ax^{2-\alpha}$

The least prime $D_{m+l}$ satisfies $D_{m+l} \leq (A+1)^{2/\alpha} D^{2/\alpha} ((D,l) = 1, 0 < \alpha \leq 1)$.

**Proposition 5**

Let $J(x) \leq Ax^{2\beta/\beta+1}$.

There is a prime $D_{m+l}$ with $N \leq m \leq [(A+1)^{\beta} D^{\beta} N]$  $(N \geq 1, \beta \geq 1 (D,l) = 1)$.

**Proposition 6**

Let $J(x) \leq Ax^{1+2\gamma}$  $(\frac{1}{2} > \gamma \geq 0, A > 0)$.

For $n > N_0$ there is a prime in $[[N-AN^{(1/2)+\gamma} ], N]$.

$p_{n+1} - p_n = O ((p_n)^{(1/2+\gamma)})$ as $n \to \infty$.

**Notes (added July 2019):**

Intuitively, since all the primes cover all the natural numbers it would be very odd if asymptotically we did not have $\gamma=0$ or $\gamma=0+\varepsilon$. i.e. if this is not the case then the primes cover all the natural numbers plus a positive proportion more. In this is the case we do get some rather familiar targets. It would also be remarkable if $\gamma=0+\varepsilon$ were decidable in rational arithmetic although an upper bound for $\gamma$ would seem quite possible.

**References**