SRQR: Primes (April 5, 20212)

In the variation of the proof of theorem 3.3 at the top of page 43, how does $\sqrt{2}ar + \sqrt{2}bs = 2br + as$?

The final part of section 3.1 is very interesting. I have re-read it several times but still have no idea of what's wrong with the proof. Can you explain why it prove that square root of 2 is an integer, and which part is wrong that leads to this conclusion?

Suppose $\sqrt{2} = a/b$ with $\gcd(a, b) = 1$. Then there exist r and s such that $ar + bs = 1$, and multiplying both sides of this equation by $\sqrt{2}$ gives

$$\sqrt{2} = \sqrt{2}ar + \sqrt{2}bs$$

is clear. Notice that $\sqrt{2} = a/b$ means that $\sqrt{2}b = a$, so we can replace the $\sqrt{2}b$ in the second summand above with a. Also, note that

$$\sqrt{2} = \frac{a}{b} \implies \frac{a}{\sqrt{2}} = b \implies \frac{2a}{\sqrt{2}} = 2b \implies \sqrt{2}a = 2b,$$

so we can replace $\sqrt{2}a$ in the first summand above with $2b$. Thus

$$\sqrt{2} = 2br + as,$$

which shows that $\sqrt{2}$ is an integer. This is a contradiction, because... The square root of a positive integer n must be positive and less than or equal to n. But the only positive integers less or equal to 2 are 1 and 2, and neither of these squares to 2.

By intuition, theorem 3.4 seems true... However, the proof looks somehow absurd to me. Could you please explain that in more detail?

I am quite confused of the whole Euclid concept about theorem 3.4. Specifically, why we add 1 after multiplying all the prime numbers?

"Proof by contradiction" is also called *reductio ad absurdum*, which is Latin for "reduction to the absurd." So somehow it is absolutely right for you to find this proof "absurd" ☺

Let me write out this proof in a slightly different way. We want to prove that there are infinitely many primes, so we suppose for a contradiction that there are only finitely many primes. Let $p_1, \ldots, p_n$ be these finitely many primes. To find a contradiction, we consider the number

$$m = p_1 \cdots p_n + 1.$$

Since $m > 1$, the fundamental theorem of arithmetic tells us that this number m must have a prime factor p. Since $p_1, \dots, p_n$ are *all* of the primes, we know that p must be among the primes in this list; in other words, we have $p = p_i$ for some i. This means that $p \mid m$ and $p \mid p - 1 \cdots p_n$, so then p also divides

$$(m - p_1 \cdots p_n) = 1.$$

This is a contradiction, since a prime number cannot divide 1.

That's the end of the proof. What's the intuition behind considering the number m? Well, the idea is to write down a number that *cannot be divisible* by any of the finitely many primes $p_1, \dots, p_n$. Adding 1 to the product does precisely that: the result cannot be divisible by any of the primes $p_1, \dots, p_n$ because we're forcing there to be a remainder of 1 whenever we divide by any of those primes! The upshot is that no finite list can possibly exhaust all of the prime numbers.

On page 47 with the modifications to "Euclid's reasoning", how does one find that $p_n \leqslant p_1 p_2 \cdots p_{n-1} + 1$, or arrive at Bonse's inequality?

To prove that $p_n \leqslant p_1 p_2 \cdots p_{n-1} + 1 \dots$ Notice that in Euclid's proof of the infinitude of primes, it's proved that the number $P = p_1 p_2 \cdots p_{n-1} + 1$ is not divisible by any of the first $n - 1$ primes (ie, it's not divisible by $p_1, \dots, p_{n-1}$). But $P \geqslant 2$, so it has *some* smallest prime factor p. Since p is not any of $p_1, \dots, p_{n-1}$, we must have $p_n \leqslant p$, since p_n is the next biggest prime after p_{n-1}. But p is a divisor of P, so clearly $p \leqslant P$ also. Putting these two inequalities together, we see that

$$p_n \leqslant p \leqslant P = p_1 p_2 \cdots p_{n-1} + 1.$$

Bonse's inequality is not proved in the text. The proof requires some new ideas.

I'm confused how the final step of the Theorem 3.5 proof works; how does it make sense that you can just replace 1 with $\$2^{\{2}$n-1}? If p_{n+1} is actually less than or equal to $2^{2^n - 1} + 1$ then it would of course be less than or equal to $2^{2^n - 1} + 2^{2^n - 1}$, but we haven't quite established that fact, so it very well could not be.

Actually, we have established this fact! Let me rewrite this proof slightly to help clarify this. We want to prove that $p_n \leqslant 2^{2^{n-1}}$ for all $n \geqslant 1$. We will use induction; actually, we'll even use strong induction (ie, the "second principle of induction' '). The base case, $n = 1$, is clear because $p_1 = 2$ and $2^{2^{1-1}} = 2^1 = 2$ also.

For the inductive step, we're assuming that we know the inequality for $n = 1, \dots, k$. In other words, we are assuming that $p_1 \leqslant 2, p_2 \leqslant 2^2, \dots, p_k \leqslant 2^{2^{k-1}}$. We want to prove the inequality for $n = k + 1$ also. By the inequality discussed above, we know that

$$\begin{aligned} p_{k+1} &\leqslant p_1 p_2 \leqslant \cdots p_k + 1 \\ &\leqslant 2 \cdot 2^2 \cdots 2^{2^{k-1}} + 1 \\ &= 2^{1+2+2^2+\cdots+2^{k-1}} + 1 \\ &= 2^{2^k - 1} + 1. \end{aligned}$$

By this point, we've used our inductive hypothesis to prove that $p_{k+1} \leqslant 2^{2^k-1} + 1$. But what we really want to show that $p_{k+1} \leqslant 2^{2^k}$. The key here is to notice that $1 \leqslant 2^{2^{2^k}-1}$ (and this is true just because 2^{2^k-1} is a positive integer power of 2, so it has to be bigger than or equal to 1), so

$$p_{k+1} \leqslant 2^{2^k-1} + 1 \leqslant 2^{2^k-1} + 2^{2^k-1} = 2^{2^k}.$$

This completes the induction, proving that $p_n \leqslant 2^{2^{n-1}}$ for all n.

> Is there a standard way for figuring out if a number is prime or not?

> I meant to ask why, in determining if a number is prime, we only need consider potential prime divisors less than or equal to the square root of that number.

The "naive" procedure for testing primality of a number n is to check divisibility by each of the primes up to $\sqrt{n}$. The reason this works is... Suppose n is composite and p is the *smallest* factor of n that's not 1. Then p must be prime, because otherwise, a prime divisor of p would be a smaller divisor of n. Also, $q = n/p$ is also a factor of n, so, since p is the smallest factor of n, we have $p \leqslant q$. But then $p^2 \leqslant pq = n$, or, in other words, $p \leqslant \sqrt{n}$. This shows that, if n is composite, it's smallest factor must be at most $\sqrt{n}$.

I say this algorithm is "naive" because, from an algorithmic perspective, this is considered a slow algorithm. Usually with these kinds of things, number theorists and computer scientists measure the complexity of an algorithm in terms of the number of binary digits that the number has. If n is a number, it has roughly $k = \log_2(n)$ binary digits. The complexity using the Sieve of Eratosthenes of finding all primes up to $\sqrt{n}$ is

$$O(\sqrt{n} \log(\log n)) = O(2^{k/2} \log(k))$$

which is exponential in k.

There are other algorithms for primality testing whose running time is polynomial in k, but they are a bit harder to describe. If you're interested in computer science, you might decide to look into algorithms for primality testing for your project!