Extra Topics 2–3.
Bertrand's Postulate

This assignment is due *Tuesday* May 1 (update: actually, Wednesday May 2).

This is a (quite long) extra assignment, worth *twice* as much as a regular homework in terms of course grade. It is not required to complete the course. If you choose to do this assignment, the grade for it will only go to the numerator of your grade.

Each problem in this set (with all sub-items) is worth 10 points. (But not necessarily all problems have the same difficulty.) There are 130 points total, 105 points is considered 100%. If you go over 105 points, you will get over 100% for this homework (up to 115%) and it will count towards your course grade.

Collaboration is welcome. If you do collaborate, make sure to write/type your own paper *and give credit to your collaborators in your pledge.* Your solutions should contain full proofs. Bare answers will not earn you much.

No extensions will be granted under any conditions. On the other hand, if you choose to, you are welcome to submit problems in this assignment one-by-one or in any combination.

All material necessary to do this assignment was covered before the Spring Break.

In this problem set we prove the following statement, known as *Bertrand's Postulate*:

Theorem. *For each integer $n \geq 2$, there is at least one prime p s.t. $n < p < 2n$. That is, between n and $2n$ there is always at least one prime.*

This was conjectured by Joseph Louis François Bertrand in 1845 and proved by Pafnuty Lvochich Chebyshev in 1850. Later (in 1932) Paul Erdős found easier proof that we follow in the problems below.

General idea is that the binomial coefficient $\binom{2n}{n}$ is divisible by each prime between n and $2n$. With that in mind, we show that if there are no primes between n and $2n$ for some *sufficiently large* n, then the binomial coefficient $\binom{2n}{n}$ cannot be as large as it should be. In fact, this "sufficiently large" condition will turn out to be $n > 2048$ at worst, and we will establish the statement for $n \leq 2048$ separately.

(1) (a) Prove that $\binom{2n}{n}$ is the largest binomial coefficient among $\binom{2n}{k}$. (*Hint:* using $\binom{m}{k} = \frac{m(m-1)\cdots(m-k+1)}{k!}$, compare $\binom{2n}{k}$ and $\binom{2n}{k+1}$.)

(b) Show the equality

$$4^n = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k},$$

then using item (a), conclude $4^n \leq (2n+1)\binom{2n}{n}$.

In the problem above we showed that the binomial coefficient can be estimated below: $\binom{2n}{n} \geq \frac{4^n}{2n+1}$. In the rest of the problem set we show that if Bertrand's postulate fails, this lower bound is not met. Do do that, we need to understand prime decomposition of $\binom{2n}{n}$.

Recall that the *integer part* $[x]$ is the greatest integer below or equal to x, i.e. $[x] = \max\{n \in \mathbb{Z} \mid n \leq x\}$. For example, $[2] = 2, [\pi] = 3, [0.9999] = 0, [-0.7] = -1$.

(2) (a) Prove that for each prime p and integer $j \geq 1$, precisely $\left[\frac{n}{p^j}\right]$ numbers from 1 to n are divisible by p^j.

(b) Using item (a), show that the exponent of prime p in prime decomposition of $n!$ is

$$\left[\frac{n}{p}\right] + \left[\frac{n}{p^2}\right] + \left[\frac{n}{p^3}\right] + \cdots = \sum_{j=1}^{\infty} \left[\frac{n}{p^j}\right].$$

(Note that the sum is actually *finite* because when $p^j > n$, $[n/p^j] = 0$.)

Let $R(n,p)$ denote the exponent of prime number p in the decomposition of $\binom{2n}{n}$, so that

$$(*) \qquad \binom{2n}{n} = p_1^{R(n,p_1)} p_2^{R(n,p_2)} \cdots p_l^{R(n,p_l)}.$$

We will estimate this product above. To do that, we estimate each $R(n,p)$.

(3) Using $\binom{2n}{n} = \frac{(2n)!}{n!n!}$ and problem 2b, show that

$$(**) \qquad R(n,p) = \sum_{j=1}^{\infty} \left[\frac{2n}{p^j}\right] - 2\sum_{j=1}^{\infty} \left[\frac{n}{p^j}\right] = \sum_{j=1}^{\infty} \left(\left[\frac{2n}{p^j}\right] - 2\left[\frac{n}{p^j}\right]\right).$$

(4) (a) Prove that for any real number x, the value $[2x] - 2[x]$ is either 0 or 1. (*Hint:* Consider two cases: $m \leq x < m + 0.5$ and $m + 0.5 \leq x < m + 1$, $m \in \mathbb{Z}$.)

(b) Prove that if $j > \frac{\ln 2n}{\ln p}$, then $\left[\frac{2n}{p^j}\right] = 0$. Conclude the largest value of j that has a chance of nonzero contribution to $R(n,p)$ in $(**)$ is $\left[\frac{\ln 2n}{\ln p}\right]$. (*Hint:* Take logarithm of the inequality $2n < p^j$.)

(c) Show that

$$R(n,p) = \sum_{j=1}^{\infty} \left(\left[\frac{2n}{p^j}\right] - 2\left[\frac{n}{p^j}\right]\right) \leq \left[\frac{\ln 2n}{\ln p}\right].$$

(*Hint:* Use (a) to estimate each term; use (b) to estimate how many terms may be nonzero.)

Now that we estimated $R(n,p)$, estimate $p^{R(n,p)}$.

(5) Prove that $p^{R(n,p)} \leq 2n$.
(*Hint:* $p^{R(n,p)} = \exp(R(n,p) \ln p)$. Use $\left[\frac{\ln 2n}{\ln p}\right] \leq \frac{\ln 2n}{\ln p}$.)

Now we estimate how large the primes p_i in $(*)$ can be.

(6) (a) Prove that if $p > 2n$, then $R(n, p) = 0$.
(b) Prove that if $n \geq p > \frac{2n}{3}$, then $R(n, p) = 0$. (*Hint:* $p^2 > \left(\frac{2n}{3}\right)^2 > n$ for all $n \geq 3$ (case $n = 2$ is trivial anyway), so the only possible nonzero term in $(**)$ is the first one. Check that the first term is $2 - 2 \cdot 1$.)
(c) Similarly to (b), show that if $\frac{2n}{3} > p \geq \sqrt{2n}$, then $R(n, p) \leq 1$. (*Hint:* Use problem 4a.)

Assume that Bertrand's postulate fails for some n. Combining this assumption with the problem above, we get that all primes p that occur in $\binom{2n}{n}$ are either $p \leq \sqrt{2n}$ or $\sqrt{2n} < p \leq \frac{2n}{3}$. Indeed,

- primes $p > 2n$ are forbidden by problem 6a,
- $p = 2n$ is not a prime,
- primes $n < p < 2n$ do not exist by assumption,
- primes $2n/3 < p \leq n$ are forbidden by 6b.

According to this, we can write the prime decomposition of $\binom{2n}{n}$:

$$(\#) \qquad \binom{2n}{n} = \prod_{p \leq \sqrt{2n}} p^{R(n,p)} \cdot \prod_{\sqrt{2n} < p \leq \frac{2n}{3}} p^{R(n,p)},$$

where the products above are taken over primes p.

(7) Prove that

$$\binom{2n}{n} \leq (2n)^{\sqrt{2n}} \prod_{\sqrt{2n} < p \leq \frac{2n}{3}} p.$$

(*Hint:* In (#), estimate the first product using problem 5, and the second product using problem 6c.)

We are almost there. Our current goal is to show that $\binom{2n}{n}$ is too small. Looking at the formula above we see that we need to put an upper bound on the product of primes $\leq 2n/3$.

We show by induction that for every $m \geq 2$,

$$(\#\#) \qquad \prod_{p \leq m} p \leq 4^m.$$

(8) (a) Verify base of induction $m = 2$ and $m = 3$.
(b) Assume that the statement is true for all $m \leq 2k - 1$, $k \geq 2$, prove that it's also true for $m = 2k$. (*Hint:* $m = 2k$ is *not* prime for $k \geq 2$.)

Case $m = 2k$ is a bit trickier. We need an auxiliary statement:

(9) (a) Prove that each prime p such that $k + 2 \leq p \leq 2k + 1$ divides $\binom{2k+1}{k}$.
(b) Prove that $\binom{2k+1}{k} \leq 2^{2k}$. (*Hint:* Write $2^{2k+1} = (1+1)^{2k+1}$ and expand brackets. Use $\binom{2k+1}{k} = \binom{2k+1}{k+1}$.)

(10) Assume that (##) is true for all $m \leq 2k$. Prove that it's also true for $m = 2k + 1$. Write

$$\prod_{p \leq 2k+1} p = \prod_{p \leq k+1} p \cdot \prod_{k+2 \leq p \leq 2k+1} p,$$

estimate the first product using induction hypothesis and the second product using problem 9. This finishes proof of (##).

Combining problem 7 and inequality (##), we have

$$\begin{aligned}\binom{2n}{n} &\leq (2n)^{\sqrt{2n}} \prod_{\sqrt{2n}<p\leq \frac{2n}{3}} p \leq \\ &\leq (2n)^{\sqrt{2n}} \prod_{p\leq \frac{2n}{3}} p \leq \\ &\leq (2n)^{\sqrt{2n}} 4^{2n/3}.\end{aligned}$$

Finally, recall that we have a lower bound for $\binom{2n}{n}$ in the problem 1b, so we get

$$\frac{4^n}{2n+1} \leq \binom{2n}{n} \leq (2n)^{\sqrt{2n}} 4^{2n/3}.$$

But... this inequality is impossible for sufficiently large n: the left hand side grows faster than any $(4-\varepsilon)^n$, and the right hand side grows slower than any $(4^{2/3}+\varepsilon)^n$, so at some point the inequality breaks.

(11) [*This problem is optional because it involves either some calculus skills or some basic programming.*] Find n_0 such that the inequality

$$(\heartsuit) \qquad \frac{4^n}{2n+1} \leq (2n)^{\sqrt{2n}} 4^{2n/3}.$$

is impossible for $n \geq n_0$. You either can do computer-aided computation (which yields $n_0 = 468$), or you can take logarithm of both sides and do some arithmetic (which yields $n_0 = 2048$ if you do rough estimates).

Recall that the inequality $(\heartsuit)$ was obtained assuming that Bertrand's postulate fails for some n, so the fact that the inequality is impossible proves Bertrand's postulate for $n \geq n_0$. The only thing that's left to check is that it holds for $n < n_0$, which is a straightforward matter:

(12) Suppose $p_1 < p_2 < \ldots < p_k$ are primes such that $p_2 < 2p_1$, $p_3 < 2p_2$ and so on: $p_{i+1} < 2p_i$, $1 \leq i \leq k$. Prove that then Bertrand's postulate holds for any n between $p_1/2 < n < p_k$.

(13) Verify that $3, 5, 7, 13, 23, 43, 83, 163, 317, 631, 1259, 2503$ are primes that satisfy the condition above. Conclude that Bertrand's postulate holds for any $2 \leq n \leq 2502$.

Some final remarks. The original Chebyshev's proof is often called analytical, but in fact at its core it's about as combinatorial as this one. The main difference is that it uses $A(n) = \mathrm{lcm}(1, 2, 3, \ldots, 2n-1, 2n)$ instead of $\binom{2n}{n}$. (Note that $A(n)$ has the same property: each prime $n < p < 2n$ appears exactly once in $A(n)$.) This leads to an argument with fewer shortcuts, but on the upside, allows to easily modify proof to obtain statements like "for each $n \geq 100$, there are at least 10 primes between n and $2n$", or like "for each $n \geq 2$, there is at least one prime between n and $1.5n$".

It is also worth mentioning that Bertrand's postulate has a number of nice consequences, which I may cover later in extra topics or in lectures.