October 2020, revised January 2021

SOLUTIONS TO CMI MILLENIUM PRIZE PROBLEMS

Jorma Jormakka

Abstract

This document of 162 pages contains my solutions to five CMI Millennium Prize problems. Solutions to the remaining two problems are not here, as they are published elsewhere. This version from January 2021 differs from the original one in October 2020 in that I found an error in the 2008 proof of the Riemann Hypothesis and removed it. The 2020 proof is still alive. The Hodge counterexample is still in a journal review. Other papers are not submitted.

Foreword

If you ever get the great idea of solving the seven millennium prize problems that Clay posed in the year 2000, so my suggestion is just forget it. It is not because those problems would be totally impossible to solve - in fact, you might very possibly find a solution that would appear to you be correct - but you might find it very hard to get your solution read and checked by any expert or journal.

Should you submit it to a journal, my guess is that the likely outcome is that the journal rejects it on whatever basis, like that they could not find a referee, or that their expert recommended not publishing it, but you would not get a review that addresses the content. Though actually I did get two-three reviews that addressed the first few pages of the submitted paper and got stuck there, though there was no error in that place. When I answered the referee comments the result was that the journal could not find a referee or that their expert recommended rejecting the paper. I have some experience on this because in 2008-2010 I wrote a set of papers that solve all seven problems and submitted four of them to journals, but from all submissions of those four papers I did not receive a single review that was correct and addressed the content. Experts mostly would not answer. Should they answer - I got four answers in total - these answers are likely to be fast written and incorrect. Their goal is just to get rid of you. But one of these four answers was different: one expert did point out an error in my original proof of the Riemann Hypothesis. That was nice of him. Now the error is corrected and the revised version of the original proof is included in this book.

Of course you may try asking some friendly mathematician who is not any special expert or a journal editor/referee, but my guess is that he will not comment your solution, unless there is some obvious error or misunderstanding in your solution. If you have an easy error, you get answers. Even I have been asked to read a few of these incorrect solutions. I broke them all in a minute or so, and answered politely pointing out the error. But this would happen only if there are obvious errors.

Solutions to well-known problems are like hot potatoes: nobody wants to touch them and they would like to stay as far away from them as possible. If you have a solution to a well-known problem, you should assume that all mathematicians you ask to read it only hope that you would just disappear. Mathematicians do not want to attack well-known hard problems. Many of them think that I would not be modest to try to solve such a problem and the work to result ratio is better if one tries other problems. It is like a sportsman who would not try to get a gold medal in olympics because that would not be modest, so he would participate only in less important competitions because it gives a better work to result ratio. Maybe. But I think one should try to hard problems.

Some people react negatively to an effort to solve the millennium problems because there is the prize and mathematics should not be made for money. Let us just say that getting the prize is highly improbable. I very much doubt that Clay has any money: the problems were chosen to be nearly impossible to solve in any foreseeable future. Of course mathematics is not done for money. There are easier ways to get a million, such as robbing a bank. I, of course, did not start solving hard problems in 2000 when Clay posed the problems. The first problem I solved (or thought I solved, if it sounds better) was the Poincaré Conjecture. In 1986 I was doing my Ph.D. in mathematics and the supervisor went for a long time to the States, so that gave me the freedom of directing my work to a more challenging direction. I decided to look at the Conjecture. It took me a large part of 1986-87, and later many months afterwards all the way to the year 2001. But I never got my solution checked. I did contact many experts and send the paper to several journals, but the result was always the same: no review of the content, only rejection. After Grigory Perelman got his solution accepted in 2002 I let the Poincaré Conjecture proof be as it is, but did make an effort with the remaining six problems in 2007-2010, just for fun. Now checked my solutions in this year

2020, not so much for fun but in order to finish unfinished tasks. In these two time periods 2070-2010, 2020, I spent about half a year of intensive work on each of these six problems, so in total four years of intensive work has been spent on these seven problems.

Four years is rather much. I doubt many people have tried these seven problems for four years, or any single problem for half a year. There is a wrong opinion that if there is a well-known unsolved mathematical problem, then certainly a large number of good mathematicians must have tried to solve it and failed. Probably the number of those who tried is quite small and it is unknown if they failed because efforts to solve well-known problems are usually not checked. Let us take the Hodge Conjecture. You might be quite wrong in thinking that most, or even many, mathematicians specializing in algebraic geometry have tried to solve this conjecture and failed. Very few of them would ever consider trying because most would think that it is a difficult problem, so they would not manage solve it and the time would be wasted. Many of them would also be keenly aware of what has happened to people who tried to solve well-known problems: the effort seriously harmed their career. And this is perfectly correct reasoning: that is what is likely to happen. All you may expect is to wake up mockers, now it is especially easy with the Internet, you might wake up some nasty trolls.

Should you try to solve all seven prize problems you face another complication. Mathematicians are in the belief that nobody can solve problems on several unrelated fields of mathematics. The prize problems are each on a different field, therefore no single person can understand and solve them. Even the idea of doing so is ludicurous and points out to mental disease.

I think this is slighty faulty type of logic. Nobody, and certainly not a mathematician, is offended because applied mathematicians go messing up in unrelated fields where they are not experts. They apply mathematical reasoning to physics, finance, medicine, even warfare, you name it: mathematics can be used in some problems in almost any field, and you do not need to be an expert of the subject field.

But applied mathematicians do not dare to apply mathematics to problems in any field of pure mathematics. I think it is because pure mathematics has no practical usage and for that reason there is no real need in solving any problems: intruders who come to solve your problems are not wellcome. They are trespassers and should be shot. But in fact, only four of the prize problems are in pure mathematics. Three are in applied mathematics. That does not change the issue: mathematicians can mess up with any other field, because as every mathematician for sure imagines, researchers on those other fields are on the level of kids playing with lego pieces, but no applied mathematician can come to mess up on the playground of a fellow mathematician. There must be some limit.

I kind of understand this logic, but do not approve it. The fact is that how do we know if a problem that is difficult on a particular field of mathematics is really difficult in essence, or is it only difficult when using the tools of the particular field? The problem may require a different way of thinking. Probably it cannot be solved with ready tools from some other field, but it may be possible to solve it by inventing a method from the scratch. This may sometimes be so. I was very happy by one answer of an expert of computational complexity of my proof of the P not NP problem. He did not want to say anything of my proof, and gave the reason that it was so totally different from what is done in that field. I guess a solution should be something totally different if the methods of the field do not solve the problem. But this is not always the case. My first proof of the Poincaré Conjecture was simply geometric topology. It seemed that one or more famous people of the field had tried something similar but failed, and judged that those methods cannot work - but used differently they did work.

As for the importance of the prize problems, it should be obvious that they do not have any importance. If they had some importance, then you would expect that somebody would be interested enough in checking if the solution is correct, but they are not interested.

Maybe that is a too negative opinion, maybe I can invent some importance to these problems. The Birch and Swinnerton-Dyer conjecture could have had clear importance. It is very difficult to find all solutions to integer elliptic curve equations, while is is easy to solve modular elliptic curve equations. The conjecture proposed that the (usually infinite) number of solutions to the integer problem (i.e., the rank of the elliptic curve) is related to the (usually infinite) number of solutions to the modular problem (through the algebraic rank of the L-function). Unfortunately this is not the case, as my proof shows. Thus, this conjecure has no importance. I also found a counterexample to the Hodge Conjecture, therefore also this problem also has no importance as the conjecture was false. This counterexample to the Hodge conjecture is in a sense a technical couterexample to what can be selected as a form: I show that a linear combination can be a sum of a form and its conjugate, or composed of a form in one side and a conjugate in another side, but not a form only without a conjugate part. Usually, one thinks that a nowhere vanishing 2-form in a K3 space can be composed without a conjugate part. This selection of a form is what I show is not a linear combination.

Am I being still too negative? I sure hope not. I found two different proofs for the Riemann Hypothesis, though one proof was finally incorrect in an inrecoverable way. This does not have any greater importance since everybody has always believed that the Riemann Hypothesis is true, but one of my proofs may have some minor usage. The new proof explains why the interesting zeros of the zeta function ζ have the real part one half. It is because the sum of poles of $\zeta'\zeta^{-1}$ must cancel in the positive infinity on the x-axis. The older proof form 2008, well, it is wrong.

The Navier-Stokes problem asks if liquid can produce an infinity starting from quite natural initial values. There are 180,000 lakes in Finland. The water in none of them produces infinities, so we knew the answer to this question already. I do not think a proof of what we know has any greater importance. But it was not so in the problem statement. The original problem statement set the initial values in such a way that one could get an infinity. This was because implicitly the problem statement used an old theorem and this theorem had an error. I did not know of this theorem, I just solved the equations in my way and the result was in conflict with the old theorem. It is sometimes good not to know the field too well: there are errors in accepted theorems. The peer-review is not capable of finding all errors. I solved this problem as it was stated in the problem statement and got the paper published. Clay modified the problem statement when my paper was already accepted but not yet published. If you do not believe that the problem statements have been changed over the years, look at the Birch and Swinnerton-Dyer problem. It refers to a paper from 2001 and not as to be published. So, was the text written in 2000? Besides, there is an error in this problem statement also. I did not like the approach of secretly modifying a problem statement and will not care to make another attempt on this problem. Solving problems of another field is essentially charity work. You put lots of effort in order to help the people on that field by solving a problem they cannot solve themselves. I expect thanks, not cheating. But as for importance, I do not think the Navier-Stokes problem has any importance.

Then there is the P not NP problem. Some people explain the importance of this problem in the following way: they say that should P equal NP, then a large range of computational problems would have a fast solution. Much of cryptographic security would be lost, while if P is not NP, then we would know that those problems do not have fast solutions and the solution time would grow exponentially as a function of the problem size. Nicely, I solved this problem and got a bound that is just slightly above a polynomial bound. This bound does not allow making any kind of conclusions if something is too difficult to compute or easy to compute. With this bound, that is, with this prize problem solved, nothing at all is clarified. The prize problem is totally without any importance.

The problem of Yang and Mills fields could have had importance if the mass gap could have been proven, but my proof gives solutions without a mass gap. The question is really of quantization. Usually in quantum gauge field theory you fix the gauge and then use the Faddeev-Popov way and get the ghost fields. The problem here is that the quantization of the field theory fails already long before this step if a sound mathematical solution is required: quantization gives divergent propagators that have to be renormalized by a step that does not tolerate day light. If this step is corrected, then you do not need to fix the gauge and proceed in the customary way. I have done this in the book Reconsidering Einstein in chapters 1 and 7. The whole book is rather destructive in character questioning many known truths in physics. I do not think any theoretical physicist would consider my results important and certainly not useful. They would preferably forget them.

But one should not forget the Poincaré Conjecture. There were many purported proofs of the Conjecture. They were not checked. Then Perelman got his proof checked and accepted. I do not know if there was any major breakthrough in geometric or low-dimensional topology because a proof was finally accepted. Possibly there was, so maybe one prize problem may have had some importance.

Of course, you may argue that the prize problems have importance because they lead to new theorems. To counter this vain hope I include a small theorem that did come as a side result of my solution to the Birch and Swinnerton-Dyer conjecture. As far as I see the theorem has no importance what-so-ever. To make the point even clearer, I add two left-over results from the the Riemann Hypothesis proofs (one was false, the other not checked). Naturally, you do get results that are finally not needed when trying to prove some chosen theorem, but there is no reason why these side results should every prove useful for anything.

I include in this book one proof of the Riemann Hypothesis, a counterexample of the Hodge Conjecture, a long version of the solution to the Birch and Swinnerton-Dyer Conjecture, a proof that P is not NP, my first proof of the

Poincaré Conjecture as far as it can be reconstructed from later versions of the proof, and two papers with totally useless results that were invented as side products.

For the Navier-Stokes problem solution see my article in EJDE from 2010 and for the Yang-Mills problem solution see my book Reconsidering Einstein Lambert Academic Publishers, 2020, chapters 1 and 7. I will not include these solutions in this book.

I hope I managed to convince you that mathematical results are usually useless. They are allowed to be useless. They are made mainly for fun, but it is not much fun if nobody agrees to read them. But fine, there should be some importance.

Let us try to find the importance elsewhere. I think it can be found from the great difficulty of getting proposed solutions to well-known problems reviewed. This difficulty gave me an idea of how to conquer the world. You need lots of money, not a million or a few millions from Clay, but millions of millions. To have billions you must control some important banks and through them finance, but money buys. You can buy much of the media with such money, and through media control the minds and souls, be the light of the nations, but if the light in you is darkness, then the nations are lead by darkness. This darkness may be essential to you because you may have your favourite story, in history for instance, and too much light might not support this story. It is not enough to control the media, you need to control people whom other people trust. Today it is the scientists, as they are the new priests. People trust that the priests know as they have the book, even though much in the book is false. You can donate lots of money to private universities so that they can hire the best experts, give scholarships to the best students, and get the best equipment. With such resources there always are good results, but as the financier you have the say-so who is given the merit for the results. It could be your man. And so it comes to the point that the prizes should go to the correct people and not to the incorrect people. But this was just a thought, let us say that it is not so.

Personally I do not care about the importance. For me the main result is that this long task will soon be finished. In this book I give four solutions to still open problems. Are the solutions correct? I just spent much of this year checking them and they seem correct, but that does not imply that they are correct. The expectation is that all solutions are incorrect, yet, no hard problems are solved without trying them. Even one correct solution is enough.

Two of the four solutions to open problems included in this book are currently submitted. Two others are almost ready to be submitted, and will be submitted. I will send only an abridged version of the long paper on the Birch and Swinnerton-Dyer conjecture to a journal. I will not check the long version carefully for typos, so be carefull, but that paper has much good material. I do not have any hopes of getting these solutions reviewed in any meaningul sense. Why should this time be any different than any other time?

In hindsight it was a totally useless exercise, but I cannot deny: it was great fun to crack those problems.

Yes. Ignore what I just wrote. Solve the seven problems. It is great fun. I warmly recommend!

Seriously. Of course. I never joke.

Content:

On the zeros of the Riemann zeta function

Abstract. Zeros and the pole of the Riemann zeta function $\zeta(s)$ correspond to simple poles of the logarithmic derivative $f(s) = \frac{d}{ds}\ln\zeta(s)$. In $Re\{s\} > 1$ the function $f(s)$ has an absolutely convergent sum expression $f(s) = \sum_{j=1}^{\infty} h_j(s)$ where $h_j(s) = h_1(js)$ and $h_1(s) = -\sum_{m=1}^{\infty}\ln(p_m)p_m^{-s}$, a sum over all primes $p_m > 1$. When the Taylor series of $f(s)$ is evaluated at a point $(l, 0)$, $l >> 1$, the absolute values of the coefficients of the Taylor series decrease in a negatively exponential manner when l increases. The function $f(s)$ has simple poles in the area $Re\{s\} < 1$. The pole gives the function $r/(s - s_k)$, which can be evaluated into a Taylor series at $(l, 0)$. The coefficients of the Taylor series of the pole decrease as l^{-1} as a function of l. This implies that in the sum of all poles of $f(s)$ poles must cancel other poles so that the negatively exponential behavior of the coefficients of the Taylor series dominates. The function of $x = l^{-1}$ arising from the pole $-1/(s - 1)$ at $s = 1$ is $-x/(1 - x)$. The poles of $f(s)$ at even negative integers give the function $-xC$. These two negative functions cannot cancel poles s_k that are on the x-axis and $0 < s_k < 1$. Thus, such poles do not exist. Pole pairs s_k, s_k^* give the function $x + x/(1-x)$ that cancels the sum $-xC - x/(1 - x)$ when $C = 1$ if only if every pole s_k has $Re\{s\} = \frac{1}{2}$. The convergence of the coefficient of every power $i > 0$ of x larger to zero at least as $O(x)$ is shown possible for this solution.

Key words: Riemann zeta function, Riemann Hypothesis, Number Theore.
AMS Mathematic Subject Classification: 11M26

1 Definitions

The Riemann zeta function is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} \tag{1}$$

where s is a complex number. The zeta function can be continued analytically to the whole complex plane except for $s = 1$ where the function has a simple pole. The zeta function has trivial zeros at even negative integers. It does not have zeros in $Re\{s\} \geq 1$. The nontrivial zeros lie in the strip $0 < x < 1$. Let

$$P = \{p_1, p_2, \dots | p_j \text{ is a prime}, p_{j+1} > p_j > 1, j \geq 1\}$$

be the set of all primes (larger than one). Let $s = x + iy$, $x, y \in \mathbb{R}$ and $x > \frac{1}{2}$. The Riemann zeta function can be expressed as

$$\zeta(s) = \prod_{j=1}^{\infty} (1 - p_j^{-s})^{-1}, \tag{2}$$

This infinite product converges absolutely if $Re\{s\} > 1$. See e.g. [1] for the basic facts of $\zeta(s)$.

2 An introductory lemma and the theorem

Lemma 1. *The functions*

$$h_j(s) = -\sum_{j=1}^{\infty} \ln(p_j) p_j^{-js} \quad , \quad j > 0 \tag{3}$$

are related by $h_j(s) = h_1(js)$. The functions $h_j(s)$ have analytic continuations to $Re\{s\} > 0$ with the exception of isolated first-order poles. The poles of $h_j(s)$ that are not on the x-axis appear in pole pairs: close to s_k, where $Im\{s_k\} > 0$, $h_j(s)$ is of the type

$$h_j(s) = \frac{r}{s - s_k} + f_1(s) \tag{4}$$

and close to s_k^*, where s_k^* is a complex conjugate of s_k, $h_J(s)$ is of the type

$$h_j(s) = \frac{r}{s - s_k^*} + f_2(s)$$

The functions $f_1(s)$ and $f_2(s)$ are analytic close to s_k and s_k^* respectively. If the pole is at the x-axis, there is only one pole of the type (4) with $Im\{s_k\} = 0$.

Proof. The claim

$$h_j(s) = h_1(js) \tag{5}$$

follows directly from (3).

The function $h_1(s)$ converges absolutely if $Re\{s\} > 1$ because

$$\sum_{j=1}^{\infty} p_j^{-s}$$

converges absolutely for $Re\{s\} > 1$ and $|\ln p_j| < |p_j^{\alpha}|$ for any fixed $\alpha > 0$ if j is sufficiently large. Therefore

$$|\ln(p_j) p_j^{-s}| < 2|p_j^{-s+\alpha}|$$

for any fixed $\alpha > 0$ if j is sufficiently large. Therefore, by (5), $h_j(s)$ converges absolutely if $Re\{s\} > \frac{1}{j}$.

From (2) follows

$$\zeta'(s)\zeta(s)^{-1} = \frac{d}{ds} \ln \zeta(s) = \sum_{j=1}^{\infty} h_j(s).$$

The derivative $\zeta'(s)$ is analytic in all points except for $s = 1$. The function $h_1(s)$ is continued analytically to $Re\{s\} > \frac{1}{2}$ by

$$h_1(s) = \zeta(s)^{-1}\zeta'(s) - g(s) \tag{6}$$

where

$$g(s) = \sum_{j=2}^{\infty} h_j(s).$$

The function $\zeta(s)^{-1}$ is analytic except for at points where $\zeta(s)$ has a zero or a pole. The function $g(s)$ is analytic for $Re\{s\} > \frac{1}{2}$ because each $h_j(s)$, $j > 1$, is analytic in $Re\{s\} > \frac{1}{j}$. Thus, the right side of (6) is defined and analytic for $\frac{1}{2} < Re\{s\}$ except for at points where $\zeta(s)$ has a zero or a pole. At those isolated points $h_1(s)$ has a pole.

At a pole s_k of $\zeta(s)$ the zeta function has the expansion

$$\zeta(s) = \frac{C}{(s-s_k)^k} + \text{higher order terms.}$$

If $Re\{s\} > \frac{1}{2}$ the function $h_1(s)$ is of the form

$$h_1(s) = \zeta'(s)\zeta(s)^{-1} - g(s) = \frac{r}{s-s_k} + f_1(s)$$

where $f_1(s)$ is analytic close to s_k and $r = -k < 0$ is an integer. The function $\zeta(s)$ has only one pole, at $s_k = 1 = (1,0)$, and it is a simple pole, thus $r = -1$.

At a zero s_k of $\zeta(s)$ the zeta function has the expansion

$$\zeta(s) = C(s-s_k)^k + \text{higher order terms.}$$

If $Re\{s\} > \frac{1}{2}$ the function

$$h_1(s) = \zeta'(s)\zeta(s)^{-1} - g(s) = \frac{r}{s-s_k} + f_1(s)$$

where $f_1(s)$ is analytic close to s_k and $r = k > 0$ is an integer. It is known that $\zeta(s)$ has many zeros with $Re\{s_k\} = 1/2$.

Thus, $h_1(s)$ has only first-order poles for $Re\{s\} > \frac{1}{2}$ and therefore $h_j(s)$ has only first-order poles for $Re\{s\} > \frac{1}{2j}$. At every pole of $h_1(s)$ in $Re\{s\} > \frac{1}{2}$ the value of r is an integer.

As $h_1(s)$ is continued to $Re\{s\} > \frac{1}{2}$ by (6), the equation (5) continues $h_j(s)$ to $Re\{s\} > \frac{1}{2j}$. Then (6) continues $h_1(s)$ to $Re\{s\} > \frac{1}{4}$. The function $h_1(s)$ has isolated poles at $Re\{s\} > \frac{1}{4}$. Each pole is a first-order pole, but the value of r at a pole does not need to be an integer.

We can repeat the procedure inductively: If $h_1(s)$ is continued to $Re\{s\} > \frac{1}{2^i}$ by (6), the equation (5) continues $h_j(s)$ to $Re\{s\} > \frac{1}{2^i j}$. Then (6) continues $h_1(s)$ to $Re\{s\} > \frac{1}{2^{i+1}}$. By induction, all $h_j(s)$ are analytically continued to $Re\{s\} > 0$.

In this inductive process $h_1(s)$ gets isolated first-order poles. In these poles s_k the values $r = r_k$ can be positive or negative, and they do not need to be integers. If $h_1(s)$ has a pole

$$h_1(s) = \frac{r}{s - s_k} + f_1(s)$$

(here $f_1(s)$ is analytic close to s_k), then $h_j(s) = h_1(js)$ has a pole at $j^{-1}s_k$ and the r value is $j^{-1}r$ since

$$h_j(s) = h_1(js) = \frac{j^{-1}r}{s - j^{-1}s_k} + f_1(js).$$

The function $h_1(s)$ is symmetric with respect to the real axis. By (4) $h_j(s)$, $j > 1$, is also symmetric with respect to the real axis. Therefore poles of each $h_j(s)$, $j > 0$, appear as pairs s_k and s_k^*. In the special case where s_k is real there is only one pole, not a pair. □

Theorem 1. *All poles of* $\sum_{j=1}^{\infty} h_j(s)$ *in* $0 < Re\{s\} < 1$ *have the real part* $\frac{1}{2}$.

Proof. Let us consider a function $f(s)$ that has a first-order pole at s_0 and write $z_1 = s - s_0$. The function $f(s)$ does not have a Taylor series at s_0, but the function $z_1 f(z1 + s_0)$ has a Taylor series at $z_1 = 0$ and $f(s)$ can be expressed as

$$f(s) = \frac{c_{-1}}{z_1} + \sum_{k=0}^{\infty} c_k z_1^k. \tag{7}$$

Let us evaluate $f(s)$ at another point at $s_0 + l$, $l > 0$, by first writing $z_1 = l - z_2$ where $|z_1| << 1$, inserting $z_1 = l - z_2$ to the series expression of $f(s)$, and then considering the result when $|z_2| << 1$. The function

$$f_1(z_1) = f(z_1 + s_0) - \frac{c_{-1}}{z_1} \tag{8}$$

has the Taylor series at $z_1 = l - z_2$ where $|z_1| << 1$ as

$$f_1(l - z_2) = \sum_{m=0}^{\infty} c_m (l - z_2)^m$$

$$= \sum_{m=0}^{\infty} \sum_{i=0}^{m} \frac{m!}{i!(m-i)!} l^i (-z_2)^{m-i} c_m$$

$$= \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} \frac{(k+i)!}{i!k!} l^i (-1)^k c_{k+i} z_2^k = \sum_{k=0}^{\infty} b_k z_2^k.$$

Thus

$$b_k = \sum_{i=0}^{\infty} \frac{(k+i)!}{i!k!} l^i (-1)^k c_{k+i}.$$

As

$$c_k = \frac{1}{k!} \frac{d^k}{dz_1^k} f_1(s)|_{z_1=0}$$

we can express

$$b_k = \left(\sum_{i=0}^{\infty} \frac{1}{i!} l^i \frac{d^i}{dz_1^i} \right) \frac{1}{k!} (-1)^k \frac{d^k}{dz_1^k} f_1(s)|_{z_1=0}. \tag{9}$$

If there is no pole of $f(s)$ at $s_0 + l$, the function

$$f_1(l - z_2) = \sum_{k=0}^{\infty} b_k z_2^k$$

is analytic and defined by its Taylor series as powers of z_2 where the series converges.

The pole of $f(s)$ at c_{-1} can be evaluated as a Taylor series of z_2 at $s_0 + l$ as

$$\frac{c_{-1}}{l - z_2} = \frac{c_{-1}}{l} \frac{1}{1 - z_2 l^{-1}} = \frac{c_{-1}}{l} \sum_{k=0}^{\infty} \left(\frac{z_2}{l} \right)^k.$$

We can subtract a set of first-order poles of $f(s)$ in points $s_j \in A$ and define

$$f_1(z_1) = f(s) - \sum_{j\in A} \frac{r_j}{s - s_j} \tag{10}$$

where $r_j = c_{-1,j}$ and express

$$s - s_k = (s - s_0) - (s_j - s_0) = z_1 - s_j + s_0 = l - z_2 - (s_j - s_0).$$

At the point $s_0 + l$ the set of poles is

$$\sum_{j\in A} \frac{r_j}{s - s_j} = \sum_{j\in A} \frac{r_j}{l - z_2 - (s - s_0)} = x \sum_{j\in A} \frac{r_j}{1 - a_j x} \tag{11}$$

where $a_j = s_j - s_0$ and $x = (l - z_2)^{-1}$.

Let us consider

$$f(s) = \sum_{k=1}^{k_1} \frac{r_k}{s - s_k} + f_1(s) \tag{12}$$

$$f_1(s) = - \sum_{j=1}^{j_{\max}} \ln(p_j) p_j^{-s}.$$

Let $l >> 1$. The Taylor series of the set of poles points s_k at s_0 in powers of z_1 is

$$- \sum_{i=0}^{\infty} \left(\sum_{k=1}^{k_1} r_k (s_k - s_0)^{-i-1} \right) z_1^i$$

and the Taylor series at $s_0 + l$ in powers of $z_2 = l - z_1$ is

$$\sum_{i=0}^{\infty} \left(\sum_{k=1}^{k_1} r_k (s_0 + l - s_k)^{-i-1} \right) z_2^i.$$

For each k the coefficient of the ith power of z_1 at s_0 is $c_i = r_k (s_k - s_0)^{-i-1}$ while the coefficient of z_2 at $s_0 + l$ is

$$b_i = r_k (s_0 + l - s_k)^{-i-1} = r_k l^{-i-1} + r_k (i+1)(s_k - s_0) l^{-i-2} + \cdots$$

The absolute value of the coefficient b_i of the Taylor series in powers of z_2 at s_0+l decreases as

$$|\sum_{k=1}^{k_1} r_k|l^{-i-1}$$

as a function of $l >> 1$.

The part $f_1(s)$ of $f(s)$ satisfies

$$|f_1(s+l)| = \left| -\sum_{j=1}^{j_{\max}} \ln(p_j)p_j^{-s-l} \right| = \left| -\sum_{j=1}^{j_{\max}} \ln(p_j)p_j^{-s}e^{-l\ln p_j} \right|$$

$$\leq \left|e^{-l\ln 2}\right| \left| -\sum_{j=1}^{j_{\max}} \ln(p_j)p_j^{-s} \right| = e^{-l\ln 2}|f_1(s)|. \tag{13}$$

The absolute value of the coefficient b_i of the Taylor series in powers of z_2 at s_0+l decreases as

$$|b_i| \leq e^{-l\ln 2}|c_i|.$$

This is negative exponential decrease and much faster than the hyperbolic decrease for the set of poles.

When $l \to \infty$, the hyperbolic contribution from the poles must vanish: every nonzero coefficient of the Taylor series of $f(s)$ at $(l,0)$ when $l \to \infty$ must decrease as a negative exponential $e^{-l\ln(2)}$. This negative exponential of l decreases faster than any negative power of l. For each power i of x, the coefficient in the power series of x coming from the sum of the poles must go to zero at least as $O(x)$ leaving the negatively exponentially decreasing coefficient from $f_1(s)$ in (12) to dominate.

The sum of the poles decreases as $O(x)$, $x = l^{-1}$, and goes to zero when $x \to 0$ assuming that the x-coordinate of every pole of $f(s)$ is smaller or equal to one, but the required convergence that each coefficient of the power series of x must go separately to zero at least as $O(x)$ is a stronger condition. The requirement that the coefficient of a power i of x the sum of poles decreases at least as $O(x)$

means that that the poles of $f(s)$ partially cancel each others when l grows. Poles cannot completely cancel: a pole at s_k with $r = r_k$ can be completely cancelled only by a pole at s_k with $r = -r_k$. The sum of poles has the poles of its terms, but at $l >> 1$ there can be partial cancellation so that the Taylor series coefficients decrease sufficiently fast as a function of l.

Let $j_{\max} \to \infty$ in (13). Then $f(s) = h_1(s)$. If $Re\{s\} = l >> 1$, the sum (13) taken to infinity converges absolutely. The inequality (13) holds when $j_{\max} \to \infty$ and the absolute values of the coefficients of the Taylor series at $s_0 + l$ for the function $h_1(s)$ must decrease in negative exponential manner as a function of l when $l \to \infty$. It follows that every $h_j(s) = h_1(js)$ also has the same negatively exponential dependence of the coefficients of the Taylor series at $(l, 0)$ on l when $l >> 1$. Consequently the sum of the poles of the functions $h_j(s) = h_1(js)$ has same negatively exponential dependence for coefficients at $(l, 0)$ on l when $l >> 1$. Therefore the sum of the poles of the function

$$f(s) = \frac{d}{ds} \ln \zeta(s) = \sum_{j=1}^{\infty} h_j(s)$$

must satisfy the requirement that the coefficient of each power of x decreases at least as $O(x)$ when l grows to infinity.

We did not continue $h_j(s)$ to the area $Re\{s\} \leq 0$ in Lemma 1, but the function $f(s)$ is analytically continued to $Re\{s\} \leq 0$ by

$$f(s) = \frac{d}{ds} \ln \zeta(s)$$

to all points where $\zeta(s) \neq 0$ and we can find all poles of $f(s)$.

The function $f(s)$ has the following poles in $Re\{s\} > 0$:

(i) There is a pole with $r = -1$ at $s = 1$.

(ii) There is a set A of pole pairs $h_1(s)$ at s_k and s_k^* where s_k has a nonzero imaginary part, and the r-value r_k is positive. All we know of s_k is that $0 < Re\{s_k\} < 1$, and that that there exist poles s_k with the real part $\frac{1}{2}$.

(iii) There may be a set A_1 of poles $s_{k,1}$ of $h_1(s)$ with $r_{k,1}$ a positive integer and the pole s_k is real, $0 < s_k < 1$. No such pole is known.

Inserting $s = s_0 + l$, $s_0 = 0$, $x = l^{-1}$ to the expression of a pole (4) on the x-axis gives (ignoring the analytic function part in (4))

$$\frac{r_k}{s - s_k} = \frac{xr_k}{1 - a_k x}.$$

Here a_k is a real number. A pole pair in the positive and negative y-axis can be written as

$$\frac{r_k}{s - s_k} = \frac{xr_k}{1 - (1 + i\alpha_k)a_k x}$$
$$\frac{r}{s - s_k^*} = \frac{xr_k}{1 - (1 - i\alpha_k)a_k x}.$$

Here $x = (l - z_2)^{-1} > 0$ is a real number and small if l is large, $a_k = Re\{s_k\}$ and α_k is chosen positive. We will always take s_0 as 0. The number l is the distance from $s_0 = 0$ to the observation point on the x-axis, $(l, 0)$, where the Taylor series with z_2 is evaluated and $|z_2| << 1$. As z_2 is the variable of the Taylor series at $(l, 0)$, the expressions are valid for any small z_2 and we select $z_2 = 0$ for easier notations. Thus, $x = l^{-1}$. The pole (i) at $s = 1$ gives the power series of x where $a_k = 1$ and $r = -1$

$$\frac{xr}{1 - (a_k x)} = \frac{-x}{1 - x} = -x \sum_{m=0}^{\infty} x^m.$$

The zeros of $\zeta(s)$ in the area $Re\{s\} \leq 0$ are the so called trivial zeros at even negative integers. They come from the formula

$$\zeta(-n) = (-1)^n \frac{B_{n+1}}{n+1}$$

where $B_m = 0$ if $m > 1$ is odd. Zeta does not have a zero at $s = 0$. From the functional equation

$$\zeta(s) = 2^s \pi^{s-1} \sin(2^{-1}\pi s) \mathit{\Gamma}(1-s)\zeta(1-s) \tag{14}$$

we can deduce that the trivial zeros are zeros of $\sin(2^{-1}\pi s)$ and therefore first-order zeros. Thus, at a point $s_k = -2k$, $k > 0$ integer, the function $f(s)$ has a first-order pole with the r-value 1.

A pole at $s_k = -2k$, $k > 0$, is

$$\frac{r_k}{s - s_k} = \frac{1}{s + 2k}.$$

We can evaluate the Taylor series of z_1 at s_0 and the Taylor series of z_2 at $s_0 + l$ for any such pole and for a finite sum of such poles:

$$\frac{1}{s_0 + z_1 + 2k} = \frac{1}{s_0 + 2k} \sum_{i=0}^{\infty} (-1)^i (s_0 + 2k))^{-1} z_1^i$$

$$\frac{1}{s_0 + l - z_2 + 2k} = \frac{1}{s_0 + l + 2k} \sum_{i=0}^{\infty} (s_0 + l + 2k))^{-1} z_2^i$$

but if sum the index k goes to infinity, the series diverges at every finite point $s_0 + l$. We will evaluate the sum of these poles at $s_0 = 0$, conclude that the contribution is negative, and present a way to move a finite but growing sum of these poles to $(l, 0)$.

First we find out the sign of the infinity of the sum of the poles $s_k = -2k$ at $s_0 = 0$ and $z_1 = 0$. Notice that for a point $s_j = -k$ the pole at that point, with the r-value r, when evaluated to a Taylor series at $s_0 = 0$ and $z_1 = 0$ is

$$\frac{r}{s - s_j} = \frac{r}{k}.$$

This is the inverse of a pole with the same r but with $s_j = k$ when evaluated to a Taylor series at $s_0 = 0$ and $z_1 = 0$. As an example, $s_j = 1$ is the pole at $s = 1$ with $r = -1$. When evaluated at $s_0 = z_1 = 0$ it is the inverse of a pole with $r = -1$ but $s_k = -1$. Thus, the pole at $s_k = -2k$ with $r = 1 > 0$ is the same at $s_0 = 0$ as a a pole at $s_k = 2k$ with $r = -1 < 0$. We see that any sum of the poles $s_k = -2k$ gives a negative infinity when evaluated at $s_0 = 0$.

The type of infinity of the sum of all poles $s = -2k$ at $s_0 = 0$ can be calculated. Using the facts that $\zeta(s)$ has a simple pole at $s = 1$

$$\zeta(s) = \frac{a}{s-1} + g(s)$$

where $g(s)$ is analytic at $s = 1$ and that $\lim_{s\to 1}(s-1)\zeta(s) = 1$, so $a = 1$, we can write

$$\zeta(1) = \lim_{s\to 1}\frac{1+(s-1)f(1)}{s-1} = \lim_{s\to 1}\frac{1}{s-1} = \lim_{s\to 0}\frac{1}{s}$$

This result gives

$$\sum_{k=1}^{\infty}\frac{1}{2k} = \frac{1}{2}\sum_{k=1}^{\infty}\frac{1}{k^1} = \frac{1}{2}\zeta(1) = \lim_{s\to 0}\frac{1}{2}\frac{1}{s}.$$

Thus, the sum of the poles at $s_k = -2k$ appears as a simple pole when evaluated at $s_0 = 0$. The pole has a negative r-value with $r = -1$ at $s_0 = 0$. However, it is not a simple pole. A simple pole with $r = -1/2$ is

$$\lim_{s\to s_0}(-1/2)1/(s-s_0).$$

It is moved to $s_0 + l$ by writing

$$\lim_{s\to s_0}(-1/2)1/(s-l-s_0) = (-1/2)/l = -x/2$$

where $x = l^{-1}$. This pole is finite for every $l > 0$, but the sum of the poles $s_k = -2k$ is infinite at every finite l. This is so because the infinity $\lim_{s\to s_0}(-1/2)/(s-s_0)$

is not caused by the pole being physically at s_0, the infinity comes from the sum of the numbers $1/(l+2k)$.

If we subtract all poles $s_k = -2k$ from $f(s)$, then $f_1(s)$ is infinite at every point. Because of this reason all poles $s_k = -2k$ cannot be moved to $(l, 0)$ at the same time. We can only move at a given $l > 0$ such a subset of poles (like a finite set) that the sum gives a finite number when moved to $(l, 0)$. All poles have to be moved at some point as the sum of all poles of $f(s)$ should be zero at $l \to \infty$. Thus, we must move more poles when l grows until all poles are moved when $l \to \infty$. The choice of which subsums of poles are moved for each l cannot influence the result. We will make a convenient choice for these sums: let us choose a suitable growing function $N(l)$ and move the subsum of poles $s_k = -2k$ satisfying $k \leq N(l)$. A finite sum up to $N(l)$ can be moved to $s_0 + l$, and when $N(l)$ increases with l, all poles $-2k$ are included in the finite sum when $k \leq N(l)$. The tail of the infinite sum that is outside the finite sum up to $N(l)$ goes to zero when $l \to \infty$.

Thus, we take a finite sum

$$\sum_{k=1}^{[N(l)]} \frac{1}{2k}.$$

As it is a finite sum, it can be moved to $(l, 0)$ without creating an infinity. If $N(l)$ is sufficiently large and fixed, and $l = 0$, the moved sum is $-x/2 - \epsilon(l)$. The number $\epsilon(l)$ depends only on $N(l)$ and we can select a function $N(l)$ such that $\epsilon < \min\{\epsilon_0, e^{-l}\}$ where $\epsilon_0 > 0$ is small. Then $\epsilon(l)$ decreases with l faster than any power of $x = l^{-l}$. The number $N(l)$ increases when l grows, and therefore the absolute value of the sum grows with l. It gives a function $-xC(l) - \epsilon(l)$. This function cannot have any higher powers of x, only the first power, because every power of x can be continued to s_0 and there would be the power of x also at s_0, but at s_0 the function is $-x/2$ when $N(0) \to \infty$. In the limit $x \to 0$ the function $xC(l)$ must be of the order $O(x)$ because all other poles give contributions of $O(x)$ and the sum of all poles must vanish when $l \to \infty$. Since $C(l) \geq 1/2$ is a growing function and limited from above, the function $-xC(l) - \epsilon$ must converge to $-xC$, where $C > 1/2$ is a finite real number. The number ϵ goes to zero, as it decreases faster than any power of x. The number C will be determined later in this proof.

The poles (iii) of A_1 sum to a series of the type $x\sum_{m=0}^{\infty} c_m x^m$ where every c_m is nonnegative. Since all of these poles are in the area $0 < s_k < 1$ and they are isolated and therefore do not have a concentration point at $s = 1$, the power series of x coming from these poles cannot be of the type $b(x + x^2 + x^3 \cdots)$, which is the type of the power series of the pole at $s = 1$. It follows that the poles of A_1 cannot be cancelled the pole at $s = 1$ giving the contribution $-x/(1-x)$. Adding the contribution $-xC$ from the the sum of poles $-2k$ does not help to cancel any poles of A_1. The poles of (ii) yield a power series of x where the coefficient of every x^i is nonnegative. They cannot cancel poles of A_1. Thus, the poles of A_1 cannot be cancelled by any set of other poles in the limit $l \to \infty$. Therefore the set A_1 must be empty.

The pole pairs in A can be cancelled by the poles in $s = 1$ and in $-2k$, as will be seen later. The coefficient of the power one of x can be cancelled by sum of the corresponding coefficient -1 of the pole at $s = 1$ and the coefficient $-C$ coming from the poles in $Re\{s_k\} \leq 0$. Higher than power one coefficients of x coming from a sum of pole pairs in A can be cancelled only by the pole at $s = 1$ since $-xC$ does not have higher powers of x.

The two poles (ii) of a pole pair have a real sum:

$$\frac{xr_k}{1 - a_k(1 + i\alpha_k)x} + \frac{xr_k}{1 - a_k(1 - i\alpha_k)x} = xr_k \frac{2(1 - a_k x)}{1 - 2a_k x + (1 + \alpha_k^2)(a_k x)^2}.$$

We expand the sum S of the poles of a pole pair omitting the multiplier xr_k for simplicity in this calculation up to (16):

$$S = \frac{2(1 - a_k x)}{1 - 2a_k x + \alpha_k^2 (a_k x)^2} = \frac{2 - 2a_k x}{1 + \alpha_k^2 (a_k x)^2} \frac{1}{1 - 2a_k x \gamma_k^{-1}}$$

where $\gamma_k = 1 + \alpha_k^2 (a_k x)^2$.

$$= \frac{2 - 2a_k x}{\gamma_k} \sum_{i=0}^{\infty} (2a_k x \gamma_k^{-1})^i.$$

Writing $\beta_{k,i} = (2a_k)^i \gamma_k^{-i-1}$ we get

$$S = 2\sum_{i=0}^{\infty} \beta_{k,i} x^i - 2a_k \sum_{i=0}^{\infty} \beta_{k,i} x^{i+1} = \sum_{i=0}^{\infty} 2\beta_{k,i} x^i - 2a_k \sum_{i=1}^{\infty} \beta_{k,i-1} x^i$$

$$= 2\beta_{k,0} + \sum_{i=1}^{\infty} (2\beta_{k,i} - 2a_k \beta_{k,i-1}) x^i.$$

For $i > 0$

$$2\beta_i - 2a_k \beta_{k,i-1} = 2\frac{(2a_k)^{i-1}}{\gamma_k^i}(2a_k \gamma_k^{-1} - a_k)$$

$$= \frac{(2a_k)^i}{\gamma_k^{i+1}}(2 - \gamma_k) = \beta_{k,i}(2 - \gamma_k).$$

This gives an equation for every $i > 0$

$$2\beta_i - 2a_k \beta_{k,i-1} = 2\beta_{k,i} - \gamma_k \beta_{k,i}.$$

Inserting $\gamma_k = 1 + (\alpha_k a_k x)$ yields for $i > 0$

$$2a_k \beta_{k,i-1} = \gamma_k \beta_{k,i} = \beta_{k,i} + x^2 (\alpha_k a_k)^2 \beta_{k,i}.$$

For every k when $l >> 1$ and thus for $0 < x = l^{-1} << 1$ and $i > 0$ holds

$$2a_k \beta_{k,i-1} = \gamma_k \beta_{k,i} = \beta_{k,i} + O(x^2).$$

The coefficient of the the power x^i, $i > 0$, is

$$2\beta_{k,i} - 2a_k \beta_{k,i-1} = \beta_{k,i} + O(x^2). \qquad (16)$$

The coefficient of the power of x^{i+1} in the power series $-x/(1-x)$ of the pole in $s=1$ is -1 for every $i>0$. The coefficient of x^{i+1} in the power series of the sum of poles (ii) is

$$\sum_{k\in A} r_k(2\beta_{k,i} - 2a_k\beta_{k,i-1})$$

where we have included the multiplier xr_k that was so far omitted. Summing the powers of i from $i=2$ to $i=i_1+1$ and inserting (16) gives the equation where the coefficients of the pole pairs (ii) must cancel the coefficients of the pole (i) to the degree of $O(x^2)$:

$$i_1 = -\sum_{i=2}^{i_1+1}(-1) = \sum_{i=2}^{i_1+1}\sum_{k\in A} r_k\beta_{k,i} + O(x^2). \tag{17}$$

For each k, when $x\to 0$ and $i>0$, holds

$$\beta_{k,i} = 2a_k\beta_{k,i-1}. \tag{18}$$

If every $a_k = \frac{1}{2}$ the recursion equation (18) gives $\beta_{k,i+1} = \beta_{k,i}$ for every k. For every k the power series of x for $i>1$ is of the form $x\beta_{k,1}(x+x^2+x^3+\cdots)$. This is the same form as the power series $-x(x+x^2+x^3+\cdots)$ for the pole $s=1$ for $i>1$. The power series for the poles k for $i>1$ add to one power series of the type $xb(x+x^2+x^3+\cdots)$. We see that if every $a_k=\frac{1}{2}$, the sum of poles (ii) cancels all powers $i>1$ in the pole in $s=1$ when $x\to 0$ and the coefficient of each power $i>1$ of x converges to the negative of the coefficient of the power i of x in the power series for the pole in $s=1$ as $O(x^2)$.

Assume that one a_k is not $\frac{1}{2}$. The functional equation (14) shows that if there exists a zero $s_0 = x_0 + iy_0$ of $\zeta(s)$ with $0 < x_0 < \frac{1}{2}$ then there exists a zero of $\zeta(s)$ at a symmetric point in $\frac{1}{2} < x < 1$. This implies that we can find $s_{k'}$ such that $2a_{k'} > 1$. The form of (18) for a nonzero x, $i>0$, is

$$\beta_{k,i} = 2a_k\beta_{k,i-1} + O(x^2). \tag{19}$$

For $a_{k'}$ the recursion (19) gives $\beta_{k',i} = \beta_{k',1}(2a_{k'})^i + O(x^2)$ From (17) we get (20):

$$i_1 = \sum_{i=2}^{i_1+1} \sum_{k \in A} r_k \beta_{k,i} + O(x^2) \geq \beta_{k',1}(2a_{k'})^{i_1} + O(x^2). \tag{20}$$

The right side in (20) grows as $\beta_{k',1}(2a_{k'})^{i_1}$ as a function of i_1 while the left side is linear in i_1. This is a contradiction. Thus, every a_k must be $\frac{1}{2}$.

By (20) each $a_k = \frac{1}{2}$. Inserting $a_k = 2^{-1}$ to (21) gives

$$\beta_i = \sum_{k \in A} r_k (1 + 2^{-2}(\alpha_k x)^2)^{-i-1}. \tag{22}$$

The recursion equation for $\beta_{k,j}$ is $\beta_{k,i} = (2a_k/\gamma_k)\beta_{k,i-1}$. As $2a_k = 1$ and since $\gamma_k \geq 1$, this implies that $\beta_{k,i-1} \geq \beta_{k,i}$ for all $i > 0$. Recursion (18) for $a_k = \frac{1}{2}$ shows that for every $i > 0$ the value $\beta_{k,i}$ is the same when $x \to 0$. Since $\gamma_k \to 1$ when $x \to 0$, β_i is the same for every $i \geq 0$. Equation (20) implies that $\beta_i = 1$ for every $i > 0$. In the limit $x \to 0$ holds $\beta_{k,0} = \beta_{k,1}$. Therefore also $\beta_i = 1$ when $x \to 0$.

The claim of Theorem 1, i.e., that each $a_k = \frac{1}{2}$ for $s_k \in A$ and A_1 is empty, is already proven. The reason for this result is that since the poles of $f(s)$ at $Re\{s\} \leq 0$ give $-xC$, all powers $i > 1$ of x in the series $-x/(1-x)$ for the pole at $s = 1$ have to be cancelled by the poles of A and A_1. This series to be cancelled by the poles of A and A_1 is $-x^2 - x^3 - x^4 - \cdots$. No sum of poles in A_1 can give this series because each pole in A_1 is smaller than one and larger than zero. A pole pair s_k, s_k^* in A gives this series if and only if $Re\{s_k\} = \frac{1}{2}$. If even one s_k has the real part not at $\frac{1}{2}$, (20) gives a contradiction. Thus, all poles of A have the real part as one half.

Let us still check if the solution is possible. We check if all

$$\beta_i = \sum_{k \in A} r_k \beta_{k,i} \tag{21}$$

can have the value 1 as the solution gives, and if all coefficients of the power series of x can go to zero at least as $O(x)$ when $l \to \infty$.

Because $x \to 0$, the values of α_k in (22) must grow to infinity with k. The set A is necessarily infinite. We renumber the poles of (ii) so that (α_k) is a growing sequence and the sum $k \in A$ is the sum $k = 1$ to infinity. Since $a_k = \frac{1}{2}$ by (20) we can evaluate

$$2\beta_{k,i} - 2a_k\beta_{k,i-1} = \beta_{k,i}(2 - \gamma_k)$$

and get

$$\beta_{k,i} = \beta_{k,i-1}\left(1 - \frac{(0.5\alpha_k x)^2}{1 + (0.5\alpha_k x)^2}\right).$$

Let $l >> 1$ be fixed. If $\alpha_k >> l = x^{-1}$, then

$$\frac{(0.5\alpha_k x)^2}{1 + (0.5\alpha_k x)^2}$$

is close to one and $\beta_{k,i}$ is close to zero. This means that large values of α_k contribute very little to the Taylor series at $s_0 + l$. The sum in (22) can be finite and there is no reason why it could not be one as the solution gives.

The contribution from the poles at $s_k = -2k$ is $-xC$, from the pole at $s = 1$ it is $-x/(1-x)$, from the poles of A_1 it is 0, and the contribution of the pole pairs of A is approaching the series $x(2 + x + x^2 + x^3 + \cdots)$ when $x \to 0$ as $O(x^2)$ separately for the coefficient of each power i of x^i. The sum of these contributions when $l \to \infty$ is

$$-Cx - \frac{x}{1-x} + 0 + x(2 + x + x^2 + x^3 + \cdots) = (-3/2 + 2\beta_0)x = (1 - C)x. \quad (24)$$

The sum (24) must be zero. This is possible when $C = 1$.

The convergence of the coefficients of the powers of x to zero in (24) when x grows is $O(x^2)$ for the coefficient of each power $i > 1$ of x^i separately, which fulfills the convergence criterion.

For the power one of x we only get the result that the convergence as at least $O(x)$ is possible. The term β_0 converges to β_1 as $O(x^2)$ since every $\beta_{k,i} = \beta_{k,i-1} + O(x^2)$. Each β_i converges to 1 as $O(x^2)$. The contribution from the poles at $-2k$ is $-xC(l) - \epsilon(l)$ where $0 < \epsilon(l) < e^{-1}$ goes to zero very fast. The sum $-xC(l) - \epsilon(l) - x + x\beta_0$ can go to zero at least as fast as $O(x)$ as the solution requires. The solution is possible. □

All known facts of the Riemann zeta function that are used in this proof can be found in [1]. The history and background of the Riemann Hypothesis are well described in the book [2]. As the problem is still open, recently published results do not add so much to the topic. As they are not needed in this proof, they are not referred to.

References

1. E. T. Whittaker and G. N. Watson, *A Course in Modern Analysis*, Cambridge, University Press, 1952.
2. K. Sabbagh, *The Riemann Hypothesis, the greatest unsolved problem in mathematics*, Farrar, Strauss and Giroux, New York, 2002.

A note on the Hodge conjecture

Abstract. The paper presents a counterexample to the Hodge conjecture.

Key words: Hodge theory, differential forms, algebraic geometry.

3 Introduction

The Hodge conjecture is one of the better known open problems in mathematics and was chosen as one of the Millennium Prize problems by the Clay Mathematics Institute [1]. The formulation of the conjecture in [1] is: On a projective non-sigular algebraic variety C, any Hodge class is a rational linear combination of classes $cl(\mathbb{Z})$ of algebraic cycles. This seems to say that every Hodge class on an algebraic projective complex manifold M is a linear combination with rational coefficients of the cohomology classes obtained by the Poincaré duality from homology classes of complex algebraic subvarieties of M.

The paper proposes a counterexample to the Hodge conjecure in the the algebraic projective complex manifold

$$\begin{gathered} M=\{(s_1:s_2:s_3:s_4:s_5:s_6:s_7:s_8)|s_1^4+s_2^4+s_3^4+s_4^4=0, \\ s_5^4+s_6^4+s_7^4+s_8^4=0, s_4^4+s_5^4=0\} \end{gathered} \tag{1}$$

M is an algebraic subvariety of codimension 3 in the complex projective space $\mathbb{P}^7$ and a closed complex manifold of dimension 4. The manifold M is composed as a product of two Fermat quadratic surfaces tied with the third homogeneous polynomial equation in order to embed M into a projective space. As the Fermat quadratic surface is a $K3$ space, the manifold M inherits a (2,0)-form that is never zero from the first copy of $K3$ and a similar (0,2)-form from the second copy of $K3$. It is shown in the presented paper that the wedge product of these forms is a (2,2)-form that cannot be represented as a $\mathbb{Q}$-linear combination of cohomology classes of deriving from algebraic subvarieties of M. This form, having only one term $f(P)dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4$, is rational (it is a linear combination of this type of forms with rational coefficients) and thus a Hodge class.

The study of what cohomology classes can be obtained by the Poincaré duality from homology classes of submanifolds was started by René Thom, but in real manifolds. Thom's original paper is in French and rather difficult to read, but we can look at [2] that is available on-line and find there the following statement (page 1 in [2]): All homology classes with integral coefficients of compact orientable differentiable manifolds of dimension < 10 are realizable by submanifolds. As a 4-dimensional complex manifold can be understood as an 8-dimensional real manifold, this statement, combined with the Poincaré duality, implies that every singular cohomology class of the 8-dimensional manifold can be realized as a linear combination with integer coefficients of classes of real submanifolds. However, these real submanifolds need not be complex submanifolds because a complex submanifold inherits the complex structure from the mother manifold.

If Z is a complex submanifold of M of codimension k then the inclusion map $i : Z \to M$ induces a cohomology class $[Z]$ that is in $H^{k,k}(M)$. In local coordinates the cohomology class $[Z]$ is then a form of the type

$$\psi = f(z, \bar{z})dz^I \wedge d\bar{z}^J \tag{2}$$

where $(z, \bar{z}) = (z^1, \ldots, z^n, \bar{z}^1, \ldots, \bar{z}^n)$ are the local coordinates. As $\psi \in H^{k,k}(M)$ there are the same number k of indices in the multi-indices I and J. This does not seem to say very much. However, if we look more carefully we see that the form (1.2) has the property: if z are the local coordinated of the submanifold, the differentials dz^j and $d\bar{z}^j$ always appear in pairs, as $dz^j \wedge d\bar{z}^j$. The reason is the following.

The tangent space of a 2-dimensional complex manifold in a chosen base point P is $\mathbb{C}^2$. There are two complex coordinates $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$.

(Let us use lower indices for a while as it is more natural.) A 2-dimensional real plane in the tangent space of the chosen basepoint can be spanned by any linear combination of four independent vectors $e_{x,1}, e_{y,1}, e_{x,2}, e_{y,2}$ but a complex line cannot separate x_j and y_j. They are not vectors, they are numbers in the field $\mathbb{C}$. Thus, a complex line through the basepoint can only be a linear combination $az_1 + bz_2 = c$, where a, b and c are complex numbers. Such a complex line defines a real plane as the x and iy coordinates of $z = x + iy$ correspond to the vectors e_x and e_y. In the local coordinates of the (real) mother manifold this (real) plane has the coordinates $e_x = ae_{x,1} + be_{x,2}$ and $e_y = ae_{y,1} + be_{y,2}$. It can be expressed by using a wedge product $xe_x \wedge ye_y$, which is a vector pointing to a direction orthogonal to the plane as the wedge product is a bit like the cross product. Instead of xe_x and ye_y we can use a linear combination of them: $z = x + iy$ and $\bar{z} = x - iy$. In this case the wedge product gets the form $z \wedge \bar{z}$. Here z is a local coordinate of the complex line. It is clearer if we denote the local coordinates of the tangent space of the complex submanifold by z'_j and the local coordinates of the tangent space of the mother manifold by z_m. Then z'_j is a linear combination (with complec coefficients) of the local coordinates z_m and a complex curve expressed as a (1,1)-form in is $f(P)dz' \wedge d\bar{z'}$ where $z' = z'_1$ is the only local coordinate of the 1-dimensional submanifold and P is the basepoint. If we take a complex surface as the submanifold, it has two local coordinates z'_1 and z'_2 and the form is of the type

$$f_1(P)f_2(P)dz_1 \wedge d\bar{z}_1 \wedge dz_2 \wedge d\bar{z}_2 \tag{3}$$

Consider a (2,2)-form of the type (changing the notation to upper indices)

$$\psi = f(z,\bar{z})dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 \tag{4}$$

Does such a form exist in the cohomology of M and can such a form be obtained as a $\mathbb{Q}$-linear combination of forms of the type (1.3)?

Lemmas 3.1-3.3 are only needed to show that a form of the type (1.4) is not exact, i.e., zero in the cohomology. Lemmas 3.1 and 3.2 show that for a particular type of a form ψ we get $*\psi$ which is quite similar to ψ in form. Lemma 3.3 shows that the form is harmonic and thus not exact. Actually, we created the form so that both $d\psi = 0$ and $d(*\psi) = 0$ for trivial reasons. There is a special condition in this part of the proof: Lemmas 3.1-3.3 needs an assumption that the submanifolds are algebraic varieties defined by polynomial equations with real coefficients. The important part is that the coefficients are real since the proof makes complex conjugation and requires that $f(P)^* = f(P^*)$, which is not always true. This assumption means that the manifold M must be selected in such a way that this assumption holds.

Lemma 3.4 shows that a form of the type (1.4) exists on $X = K3 \times K3$, which is a compact complex manifold of complex dimension 4. Lemma 3.3 shows that the form is not zero in $H^{2,2}(X)$. This is where we need the surface to be K3 and e.g. an Abelian surface would not work. Complex conjugation sends a harmonic form to a harmonic form, and thus we have a wedge of two harmonic forms, but the wedge of two harmonic forms is not necessarily harmonic - it can be exact. The nowhere vanishing (2,0)-form is needed in order to conclude that the wedge product is harmonic.

Lemma 3.5 proves that a form of the type (1.4) cannot arise as a $\mathbb{Q}$-linear combination from the cohomology classes of submanifolds. Here we again need a similar condition as in Lemma 3.1 that $f(P)^* = f(P^*)$. This condition follows if the mother manifold and all submanifolds are algebraic varieties defined by homogeneous polynomial equations with real coefficients. A restrictive assumption must be added to the lemma and fulfilling it requires a special form for M.

The manifold X in Lemma 3.4 is not quite what we want. It is a submanifold of $\mathbb{P}^3 \times \mathbb{P}^3$ but we want an algebraic subvariety of $\mathbb{P}^n$ for some n. Thus, we replace X by M. No changes are needed to Lemma 3.4.

The (2,2)-form from Lemma 3.4 is rational and represents a Hodge class. This fact seems obvious considering that it is like (1.4) and has only one term. However, in 2011 I received a comment from an expert of the field stating that the form is not rational. Therefore this issue must be addresses.

Finally I comment the apparent conflict of of the presented result with a published Ph.D. thesis [4]. The thesis proves that for some K3xK3 manifolds the Hodge conjecture holds. The conflict is only apparent since the proofs of Lemmas 3.1 and 3.5 require special conditions that $f(P)^* = f(P^*)$ and do not state anything for the general case of K3xK3.

4 Notations and concepts

We will use the notations in Kodaira [7] page 147. Local coordinates of a complex manifold M of complex dimension n at the base point z_0 are denoted by $z^1, \ldots, z^n$, $\bar{f}(z^1, \ldots, z^n)$ is the complex conjugate of $f(z^1, \ldots, z^n)$, and the Hodge star operation is denoted by $*$.

Let φ and ψ be $C^\infty(p,q)$-forms in a complex manifold M.

$$\begin{aligned}\varphi &= \frac{1}{p!q!}\sum \varphi_{\alpha_1,\ldots,\alpha_p,\bar{\beta}_1,\ldots,\bar{\beta}_q}(z)dz^{\alpha_1}\wedge\cdots\wedge dz^{\alpha_p}\wedge d\bar{z}^{\beta_1}\wedge\cdots\wedge d\bar{z}^{\beta_q}\\ \psi &= \frac{1}{p!q!}\sum \psi_{\alpha_1,\ldots,\alpha_p,\bar{\beta}_1,\ldots,\bar{\beta}_q}(z)dz^{\alpha_1}\wedge\cdots\wedge dz^{\alpha_p}\wedge d\bar{z}^{\beta_1}\wedge\cdots\wedge d\bar{z}^{\beta_q}\end{aligned} \tag{5}$$

The inner product is defined as

$$(\varphi,\psi)(z) = \frac{1}{p!q!}\sum \varphi_{\alpha_1,\ldots,\alpha_p,\bar{\beta}_1,\ldots,\bar{\beta}_q}(z)\bar{\psi}_{\alpha_1,\ldots,\alpha_p,\bar{\beta}_1,\ldots,\bar{\beta}_q}(z) \tag{6}$$

and

$$(\varphi,\psi) = \int_M (\varphi,\psi)(z)\frac{\omega^n}{n!} \tag{7}$$

where the volume element $\frac{\omega^n}{n!}$ is

$$\frac{\omega^n}{n!} = (-1)^{n(n-1)/2} g(z) dz^1\wedge\cdots\wedge dz^n\wedge d\bar{z}^1\wedge\cdots\wedge d\bar{z}^n \tag{8}$$

Here

$$g(z) = det(g_{\alpha\bar{\beta}}(z))_{\alpha,\beta=1,\dots,n} \tag{9}$$

is given by a Hermitian metric

$$\sum_{\alpha,\beta=1}^{n} g_{\alpha\bar{\beta}} dz^{\alpha} \otimes d\bar{z}^{\beta} \tag{10}$$

and

$$\omega = i \sum_{\alpha,\beta=1}^{n} g_{\alpha\bar{\beta}} dz^{\alpha} d\bar{z}^{\beta} \tag{11}$$

is the associated (1,1)-form, and $\omega^n = \omega \wedge \cdots \wedge \omega$ is the n-fold product. Then

$$\begin{gathered}(\varphi, \psi)\frac{\omega^n}{n!} = (i)^n(-1)^{n(n-1)/2} g(z) \frac{1}{p!q!} \sum \varphi_{\alpha_1\dots\alpha_p\bar{\beta}_1\dots\bar{\beta}_q}(z) \\ \times \bar{\psi}_{\alpha_1\dots\alpha_p\bar{\beta}_1\dots\bar{\beta}_q} dz^1 \wedge \cdots \wedge dz^n \wedge d\bar{z}^1 \wedge \cdots \wedge \bar{z}^n \\ = \frac{1}{p!q!} \sum \varphi_{\alpha_1\dots\alpha_p\bar{\beta}_1\dots\bar{\beta}_q}(z) dz^{\alpha_1} \wedge \cdots \wedge dz^{\alpha_p} \wedge d\bar{z}^{\beta_1} \wedge d\bar{z}^{\beta_q} \wedge *\psi(z) \\ = \varphi(z) \wedge *\psi(z)\end{gathered} \tag{12}$$

where

$$\begin{gathered}*\psi(z) = (i)^n(-1)^{\frac{1}{2}n(n-1)+(n-p)q} \sum_{A_p, B_q} sgn \begin{pmatrix} A_p & A_{n-p} \\ B_q & B_{n-q} \end{pmatrix} g(z) \bar{\psi}^{A_p B_q}(z) \\ \times dz^{B_{n-q}} \wedge d\bar{z}^{A_{n-p}}\end{gathered} \tag{13}$$

The multi-indices are $A_p = \alpha_1 \dots \alpha_p$, $A_{n-p} = \alpha_{p+1} \dots \alpha_n$, $B_q = \beta_1 \dots \beta_q$, $B_{n-q} = \beta_{q+1} \dots \beta_n$. Notice, that Kodaira on page 117 defines the Hodge star differently as

$$(\varphi, \psi)\frac{\omega^n}{n!} = \varphi(z) \wedge *\bar{\psi}(z) \tag{14}$$

but the usual definition in the literature is

$$(\varphi, \psi) vol_x = \varphi(z) \wedge *\psi(z) \tag{15}$$

i.e.,

$$(\varphi, \psi)\frac{\omega^n}{n!} = \varphi(z) \wedge *\psi(z) \tag{16}$$

as we have defined.

Let us select local coordinates such that $g_{\alpha\bar{\beta}}(z_0) = \delta_{\alpha\beta}$ where z_0 is the base point. Then also $g^{\alpha\bar{\beta}}(z_0) = \delta_{\alpha\beta}$.

We will usually mark indices as upper indices but in a calculation in Lemma 3.5 upper indices get confused with powers and indices are marked as lower indices.

5 Lemmas and a Theorem

Lemma 1. *Let* $p = q = 2$*,* $n = 4$*, and*

$$\begin{gathered} \psi_{12\bar{3}\bar{4}}(z) = 4f(z^1, z^2, \bar{z}^3, \bar{z}^4) \\ \psi_{\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2}(z) = 0 \quad if \quad (\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2) \neq (12\bar{3}\bar{4}) \end{gathered} \tag{17}$$

and let $f(z^1, z^2, z^3, z^4)$ *be a holomorphic function satisfying* $f(z^1, z^2, z^3, z^4)^* = f(\bar{z}^1, \bar{z}^2, \bar{z}^3, \bar{z}^4)$*. Then in local coordinates*

$$\begin{gathered} \psi = f(z^1, z^2, \bar{z}^3, \bar{z}^4) dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 \\ *\psi = g(z) f(\bar{z}^1, \bar{z}^2, z^3, z^4) dz^3 \wedge dz^4 \wedge d\bar{z}^1 \wedge d\bar{z}^2 \end{gathered} \tag{18}$$

Proof. The first claim is obvious since

$$\psi = \frac{1}{4}\psi_{12\bar{3}\bar{4}}(z) dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 \tag{19}$$

For the second claim we calculate

$$*\psi = (i)^4(-1)^{\frac{1}{2}4\cdot 3+2\cdot 2} sgn \begin{pmatrix} 1\;2\;3\;4 \\ 3\;4\;1\;2 \end{pmatrix} g(z)\psi^{12\bar{3}\bar{4}}(z) dz^3 \wedge dz^4 \wedge d\bar{z}^1 \wedge d\bar{z}^2 \tag{20}$$

Since

$$\begin{aligned} \psi^{\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2}(z) &= \sum_{\lambda_1,\lambda_2,\mu_1,\mu_2} g^{\bar{\lambda_1}\alpha_1} g^{\bar{\lambda_2}\alpha_2} g^{\bar{\beta}_1\mu_1} g^{\bar{\beta}_2\mu_2} \bar{\psi}_{\lambda_1\lambda_2\bar{\mu_1}\bar{\mu_2}}(z) \\ &= \sum_{\lambda_1,\lambda_2,\mu_1,\mu_2} \delta_{\lambda_1\alpha_1}\delta_{\lambda_2\alpha_2}\delta_{\beta_1\mu_1}\delta_{\beta_2\mu_2}\bar{\psi}_{\lambda_1\lambda_2\bar{\mu_1}\bar{\mu_2}}(z) \\ = \psi_{\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2}(z) &= \begin{cases} \bar{\psi}_{12\bar{3}\bar{4}}(z) \quad if \quad (\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2) = (12\bar{3}\bar{4}) \\ 0 \quad otherwise \end{cases} \end{aligned} \tag{21}$$

Thus

$$\begin{aligned} *\psi &= g(z)\bar{\psi}_{12\bar{3}\bar{4}}(z)dz^3 \wedge dz^4 \wedge d\bar{z}^1 \wedge \bar{z}^2 \\ &= g(z)\bar{f}(z^1, z^2, \bar{z}^3, \bar{z}^4)dz^3 \wedge dz^4 \wedge d\bar{z}^1 \wedge \bar{z}^2 \end{aligned} \tag{22}$$

and the assumption on $f(z^1, z^2, z^3, z^4)$ gives

$$\bar{f}(z^1, z^2, \bar{z}^3, \bar{z}^4) = f(z^1, z^2, \bar{z}^3, \bar{z}^4)^* = f(\bar{z}^1, \bar{z}^2, z^3, z^4) \tag{23}$$

Lemma 2. *Let $p = q = 2$, $n = 4$, and*

$$\begin{aligned} \psi_{12\bar{3}\bar{4}}(z) &= 4f(z^1, z^2, \bar{z}^3, \bar{z}^4) \\ \psi_{\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2}(z) &= 0 \quad if \quad (\alpha_1\alpha_2\bar{\beta}_1\bar{\beta}_2) \neq (12\bar{3}\bar{4}) \end{aligned} \tag{24}$$

and let $f(z^1, z^2, z^3, z^4)$ be a holomorphic function satisfying $f(z^1, z^2, z^3, z^4)^* = f(\bar{z}^1, \bar{z}^2, \bar{z}^3, \bar{z}^4)$. We can select the metric such that in local coordinates

$$\begin{aligned} \psi &= f(z^1, z^2, \bar{z}^3, \bar{z}^4)dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 \\ *\psi &= f(\bar{z}^1, \bar{z}^2, z^3, z^4)dz^3 \wedge dz^4 \wedge d\bar{z}^1 \wedge d\bar{z}^2 \end{aligned} \tag{25}$$

Proof. At the base point z_0 we can select the metric $g(z_0) = \det(g_{\alpha\bar{\beta}})_{\alpha,\beta=1,\ldots,n}$ such that $g_{\alpha\bar{\beta}} = \delta_{\alpha\beta}$. Then $g(z_0) = \det(g_{\alpha\bar{\beta}})_{\alpha,\beta=1,\ldots,n} = 1$. We can make the same selection at all points z and thus $g(z) = 1$.

Lemma 3. *Let ψ be a (2,2)-form in a complex manifold M. Let ψ be expressed in local coordinates as*

$$\psi = f(z^1, z^2, \bar{z}^3, \bar{z}^4) dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 \tag{26}$$

where $f(z^1, z^2, \bar{z}^3, \bar{z}^4)$ is holomorphic satisfying $f(z^1, z^2, z^3, z^4)^* = f(\bar{z}^1, \bar{z}^2, \bar{z}^3, \bar{z}^4)$. Let us assume that the complex dimension of M is four, M is compact and without boundary (i.e., closed). Then $\triangle\psi = 0$.

Proof. For compact manifolds without boundary

$$\triangle\psi = 0 \quad \Leftrightarrow \quad d\psi = 0 \quad \mathit{and} \quad \delta\psi = 0 \tag{27}$$

Let us calculate $d\psi$

$$d\psi = \sum_{j=1}^{4} \frac{\partial f}{\partial z^j}(z^1, z^2, \bar{z}^3, \bar{z}^4) dz^j \wedge dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 = 0 \tag{28}$$

since $dz^j \wedge dz^1 = 0$ if $j = 1$, $dz^j \wedge dz^2 = 0$ if $j = 2$, and

$$\begin{aligned} \frac{\partial f}{\partial z^3}(z^1, z^2, \bar{z}^3, \bar{z}^4) = 0 \\ \frac{\partial f}{\partial z^4}(z^1, z^2, \bar{z}^3, \bar{z}^4) = 0 \end{aligned} \tag{29}$$

since in the coordinate system $(z^1, \ldots, z^n, \bar{z}^1, \ldots, \bar{z}^n)$ the coordinates z^j and $\bar{z}^j$ are considered independent. The second assertion follows from the definition of the codifferential: if $* : \Omega^k \to \Omega^{n-k}$ is the Hodge star operator then $\delta : \Omega^k \to \Omega^{k-1}$ is defined by

$$\delta\psi = (-1)^{n(k+1)+1}(*d*)\psi \tag{30}$$

Inserting $n = 4$, $k = 2$, yields

$$\delta\psi = -(*d*)\psi = - * d(*\psi) \tag{31}$$

We may assume that the metric is chosen such that $g(z) = 1$. By Lemma 3.2

$$*\psi = f(\bar{z}^1, \bar{z}^2, z^3, z^4) dz^3 \wedge dz^4 \wedge d\bar{z}^1 \wedge d\bar{z}^2 \tag{32}$$

As in the previous case, we conclude that

$$d(*\psi) = \sum_{j=1}^{4} \frac{\partial f}{\partial z^j}(\bar{z}^1, \bar{z}^2, z^3, z^4) dz^j \wedge dz^3 \wedge dz^4 \wedge d\bar{z}^1 \wedge d\bar{z}^2 = 0 \tag{33}$$

Lemma 4. *Let X be a product of two K3 surfaces. Then X allows a (2,2)-form of the type*

$$f(z^1, z^2) f(\bar{z}^3, \bar{z}^4) dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 \tag{34}$$

where $f(z^1, z^2)$ is holomorphic and nowhere vanishing.

Proof. The existence of a nowhere vanishing 2-form is often taken as the definition of a K3 surface, the additional condition guaranteeing that a 2-dimensional complex manifold X is a K3 surface is that the manifold X is connected. This nowhere vanishing 2-form λ is the generator of $H^{2,0}(X)$, i.e., every other element $\alpha \in H^{2,0}(X)$ can be expressed as $\alpha = c\lambda$ for some $c \in \mathbb{C}$. The complex conjugate $\bar{\lambda}$ of λ is the generator of $H^{0,2}(X)$ as is shown by the Hodge duality pairing. Let the (2,0)-form λ be expressed in local coordinates as

$$\lambda = f(z^1, z^2) dz^1 \wedge dz^2 \tag{35}$$

Then $f(z^1, z^2)$ is holomorphic and nowhere vanishing. If $M = K3 \times K3$ then there are two nowhere vanishing 2-forms λ_1 and λ_2. Let us remember that a complex K3-surface is compact and as a real manifold it is a 4-dimensional closed manifold. Thus, M is closed and we can use Lemma 3.3. We can make a (2,2)-form as a wedge product of

$$\lambda_1 = f(z^1, z^2) dz^1 \wedge dz^2 \tag{36}$$

and

$$\bar{\lambda}_2 = \bar{f}(z^3, z^4) d\bar{z}^3 \wedge d\bar{z}^4 = f(\bar{z}^3, \bar{z}^4) d\bar{z}^3 \wedge d\bar{z}^4 \tag{37}$$

as $f(z^3, z^4)$ is holomorphic. Thus, we define

$$\begin{aligned} \psi &= f(z^1, z^2) dz^1 \wedge dz^2 \wedge f(\bar{z}^3, \bar{z}^4) d\bar{z}^3 \wedge d\bar{z}^4 \\ &= f(z^1, z^2) f(\bar{z}^3, \bar{z}^4) dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4 \end{aligned} \tag{38}$$

The form ψ satisfies $\triangle\psi = 0$ by Lemma 3.3. We still have to show that it is not zero in $H^{2,2}(X)$. Typical ways to show this are calculating periods or intersection products, but we will show it differently. For any complex manifold M the Hodge star gives an isomorphism from $H^k(M)$ to $H^{n-k}(M)$ and the Poincaré duality gives an isomprphism from $H_k(M)$ to $H^{n-k}(M)$. Especially, when $n = 4$, $H^2(X)$ and $H_2(X)$ are isomorphic and if $X = S_1 \times S_2$ $H^2(S_i)$ and $H_2(S_i)$, $i = 1, 2$ are isomorphic (for any coefficients, so coefficients are supressed in the notations). Let this isomorphism be $\Psi : H_2(X) \to H^2(X)$. Let C be a 2-chain in S_1 and let $(C, p_t) \in X$ be a family of pairs where p_t is a path in S_2. This family $C \times [0, 1]$ defines a homotopy from (C, t_0) to (C, t_1). By the homomorphism we have also a family $\lambda_1 \wedge \bar{\lambda}_2(P_t)$ parametrized by $t \in [0, 1]$. Then P_t defines a path in S_2. This is where we need the nowhere vanishing (2,0)-form. If there existed a point P_1 such that $\bar{\lambda}_2(P_1) = 0$ then $\psi = 0$ at any point $(Q, P_1) \in S_1 \times S_2$. As the preimage of zero is a point, this would yield a homotopy from C to a point, i.e., $(C, p_0) = 0$ in $\pi_2(X)$. Consequently $(C, p_0) = 0$ in $H_2(X)$ and $\Psi(C, p_0) = (\lambda_1, P_0) = 0$ in $H^2(X)$. As there is no such point P_1, (C, p_0) is not contractible in X and is therefore nonzero in $H_2(X)$. Its image under Ψ is therefore nonzero in $H^2(X)$. Thus, ψ is a harmonic form in $H^{2,2}(X)$.

Lemma 5. *A (2,2)-form of the type $f(z, \bar{z}) dz^1 \wedge dz^2 \wedge d\bar{z}^3 \wedge d\bar{z}^4$ is not a linear combination with rational coefficients of cohomology classes deriving from complex submanifolds of complex codimension 2 in a complex submanifold of complex dimension 4 assuming two conditions: 1) the mother manifold is an algebraic vari-*

ety defined by homogeneous polynomial equations with real coefficients, and 2) all submanifolds are algebraic varieties defined by homogeneous polynomial equations with real coefficient.

Proof. Let N be a complex manifold of dimension n and let M be a complex submanifold of N of complex dimension k. Let $i : M \to N$ be the inclusion map and the complex structure of M be induced by the complex structure of N. Let $P \in M$ be a point in M and $(x'^1, y'^1, \ldots, x'^k, y'^k)$ be a local coordinate system in TM_P where M is considered as a real manifold of dimension $2k$. The coordinate system can be completed to a coordinate system of $TN_{i(P)} = TN_P$ by adding $2(n-k)$ coordinate vectors. This yields a coordinate system

$$(x', y') = (x'^1, y'^1, \ldots, x'^k, y'^k, x'^{k+1}, y'^{k+1}, \ldots, x'^n, y'^n)$$

to TN_P where $P = (0, 0) = (0, \ldots, 0)$. We can assume that the coordinates are orthonormal. The manifold N can be considered as a real manifold of dimension $2n$ and TN_P be given a local coordinate system

$$(x, y) = (x^1, y^1, \ldots, x^n, y^n) \quad , \quad P = i(P) = (0, 0)$$

There is an orthonormal coordinate transform $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, $A \in SO(2n)$, such that

$$[x' \; y']^T = A[x \; y]^T \tag{39}$$

The coordinate systems at TN_P can be chosen such that

$$\begin{aligned} z^j &= x^j + iy^j \quad , \quad \bar{z}^j = x^j - iy^j \quad j = 1, \ldots, n \\ z'^j &= x'^j + iy'^j \quad , \quad \bar{z}'^j = x'^j - iy'^j \quad j = 1, \ldots, n \end{aligned} \tag{40}$$

and there is an orthonormal coordinate transform $B : \mathbb{C}^n \to \mathbb{C}^n$ such that for

$$\begin{aligned} (z, \bar{z}) &= (z^1, \bar{z}^1, \ldots, z^n, \bar{z}^n) \\ (z, \bar{z}') &= (z'^1, \bar{z}'^1, \ldots, z'^n, \bar{z}'^n) \end{aligned} \tag{41}$$

the transform takes $(z, \bar{z})$ to $(z', \bar{z}')$, $[z' \ \bar{z}']^T = B[z \ \bar{z}]^T$. As the complex structure of M is induced from the complex structure of N, the following statement holds

$$\begin{aligned} if \quad (z', \bar{z}') = (x'^1, x'^1, \ldots, x'^n, x'^n) \quad i.e. \quad y'^j = 0 \ for \ j = 1, \ldots, n \\ then \quad (z, \bar{z}) = (x^1, x^1, \ldots, x^n, x^n) \quad i.e. \quad y^j = 0 \ for \ j = 1, \ldots, n \end{aligned} \tag{42}$$

It follows that if

$$z'^j = \sum_m a_{m,j} z^m \tag{43}$$

then

$$\bar{z}'^j = \sum_m a_{m,j} \bar{z}^m \tag{44}$$

that is, $a_{m,j} \in \mathbb{R}$ for all j and m. The Poincaré dual $[M]$ of M considered as a $2k$-dimensional real manifold in the $2n$-dimensional real manifold N satisfies the condition that $[M]$ capped with the fundamental class of N is the homology class of M. Thus, $[M]$ is a form of the type

$$\varphi(x', y') dx'^{k+1} \wedge dy'^{k+1} \wedge \cdots \wedge dx'^n \wedge dy'^n \tag{45}$$

Since x'^j and y'^j are independent, the class $[M]$ is represented at p by a form of the type

$$\varphi(z', \bar{z}') dz'^{k+1} \wedge d\bar{z}'^{k+1} \wedge \cdots \wedge dz'^n \wedge d\bar{z}'^n \tag{46}$$

In coordinates $\{z^1, \bar{z}^1, \ldots, z^n, \bar{z}^n\}$ the form can be expressed by inserting dz'^j and $d\bar{z}'^j$ as linear combinations (with real coefficients) of dz^k and $d\bar{z}^k$, $k = 1, \ldots, n$.

Let us now take a two dimensional complex plane in a four dimensional complex manifold. We will use lower indices in this calculation as we obtain squares at some places and upper indices get easily confused with powers. The two local coordinates z''_1, z''_2 of a complex plane in a four dimensional complex manifold can be expressed in the local coordinates z_1, z_2, z_3, z_4 of the mother manifold as

$$\begin{aligned} z''_1 &= b_{11}z_1 + b_{12}z_2 + b_{13}z_3 + b_{14}z_4 \\ z''_1 &= b_{21}z_1 + b_{22}z_2 + b_{23}z_3 + b_{24}z_4 \end{aligned} \tag{47}$$

where b_{ij} are complex numbers. We want to see what terms come from

$$dz''_1 \wedge d\bar{z}''_1 \wedge dz'_2 \wedge d\bar{z}'_2 \tag{48}$$

when the local coordinates z''_i of the submanifold are replaced by the local coordinates z_i of the mother manifold. We get 36 terms the form $dz_{m_1} \wedge dz_{m_2} \wedge d\bar{z}_{m_3} \wedge d\bar{z}_{m_4}$ where $m_1 < m_2$ and $m_3 < m_4$ because if f $m_1 = 1$, m_1 can be $2, 3, 4$, if $m_1 = 2$, m_2 can be $3, 4$, and if $m_1 = 3$ then $m_2 = 4$. This means that there are six possibilities for m_1, m_2. There are also six possibilities for m_3, m_4. Together there are 36 possibilities and thus 36 different terms. While the terms can be calculated we get simpler expressions by making a linear transform of the local coordinates z''_1, z''_2

$$\begin{aligned} z'_1 &= \alpha(-b_{22}z''_1 + b12z''_2) \\ z'_2 &= \alpha(b_{21}z''_1 - b11z''_2) \\ \alpha &= (b_{21}b_{12} - b_{11}b_{22})^{-1} \end{aligned} \tag{49}$$

Then

$$\begin{aligned} z'_1 &= z_1 + a_{13}z_3 + a_{14}z_4 \\ z'_2 &= z_2 + a_{23}z_3 + a_{24}z_4 \\ a_{13} = \alpha(b_{12}b_{23} - b_{13}b_{22}) \quad &, \quad a_{14} = \alpha(b_{12}b_{24} - b_{14}b_{22}) \\ a_{23} = \alpha(b_{21}b_{13} - b_{11}b_{23}) \quad &, \quad a_{24} = \alpha(b_{21}b_{14} - b_{11}b_{24}) \end{aligned} \tag{50}$$

These new coordinates are not orthogonal, but that does not matter here: we just want to see what terms come from

$$dz_1' \wedge d\bar{z}'_1 \wedge dz_2' \wedge d\bar{z}'_2 \tag{51}$$

when the z_i coordinates are inserted. Again there are 36 terms, but the expressions are shorter than for the original coordinates z_i''. The first six terms are:

$$\begin{gathered}
= -dz_1 \wedge dz_2 \wedge d\bar{z}_1 \wedge d\bar{z}_2 - |a_{13}|^2 dz_2 \wedge dz_3 \wedge d\bar{z}_2 \wedge d\bar{z}_3 - |a_{14}|^2 dz_2 \wedge dz_4 \wedge d\bar{z}_2 \wedge d\bar{z}_4 \\
-|a_{23}|^2 dz_1 \wedge dz_3 \wedge d\bar{z}_1 \wedge d\bar{z}_3 - |a_{24}|^2 dz_1 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_4 \\
-|a_{23}a_{14} - a_{24}a_{13}|^2 dz_3 \wedge dz_4 \wedge d\bar{z}_3 \wedge d\bar{z}_4
\end{gathered} \tag{52}$$

Terms 7-16 numbered in this order:

$$\begin{gathered}
+a_{13}^* dz_1 \wedge dz_2 \wedge d\bar{z}_2 \wedge d\bar{z}_3 + a_{14}^* dz_1 \wedge dz_2 \wedge d\bar{z}_2 \wedge d\bar{z}_4 \\
+a_{13} dz_2 \wedge dz_3 \wedge d\bar{z}_1 \wedge d\bar{z}_2 + a_{14} dz_2 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_2 \\
-a_{13}a_{14}^* dz_2 \wedge dz_3 \wedge d\bar{z}_2 \wedge d\bar{z}_4 - a_{13}^* a_{14} dz_2 \wedge dz_4 \wedge d\bar{z}_2 \wedge d\bar{z}_3 \\
+a_{23}(a_{23}^* a_{14}^* - a_{24}^* a_{13}^*) dz_1 \wedge dz_3 \wedge d\bar{z}_3 \wedge d\bar{z}_4 \\
+a_{24}(a_{23}^* a_{14}^* - a_{24}^* a{\ast}_{13}) dz_1 \wedge dz_4 \wedge d\bar{z}_3 \wedge d\bar{z}_4 \\
+a_{23}^*(a_{23}a_{14} - a_{24}a_{13}) dz_3 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_3 \\
+a_{24}^*(a_{23}a_{14} - a_{24}a_{13}) dz_3 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_4
\end{gathered} \tag{53}$$

Terms 17-22:

$$\begin{gathered}
-a_{24}a_{23}^* dz_1 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_3 - a_{24}^* a_{23} dz_1 \wedge dz_3 \wedge d\bar{z}_1 \wedge d\bar{z}_4 \\
-a_{13}(a_{23}^* a_{14}^* - a_{24}^* a_{13}^*) dz_2 \wedge dz_3 \wedge d\bar{z}_3 \wedge d\bar{z}_4 \\
-a_{14}(a_{23}^* a_{14}^* - a_{24}^* a_{13}^*) dz_2 \wedge dz_4 \wedge d\bar{z}_3 \wedge d\bar{z}_4 \\
-a_{13}^*(a_{23}a_{14} - a_{24}a_{13}) dz_3 \wedge dz_4 \wedge d\bar{z}_2 \wedge d\bar{z}_3 \\
-a_{14}^*(a_{23}a_{14} - a_{24}a_{13}) dz_3 \wedge dz_4 \wedge d\bar{z}_2 \wedge d\bar{z}_4
\end{gathered} \tag{54}$$

Terms 23-34:

$$\begin{aligned}
&+a^*_{23}a_{13}dz_2\wedge dz_3\wedge d\bar{z}_1\wedge d\bar{z}_3+a^*_{23}a_{14}dz_2\wedge dz_4\wedge d\bar{z}_1\wedge d\bar{z}_3\\
&+a^*_{24}a_{13}dz_2\wedge dz_3\wedge d\bar{z}_1\wedge d\bar{z}_4+a^*_{24}a_{14}dz_2\wedge dz_4\wedge d\bar{z}_1\wedge d\bar{z}_4\\
&+a_{23}a^*_{13}dz_1\wedge dz_3\wedge d\bar{z}_2\wedge d\bar{z}_3+a_{23}a^*_{14}dz_1\wedge dz_3\wedge d\bar{z}_2\wedge d\bar{z}_4\\
&+a_{24}a^*_{13}dz_1\wedge dz_4\wedge d\bar{z}_2\wedge d\bar{z}_3+a_{24}a^*_{14}dz_1\wedge dz_4\wedge d\bar{z}_2\wedge d\bar{z}_4\\
&\quad-a^*_{23}dz_1\wedge dz_2\wedge d\bar{z}_1\wedge d\bar{z}_3-a^*_{24}dz_1\wedge dz_2\wedge d\bar{z}_1\wedge d\bar{z}_3\\
&\quad-a_{23}dz_1\wedge dz_3\wedge d\bar{z}_1\wedge d\bar{z}_2-a_{24}dz_1\wedge dz_4\wedge d\bar{z}_1\wedge d\bar{z}_2
\end{aligned}\tag{55}$$

We are mainly interested in these two last terms, terms 35 and 36:

$$\begin{aligned}
&+(a_{23}a_{14}-a_{24}a_{13})dz_3\wedge dz_4\wedge d\bar{z}_1\wedge d\bar{z}_2\\
&+(a^*_{23}a^*_{14}-a^*_{24}a^*_{13})dz_1\wedge dz_2\wedge d\bar{z}_3\wedge d\bar{z}_4
\end{aligned}\tag{56}$$

We assume that the mother manifold is an algebraic variety defined by a finite number of homogeneus polynomial equations of a finite number of complex parameters s_i and real coefficients. An example of such a variety is the Fermat quadratic surface defined by the homogeneous polynomial equation

$$s_1^4+s_2^4+s_3^4+s_4^4=0\tag{57}$$

In thie polynomial equation the coefficients are all integers (all are 1s).

If $P_o=(s_{10},s_{20},s_{30},s_{40})$ is a chosen basepoint, the local coordinates of the manifold can be chosen as $z_i=s_i-s_{i0}$. Let us consider a (2,2)-form

$$f(P)dz_1\wedge dz_2\wedge d\bar{z}_3\wedge d\bar{z}_4\tag{58}$$

that is linear combination with rational coefficients of classes of n complex submanifolds. Each submanifold gives 36 terms i of the type

$$\varphi_j(P)\varphi_k(P)A_{j,k,i}dz_{m_1} \wedge dz_{m_2} \wedge d\bar{z}_{m_3} \wedge d\bar{z}_{m_4} \tag{59}$$

Here $A_{j,k,i}$ is the coefficient of $dz_{m_1} \wedge dz_{m_2} \wedge d\bar{z}_{m_3} \wedge d\bar{z}_{m_4}$ for two (1,1)-forms j and k and i is numbered as in (3.36-3.40) where we have explicitly written the coefficients $A_{1,2,i}$, $i = 1, \ldots, 36$. Thus, for $j = 1$ and $k = 2$ the coefficients of the last two terms are

$$\begin{aligned} A_{1,2,35} &= a_{23}a_{14} - a_{24}a_{13} \\ A_{1,2,36} &= a^*_{23}a^*_{14} - a^*_{24}a^*_{13} \end{aligned} \tag{60}$$

where a_{mn} are calculated for the submanifold corresponding to two (1,1)-forms that we numbered as $j =$ and $k = 2$. For different values of j and k we get different a_{mn} but it was inconvenient to add the indices j, k to (3.36)-(3.40).

If we use the coordinates z''_i instead of the simpler coordinates z'_i and calculate $A_{1,2,35}$ and $A_{1,2,36}$, they are

$$\begin{aligned} A_{1,2,35} &= -b_{13}b_{24}b^*_{11}b^*_{22} + b_{14}b_{23}b^*_{11}b^*_{22} + b_{13}b_{24}b^*_{12}b^*_{21} - b_{14}b_{23}b^*_{12}b^*_{21} \\ A_{1,2,36} &= -b^*_{13}b^*_{24}b_{11}b_{22} + b^*_{14}b^*_{23}b_{11}b_{22} + b *_{13} b^*_{24}b_{12}b_{21} - b^*_{14}b^*_{23}b_{12}b_{21} \end{aligned} \tag{61}$$

Clearly, always holds

$$A_{j,k,36} = {A_{j,k,35}}^* \tag{62}$$

In order for a (2,2)-form of the type (3.42) to be a $\mathbb{Q}$-linear combination forms corresponding to submanifolds (of this type) in each point P must hold 35 equations of the type

$$\sum_{j,k} \varphi_j(P)\varphi_j(P)A_{j,k,i} = 0 \quad i = 1, \ldots, 35 \tag{63}$$

and the last sum must be non-zero

$$\sum_{j,k} c_{jk}\varphi_j(P)\varphi_j(P)A_{j,k,36} \neq 0 \tag{64}$$

where c_{jk} are rational coefficients. This is a set of linear equations for the unknowns c_{jk}. It is clear that in a chosen basepoint P_0 we can solve these linear equations and find complex numbers c_{jk} that satisfy these 36 equations, provided that there are at least 36 independent submanifolds. It is not clear if we can find rational numbers c_{jk} that satisfy all equations, but let us assume that we have found rational c_{jk} that satisfy these equations at the point P_0. The question is if these same c_{jk} can satisfy these 36 equations in every other point P.

It is easily shown that they cannot, assuming what we have assumed that the homogeneous polynomial equations defining the mother manifold and the submanifolds have real coefficients. With real coefficients we can find another point P_1 as a complex conjugate of P_0.

Thus, let us take the second point as $P_1 = (s_{10}^*, s_{20}^*, s_{30}^*, s_{30}^*)$. Then $P_1 = P_0^*$. The mother manifold and submanifolds are assumed to be algebraic varieties defined by a finite number of homogeneous polynomials with real coefficients. Therefore if P_0 is a solution to the homogeneous polynomial equations, then so it P_0^*. The functions $\varphi_j(P)$ have the same form at $P = P_0$ and $P = P_1$ but the values differ: $\varphi_j(P_1) = \varphi_j(P_0^*) = \varphi_j(P_0)^*$ because the coefficients of the homogeneous polynomial equations are real. This is the essence: in order to show that the 36 equations cannot be satisfied at every point P we need to find another point P_1 which is far from P_0 (in the close vicinity of P_0 the 36 equations are satisfied) and be able to calculate what the 36 equations are at the point $P-1$. It is not likely that the 36 equations could be satisfied at each point P even if we would not make the restrictive assumption, but it is difficult to find a point P_1 without making this assumption.

There are two ways to transfer the local coordinates from P_0 to P_1. One way is to choose similar local coordinates z_i, $i = 1, \ldots, 4$, and z_i'', $j = 1, 2$, in the basepoint P_1 as in P_0. In this case the form $dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4$ does not change in the transformation and the numbers $A_{jk,i}$, $i = 1, \ldots, 36$, do not change:

$$A_{j,k,36}(P_1) = A_{j,k,36}(P_0) \tag{65}$$

where $A_{j,k,i}(P)$ means $A_{j,k,i}$ calculated at the point P. The other method is to take conjugates. In this case both the form $dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4$ and the numbers $A_{jk,i}$, $i = 1, \ldots, 36$ change to complex conjugates.

Let us follow the first way. The equation 36 at the point $P = P_1$ gives the coefficient of the term $dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4$. The coefficient is

$$\begin{aligned}\sum_{j,k} c_{jk}\varphi_j(P_1)\varphi_j(P_1)A_{j,k,36}(P_1) &= \sum_{j,k} c_{jk}\varphi_j(P_0)^*\varphi_j(P_0)^*A_{j,k,36}(P_0)\\ &= \left(\sum_{j,k} c_{jk}\varphi_j(P_0)\varphi_j(P_0)\left(A_{j,k,36}(P_0)\right)^*\right)^*\\ &= \left(\sum_{j,k} c_{jk}\varphi_j(P_0)\varphi_j(P_0)A_{j,k,35}(P_0)\right)^* = 0\end{aligned} \tag{66}$$

That is, at the point P_1 the form $dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4$ vanishes. The coefficient that does not vanish at P_1 is the coefficient of the term $dz_3 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_2$.

The second way corresponds to finding the form at P_1 by conjugating the form at P_0. In this case

$$\begin{aligned}&\left(\sum_{j,k} c_{jk}\varphi_j(P_0)\varphi_j(P_0)A_{j,k,36}(P_0)dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4\right)^*\\ &= \sum_{j,k} c_{jk}\varphi_j(P_1)\varphi_j(P_1)A_{j,k,36}(P_1)\left(dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4\right)^*\\ &= \sum_{j,k} c_{jk}\varphi_j(P_1)\varphi_j(P_1)A_{j,k,36}(P_1)d\bar{z}_1 \wedge d\bar{z}_2 \wedge dz_3 \wedge dz_4\\ &= \sum_{j,k} c_{jk}\varphi_j(P_1)\varphi_j(P_1)A_{j,k,36}(P_1)dz_3 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_2\end{aligned} \tag{67}$$

In this method the coefficient of the form stays as nonzero, but it is not the coeffient of $dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4$ at P_1. It is the coefficient of $dz_3 \wedge dz_4 \wedge d\bar{z}_1 \wedge d\bar{z}_2$.

The coeffient of $dz_1 \wedge dz_2 \wedge d\bar{z}_3 \wedge d\bar{z}_4$ at P_1 is zero also in this method. Naturally, both ways give the same result.

We conclude that the form (3.42) cannot be created as a linear combination of terms (3.43). There should be at least two nonzero terms, those corresponding to $A_{j,k,36}$ and to $A_{j,k,35}$.

Theorem 1. *The algebraic variety*

$$\begin{gathered} M = \{(s_1 : s_2 : s_3 : s_4 : s_5 : s_6 : s_7 : s_8) | s_1^4 + s_2^4 + s_3^4 + s_4^4 = 0, \\ s_5^4 + s_6^4 + s_7^4 + s_8^4 = 0, s_4^4 + s_5^4 = 0\} \end{gathered} \tag{68}$$

is an algebraic subvariety of the complex projective space $\mathbb{P}^7$ of codimension 3 and a complex manifold of dimension 4. There is a Hodge class in $H^{2,2}$ that cannot be represented as a Q-linear combination of the classes of algebraic subvarieties of M.

Proof. Clearly M is a submanifold of $\mathbb{P}^7$ and it is an algebraic variety. It contains two copies of the Fermat quadratic surface, which is a K_3 surface. The submanifolds of dimension 1 for the Fermat quadratic surface are defined by adding one homogeneous polynomial equation. The dimension of the space of (1,1)-forms in K_3 is 20, thus if we find 20 polynomial equations that give an essentially different submanifold, we have represented all (1,1)-forms by an algebraic subvarity. By inspectation, the homogeneous polynomial equations are

$$\begin{gathered} s_i^4 = 0 \quad i = 1, 2, 3, 4 \\ s_i^4 = 1 \quad i = 1, 2, 3, 4 \\ s_i^4 + s_j^4 = 0 \quad (i, j) = (1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4) \\ s_i^4 + s_j^4 = 1 \quad (i, j) = (1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4) \end{gathered} \tag{69}$$

Clearly, fixing one of s_i to zero gives a submanifold and the submanifold is different for each $i = 1, 2, 3, 4$. In the complex projective space $\mathbb{P}^1$ there are only two numbers, zero and one. Thus, setting $s_i = 1$ gives another set of four different submanifolds and there are no more equations tying only one s_i. Setting the sum of two s_i variable terms together gives six different submanifolds if set the sum to zero and another six if we set the sum to one. There could be more equations of this type, but as these 20 equations give 20 different classes, we need not look further. There cannot be more classes: all other equations yield linear combinations of the classes. All of these 20 equations have real coefficients and we can use Lemmas 3.1-3.3 and 3.5.

Lemma 3.2 holds without any special considerations. In Lemma 3.3 we notice that a complex K_3 surface is compact and as a real 4-manifold it is closed. Thus, the product of two K_3 surfaces is closed. M is obtained from $K3 \times K3$ by adding an equation and is also closed. Lemma 3.3 can be used.

Lemma 3.4 applies also to M and gives the (2,2)-form.

The conditions for Lemma 3.5 are filled by M and the lemma shows the (2,2)-form of Lemma 3.4 cannot be represented as a $\mathbb{Q}$-linear combination of classes of submanifolds.

The (2,2)-form in Lemma 3.4 has only one term in the local coordinates of M everywhere. Thus, it is not a linear combination. It is a single term and therefore can be considered as a rational class.

I wrote the first version of this paper in 2011 and sent it to arXiv for discussion purposes. I was not at all sure if my result was correct since the solution seemed too easy and there was a conflict with a statement in [4] on page 54. I got an answer from an expert of the field, Bert van Geemen. His comment was that the (2,2)-form in Lemma 3.4 does not represent a Hodge class since we cannot show that it is rational.

The explanation by Bert van Geeman is the following:

$''$ We remember that singular cohomology of a complex manifold of complex dimension n is defined as the singular cohomology of the underlying real manifold of dimension $2n$. This yields $H^*(X;\mathbb{Z})$. In order to get $H^*(X;k)$, where the field

k is $\mathbb{Q}$ or $\mathbb{R}$, or $\mathbb{C}$, we form the tensor product of $H^*(X;\mathbb{Z})$ with k. When the base classes of $\mathbb{Z}$ are selected, the classes of the tensor products are linear combinations of the base classes with coefficients in k. We can multiply the created (2,2)-class with any number in k and get a harmonic class. If the class is a multiple of a base vector, we can always multiply it with a suitable number to get a class in $H^4(X;\mathbb{Q})$. However, if it is a linear combination with generic coefficients, multiplication by one number does not give a rational class. This seems to be the case with this (2,2)-form.

Theorem 2 in [4] on page 54, also published in [5], shows that the Hodge conjecture holds for certain K3xK3 spaces, showing that in the general case the constructed (2,2)-form cannot be a Hodge class. The space of Hodge classes $B(H^2(S,\mathbb{Q})\otimes H^2(S,\mathbb{Q}))$ is identified up to a Tate twist with $End_{Hdg}(H^2(S,\mathbb{Q}))$ on page 14 in [4]. Zarhin's theorem is used to characterize $End_{Hdg}(T)$ on page 16, and Mukai's theorem is used in the proof of Theorem 2, as in the proofs of other theorems in [4]. Mukai's theorem requires that the endomorphism $\varphi : T(S_1) \to T(S_1)$ is a Hodge isometry. A Hodge isometry maps $H^{2,0}(X)$ to $H^{2,0}(X)$, as is clear e.g. from [6] page 211. This method in [4] seems to derive from a paper of D. Morrison [7]. The endomorphism in the presented paper is complex conjugation, which sends $H^{2,0}(X)$ to $H^{0,2}(X)$ and is thus not a Hodge isometry. Mukai's theorem thus cannot be used, but the identification of the space of Hodge classes with endomorphisms preserving the Hodge structure still holds, and consequently the (2,2)-form created in Lemma 3.4 does not represent a Hodge class.″

I accepted this answer from an expert at that time and wrote in revised version of the arXiv paper that the proof does not give a counterexample to the Hodge conjecture. However, when I have now checked the paper, I find that the presented (2,2)-form is not a linear combination of (2,2)-forms. It has a single term and therefore a multiplication with a number does give a rational form. Thus, it is a Hodge class. The complex manifold M can be understood as a real manifold of dimension eight. The form corresponds to a real submanifold of dimension four that is obtained by selecting local coordinates as $v_1 = e_{x,1} + e_{y,1}$, $v_2 = e_{x,1} - e_{y,1}$, $v_3 = e_{x,2}+e_{y,2}$ and $v_4 = e_{x,2}-e_{y,2}$. This submanifold gives a homology class that is the class of the submanifold. By duality there is a cohomology class corresponding to it. This cohomology class is inherited to singular cohomology of M and it is not a linear combination of anything.

The issue with the apparent conflict with [4] is avoided by noticing that the proofs of Lemma 3.1 and 3.5 require a special condition on the K3-surface. Therefore they do not state anything of K3xK3 spaces in the general case. Yet, if it follows from those results that the (2.2)-form in Lemma 3.4 cannot be rational, then the logic of that argument should be checked.

References

1. P. Deligne: The Hodge onjecture, 2001. available on-line at www.claymath.org.
2. H. Suzuki: On the realization of homology classes by submanifolds, Tr. American Math. Soc., Vol. 87, No. 2, 1958, pp. 541-550. available on-line.
3. K. Kodaira: Complex Manifolds and Deformation of Complex Structures. Springer-Verlag, New-York, 1986.
4. U. Schlickewei: Hodge classes on self-products of K3 surfaces. Ph.D. Thesis, available on-line at hss.ulb.uni-bonn.de/2009/1805/1805.htm.
5. U. Schlickewei: The Hodge conjecture for self-products of certain K3 surfaces. Journal of Algebra, 324 (2010), 507-529.
6. B. van Geemen and J. Top: An isogeny of K3 surfaces, Bull. London Math. Soc. 38 (2008), 209-223. available on-line.
7. D. Morrison: Algebraic cycles on products of surfaces, Proc. Algebraic Geometry Symposium, Tohoku University (1984), 194-210. available on-line.

On the rank of elliptic curves

Abstract. The paper proves that the Birch and Swinnerton-Dyer conjecture is false. This is the long version of the paper. A short version was submitted to a journal.

Key words: Elliptic curves, Euler product, Birch and Swinnerton-Dyer conjecture.

6 Introduction

Let $P = \{p_1, p_2, \dots | p_j$ is a prime, $p_{j+1} > p_j > 1, j \geq 1\}$ be the set of all primes larger than one. In [1] an elliptic curve C over the field of rational numbers Q is a curve defined by the Weierstrass equation

$$y^2 = x^3 + ax + b$$

where $a, b \in Z$ and $x, y \in Q$. The discriminant of the cubic equation is $\varDelta = -16(4a^3 - 27b^2) \neq 0$. Let N_{p_j} denote the number of solutions to $y^2 = x^3 + ax + b$ mod p_j and let $a_{p_j} = p_j - N_{p_j}$.

The incomplete L-function of the curve C is

$$L(C, s) = \prod_{j \in A_C} (1 - a_{p_j} p_j^{-s} + p_j^{1-2s})^{-1}, \tag{1}$$

where

$$A_C = \{j \in \mathbb{N}, j > 0, p_j \textit{ does not divide } \varDelta\}.$$

The Euler product (1) converges absolutely at least if $Re\{s\} > 2$ because $|a_p| \leq 2p$. This upper bound for $< a_p|$ is obvious since x takes p values and y can take two values for each x. Hasse's statistical bound $|a_p| \leq 2p^{\frac{1}{2}}$ improves the area of absolute convergence to $Re\{s\} > 3/2$ and [1] gives this area. The problem statement [1] tells that $L(C, s)$ has a holomorphic continuation to the whole complex plane, thus it does not have poles.

The Birch and Swinnerton-Dyer conjecture says that the Taylor expansion of $L(C, s)$ at $s = 1$ has the form

$$L(C, s) = c(s-1)^r + \text{ } \textit{higher order terms} \tag{2}$$

with $c \neq 0$ and r the rank of C. The rank of an elliptic curve is defined as the rank of the group of solutions in the rational numbers. The number r in the Taylor expansion of $L(C, s)$ is called the algebraic rank of the curve. The conjecture is thus that the rank and the algebraic rank are equal.

Let $p > 2$ be prime, Z_p the cyclic group of integers modulo p, and $Z_p^* = \{1, \ldots, p-1\}$. The set of quadratic residues modulo p is the set

$$QR_p = \{x \in Z_p^* | \exists y \in Z_p^* \text{ such that } y^2 \equiv x \pmod p\}$$

and the set of nonresidues modulo p is

$$QNR_p = \{x \in Z_p^* | x \notin QR_p\}.$$

If g is a primitive root of Z_p^*, then

$$Z_p^* = \{g^0, g^1, \ldots, g^{p-2}\}.$$

The set QR_p is the subset where g has even powers:

$$QR_p = \{g^0, g^2, \ldots\}.$$

Thus, $\#QR_p = \#QNR_p$, the sets QR_p and QNR_p have equally many elements. If the integer a divides integer b it is written as $a|b$. For brevity, we write $y \equiv x$ as a shorthand of $y \equiv x \pmod p$ when there is no chance of confusion.

There is a recursion formula for deriving rational solutions from a rational base point (x, y)

$$x_{i+1} = S_i^2 - 2x_i \ , y_{i+1} = y_i + S_i(x_{i+1} - x_i) \tag{3}$$

$$S_i = \frac{a + 3x_i^2}{2y_i}$$

This recursion gives a new rational solution in the following way:

$$y_{i+1}^2 = y_i^2 + 2y_iS_i(x_{i+1} - x_i) + S_i^2(x_{i+1} - x_i)^2$$

$$= x_i^3 + ax_i + b + 2y_iS_i(x_{i+1} - x_i) + S_i^2(x_{i+1} - x_i)^2$$

$$= x_{i+1}^3 + ax_{i+1} + b$$

yielding

$$x_{i+1}^2 + x_{i+1}x_i + x_i^2 + a = 2y_iS_i + (x_{i+1} + 2x_i)(x_{i+1} - x_i)$$

which gives

$$3x_i^2 + a = 2y_iS_i.$$

The recursion may end or it may generate an infinite number of rational solutions. An example where the recursion ends is the following:

We define the curve C_1 by a Weiestrass form with $a = 33$ and $b = -26$. As the base point we take $x_0 = 3$ and $y_0 = 10$. The recursion (3) shows that $x_1 = x_0$, $y_1 = y_0$. It follows that recursion generates only one solution $(3, 10)$. The curve C_1 is a special case of

$$x_0 = 3s^2 \qquad y_0 = 9s^2 \pm s \tag{4}$$

$$a = 27s^4 \pm 6s^2 \qquad b = s^2 - 27s^6$$

with $s = 1$. For every nonzero integer value s the solution (4) gives $x_1 = x_0$, $y_1 = y_0$. These solutions are found by setting $x_1 = x_0$ in (3).

The rank of an elliptic curve is the number of independent base points from which the recursion derives an infinite number of rational solutions. For C_1 the recursion gives only finitely many points, but for that special elliptic curve there may be other base points that give infinitely many different points. An example of an elliptic curve having infinitely many rational solutions is $y^2 = x^3 - 5^2x$. This is known since 5 is a congruent number. If d is a noncongruent number, such as r^2 for any integer r there are only three solutions: $(0, 0)$ and $(\pm d, 0)$.

The recursion formula (3) has a corresponding operation in integers modulo p in the form

$$x_{i+1} \equiv S_i^2 - 2x_i \pmod p \tag{5}$$

$$S_i \equiv (s + 3x_i^2)(2t_i)^{-1} \pmod p$$

$$t_i \equiv x^3 + ax + b \pmod p.$$

If (x_i, y_1) is a solution in Z_p^* then the recursion formula in Z_p gives another solution (x_{i+1}, y_{i+1}), $x_{i+1}, y_{i+1} \in Z_p^*$, where

$$y_{i+1}^2 \equiv t_{i+1} \pmod p.$$

The operation also takes a pair (x_i, t_i) where $t_i \in QNR_p$ into a pair (x_{i+1}, t_{i+1}) where $t_{i+1} \in QNR_p$. Iterating the operation gives classes of pairs (x_i, y_i). If there is a solution in Q, then all of the iterated solutions map to the same set of (x_i, y_i) in Z_p.

The claim that the Birch and Swinnerton-Dyer conjecture should hold seems to be based on the idea that an infinite number of solutions in rationals for an elliptic curve C would give more solutions in the modular case. This is a very strange idea because there are very many solutions for a modular equation e.g. in knapsack problems and it is very difficult to find integer solutions to knapsack problems. The modular problem and the integer problem are quite different issues. The same should be the case with the modular problem and the rational problem in elliptic curves. The natural expectation is that these problems are very different and one does not give information of the other.

Two elliptic curves over rationals are known to have very high rank (one exactly or rank 20 and the other of rank at least 28). They are of the form

$$y^2 + xy + y = x^3 - x^2 - b$$

where $b \in \mathbb{N}$. Writing this curve in the Weierstrass form gives

$$y_1^3 = x_1^3 + (-\frac{49}{48})x_1 + (-\frac{2149}{576} - b)$$

$$y_1 = y + \frac{1}{2}(x+1) \qquad x_1 = x - \frac{5}{12}$$

and in the form where coefficients are integers is

$$y_2^3 = 9x_1^3 - 147x_1 - 2149 - 576b$$

$$y_2 = 24y + 12x + 12 \qquad x_2 = 4x - \frac{5}{3}$$

As a and b in the Weierstrass form are not integers in these elliptic curves, they are not elliptic curves considered in [1] and in this paper.

7 Calculation of a_{p_j} for $y^2 = x^3 - d^2x$

Lemma 1. *Let $p > 2$ be prime and a an integer. Assume $-1 \in QNR_p$ and $a \not\equiv 0 \pmod{p}$. The number N_p of solutions to the modular equation*

$$y^2 \equiv x^3 + ax \pmod{p} \tag{6}$$

is $N_p = p$.

Proof. Let

$$A = \{x \in Z_p^* | t_1 \in QR_p, t_1 \equiv x(x^2+a) \pmod p\},$$

$$B = \{x \in Z_p^* | t_1 \in QNR_p, t_1 \equiv x(x^2+a) \pmod p\},$$

and $m_1 = \#A$, $m_2 = \#B$. We can write

$$A_1 = \{x = 1, \dots, \frac{p-1}{2} | t_1 \in QR_p, t_1 \equiv x(x^2+a) \pmod p\},$$

$$A_2 = \{x = \frac{p+1}{2}, \dots, p-1 | t_1 \in QR_p, t_1 \equiv x(x^2+a) \pmod p\},$$

$$B_1 = \{x = 1, \dots, \frac{p-1}{2} | t_1 \in QNR_p, t_1 \equiv x(x^2+a) \pmod p\},$$

$$B_2 = \{x = \frac{p+1}{2}, \dots, p-1 | t_1 \in QNR_p, t_1 \equiv x(x^2+a) \pmod p\},$$

and $m_{1,i} = \#A_i$, $m_{2,i} = \#B_i$, $i = 1, 2$. The sets A_1 and A_2 are disjoint and $A = A_1 \cup A_2$. Similarly, the sets B_1 and B_2 are disjoint and $B = B_1 \cup B_2$. Calculating

$$A_2 = \{-x = -\frac{p+1}{2}, \dots, -p-1 | t_1 \in QR_p, t_1 \equiv x(x^2+a) \pmod p\}$$

$$= \{-x = p - \frac{p+1}{2}, \dots, p-p+1 | t_1' \in QR_p$$

$$t_1' = p - t_1 \equiv (-x)((-x)^2+a) \pmod p\}$$

$$= \{-x = 1, \dots, \frac{p-1}{2} | t_1' \in QR_p, t_1' \equiv (-x)((-x)^2+a) \pmod p\}.$$

If it were true that $-1 \in QR_p$, then there would exists ϵ such that $-1 \equiv \epsilon^2$. Then for any y holds $-y^2 \equiv (\epsilon y)^2 \in QR_p$. But as we require that $-1 \in QNR_p$ it is not possible that $-y^2 \equiv h^2$ for any h because if it is $-1 \equiv (y^{-1}h)^2 \in QR_p$. Thus,

$-y^2 \in QNR_p$ for every y. Therefore

$$A_2 = \{x' = 1, \dots, \frac{p-1}{2} | t_1' \in QNR_p, t_1' \equiv x'(x'^2 + a) \pmod{p}\} = B_1.$$

Similarly, $A_1 = B_2$. It follows that

$$m_1 = m_{1,1} + m_{1,2} = m_{1,1} + m_{2,1},$$

$$m_2 = m_{2,1} + m_{2,2} = m_{2,1} + m_{1,1}.$$

Thus, $m_1 = m_2$. Let $a \in QR_p$. Then there are two values $x \in Z_p^*$ that yield $t_1 \equiv 0 \pmod{p}$. Therefore

$$m_1 + m_2 = p - 3 \ \Rightarrow m_1 = \frac{p-3}{2}.$$

Every $x \in A$ yields two solutions $y, p - y$ to (6). Every x giving $t_1 \equiv 0 \pmod{p}$ yields one solution $y = 0$ to (6). The number of solutions is

$$N_p = 2\frac{p-3}{2} + 3 = p.$$

If $a \in QNR_p$ then $m_1 + m_2 = p - 1$ and

$$N_p = 2\frac{p-1}{2} + 1 = p.$$

The lemma is proved. □

Let us give an example of Lemma 1. Let $d = 1$ and $p = 11$. Then $QR_{11} = \{1, 3, 4, 5, 9\}$. When x ranges from 0 to 10 the values of $x(x^2 - 1)$ give the sequence $0, 0, 6, 2, 5, 10, 1, 6, 9, 5, 0$. Removing zeros from this sequnce as they are neither in QR_p nor in QNR_p we notice that $-6 \equiv 5$. Because $-1 \in QNR_p$ we have $6 \in QNR_p$ and $-5 \in QR_p$. Likewise $-2 \equiv 9$, so $2 \in QNR_p$ and $9 \in QR_p$; $-10 \equiv 1$. The same is with $-1 \equiv 10 \in QNR_p$ and $1 \in QR_p$. We get $2(p-3)/2 = 8$

solutions: $(4,4)$, $(4,7)$, $(6,1)$, $(6,10)$, $(7,3)$, $(7,8)$, $(9,4)$, $(9,7)$, that is, for each x there are two y values. Additionally we have the zeros. They give three solutions $(1,0)$ and $(10,0)$ from $x^2-1 \equiv 0$ and $(0,0)$ is a solution. Together there are $11 = p$ solutions.

Lemma 2. *Let $p > 2$ be prime. The number of solutions y^2 to the equation*

$$y^2 - c \equiv x^2 \pmod p \tag{7}$$

satisfying $y^2, x^2 \in Z_p^*$ is

$\frac{p-5}{4}$ if $-1 \in QR_p$ and $c \in QR_p$,

$\frac{p-3}{4}$ if $-1 \in QNR_p$,

$\frac{p-1}{4}$ if $-1 \in QR_p$ and $c \in QNR_p$.

Proof. Let us assume that (7) holds. Thus there exists $z \in Z_p^*$ such that the modular equation

$$y^2 - x^2 = (y-x)(y+x) \equiv c$$

can be written as

$$y - x \equiv z \ , \ y + x \equiv z^{-1}c.$$

Then

$$y \equiv 2^{-1}x^{-1}(z^2+c) \ , \ x \equiv 2^{-1}z^{-1}(z^2-c).$$

Let $\pm\epsilon$ denote the two roots of $z^2 \equiv -1$ if $-1 \in QR_p$. If $-1 \in QNR_p$ there are no such roots.

If $c \in QNR_p$ and $-1 \in QR_p$ there are no solutions to the equations

$$z^2 \equiv c \qquad , \qquad (\epsilon z)^2 \equiv -c. \tag{8}$$

In this case we let z range over the $p-1$ numbers in Z_p^* in the equation for y. If two values z_1 and z_2 give the same y, then

$$z_1^{-1}(z_1^2+c) \equiv z_2^{-1}(z_2^2+c)$$

i.e.,

$$z_1 + cz_1^{-1} \equiv z_2 + cz_2^{-1},$$

$$z_1 - z_2 \equiv c(z_2^{-1} - z_1^{-1}).$$

Multiplying by z_1z_2

$$z_1z_2(z_1-z_2) \equiv z_1z_2c(z_2^{-1}-z_1^{-1}) \equiv c(z_1-z_2)$$

and $z_1z_2 \equiv c$, i.e, $z_2 \equiv cz_1^{-1}$. When z ranges over all values in Z_p^* the number y gets all values it can get and exactly two values z map to the same y. The number of different y is therefore $\frac{p-1}{2}$.

If some value of z gives y, another value of z gives $-y$. As $\pm y$ yield the same y^2 the number of different y^2 is half of the numbers of y, that is, $\frac{p-1}{4}$.

If $c \in QR_p$ and $-1 \in QR_p$ then there are two solutions z to both of the equations in (8). These four values of z are all different. Removing them gives $p-5$ values for the range of z. The number of different values y^2 is $\frac{p-5}{4}$.

If $c \in QR_p$ and $-1 \in QNR_p$ there are two solutions for $z^2 \equiv c$ but no solutions to $z^2 \equiv -c$. The number of different y^2 is $\frac{p-3}{4}$.

If $c \in QNR_p$ and $-1 \in QNR_p$ there are no solutions for $z^2 \equiv c$ but two solutions to $z^2 \equiv -c$. The number of different y^2 is $\frac{p-3}{4}$.□

Lemma 3. *Let $p > 2$ be prime and a an integer. Let $-1 \in QR_p$, $a \not\equiv 0$ (* mod p) *and g a primitive root of Z_p^*. The number N_p of solutions to the modular equation*

$$y^2 \equiv x^3 + ax \pmod{p} \tag{9}$$

is

$N_p = 8n_1 + 7$ if $-a \equiv g^{2i}$ and i is even,

$N_p = 2p - 8n_1 - 7$ if $-a \equiv g^{2i}$ and i is odd,

$N_p = 8n_g + 3$ if $-a \equiv g^{2i+1}$ and i is even,

$N_p = 2p - 8n_g - 3$ if $-a \equiv g^{2i+1}$ and i is odd.

Here n_c is the number of solutions $y^4 \in Z_p^*$ yielding $y^4 - c \in QR_p$, $c = 1$ or $c = g$.

Proof. Let

$$B = \{x' \in Z_p^* | t' \in QR_p, t' \equiv x'^3 + ax' \pmod{p}\}. \tag{10}$$

If $-a \equiv g^{2i}$ we insert $t \equiv g^{-3i}t'$ and $x \equiv g^{-1}x'$. Then $t' \equiv x'^3 + ax'$ changes to $g^{3i} \equiv g^{3i}x^3 - g^{2i}g^i x$, i.e., to $t \equiv x^3 - x$. We reduced $-a$ to $c = 1$.

If $-a \equiv g^{2i+1}$ we insert the same t and x as above. Then $t' \equiv x'^3 + ax'$ changes to $g^{3i} \equiv g^{3i}x^3 - g^{2i+1}g^i x$, i.e., to $t \equiv x^3 - gx$. We reduced $-a$ to $c = g$.

We write both of these cases as $t \equiv x^3 - cx$ where $c = 1$ if $-a = g^{2i}$ and $c = g$ if $-a = g^{2i+1}$.

Let

$$A = \{x \in Z_p^* | t_1 \in QR_p, t_1 \equiv x^3 - cx \pmod{p}\} \tag{11}$$

$$A' = \{x \in Z_p^* | t_1 \in QNR_p, t_1 \equiv x^3 - cx \pmod{p}\}.$$

If i is even then g^i is in QR_p and in the substitution $t = g^{3i}t'$ holds: if $t \in QR_p$ then $t' \in QR_p$. If i is odd, then $t \in QR_p$ implies that $t' \in QNR_p$. Thus, for even i $B = A$ while for odd i $B = A'$.

Let us write the sets A and A' differently

$$A = \{g^k | g^k(g^{2k} - c) \in QR_p, k = 0, \ldots, p-2\} \tag{12}$$

$$A' = \{g^k | g^k(g^{2k} - c) \in QNR_p, k = 0, \ldots, p-2\}$$

and let us divide them into subsets of even and odd indices of k

$$A_1 = \{g^{2k} | g^{2k}(g^{4k} - c) \in QR_p, k = 0, \ldots, \frac{p-3}{2}\}$$

$$A_2 = \{g^{2k} | g^{2k+1}(g^{2(2k+c)} - c) \in QR_p, k = 0, \ldots, \frac{p-3}{2}\}$$

$$A_1' = \{g^{2k+1} | g^{2k}(g^{4k} - c) \in QNR_p, k = 0, \ldots, \frac{p-3}{2}\}$$

$$A_2' = \{g^{2k+1} | g^{2k+1}(g^{2(2k+c)} - c) \in QNR_p, k = 0, \ldots, \frac{p-3}{2}\}.$$

Then $A = A_1 \cup A_2$, $\#A = \#A_1 + \#A_2$ and $A' = A_1' \cup A_2'$, $\#A' = \#A_1' + \#A_2'$.

We also define sets that do not have the $x = g^k$ term in $t = x(x^2 - c)$.

$$C = \{g^{2k} | g^{2k} - c \in QR_p, k = 0, \ldots, \frac{p-3}{2}\} \tag{13}$$

$$C' = \{g^{2k} | g^{2k} - c \in QNR_p, k = 0, \ldots, \frac{p-3}{2}\}$$

and divide these sets into subsets where a set with a running index $2k$ is divided into two sets with running indices $4k$ and $4k + 2$:

$$C_1 = \{g^{4k} | g^{4k} - c \in QR_p, g^{4k} \leq \frac{p-1}{2}, k = 0, \ldots, \frac{p-3}{2}\} \tag{14}$$

$$C_2 = \{g^{4k+2} | g^{4k+2} - c \in QR_p, g^{4k+2} \leq \frac{p-1}{2}, k = 0, \ldots, \frac{p-3}{2}\},$$

$$C_1' = \{g^{4k} | g^{4k} - c \in QNR_p, g^{4k} \leq \frac{p-1}{2}, k = 0, \ldots, \frac{p-3}{2}\},$$

$$C_2' = \{g^{4k+2} | g^{4k+2} - c \in QNR_p, g^{4k+2} \leq \frac{p-1}{2}, k = 0, \ldots, \frac{p-3}{2}\}.$$

The rule $g^{4k} \leq \frac{p-1}{2}$ and $g^{4k+2} \leq \frac{p-1}{2}$ removes half of the values of the running index. Then $C = C_1 \cup C_2$, $\#C = \#C_1 + \#C_2$ and $C' = C_1' \cup C_2'$, $\#C' = \#C_1' + \#C_2'$.

The idea is to map the solutions of $t \equiv g^k(g^{2k} - c)$ bijectively to solutions of $t' \equiv g^{2k} - c$. Clearly, if $g^k \in QNR_p$ multiplying with it changes $t' \in QR_p$ to $t \in QNR_p$ and if $g^k \in QR_p$ multiplying by it does not change the set. This is why we divided the sets to A_i, A'_i, $i = 1, 2$. In $i = 2$ sets $g^k \in QNR_p$, so if an element of C'_2 is multiplied by g^k we get an element of A_2. Likewise, C_2 and A'_2 correspond to each other.

The following relations hold

$$\begin{aligned} \#A =& \#2C_1 + \#2C'_2 \\ \#A' =& \#2C'_1 + \#2C_2 \\ \#C_2 =& \#C - \#C_1 \\ \#C'_2 =& \#C' - \#C'_1. \end{aligned}$$

Solving $\#A$ yields

$$\begin{aligned} \#A &= 2\#C_1 + 2\#C'_2 \\ &= 2\#C_1 + 2\#C' - 2\#C'_1. \end{aligned}$$

The value a is used in the proof of this lemma in two places only. One is in (10): if $-a \equiv g^{2i}$ or $-a \equiv g^{2i+1}$ and the index i is even, then $B = A$. If i is odd, then $B = A'$. The other place is in Lemma 2 where the numbers of solutions in the different cases depend on if whether $-a \equiv g^{2i}$ or $-a \equiv g^{2i+1}$, i.e., if $c = 1$ or $c = g$.

Case 1: $-a = g^{2i}$, i even. Then $c = 1$ and the relation

$$\#C' = \tfrac{p-3}{2} - \#C.$$

holds. In this relation we have counted the values of k in $C \cup C'$ and excluded the one value of k that gives $g^{2k} - 1 \equiv 0 \pmod{p}$ because $0 \notin QR_p \cup QNR_p$. Thus, the number of valid indices k is one less than the number $\frac{p-1}{2}$ of indices $k = 0, \ldots, \frac{p-3}{2}$ in (13). The correct value of valid indices $\frac{p-3}{2}$.

Counting indices k in $C_1 \cup C'_1$ in (14) gives

$$\#C'_1 = \tfrac{p-5}{4} - \#C_1.$$

In this relation we have excluded the solution to $g^{2k} - c \equiv 0 \pmod{p}$. In C_1 and C'_1 the counted element is not the number of indices k. It is the number of values g^{4k}. This number must be reduced by one. The result follows as $\frac{p-1}{4} - 1 = \frac{p-5}{4}$.

Calculating $\#A$ gives

$$\begin{aligned}\#A &= 2\#C_1 + p - 3 - 2\#C - \tfrac{p-5}{2} + 2\#C_1\\ &= 4\#C_1 - 2\#C + \tfrac{p-1}{2}.\end{aligned}$$

Writing $\#C_1 = n_1$ and inserting from Lemma 2 the case $-1 \in QR_p$ and $c \in QR_p$ where $\#C = \frac{p-5}{4}$ yields

$$\#\text{A}= 4n_1 - \tfrac{p-5}{2} + \tfrac{p-1}{2} = 4n_1 + 2.$$

In Case 1 holds $N_p = 2\#A + 3$ because if there is a solution $y^2 \equiv x(x^2 - 1)$, then it is satisfied by two y values, $\pm y$, and there are three solutions where $y \equiv 0$, namely $x \equiv 0$, $x^2 \equiv \pm 1$. Thus $N_p = 8n_1 + 7$.

Case 2: $-a = g^{2i}$, i odd. Then $B = A'$. Thus

$$\begin{aligned}\#A' &= 2\#C_1' + 2\#C_2\\ &= \tfrac{p-5}{2} - 2\#C_1 + 2\#C - 2\#C_1\\ &= \tfrac{p-5}{2} - 4n_1 + \tfrac{p-5}{2} = p - 4n_1 - 5\end{aligned}$$

In Case 2 $N_p = 2\#A' + 3$, thus $N_p = 2p - 8n_1 - 7$.

Case 3: $-a = g^{2i+1}$, i even. The differences to Case 1 are

$$\begin{aligned}N_p &= 2\#A + 1\\ \#C' &= \tfrac{p-1}{2} - \#C\\ \#C_1' &= \tfrac{p-3}{4} - \#C_1\\ \#C &= \tfrac{p-3}{4}\end{aligned}$$

because $c = g$ and $g^{2k} - g \equiv 0$ is not possible.

We denote $\#C_1 = n_g$ and insert from Lemma 2 the case $-1 \in QR_p$ and $c \in QNR_p$ where $\#C = \frac{p-1}{4}$. Making these changes to the calculation of Case 1 gives $N_p = 8n_g + 3$.

Case 4: $-a = g^{2i+1}$, i odd. Analogically with Cases 2 and 3 we get $N_p = 2p - 8n_g - 3$.

The definition of C_1 is a bit complicated as the running index k loops over twice as many indices than are needed and the set has a test to discard half of the values values k because in this way C is clearly the union of C_1 and C_2. It is good to notice that the set C_1 has as many members as the set

$$\{y^4 | y^4 - c \in QR_p\}$$

where $c = 1$ for Cases 1 and 2 and $c = g$ for Cases 3 and 4.

The proof of the lemma is completed. □

Let us look at an example of Lemma 2. Let $a = -d^2$ for $d = 1$ in (9) and $p = 13$. As g we choose 2, which is a primitive root for Z_{13}^*. Then $1 \equiv 2^0$, $2 \equiv 2^1$, $3 \equiv 2^4$, $4 \equiv 2^2$, $5 \equiv 2^9$, $6 \equiv 2^5$, $7 \equiv 2^{11}$, $8 \equiv 2^3$, $9 \equiv 2^8$, $10 \equiv 2^{10}$, $11 \equiv 2^7$ and $12 \equiv 2^6$. The sets are

$$A = \{2^3, 2^9\} \qquad A' = \{2^1, 2^2, 2^4, 2^5, 2^7, 2^8, 2^{10}, 2^{11}\}$$

$$A_1 = \emptyset \qquad A_2 = \{2^3, 2^9\}$$

$$A_1' = \{2^2, 2^4, 2^8, 2^{10}\} \qquad A_2' = \{2^1, 2^5, 2^7, 2^{11}\}$$

$$C = \{2^1, 2^5\} \qquad C' = \{2^2, 2^3, 2^4, 2^9\}$$

$$C_1 = \emptyset \qquad C_2 = \{2^1, 2^5\}$$

$$C_1' = \{2^2, 2^4\} \qquad C_2' = \{2^9\}$$

There is a direct correspondence between C_1 and the first half of A_1, as there is between C_1' and the second half of A_1'. This is because if $g^{2k}(g^{4k} - 1) \in QR_p$ then $g^{4k} - 1 \in QR_p$ and if $g^{4k} - 1 \in QR_p$ then $\pm g^{2k}(g^{4k} - 1) \in QR_p$ since $-1 \in QR_p$. There is also a direct correspondence between C_2' and the first half of A_2, as there is between C_2 and A_2'. This is because if $g^{2k+1}(g^{2(2k+1)} - 1) \in QR_p$ then $g^{2(2k+1)} - 1 \in QNR_p$ as $g \in QNR_p$, and if $g^{2(2k+1)} - 1 \in QNR_p$ then $\pm g^{2k+1}(g^{2(2k+1)} - 1) \in QR_p$ since $-1 \in QR_p$. This gives the relations between the sizes of the sets.

The case of Lemma 1 covers half of all p because of Lemma 4.

Lemma 4. *The following statements hold:*

(i) $-1 \in QR_p$ if and only if $4|(p-1)$

(ii) The number of $p < N$ such that $4|(p-1)$ approaches half when N grows to infinity.

Proof. Let g be a primitive root of Z_p^*. If $4|(p-1)$, then $a \equiv g^{\frac{p-1}{4}}$ is in Z_p^* and $a^2 \equiv -1$. If $-1 \in QR_p$, then $-1 \equiv g^{2i}$ for some i, $0 \leq i \leq p-2$. Since $-1 \not\equiv 1$ holds $2i \not\equiv 0 \ (\mod (p-1))$. Thus $2i \neq 0$ and $2i \neq p-1$. As $(-1)^2 \equiv 1 \equiv g^{p-1} \equiv g^{4i}$ holds $4i = k(p-1)$ for some k where k has the possible values $1, 2, 3$. If $k = 2$, then $-1 \equiv g^{2i} \equiv g^{p-1} \equiv 1$, which is impossible. Thus, $k \in \{1, 3\}$. Then $gcd(4, k) = 1$ and therefore $4|(p-1)$. This proves the claim (i).

Claim (ii) is shown true by considering the Sieve of Eratosthenes. In this algorithm primes are found by reserving a memory vector for all numbers and marking the place of 1 as full and all other places empty at the beginning. On each step the first unmarked place is taken as the next prime p. The place of p is marked and all multiples of p are marked. In this algorithm the first step takes $p = 2$ and marks all multiples of 2. The unmarked numbers are all odd. The next prime is $p = 3$, the first unmarked number. All multiples of 3 are marked. The numbers that are marked for p, i.e., multiples of p, are are all odd and equally distributed modulo 4. Consequently, the numbers that remain unmarked are all odd and equally distributed modulo 4. This continues in each step, thus the numbers that remain unmarked are all odd and equally distributed between 1 (mod 4) and 3 (mod 4).

In each step the first unmarked number is the next prime p. It is selected as the smallest number in a set of unmarked numbers that are always odd and distributed equally between two sets 1 (mod 4) and 3 (mod 4). The next prime p has half a chance in belonging to either set. The number $p - 1$ is always even and if $p \equiv 1$ modulo 4, then $4|(p-1)$. This is so in half of the cases when N approaches infinity. □

The numbers n_1 and n_2 in Lemma 3 are not easily evaluated for statistical purposes as we did not check how many numbers x give the same $t \equiv x(x^2+a)$. This is done in Lemma 5.

Lemma 5. *Let $p > 2$ be a prime and $-1 \in QR_p$. If $c = 1$ the solutions are divided into singlets and multiplets. In the set of singlets each x is mapped to a unique y^2. In multiplets three, in maximum size cases two, values of x are mapped to the same y^2.*

(i) There is a running index t such that in singlets t runs from 1 to $p-1$ and two values of t map to the same x.

(ii) In multiplets there is a running index t such that t runs from 1 to $p-1$ and two values of t map to the same x. For $p-m$ values of t three values of x map to the same y^2 and three values of $-x$ map to the same $-y^2 \in QR_p$. Thus, twelve indices t map together either to QR_p or to QNR_p. That is, two t map to the same x. Six t map to the same (h, x) where $h \equiv x(x^2-c)$. If $h \in QR_p$, then $-h \in QR_p$ and twelve values of t giving (h, x) or $(-h, -x)$ are all in QR_p. If $h \in QRP_p$, then twelve values of t map to (h, x) or $(-h, -x)$ are all in QNR_p. The value of m is as follows:

Case $c = 1$:

$$m = 13 \text{ if } 3 \in QR_p \text{ and } 3^{-1}+1 \in QR_p$$

$$m = 9 \text{ if } 3 \in QNR_p \text{ and } 3^{-1}+1 \in QR_p$$

$$m = 9 \text{ if } 3 \in QR_p \text{ and } 3^{-1}+1 \in QNR_p$$

$$m = 5 \text{ if } 3 \in QNR_p \text{ and } 3^{-1}+1 \in QNR_p$$

Case $c = g$:

If $3 \in QR_p$ and $g+1 \in QR_p$ then $m = 5$.

If $3 \in QR_p$ and $g+1 \in QNR_p$ then $m = 1$.

If $3 \in QR_p$ then $m = 1$.

(iii) In multiplets there are m_1 values of t that map to x such that two values of x map to the same $h \equiv x(x^2 - c)$ because one of the three values of x is equal to one of the other two. In this case eight values of t map to (h, x) or $(-h, -x)$, all are either in QR_p or all in QNR_p. The value of m_1 is as follows:

Case $c = 1$:

$$m_1 = 6 \text{ if } 3 \in QR_p \text{ and } 3^{-1} + 1 \in QR_p$$

$$m_1 = 4 \text{ if } 3 \in QNR_p \text{ and } 3^{-1} + 1 \in QR_p$$

$$m_1 = 2 \text{ if } 3 \in QR_p \text{ and } 3^{-1} + 1 \in QNR_p$$

$$m_1 = 0 \text{ if } 3 \in QNR_p \text{ and } 3^{-1} + 1 \in QNR_p$$

Case $c = g$:

$$\text{If } 3 \in QR_p \text{ and } g + 1 \in QR_p \text{ then } m_1 = 4.$$

$$\text{If } 3 \in QR_p \text{ and } g + 1 \in QNR_p \text{ then } m_1 = 0.$$

$$\text{If } 3 \in QR_p \text{ then } m_1 = 0.$$

Proof. As in Lemma 3 we can reduce the equation $h = x(x^2 + a)$ to $h = x(x^2 - c)$ where $c = 1$ or $c = g$, where g is a primitive root of Z_p^*. We will consider only the cases $c = 1$ and $c = g$.

Let us first consider how the equation

$$x^2 - r^2 \equiv y^2$$

where x and y belong to $QR_p \cup QNR_p$ and $r \not\equiv 0$ is fixed can be solved with a running index t. We have two ways of solving it.

In the first way we write $z^2 \equiv y^2$ and

$$x^2 - r^2 \equiv (x + r)(x - r) \equiv z^2.$$

There exist t such that

$$x + r \equiv zt \ , \ x - r \equiv zt^{-1}.$$

Then

$$x \equiv 2^{-1}z(t + t^{-1}) \ , \ z \equiv 2r(t - t^{-1})^{-1}.$$

Thus

$$x \equiv r(t - t^{-1})^{-1}(t + t^{-1}) \ , \ y \equiv \pm 2r(t - t^{-1})^{-1}.$$

If $x \equiv r(t-t^{-1})^{-1}(t+t^{-1})$ for some t, then the equation is automatically satisfied. We can count the indices t. If two values t_1 and t_2 give the same x, then

$$(t_1 - t_1^{-1})^{-1}(t_1 + t_1^{-1}) \equiv (t_2 - t_2^{-1})^{-1}(t_2 + t_2^{-1}).$$

Solving gives $t_2^2 \equiv t_1^2$. There are three different values of t, namely $t_1, t_1^{-1}, -t_1^{-1}$, that could give x, but checking shows that t_1^{-1} gives $-x$ and only two different values of t, namely t_1 and $-t_1^{-1}$ give x. Additionally x and $-x$ yield the same y^2 and $\pm y$ give the same y^2. Thus, eight indices t yield the same solution.

The second way to solve the equation is to write it as

$$x^2 - z^2 \equiv r^2 \ , \ z^2 \equiv y^2$$

and find a running index t such that

$$x + z \equiv rt \ , \ x - z \equiv rt^{-1}.$$

Then $x = 2^{-1}r(t + t^{-1})$ and $y = \pm 2^{-1}r(t - t^{-1})$. If $x \equiv 2^{-1}r(t + t^{-1})$ for some t, then the equation is automatically satisfied. Two values of t map to the same x because if

$$t_1 + t_1^{-1} \equiv t_2 + t_2^{-1}$$

then $t_2 \equiv t_1^{-1}$. The two values ±x give the same x^2 and two values ±y give the same y^2. Also in this way of solving eight indices of t give the same solution. The two solution methods give the same result.

In both ways we have to exclude indices t for which $t - t^{-1} \equiv 0$ or $t + t^{-1} \equiv 0$ because for suc t the number x or y is not in $QR_p \cup QNR_p$.

Next, consider the case of

$$x^2 + a \equiv b$$

where $a \in QNR_p$ and $b \in QNR_p$. Multiplying by a primitive root g gives the equation

$$c^2 - d^2 \equiv (c+d)(c-d) \equiv gx^2$$

where $c^2 \equiv gb \in QR_p$, $d^2 \equiv ga \in QR_p$ and we can use the running index method as

$$c + d \equiv gxt \ , \ c - d \equiv xt^{-1}.$$

If $b \in QNR_p$ and $-a \in QR_p$ we can write $-a = c^2$ and proceed as

$$x^2 - c^2 \equiv (x+c)(x-c) \equiv b.$$

The running index solution method fails only if $b \in QNR_p$ and $a \in QR_p$ but $-1 \in QNR_p$.

The second general issue we have to look is how to estimate the number of solutions to $h \in QR_p$ for $h \equiv x(x^2 - c)$. We can solve the equation

$$z^2 \equiv x^2 - c$$

by a running index s as

$$z + x \equiv -cs \ , z - x \equiv s^{-1}$$

i.e., if $x \equiv -2^{-1}(cs + s^{-1})$ for some s, then $x^2 - c \in QR_p$. If x is not of that form for any s, then $x^2 - c \in QNR_p$. If ($x \in QR_p$ and $x^2 - c \in QR_p$) or ($x \in QNR_p$ and $x^2 - c \in QNR_p$), then $h \equiv x(x^2 - c) \in QR_p$, else $h \in QNR_p$. The numbers x are produced by a running index t that gives every number x twice. Thus, x has half of the values in Z_p^*, or possibly a few less if some t must be discarded. The possible values that x should have in order that $x^2 - c$ is a square z^2 are given by another running index s. Also here some values of s may need to be discarded. The running indes s also gives each x twice and can reach (about) half of the values in Z_p^*.

If these two sets of numbers x can be considered independent, then for one fourth of the indices t holds $x \in QR_p$ and $x^2 - c \in QR_p$. For one fourth of the indices t holds $x \in QNR_p$ and $x^2 - c \in QNR_p$. Thus, for half of the indices t holds $h \in QR_p$. If so, then we can think of indices t as probabilistic trials where half of the time hit a success, $h \in QR_p$, and half of the time fail, $h \in QNR_p$. We will later formulate this condition as the Statistical Assumption.

The issue in this statistical method to know is how many of these trials can be considered independent. It means, how many indices t act as a group where all values of t in the group give either $h \in QR_p$ or $h \in QNR_p$. We alreasy know that at least four indices t act as a group: each x is given by two values of t and $\pm x$ map to $\pm h$, which are either both in QR_p or both in QNR_p when $-1 \in QR_p$. There can be more indices t acting as a group.

Let us start from the case where more than one x map to the same $h \equiv x(x^2 - c)$. If so, then we can find x_1 and x_2, $x_1 - x_2 \not\equiv 0$, such that

$$x_1^3 - cx_1 \equiv x_2^3 - cx_2.$$

Then

$$x_1^2 + x_1x^2 + x_2^2 \equiv c$$

i.e.,

$$(2x_1 + x_2)^2 + 3x_2^2 \equiv 4c. \tag{15}$$

If there are solutions $x_1 \not\equiv x_2$ to (15), we say that x_2 belongs to the case of multiples. If x_2 is the only value of x that gives t, then we say that x_2 belongs to the case of singletons. We have to look separately at the cases $c = 1$ and $c = g$.

Case $c = 1$. Let us write

$$z \equiv \pm(2x_1 + x_2)$$

and

$$2^2 - z^2 \equiv 3x_2^2.$$

Then there exists t such that

$$2 + z \equiv 3tx_2 \ , \ 2 - z \equiv t^{-1}x_2.$$

Solving these equations yields

$$x_2 \equiv 4(3t + t^{-1})^{-1} \ , z \equiv 2^{-1}x_2(3t - t^{-1}).$$

The value of $x_{1\pm}$ is then

$$x_{1\pm} = 2^{-1}(z - x_2) \equiv 2^{-2}x_2(\pm(3t - t^{-1}) - 2).$$

We can write

$$h \equiv x_2(x_2^2 - c) \equiv 4(3t + t^{-1})^{-2}(3t + t^{-1})^{-1}(16 - (3t + t^{-1})^2). \tag{16}$$

Let us do the former general consideration with two running indices t and s explicitly, just for clarity. We take another running index s that assures that the term $16 - (3t + t^{-1})^2$ is a square. As we assume that $-1 \in QR_p$ in this lemma, we can write the square as $-r^2$ for some r. Thus

$$16 - (3t + t^{-1})^2 \equiv -r^2 \tag{17}$$

which yields

$$(3t + t^{-1})^2 - r^2 \equiv 16.$$

There exists s such that

$$3t + t^{-1} + r \equiv 4s \ , \ 3t + t^{-1} - r \equiv 4s^{-1}.$$

Solving gives

$$3t + t^{-1} \equiv 2(s + s^{-1}). \tag{18}$$

If s loops over all numbers in Z_p^* and t satisfies (18), then the term (17) is a square in (16) and

$$h \equiv x_2(x_2^2 - c) \equiv m^2(3t + t^{-1})^{-1}$$

where m^2 is a square:

$$m^2 \equiv 4(3t + t^{-1})^{-2}(16 - (3t + t^{-1})^2) \equiv -4(s + s^{-1})^{-2}(s - s^{-1})^2.$$

Writing $3t + t^{-1}$ as $2(s + s^{-1})$ gives

$$h \equiv x_2(x_2^2 - c) \equiv m^2 2(s + s^{-1})^{-1}$$

showing very clearly that $h \in QR_p$ if and only if $2(s + s^{-1}) \in QR_p$.

If $16 - (3t + t^{-1})^2$ is not a square, then $3t + t^{-1}$ belongs to the numbers that are not reached by $s + s^{-1}$. This set of numbers is (about) half of all numbers in Z_p^*. If $3t + t^{-1}$ is in that set and if $(3t + t^{-1}) \in QNR_p$, then from (9) follows that

$$b \equiv (3t + t^{-1})^{-1}(16 - (3t + t^{-1})^2)$$

is in QR_p and

$$h \equiv x_2(x_2^2 - c) \equiv 4(3t + t^{-1})^{-2}b$$

is in QR_p. This explicit calculation agrees with the general consideration that $h \in QR_p$ for (about) half of the indices t, but there is some probabilistic variation, the trials of t can be seen as probabilistic trials.

In the case of multiplets when $c = 1$ we may have to discard a few cases of the running index t: if we do not then some cases may do not give three different values of x, or any values at all. As $-1 \in QR_p$ is assumed in the lemma, there exists ϵ such that $\epsilon^2 \equiv -1$. If $3 \in QR_p$, then $3 \equiv \beta^2$ for some β. Then t values $\pm\epsilon\beta$ and $\pm\beta$ need to be discarded. If $3 \in QNR_p$, there are no solutions to $t^2 \equiv \pm 3$.

The number x_2 cannot be modulo zero because zero is not in QR_p or QNR_p. Thus, the roots of $3t + t^{-1} \equiv 0$ must be discarded. This means that if $3 \in QR_p$, two values of t must be discarded.

There are cases when one of the three x values equals one of the two. If $x_{1+} \equiv x_{1-}$, then $z \equiv 0$. This means that $3t - t^{-1} \equiv 0$, i.e., $t^2 \equiv 3^{-1}$. For $3 \in QNR_p$ this cannot happen, but for $3 \in QR_p$ there are two values of t filling this condition. If so, we get only two different values of x.

If $x_2 \equiv x_{i\pm}$, then $\pm 6 \equiv 3t - t^{-1}$, i.e., $(t \pm 1)^2 \equiv 3^{-1} + 1$. If $3^{-1} + 1 \in QR_p$, there are four values of t that give only two different values of x.

Solutions $x \equiv \pm 1$ give $x^2 - c \equiv 0$ and thus $h \equiv 0$. We have to discard four values of t that give $x \equiv \pm 1$.

Counting the numbers m and m_1 gives:

$$p - m = p - 1 - 2 - 2 - 4 - 4 = p - 13 \text{ if } 3 \in QR_p \text{ and } 3^{-1} + 1 \in QR_p$$

$$p - m = p - 1 - 4 - 4 = p - 9 \text{ if } 3 \in QNR_p \text{ and } 3^{-1} + 1 \in QR_p$$

$$p - m = p - 1 - 4 - 4 = p - 9 \text{ if } 3 \in QR_p \text{ and } 3^{-1} + 1 \in QNR_p$$

$$p - m = p - 1 - 4 = p - 5 \text{ if } 3 \in QNR_p \text{ and } 3^{-1} + 1 \in QNR_p$$

$$m_1 = 6 \text{ if } 3 \in QR_p \text{ and } 3^{-1}+1 \in QR_p$$

$$m_1 = 4 \text{ if } 3 \in QNR_p \text{ and } 3^{-1}+1 \in QR_p$$

$$m_1 = 2 \text{ if } 3 \in QR_p \text{ and } 3^{-1}+1 \in QNR_p$$

$$m_1 = 0 \text{ if } 3 \in QNR_p \text{ and } 3^{-1}+1 \in QNR_p$$

For singletons Lemma 3 gives a calculation of n_1 where the looping index t goes from $t = 1$ to $t = p - 1$ and gives each x two times. Exactly half of the numbers t give multiplets and are first removed. The remaining (exact) half of the numbers t give singletons. The cases to be excluded are already in the multiplet numbers. Thus, singletons give exactly $(p-1)/2$ values of x. The values $\pm x$ yield the same h and two values of y give $h \equiv y^2 \in QR_p$. Thus, eight values of the running index t map as a group. This is shown in Lemma 3 where $N_p = 8p+7$ for Case 1. About half of the groups map to $h \in QR_p$ and about half to $h \in QNR_p$.

Case $c = g$. For x_2 belonging to the case of multiplets (15) gives the equation

$$(2x_1 + x_2)^2 + 3x_2^2 \equiv 4g. \qquad (19)$$

We split the analysis into two cases:

Case $3 \in QNR_p$. Then we write $z^2 \equiv (2x_1 + x_2)^2$ and thus

$$g^{-1}z^2 \equiv 2^2 - (\eta x_2)^2$$

where $\eta^2 \equiv 3g^{-1}$. Then

$$2 + \eta x_2 \equiv g^{-1}zt \ , \ 2 - \eta x_2 \equiv zt^{-1}.$$

Thus $4 \equiv z(g^{-1}t + t^{-1})$ and $x_1 \equiv 2^{-1}(\pm z - x_2)$, i.e.,

$$x_2 \equiv (2\eta)^{-1}4(g^{-1}t + t^{-1})^{-1}(g^{-1}t - t^{-1}).$$

There are two values of t that map to the same x_2, namely t and $-t^{-1}$, solved as before. The third value t^{-1} maps to $-x_2$.

There are no values of t giving $g^{-1}t + t^{-1} \equiv 0$ which would yield no z or no x_2 and would have to be discarded. Thus, all values of t yield x_2 and $x_{1\pm}$. There are no values of t for which $x_2^2 - g \equiv 0$ and consequently $h \equiv 0$. Thus, all values of t yield a valid h. There are no values of t for which $z \equiv 0$ and $x_{1+} \equiv x_{1-}$. All these special cases are missing if $c = g$ and $3 \in QNR_p$, but there is still one special case left:

If $x_2 \equiv x_{1\pm}$ then $3x_2 \equiv \pm z$ and solving it gives

$$t^2 \pm 2\eta 3^{-1} g t - g \equiv 0.$$

Inserting $3g^{-1} \equiv \eta^2$ yields

$$t^2 \pm 2t + 1 \equiv (t \pm 1)^2 \equiv g + 1.$$

If $g + 1 \in QR_p$ there are four indices of t that give solutions where $x_2 \equiv x_{1+}$ or $x_2 \equiv x_{1-}$. The other values of t give three different values of x. If $g + 1 \in QR_p$ we get $m = 5$ and $m_1 = 4$. If $g + 1 \in QNR_p$ then $m = 1$, $m_1 = 0$.

Case $g = 1$ and $3 \in QR_p$. Then there exists β such that $\beta^2 \equiv 3$ and since we assume that $-1 \in QR_p$ there exists ϵ such that $\epsilon^2 \equiv -1$. Then $-3 \equiv (\epsilon\beta)^2$ and (12) can be written as

$$z^2 - (\epsilon\beta x_2)^2 \equiv 4g$$

where

$$z \equiv \pm 2x_1 + x_2.$$

There exists t such that

$$z + \epsilon\beta x_2 \equiv 2gt \ , z - \epsilon\beta x_2 \equiv 2t^{-1}.$$

Then

$$z \equiv gt + t^{-1} \ , \ x_2 \equiv (\epsilon\beta)^{-1}(gt - t^{-1}).$$

Thus

$$x_1 \equiv 2^{-1}(\pm z - x_2) \equiv 2^{-1}(\pm(gt + t^{-1}) - (\epsilon\beta)^{-1}(gt - t^{-1})).$$

There are three values of x that map to the same $h = x(x^2 - g)$, except for in possible special cases.

Here $gt - t^{-1} \not\equiv 0$ and $gt + t^{-1} \not\equiv 0$ as $-1 \in QR_p$. Therefore z and x_2 are always in Z_p^*. It means that $x_{1+} \not\equiv x_{1-}$. It is also not possible that $x^2 - g \equiv 0$.

The only special case that can appear here is that $x_2 \equiv x_{1\pm}$. If so, then $x_1 = 2^{-1}(\pm z - x_2) = x_2$. Thus $3x_2 = \pm z$. We get the equation

$$3(\epsilon\beta)^{-1}(gt - t^{-1}) \equiv \pm(gt + t^{-1})$$

which simplifies to

$$3(\epsilon\beta)^{-1}(gt - t^{-1}) \equiv \pm(gt + t^{-1}).$$

This leads to

$$t^2 \equiv g^{-1}(1 \pm 3(\epsilon\beta)^{-1})^{-1}(1 \mp 3(\epsilon\beta)^{-1})$$

$$t^2 \equiv g^{-1}4(1 \pm 3(\epsilon\beta)^{-1})^{-2}.$$

Clearly, there are no solutions t to this equation.

The result is that if $c = g$ and $3 \in QR_p$ (and $-1 \in QR_p$ as assumed in the lemma), then $m = 1$ and $m_1 = 0$.

The proof of Lemma 5 is complete. □

Let us give two examples of Lemma 5. For $p = 13$ we get for the running index $t = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12$ the following sequences $x_2 = 1, -, 6, 12, 3, 10, 3, 10, 1, 7, -, 12$, $x_{1+} = 0, -, 10, 0, 7, 6, 3, 10, 12, 3, -, 1$, $x_{1-} = 12, -, 10, 1, 3, 10, 7, 6, 0, 3, -, 0$, $h \equiv x_2(x_2^2 - 1) = 0, -, 2, 0, 11, 2, 11, 2, 0, 11, -, 0$.

The two values of t that do not give x_2 are 2 and $-2 = 12$. They are solutions to $3t + t^{-1} \equiv 0$, i.e., $t^2 \equiv -3^{-1} \equiv -9 \equiv 4$, thus $t \equiv \pm 2$. There are two values of t giving $x_{1+} \equiv x_{1-}$. They are solutions to $t^2 \equiv 3^{-1} \equiv 9$, i.e., $t = 3, 10$. We get these four special values of t because $3 \equiv 2^4 \in QR_{13}$.

The four values of t that give $h \equiv x_2(x_2^2 - 1)$ are values $t = \pm 1, \pm 4$. They give $x_2 \equiv \pm 1$ and thus $x_2^2 - 1 \equiv 0$. These special values of t appear because $c = 1$. For $c = g$ there are no such special values of t.

There are four values of t giving $x_2 \equiv x_{1\pm}$. They are solutions to $(t \pm 1)^2 \equiv 3^{-1} + 1 \equiv 10$. Thus, $t \pm 1 = \pm 6$. That yields $t = 5, 6, 7, 8$. We get these special values of t because $3^{-1} + 1 \equiv 10 = 2^{10} \in QR_{13}$.

The sum of the special values of t is $4 + 4 + 4 = 12$. This is subtracted from $p - 1 = 12$ and the result is zero showing that there are no cases when three different values of x map to the same h for $p = 13$.

As the second example consider $p = 29$. There are four groups of triplet x values that map to the same h:

$x = 5, 10, 14 = -24, -19, -15$ map to 4, and

$x = 19, 24, 15 = -10, -5, -14$ map to $25 \equiv -4$. While

$x = 12, 20, 26 = -17, -9, -3$ map to 5, and

$x = 3, 9, 17 = -26, -20, -12$ map to $24 \equiv -5$.

These 12 values of x need 24 values of the running index t. Additionally we have two values of x that give $h \equiv 0$, that is, $x \equiv \pm 1$. These need four values of t. This means that all $p - 1 = 28$ values of t are already used in the case of multiples and there cannot be any more special values of t. This is indeed true since $3 \in QNR_{29}$ and $3^{-1} + 1 \equiv 11 \in QNR_{29}$. Therefore $m = 5$, as Lemma 5 says, and the number of triplets of values of x mapping to the same h is $(p - 5)/6 = 4$. Both $h \equiv \pm 5$ are in QR_{29} and both $h \equiv \pm 4$ are in QR_{29}. The expectation value is that $(p - 5)/24 = 1$ are in QR_{29}. We got one more in $p = 29$: there is random variation around the mean.

There are exactly $14 = (p-1)/2$ singletons:

$$x = 2, 4, 6, 7, 8, 11, 18, 21, 23, 24, 25, 27, 13, 16$$

map to

$$h = 6, 2, 7, 17, 11, 15, 14, 18, 22, 12, 27, 23, 9, 20.$$

Of these 14 singleton h values six are in QR_{29}, namely $6, 7, 22, 23, 9, 20$. It is about half of $14 = (p-1)/2$ numbers in QR_{13}. It is not exactly half, there is random variation around the half.

Let us formulate an assumption that allows us to estimate averages and variances of the numbers a_p.

Statistical Assumption. *Let the numbers h be defined as*

$$h \equiv s + s^{-1} \pmod{p_j} \;\; s = 1, \ldots, p_j - 1$$

and let n_j be the number of s that give $h \in QR_{p_j}$. Then $n_j/2$ is binomially distributed with the mean at $(p_j - 1)/2$ and the variance $(p_j - 1)/4$. Notice that $s + s^{-1}$ gives each value h exactly twice.

Lemma 6. *Let $a_{p_j} = p_j - N_{p_j}$ and N_{p_j} be the number of solutions to*

$$y^2 = x^3 + ax + b. \tag{20}$$

The following claims hold assuming that the Statistical Assumption holds.

(i) In Cases 1 and 2 of Lemma 3 the expectation values of a_{p_j} are 2 and -2 respectively when p_j ranges over all primes. For Cases 3 and 4 the expectation value of a_{p_j} is zero.

(ii) The variance of a_{p_j} is $2p_j$ when p_j ranges over all primes.

(iii) The statistical bound by the standard deviation is: $|a_{p_j}| \le 2p_j^{\frac{1}{2}}$.

Proof. By the Statistical Assumption the expectation value of n_1 in Lemma 3 for Cases 1 and 2 is $E[n_1] = \frac{p-5}{8}$. Thus $E[a_p] = p - (p - 5 + 7) = -2$ for Case 1 and $E[a_p] = p - (2p - p + 5 - 7) = 2$ for Case 2. In the Cases 3 and 4 of Lemma 3 $E[n_2] = \frac{p-3}{8}$. Thus $E[a_p] = p - (p - 3 + 3) = 0$ for Case 3 and $E[a_p] = p - (2p - p + 3 - 3) = 0$ for Case 4.

The expressions of N_p in Lemma 3 are not convenient for calculating variances needed in (ii) as Lemma 3 does not count how many times the running index hits the same x. We have to use the forms in Lemma 5.

Lemma 5 lists numbers $m \le 13$ and $m_1 \le 6$, which are different for different cases of multiple x values. These few special cases are not important when estimating the variance of a_p when p ranges over all primes. For that reason we assume that p is so large that $13 << p$ and all values of t in the case of multiples give a value h that comes from three values of x.

Case 1: $b = 0$. Lemma 4 shows that $-1 \in QNR_{p_j}$ in half of the cases of p_j. By Lemma 1 $a_{p_j} = 0$ in these cases. In the remaining half of the cases $-1 \in QR_{pj}$. Exactly half of them are singleton cases of Lemma 5 and exactly half are multiplet cases of Lemma 5.

In Lemma 5 is used a running index $t = 1, .., p - 1$. In singleton cases two values of t map to the same x. Only one x maps to a given $h = x(x^2 - c)$, $c = 1$ or $c = g$. Both $h = x(x^2 - c)$ and $-h$ either belong to QR_{p_j} or both belong to QNR_{p_j}. For each $h \in QR_{p_j}$ there are two values y, (i.e., $\pm y$) that give the same $y^2 \equiv h$. Thus, eight values of t map together as a unit.

In multiplet cases two values of t map to the same x. Ignoring the $m \le 13$ special cases, three values of x map to a given h. Both $h = x(x^2 - c)$ and $-h$ either belong to QR_{p_j} or both belong to QNR_{p_j}. For each $h \in QR_{p_j}$ there are two values y, (i.e., $\pm y$) that give the same $y^2 \equiv h$. Thus, twenty four values of t map together as a unit. There is no difference between the cases $c = 1$ and $c = g$ except for in the small numbers m and m_1 in Lemma 5.

By the Statistical Assumption we can treat the situation as trials in a binomial distribution. In the binomial distribution there are n trials and the success probability is p and $q = 1 - p$. Then the expectation value is $E[$ number of successes $] = np$ and the variance is npq. Here we have p_j/k independent trials, where for singletons $k = 8$ and for multiplets $k = 24$. Each success gives k units and the probability $p = 0.5$. Thus, the average np must give p_j as $a_{p_j} = p_j - N_{p_j}$ should have the average at zero, or close to zero, for us to use the binomial distribution assumption. We have a better match to the binomial distribution if we count not (x, y) but (x, y^2) cases. Then the expected number of successes from p_j trials is $p_j/2 = p_j \cdot 0.5 = p_j/2$. In the binomial distribution the variance would be $p_j \cdot 0.5 \cdot 0.5 = p_j/4$ as $p = q = 0.5$, but in our distribution $k/2$ units of success are in a group. The number of independent trials is $p_j/(k/2)$ and the unit size is multiplied by $k/2$. The effect of this grouping is that the variance is multiplied by $k/2$, i.e., the variance $\sigma^2 = \frac{1}{n}\sum(y_i - \bar{y}_i)^2$ is multiplied by a if n is chaged to n/a and y_i changed to ay_i.

The variance of the singleton cases is therefore $(p_j/4) \cdot (8/2) = p_j$ and the variance of the multiplet case is $(p_j/4) \cdot (24/2) = 3p_j$. Singletons represent one fourth of all cases and multiplets represent one fourth of all cases, while $-1 \in QNR_{p_j}$ are half of all cases. Adding the cases with their probabilities of occurance gives

$$\sigma_{y^2}^2 = \frac{1}{2} \cdot 0 + \frac{1}{4} \cdot p_j + \frac{1}{4} \cdot 3p_j = p_j.$$

Multiplying the result by 2 to count $\pm y$ instead of y^2 gives

$$\sigma = 2p_j.$$

Case $b \neq 0$. In this case the addition of b removes the condition that guaranteed that $a_{p_j} = 0$ if $-1 \in QNR_{p_j}$. For the case $-1 \in QR_{p_j}$ this same condition guaranteed that x and $-x$ both give h and $-h$ in QR_{p_j} or both give h and $-h$ in QNR_{p_j}. This condition is removed from both cases when $b \neq 0$ because if

$h_1' - b \equiv h$ and $h_2' - b \equiv -h$, then the relation between h_1' and $h_2' \equiv -h_1' + 2b$ does not say anything of $h_2' \in QR_{p_j}$ if $h_1' \in QR_{p_j}$.

In other ways the case $b \neq 0$ does not differ. Having $b \neq 0$ does not even change the special cases of t. It only removes the condition that was discusses above. Exactly half of the cases are singletons and exactly half of the cases are multiplets. The effect of removing the coupling of x and $-x$ causes that in the singleton case $k = 4$ and in the multiplet case $k = 12$. As a compensation, there is no case of $-1 \in QNR_{p_j}$ when there is zero variation of a_{p_j}. The case where $-1 \in QNR_{p_j}$ is the same as the case where $-1 \in QR_{p_j}$. These cases give the same variance and the total variance is

$$\sigma_{y^2}^2 = \frac{1}{2}(\frac{1}{2}p_j + \frac{1}{2}3p_j) + \frac{1}{2}(\frac{1}{2}p_j + \frac{1}{2}3p_j) = p_j$$

Counting the solutions (x, y) instead of (x, y^2) gives

$$\sigma = 2p_j.$$

The statistical bound (iii) is a bound by standard deviation and it is known as the Hasse bound. The average variance over all p_j is $2p_j$, but if $b = 0$, then for the part $-1 \in QR_{p_j}$ the variance is $4p_j$. Thus the standard deviation for these p_j is $2p_j^{\frac{1}{2}}$.

The proof of Lemma 6 is finished. □

8 On the zeros of the Taylor series

Lemma 7. *Let $\phi(s)$ have a Taylor series at s_0 of the form*

$$\phi(s) = C_0 + C_r(s - s_0)^r + C_{r+1}(s - s_0)^{r+1} \cdots$$

Let

$$h(s) = \frac{d}{ds} \ln \phi(s).$$

Then

$$h(s) = \frac{rC_r}{C_0}(s - s_0)^{r-1} + O((s - s_0)^r) \qquad \text{if} \quad C_0 \neq 0$$

$$h(s) = \frac{r}{s - s_0} + O(1) \qquad \text{if} \quad C_0 = 0.$$

Proof. The claim comes directly from calculating

$$\phi'(s) = rC_r(s - s_0)^{r-1} + (r - 1)C_{r+1}(s - s_0)^r \cdots$$

$$h(s) = \frac{d}{ds} \ln \phi(s) = \frac{\phi'(s)}{\phi(s)}.$$

□

Why Lemma 7 is written down here is that it may not be obvious that the only singularity $h(s)$ can have at $s = s_0$ is a first order pole. The algebraic rank of $\phi(s)$ is a multiplier (i.e., residue) in the divergent part of $h(s)$. If the residue of $h(s)$ negative, then the function $\phi(s)$ has a pole of the order ot the residue.

Lemma 8. *Consider an infinite product*

$$\phi(s) = \prod_{j \in A} (1 - f_j(s))^{-1} \tag{21}$$

where A is an infinite subset of $\mathbb{N}$. Let us assume there is $\alpha > 0$ and $C > 0$ such that

$$|f_j(s)| < Cp_j^{-\alpha x} \tag{22}$$

and

$$|f_j'(s)| < C \ln(p_j) p_j^{-\alpha x}. \tag{23}$$

Here $s = x + iy$ and $f_j'(s)$ denotes the derivative of $f_j(s)$. The following claims hold:

(i) The function $\phi(s)$ is finite and nonzero if $Re\{s\} > \frac{1}{2\alpha}$.

(ii) Let s_0 satisfy $\frac{1}{3\alpha} < Re\{s_0\} \leq \frac{1}{2\alpha}$. If the function

$$g_1(s) = \sum_{j \in A} f_j'(s) f_j(s) \tag{24}$$

is finite at s_0, then

$$h(s) = \sum_{j \in A} f_j'(s). \tag{25}$$

is finite at s_0 and $\phi(s_0)$ is finite and nonzero.

Proof. From the assumption (22) follows that the infinite product (21) is absolutely convergent if $Re\{s\} = x > \alpha^{-1}$. We expand

$$\frac{d}{ds} \ln \phi(s) = \sum_{j \in A} \left(f_j'(s) \sum_{k=0}^{\infty} f_j(s)^k \right). \tag{26}$$

By (22) $|f_j(s)| < C p_j^{\alpha x}$. Because

$$\sum_{j=1}^{\infty} p_j^{-x}$$

is absolutely convergent if $x > 1$, the derivative of the series

$$\sum_{j=1}^{\infty} \ln(p_j) p_j^{-x}$$

is absolutely convergent if $x > 1$. By (23) holds

$$|f_j'(s)| < C \ln(p_j) p_j^{-\alpha x}.$$

Thus the series

$$\sum_{j \in A} f_j'(s)$$

is absolutely convergent if $Re\{s\} > \alpha^{-1}$.

In the right side series of (26) the terms for each k converge absolutely if $Re\{s\} > \frac{1}{(k+1)\alpha}$. The terms for $k > 0$ define an analytic function

$$g(s) = \sum_{j\in A} \left(f_j'(s) \sum_{k=1}^{\infty} f_j(s)^k \right)$$

This function is finite if $Re\{s\} > \frac{1}{2\alpha}$. The function

$$g_1(s) = \sum_{j\in A} f_j'(s) f_j(s)$$

in (11) is the first part of this series. The other parts converge if $Re\{s\} > \frac{1}{3\alpha}$. The function $g_1(s)$ can be analytically continued to the area $\frac{1}{2\alpha} > Re\{s\} \geq \frac{1}{3\alpha}$. The continuation is analytic with the exception of possible isolated singularities, poles. The series for $k = 0$ converges if $Re\{s\} > \frac{1}{\alpha}$ and defines an analytic function

$$h(s) = \sum_{j\in A} f_j'(s).$$

This function can be analytically continued to $Re\{s\} > \frac{1}{3\alpha}$. The continuation is analytic with the exception of possible isolated singularities, poles.

Whether $h(s)$ or $g_1(s)$ is infinite in a given point s or not, we can formally write

$$\frac{d}{ds} \ln \phi(s) = \phi(s)^{-1} \frac{d}{ds} \phi(s) = h(s) + g(s)$$

$$\frac{d}{ds} \phi(s) = \phi'(s) = h(s)\phi(s) + g(s)\phi(s)$$

Assume $\phi(s_0) = 0$ for $s_0 = x_0 + iy_0$, $\frac{1}{\alpha} > x_0 \geq \frac{1}{2\alpha}$. The function $g(s)$ is finite at s_0, thus $g(s_0)\phi(s_0) = 0$. If $h(s)$ is finite at s_0, then $h(s_0)\phi(s_0) = 0$ and $\phi'(s_0) = 0$. All derivatives of $g(s)$ are finite at s_0.

If $h(s)$ is finite at s_0, then all derivatives of $h(s)$ are finite at s_0. If so, then by induction all derivatives of $\phi(s)$ are zero at s_0: assuming that it is proved that $\frac{d^j}{ds^j}\phi(s)$ for $j < n$, then $\frac{d^n}{ds^n}\phi(s)$ is given by a sum where every term is a finite value from derivatives of $g(s)$ or $h(s)$ multiplied by zero, a derivative $j \geq 0$ of $\phi(s)$ at s_0. If all derivatives of $\phi(s)$ are zero at s_0, then $\phi(s)$ is zero everywhere. This is a contradiction as $\phi(s)$ is described by an infinite product (21) that converges and is not zero when $Rs\{s\} > \alpha^{-1}$. It follows that $\phi(s_0)$ cannot be zero if $h(s_0)$ is finite.

Next, consider the infinite product of two complex variables s and z

$$\psi(s,z) = \prod_{j\in A} (1 - f_j(s) + f_j(s+z))^{-1}. \tag{27}$$

The infinite product is absolutely convergent if $Re\{s\} > \alpha^{-1}$. Let us expand

$$\frac{\partial}{\partial z} \ln \psi(s,z) = -\sum_{j\in A} \left(f_j'(s+z) \sum_{k=0}^{\infty} (f_j(s) - f_j(s+z))^k \right) \tag{28}$$

The terms $k > 0$ define a function of two complex variables

$$u(s,z) = -\sum_{j\in A} \left(f_j'(s+z) \sum_{k=1}^{\infty} (f_j(s) - f_j(s+z))^k \right)$$

This function is finite if $Re\{s\} > \frac{1}{2\alpha}$.

The series for $k = 0$ converges absolutely if $Re\{s\} > \frac{1}{\alpha}$ and defines a complex analytic function

$$h(s+z) = \sum_{j\in A} f_j'(s+z).$$

This function is the same function $h(s)$ as earlier. Now it only has the variable $s+z$ instead of s. We assume it is already continued analytically to $Re\{s\} > \frac{1}{3\alpha}$. It may have isolated poles when $Re\{s\} > \frac{1}{\alpha}$.

We can write

$$\frac{\partial}{\partial z} \ln \psi(s,z) = \psi(s,z)^{-1} \frac{\partial}{\partial s} \psi(s,z) = h(s+z) + u(s,z)$$

$$\frac{\partial}{\partial z} \psi(s,z) = h(s+z)\psi(s,z) + u(s,z)\psi(s,z) \tag{29}$$

The functions $g(s)$ and $u(s,z)$ are analytic (implying finite) at an open environment of the point $s = s_0$ because $\frac{1}{\alpha} \geq Re\{s\} > \frac{1}{2\alpha}$.

The function $\psi(s, s_0 - s)$ is an analytic function of one complex variable s and $\psi(s_0, 0) = 1$. The function $\psi(s, s_0 - s)$ cannot be zero in every point in an open environment of s_0 because else it is zero everywhere. Likewise, $\psi(s, s_0 - s)$ cannot be infinite in every point in an open environment of s_0 because then $\psi(s, s_0 - s)^{-1}$ is zero everywhere. It follows that we find $z \neq 0$ such that $|z| << 1$, $s_0 = s + z$ and $\psi(s, s_0 - s) \neq 0$. Let this point be (s_2, z_2). It follows that $s_2 + z_2 = s_0$. Keeping s fixed at s_2 and letting z vary, $\psi(s_2, z)$ defines a complex analytic function of z. As $\psi(s_2, z_2)$ is finite, the partial derivative of $\psi(s_2, z)$ with respect to z at $z = z_2$ is finite. Thus, the left side of (29) is finite. The right side contains a finite term $u(s_2, z_2)\psi(s_2, z_2)$ and $\psi(s_2, z_2)$ is finite and nonzero. It follows that $h(s_2 + z_2) = h(s_0)$ is finite. Because $h(s_0)$ is finite $\phi(s_0)$ is nonzero. The first claim (i) of the lemma is proven.

In the second claim (ii) of the lemma we select a point $s = s_0$ such that $\frac{1}{2\alpha} \geq Re\{s\} > \frac{1}{3\alpha}$. The function $g(s) - g_1(s)$ is finite in this area because the series defining this function converges absolutely. The function $u(s,z)$ converges with the exception of the first term. Thus

$$u_2(s,z) = -\sum_{j \in A} \left(f_j'(s+z) \sum_{k=2}^{\infty} (f_j(s) - f_j(s+z))^k \right)$$

is finite in this area. In the first term

$$u_1(s,z) = -\sum_{j \in A} f_j'(s+z)(f_j(s) - f_j(s+z))$$

we insert $z = s_0 - s$ and get a complex analytic function of one variable s. This function can be continued analytically to the area $\frac{1}{2\alpha} \geq Re\{s\} > \frac{1}{3\alpha}$. If $g_1(s_0)$ is finite, then also $u_1(s_0, s_0 - s_0) = u_1(s_0, 0)$ is finite. The argument that $h(s_0)$ is finite goes in the same way as in the case where $\frac{1}{2\alpha} < Re\{s\} \leq \frac{1}{\alpha}$. This proves the second claim. □

9 Application to $L(C, s)$

Lemma 9. *Let* $a_{p_j} = p_j - N_{p_j}$ *and let* N_p *be the number of solutions to*

$$y^2 \equiv x^3 - d^2x \quad (\mod p).$$

(i) The function $g_1(s)$ with the series expression

$$g_1(s) = \sum_{j \in A} \ln(p_j) a_{p_j}^2 p_j^{-2s}$$

in the area where the series converges does not depend on d.

(ii) If $d = k^2$ the function $h(s)$ with the series expression

$$h(s) = \sum_{j \in A} \ln(p_j) a_{p_j} p_j^{-s}$$

in the area where the series converges does not depend on k.

Proof. From Lemma 1 follows that if $-1 \in QNR_{p_j}$ then $a_{p_j} = 0$. If $-1 \in QR_{p_j}$ Lemma 3 gives two possible values for a_{p_j}:

$$a_{p_j} = p_j - N_{p_j} = p_j - 8n_{1,j} - 7 \text{ if } d \in QR_{p_j}$$

$$a_{p_j} = p_j - N_{p_j} = -p_j + 8n_{1,j} + 7 \text{ if } d \in QNR_{p_j}$$

where $n_{1,j}$ is the number of solutions $y^4 \in Z_{p_j}^*$ yielding $y^4 - 1 \in QR_{p_j}$. The number $n_{1,j}$ does not depend on d, thus the square of either of s_{p_j} does not depend on d. It follows that $g_1(s)$ does not depend on d.

In $h(s)$ the number d is a square k^2. Therefore d is a square in every $Z_{p_j}^*$. It follows that for every p_j the case in Lemma 3 is always Case 1. The value

$N_{p_j} = 8n_{1,j} + 7$ for every p_j. Thus, the function $h_0(s)$ does not depend on what square number $d = k^2$ is used: in Lemma 2 the number $a = -d^2$ is removed at the beginning by a substitution and $c = 1$ in Case 1 of Lemma 2. □

The problem statement [1] says that the product (1) converges absolutely for $Re\{s\} > \frac{3}{2}$. This claim assumes the statistical Hasse bound $|a_{p_j}| < 2p_j^{\frac{1}{2}}$. Therefore this bound must be considered accepted in the context of the problem statement of [1]. The Hasse bound follows from the Statistical Assumption and implies that some similar statistical assumption is used in the bound.

Lemma 10. *If $\phi(s)$ is Lemma 8 is $L(C,s)$ as defined in (1), then the term of the infinite product in Lemma 8 is*

$$f_j(s) = a_{p_j}p_j^{-s} - p_j^{1-2s}.$$

Then

$$h(s) = \sum_{j\in A} f_j'(s) = h_1(s) + h_2(s)$$

where

$$h_1(s) = -\sum_{j\in A} a_{p_j}\ln(p_j)p_j^{-s}.$$

The function $h_2(s)$ is

$$h_2(s) = 2\sum_{j\in A}\ln(p_j)p_j^{1-2s}.$$

It diverges at $s = 1$ and has a simple pole

$$h_2(s) = \frac{1}{s-1} + \text{finite terms}$$

at $s = 1$. The function $g_1(s)$ we define as

$$g_1(s) = \sum_{j\in A} f_j'(s)f_j(s) = g_{1,1}(s) + g_{1,2}(s)$$

where

$$g_{1,1}(s) = \sum_{j \in A} \ln(p_j) a_{p_j}^2 p_j^{-2s}.$$

It is the the part of $g(s)$ that diverges if $s = 1$ and $g_{1,2}(s)$ and the other parts of $g(s)$ converge at $s = 1$.

Proof. A simple calculation shows that the forms of $h(s)$ and $g_1(s)$ are as in the claim of the lemma.

The pole of the function $h_2(s)$ at $s = 1$ is derived from the pole of the Riemann zeta function $\zeta(s)$ at $s = 1$. Zeta has a simple pole of residue 1 at $s = 1$. Thus close to $s = 1$ zeta is

$$\zeta(s) = \prod_{p_j} (1 - p_j^{-s})^{-1} = \frac{1}{s-1} + \text{finite terms.}$$

Derivating gives $h(s)$ for the zeta function:

$$\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{(s-1)^2} \cdot \frac{s-1}{1} + \text{finite terms} = h(s) + g(s).$$

Calculating from the infinite product we get

$$\frac{d}{ds} \ln(\zeta(s)) = h(s) + g(s)$$

where

$$h(s) = -\sum_j \ln(p_j) p_j^{-s}$$

$$g(s) = -\sum_j \ln(p_j) p_j^{-2s} + \cdots$$

and $g(s)$ converges when $Re\{s\} > \frac{1}{2}$. Thus, the divergent part is

$$\sum_j \ln(p_j) p_j^{-s} = \frac{1}{s-1} + \text{finite terms}$$

close to $s = 1$. Changing $-s = 1 - 2z$ gives

$$2\sum_j \ln(p_j)p_j^{1-2z} = 2\frac{1}{2z-2} + \text{finite terms}$$

that is

$$2\sum_j \ln(p_j)p_j^{1-2s} = \frac{1}{s-1} + \text{finite terms.}$$

In Lemma 6 we derived the Hasse bound $|a_{p_j}| \le 2p_j^{\frac{1}{2}}$. It is a statistical bound where a random variable is estimated by the standard deviation. This statistical bound can be used for investigation of the convergence of the $L(C,s)$ function because any looser bound $|a_{p_j}| \le 2p_j^{\frac{1}{2}+\alpha}$, $\alpha > 0$, on a general p_j can be improved by adding the probability that $|a_{p_j}|$ is higher than (a chosen multiple) of the standard deviation of $|a_{p_j}|$.

Lemma 8 shows that the bound on $< a_{p_j}|$ cannot be a stricter bound of the type $|a_{p_j}| \le p_j^{\frac{1}{2}-\alpha}$ where $\alpha > 0$ for a general p_j because if there is such a bound, then $g(s)$ converges at $s = 1$ and Lemma 8 shows that $L(C,s)$ is always finite and nonzero. Yet we know that $\ln(L(C,s))$ is infinite at $s = 1$ for some elliptic curves C.

From the Hasse bound it follows that $g_1(s)$ converges for $Re\{s\} > 1$. By Lemma 8 $g_1(s)$ must diverge at least for some C at $s = 1$ because there are elliptic curves C such that $L(C,1)$ is not finite and nonzero. The Hasse bound guarantees that $g(s) - g_1(s)$ converge absolutely if $Re\{s\} > \frac{1}{2}$ and that $g_{1,2}(s)$ also converges absolutely if $Re\{s\} > \frac{1}{2}$. The part that cannot be shown to converge is $g_{1,1}(s)$. It must diverge at $s = 1$ because $g_1(s)$ is the same function for all elliptic curves C of the type $y^2 = x(x^2 - d^2)$ (see Lemma 9) and for some values of d the L-function has a zero value at $s = 1$. If $g_{1,1}(s)$ would converge at $s = 1$, then $L(C,1)$ would be finite and nonzero for every d.

All claims of the lemma are proven.□

With these results we can discuss the Birch and Swinnerton-Dyer conjecture.

10 The Birch and Swinnerton-Dyer conjecture

The original reason why Birch and Swinnerton-Dyer formulated their conjecture was that the function

$$\log \prod_{j=1}^{n} \frac{N_{p_j}}{p_j}$$

was growing approximatively linearly for some values of d for elliptic curves of the type $y^2 = x(x^2 - d^2)$, while for some other values of d the function did not tend to infinity when n grows.

The function they studied in [2] can be written as

$$\log \prod_{j=1}^{n} \frac{N_{p_j}}{p_j} = -\log \prod_{j=1}^{n} (1 - a_{p_j} p_j^{-1})^{-1}$$

$$= \sum_{j=1}^{n} \log(1 - a_{p_j} p_j^{-1}) = -\sum_{j=1}^{n} a_{p_j} p_j^{-1} + \sum_{j=1}^{n} a_{p_j}^2 p_j^{-2} + \cdots$$

As a_{p_j} is on the range of $p_j^{\frac{1}{2}}$ by the Hasse bound, the higher terms converge. The first two terms may diverge at $s = 1$.

The function studied can be completed into a complex function that is evaluated at $s = 1$:

$$\prod_{j=1}^{\infty} (1 - a_{p_j} p_j^{-1})^{-1} \to \prod_{j=1}^{\infty} (1 - a_{p_j} p_j^{-s})^{-1}.$$

Let the sum to be over the set A

$$L_2(C, s) = \prod_{j \in A} (1 - a_{p_j} p_j^{-s})^{-1}.$$

The difference with this function and $L(C, s)$ in (1) is that the term p_j^{1-2s} in (1) is missing.

For $L_2(C, s)$ the function in Lemma 10 is $f_j(s) = a_{p_j} p^{-s}$ and $h(s)$ has the expression

$$h(s) = h_1(s) = \sum_{j \in A} f_j'(s) = -\sum_{j \in A} a_{p_j} \ln(p_j) p_j^{-s}.$$

The sum diverges at $s = 1$ and $h(s)$ may have a simple pole at $s = 1$. The function $g_1(s)$ is

$$g_1(s) = -\sum_{j\in A} \ln(p_j) a_{p_j}^2 p_j^{-2s}.$$

The part of the sum expression of $g_1(s)$ diverging at $s = 1$ is

$$g_{1,1}(s) = g_1(s).$$

The sum diverges at $s = 1$ and $g_{1,1}(s)$ may have a simple pole at $s = 1$.

For $L(C, s)$ the function $f_j(s) = a_{p_j} p^{-s} - p^{1-2s}$ and $h(s)$ has the expression

$$h(s) = \sum_{j\in A} f'_j(s) = -\sum_{j\in A}(a_{p_j} \ln(p_j)p_j^{-s} - 2\ln(p_j)p_j^{1-2s}) = h_1(s) + h_2(s).$$

The sum diverges at $s = 1$ and $h(s)$ may have a simple pole at $s = 1$. The function $g_1(s)$ is

$$g_1(s) = -\sum_{j\in A} \ln(p_j) a_{p_j}^2 p_j^{-2s} + \sum_{j\in A} \ln(p_j) a_{p_j} \left(3p_j^{1-3s} - 2p_j^{2-4s}\right).$$

The part of the sum expression of $g_1(s)$ diverging at $s = 1$ is

$$g_{1,1}(s) = \sum_{j\in A} \ln(p_j) a_{p_j}^2 p_j^{-2s}.$$

The sum diverges and $g_{1,1}(s)$ may have a simple pole at $s = 1$.

Thus, the (possibly) divergent part $g_{1,1}(s)$ of $g_1(s)$ is the same for $L_2(C, s)$ and for $L(C, s)$, but the function $h(s)$ for $L_2(C, s)$ lacks the second part $h_2(s)$ in $h(s)$ for $L(C, s)$. This missing function has a first order pole and residue one at $s = 1$ as the Lemma 10 shows.

The Theorem on page 4 in [1] states that it is proven that if the elliptic curve C has rank zero then the L-function $L(C, s)$ has the algebraic rank zero. There is a conflict: Theorem 1 proves that Theorem on page 4 in [1] is in contradiction with the initial experiments of Birch and Swinnerton-Dyer, assuming they have been correctly described in the literature:

Theorem 1. *Assuming that $L_2(C, s)$ satisfies the Birch and Swinnerton-Dyer conjecture for rank zero, then $L(C, s)$ cannot satisfy the conjecture for rank zero.*

Proof. Birch and Swinnerton-Dyer studied elliptic curves of the form $y^2 = x(x^2 - d^2)$ with several values of d including values $d = 1$ and $d = 5$. These curves have rank zero if $d = k^2$ for some integer k. Assuming that Birch and Swinnerton-Dyer concluded that the function they studied, $L_2(C, s)$, has a finite nonzero value at $s = 1$ for $d = 1$, then it follows that $h(s) + g_{1,1}(s)$ must be finite at $s = 1$ for $d = 1$. For $L_2(C, s)$ holds $h(s) = h_1(s)$, thus $h_1(s) + g_{1,1}(s)$ must be finite for $d = 1$.

Consequently, if $L_2(C, s)$ fills the conjecture for rank zero, then the function $L(C, s)$ cannot have a finite nonzero value at $s = 1$ for elliptic curves with $a = -d^2 = -1$ and $b = 0$. The function $h_2(s)$ has a pole at $s = 1$ and is of the form

$$h_2(s) = \frac{1}{s-1} + \text{finite terms.}$$

Therefore

$$h_1(s) + h_2(s) + g_{1,1}(s) = \frac{1}{s-1} + \text{finite terms.}$$

Because of this pole, the function $L(C, s)$ has a zero at $s = 1$. Thus, for an elliptic curve of rank zero the L-function $L(C, s)$ has a zero at $s = 1$ and has the algebraic rank one. □

Errors happen, but if Theorem on page 4 in [1] (and the literature results that it is based on) is correct, then Birch and Swinnerton-Dyer were a bit mixed up: they thought that a pole of $L_2(C, s)$ at $s = 1$ for $d = 1$ is a nonzero value of $L_2(C, s)$ and that a nonzero value of $L_2(C, s)$ at $s = 1$ for $d = 5$ is a zero.

Though it seems that the problem as stated in [1] is solved in negative without any use of Lemmas 1-4 simply because it contradicts what Birch and Swinnerton-Dyer found in [2], these lemmas were given for the purpose of solving the problem statement in [1]. Thus, we will assume that the problem statement [1] is correct in its claims and show that it leads to a contradiction without referring to the results of Birch and Swinnerton-Dyer in [2].

The problem statement [1] says that $L(C, s)$ is analytic in the whole plane implying that it does not have a pole at $s = 1$ for any d. Theorem on page 4 in [1] says that for rank one and zero the Birch and Swinnerton-Dyer conjecture is true.

The function $h_2(s)$ has a pole with residue 1 at $s = 1$.

Let us consider the values $a = -d^2$ with $d = 1$, $d = 5$ and $d = 19$ for the simple elliptic curves that Birch and Swinnerton-Dyer studied: thus $a = -d^2$ and $b = 0$. The numbers 1 and 19 are noncongruent numbers and give an elliptic curve with zero rank, while $d = 5$ is a congruent number and gives an elliptic curve with rank one. I verified that 19 is a noncongruent number in [3] as it seems necessary to verify literature results in this field.

The value $d_1 = 1$ is a square number and picks up only the first case with $i = 0$ in Lemma 3. Thus, $a_p = N_p - p = 8n_1 + 7 - p$. The other two values of d give a choice of the two first cases in Lemma 3: if $d \in QR_p$, then $a_{p_j,d} = a_{p_j,1}$, while if $d \in QNR_p$, then $a_{p_j,d} = -a_{p_j,1}$.

The divergent part $g_{1,1}(s)$ of $g_1(s)$ has a square of a_{p_j} and therefore it is the same function for all three values of d. Let the residue of the (possible) pole of $g_{1,1}(s)$ at $s = 1$ be r. That is, if there is no pole, then $r = 0$.

The function $h_1(s)$ is different for different values of d. Let us write $h_{1,d}(s)$ for the function $h_1(s)$ for the value d. We will denote the value of the residue of the (possible) pole of $h_{1,d}(s)$ at $s = 1$ by $-r_d$. If there is no pole, then $r_d = 0$.

Assuming that the conjecture holds for ranks zero and one, the (possible) poles of $h(s) = h_{1,1}(s) + h_2(s)$ and $g_{1,1}(s)$ must cancel at $s = 1$. We have $-r_1 + 1 + r = 0$.

For $d = 5$ the function $h_{1,5}(s) + h_2(s) + g_{1,1}(s)$ has a simple pole with residue one. Thus $-r_5 + 1 + r = 1$. These two equations give

$$r_5 = r_1 - 1 < r_1. \tag{30).}$$

The value $d = 19$ is a noncongruent number. Since Theorem on page 4 in [1] says that the conjecture holds, the value $d = 19$ must give zero algebraic rank. Thus,

$$r_{19} = r_1. \tag{31}$$

Theorem 2. *If statistical arguments of the distribution of prime numbers are not allowed in the proof, then the problem statement in [1] is not well-defined and cannot be answered. If statistical arguments of the distribution of prime numbers are allowed in the proof, then the Birch and Swinnerton-Dyer conjecture for $L(C, s)$ fails in rank one.*

Proof. The Hesse bound is a statistical bound and requires treating prime numbers in a statistical manner. If statistical arguments are not allowed, then the problem statement should not use the Hesse bound. As it does make an argument with the Hesse bound, the problem statement is not well formulated and a poorly formulated problem statement cannot be answered.

No statistical arguments are needed in Lemmas 1-3 and Lemma 4 claim (i). In Lemma 4 claim (ii) there is a statistical argument using the Sieve of Eratosthenes. Prime numbers can be deterministically generated by the Sieve of Eratosthenes. Because of the construction of this sieve we can make some statistical observations, such as the claim (ii) in Lemma 4.

The number of solutions for $y^2 - 1 \in QR_p$ is $(p - 5)/4$ for the case $-1 \in QR_p$ according to Lemma 2. Exactly half of the numbers y are in QR_p and half are in QNR_p. We will now make a statistical argument: for a randomly chosen prime p (about) half of the solutions $y^2 - 1 \in QR_p$ have $y \in QR_p$ and if p_j ranges over all values, the probability that $y \in QR_p$ is exactly 0.5. Thus, the expectation value for n_1 is $(p-5)/8$ where the expectation value means that the number p is randomly chosen. It follows that the expectation value of a_{p_j} for $d = 1$ is $E[a_{p_j}] = p_j - 5 + 7 - p_j = 2$. Because the mean of a_{p_j} is 2, the function $h_{1,1}(s)$ for diverges at $s = 1$. It has a simple pole with residue $-r_1$ for some $r_1 > 0$.

We will make another statistical assumption: the prime d, where $d = 5$ or $d = 19$, does not have any special relationship with a randomly chosen large prime p_j. Therefore the expectation value of a_{p_j} is 2 if the randomly chosen prime p_j is chosen from the set satisfying $d \in QR_p$. Likewise, the expectation value of a_{p_j} is 2 if the randomly chosen prime p_j is chosen from the set satisfying $d \in QNR_p$.

We need another statistical assumption: because the prime d, where $d = 5$ or $d = 19$, does not have any special relationship with a randomly chosen large prime p_j, the probability that $d \in QR_p$ for a randomly chosen p_j is 0.5.

If these statistical assumptions are accepted in a proof, then we can make the following observations. Let $a_{p_j,d}$ denote the number a_{p_j} for the value d.

By Lemma 1, if $-1 \in QNR_p$, then $a_p = N_p - p = 0$ for every d.

By the claim (ii) in Lemma 4 holds $-1 \in QNR_p$ for half of the randomly chosen primes p_j. It follows that $a_{p_j,d}$ is zero for every d for half of the randomly chosen primes p_j.

For the other half of the randomly chosen primes p_j holds $-1 \in QR_{p_j}$. Of these values half have $d \in QR_{p_j}$ for $d \neq 1$ and then $a_{p_j,d} = a_{p_j,1}$. The rest (i.e., half) have $d \in QNR_p$ and thus $a_{p_j,d} = -a_{p_j,1}$. Let $d \in \{1, 5, 19\}$ and

$$A_{1,d} = \{j \in A | d \in QR_{p_j}, -1 \in QR_{p_j}\}$$

$$A_{2,d} = \{j \in A | d \in QNR_{p_j}, -1 \in QR_{p_j}\}.$$

The statistical assumptions mean that for $d \in \{5, 19\}$ we have

$$\sum_{j \in A_{1,d}} a_{p_j,d} p^{-s} = \sum_{j \in A_{1,1}} a_{p_j,1} p^{-s} \tag{17}$$

$$\sum_{j \in A_{2,d}} a_{p_j,d} p^{-s} = - \sum_{j \in A_{2,1}} a_{p_j,1} p^{-s}$$

$$\sum_{j \in A_{1,d}} a_{p_j,d} p^{-s} = - \sum_{j \in A_{2,d}} a_{p_j,d} p^{-s}.$$

Therefore for $d = 5$ we get

$$h_{1,5}(s) = \sum_{j \in A_{1,5}} a_{p_j,5} p^{-s} + \sum_{j \in A_{2,5}} a_{p_j,5} p^{-s} = 0$$

and $r_5 = 0$. Thus, by (30) follows $r_1 = 1$ and thus $r = 0$.

But by the same argument, for $d = 19$ holds

$$h_{1,19}(s) = \sum_{j \in A_{1,19}} a_{p_j,19} p^{-s} + \sum_{j \in A_{2,19}} a_{p_j,19} p^{-s} = 0$$

and $r_{19} = 0$. This is in contradiction with (31).

Thus, if statistical arguments are allowed, the Birch and Swinnerton-Dyer conjecture fails for rank one for $L(C, s)$. □

It is not any better for $L_2(C, s)$:

Corollary 1. *If statistical arguments of the distribution of prime numbers in Theorem 2 are allowed, then the Birch and Swinnerton-Dyer conjecture for $L_2(C, s)$ fails in rank one.*

Proof. If statistical assumptions of Theorem 2 are allowed, then (17) holds. We get the same contradiction for $L_2(C, s)$ as for $L(C, s)$. The only difference is that $r = 1$. □

We can make the following two observations.

(1) If $a = -d^2$ and $b = 0$ the function $g_{1,1}(s)$ has at $s = 1$ a simple pole of the form

$$g_{1,1}(s) = \frac{-1}{s-1} + \text{finite terms} \tag{32}$$

If we replace $a_{p_j}^2$ in $g_{1,1}(s)$

$$g_{1,1} = -\sum_{j \in A} \frac{a_{p_j}^2}{p_j} \ln(p_j) p_j^{1-2s}$$

by its average value $2p_j$ from Lemma 6 we get the approximation

$$g_{1,1} = -\sum_{j \in A} 2\ln(p_j) p_j^{1-2s} = \frac{-1}{s-1} + \text{finite terms}$$

(2) For $d = k^2$ the function $h(s)$ with the series expression

$$h(s) = \sum_{j \in A} \ln(p_j) a_{p_j} p_j^{-s}$$

in the area where the series converges is of the form

$$h_1(s) = \frac{1}{s-1} + \text{finite terms} \tag{33}$$

when s is close to 1. If $-1 \in QNR_{p_j}$ the value of a_{p_j} is zero by Lemma 1. If for $-1 \in QR_p$ we replace a_{p_j} by its average value -2 from Lemma 6 we get the approximation

$$h(s) = \frac{1}{2} \sum_{j \in A} 2\ln(p_j) p_j^{-s}$$

which is exactly of the correct form (31).

Both observations support the idea that the Statistical Assumption can be used in estimating the values of a_{p_j}. If the Statistical Assumption holds, the function $g_{1,1}(s)$ has a simple pole of the same form (32) for any elliptic curve. The function $h_1(s)$ also cannot be different: either it is finite or it has the form

(33). Therefore it is not possible that the L-function can give algebraic ranks of six or so on, though elliptic curves can have such ranks.

References

1. A. Wiles, The Birch and Swinnerton-Dyer Conjecture. *available on-line at* www.claymath.org.
2. B. Birch and P. Swinnerton-Dyer, Notes on Elliptic Curves (II). *J. Reine Math.* 165 (218): 79-108, 1965.
3. J. Jormakka, A Theorem of Congruent Primes, June 2020, available ResearchGate.

A Theorem of Congruent Primes

Abstract. The paper presents a theorem when a prime number is not a congruent number. This theorem does not add to the present knowledge of congruent primes since all primes fulfilling the conditions of the theorem can already be classified into congruent and noncongruent numbers, but the proof of the theorem has certain own interest and this is why I decided to write it into a paper.

Key words: Congruent numbers, elliptic curves, number theory.

11 Introduction

Consider an elliptic curve of the form:

$$y^2 = x^3 - d^2x \tag{1}$$

where d is an integer. A rational solution (x, y) to the elliptic curve (1) is a solution where x and y are rational numbers.

The substitution $x = d(a+b)/b$, $y = 2d^2(a+c)/b^2$ changes $y^2 = x^3 - d^2x$ to $a^2 + b^2 = c^2$ with $ab = 2d$. Then $4d^2 = a^2(c^2 - a^2)$. Integers d that give rational number solutions to $a^2 + b^2 = c^2$, $ab = 2d$ are called congruent numbers. If d is a congruent number the elliptic curve (1) has a rational solution where y is not zero. In that case it has infinitely many rational solutions.

If there is a solution for $d = s^2$, then there is a solution for $d = 1$ because the substitution $y = s^3y'$, $x = s^2x'$ changes $y^2 = x^3 - d^2x$ to $y'^2 = x'^3 - x'$. It is known that every $d = s^2$ is a congruent number. The case where d is a prime number is amost solved.

For notations the following concepts suffice: The condition that the integer a divides integer b is written as $a|b$. If $p > 2$ is a prime, the cyclic group of integers modulo p is denoted by Z_p and $Z_p^* = \{1, \ldots, p-1\}$. The set of quadratic residues modulo p is the set

$$QR_p = \{x \in Z_p^* | \exists y \in Z_p^* \text{ such that } y^2 \equiv x \pmod{p}\}.$$

The set of quadratic nonresidues modulo p is the set

$$QNR_p = \{x \in Z_p^* | x \notin QR_p\}.$$

Let us start by two very simple lemmas.

Lemma 1. *Let $c^2 = a^2 + b^2$, $a, b, c \in Z$, then $\exists h, m, e \in \mathbb{N}$ such that*

$$a = \pm hem \ , \ b = \pm \frac{1}{2} h(m^2 - e^2) \ , \ c = \pm \frac{1}{2} h(m^2 + e^2).$$

Proof. Without loss of generality we can assume that $a, b, c \in N$. We can write $c^2 - b^2 = (c-b)(c+b) = a^2$. Let $h = gcd(c+b, c-b)$. Then there exists m and e, $m > e$, $gcd(m, e) = 1$, such that $c + b = hm^2$, $c - b = he^2$. The claim follows. □

With Lemma 1 we can characterize congruent numbers.

Lemma 2. *Let $d \in Z$, $d > 0$. Rational solutions (x, y) with $x \neq 0, y \neq 0$ to*

$$y^2 = x^3 - d^2 x$$

are of the form

$$(x_1, y_1) = \left(d\frac{m+e}{m-e}, \pm\frac{k}{j} d\frac{m+e}{m-e} \right),$$

$$(x_2, y_2) = \left(d\frac{m-e}{m+e}, \pm\frac{k}{j} d\frac{m-e}{m+e} \right),$$

where $k, j, e, m \in N$, $m > e$, $gcd(m, e) = 1$, $gcd(k, j) = 1$, satisfy

$$d = \left(\frac{k}{2j} \right)^2 \frac{m^2 - e^2}{em}. \tag{2}$$

Proof. Let $x, y \in Q$, $x \neq 0, y \neq 0$. Let us write $\alpha = \frac{d}{x} + 1 \in Q$, $\beta = \frac{y}{x} \in Q$. Solving (10) for x and solving x from the definition of α yields

$$x = \frac{\beta^2}{2\alpha - \alpha^2} = \frac{d}{\alpha - 1}.$$

Writing $\beta = \frac{k}{j}$ for some $k, j \in N$ gives

$$\alpha_{1,2} = 1 - \frac{k^2}{j^2 2d} \pm \frac{\sqrt{(2dj^2)^2 + (k^2)^2}}{j^2 2d}.$$

As $y \neq 0$, $k \neq 0$. By Lemma 1, $\alpha_{1,2} \in Q$ if and only if there exist $h, e, m \in N$, $gcd(e, m) = 1$, $m > e$, such that

$$k^2 = hem \ , \ 2dj^2 = \frac{1}{2}h(m^2 - e^2) \ , \ c = \frac{1}{2}h(m^2 + e^2).$$

If $em = 0$, then $k = 0$ and $y = 0$. This solution gives $j = 2dj^2$

$$\alpha_{1,2} = 1 \pm \frac{2dj^2}{2dj^2} = 1 \pm 1 \ , \ \alpha_1 = 2, \alpha_2 = 0,$$

$$x_1 = \frac{d}{\alpha - 1} = d \ , \ x_2 = -d \ , y = 0$$

but we have excluded this case in the assumptions. Since $em \neq 0$, let us write $h = \frac{k^2}{em}$. Eliminating h yields

$$d = \left(\frac{k}{2j}\right)^2 \frac{m^2 - e^2}{em},$$

$$c = \frac{k^2}{2}(m^2 + e^2).$$

Simplifying $\alpha_{1,2}$ yields

$$\alpha_{1,2} = \frac{1}{m^2 - e^2}\left(m^2 - e^2 - 2em \pm (m^2 + e^2)\right),$$

i.e.,

$$\alpha_1 = \frac{2m}{m+e} \ , \ \alpha_2 = -\frac{2e}{m-e}$$

$$x_1 = \frac{d}{\alpha_1 - 1} = d\frac{m+e}{m-e} \ , \ x_2 = -d\frac{m-e}{m+e},$$

$$y = \beta x \ , \beta^2 = \left(\frac{k}{j}\right)^2 = 4d\frac{em}{m^2-e^2}.$$

This gives the claim. □

As two examples of Lemma 2

$$d = 5 = \left(\frac{3}{2\cdot 2}\right)^2 \frac{9^2-1^2}{9\cdot 1}$$

$$d = 7 = \left(\frac{24}{2\cdot 5}\right)^2 \frac{16^2-9^2}{16\cdot 9}$$

are both congruent numbers. Notice that $gcd(k,j) = 1$ but it is allowed that $2|k$.

If d is a square, there are no rational solutions to (1) with $y \neq 0$. There are the three solutions $(0,0), (d,0), (-d,0)$ to (1), so the number of rational solutions of (1) is finite, the rank of the elliptic curve is zero.

In the next theorem gives a set of values where d is a prime number and (1) has no rational solutions, i.e., the elliptic curve has rank zero. The case of prime numbers d is rather well known: if $p \equiv 5 \pmod{d}$ or $p \equiv 7 \pmod{d}$ the number d is a congruent number and there are solutions to (1). If $p \equiv 3 \pmod{d}$ there are no solutions and d is not a congruent number. The only case remaining is $p \equiv 1 \pmod{d}$. For that case it is known that e.g. $p = 41$ is a congruent number, while e.g. $p = 17$ is not.

The next theorem does not solve the problem for any prime p that is one modulo eight because if $p \equiv 1 \pmod{8}$ it is necessarily true that $-1 \in QR_p$, i.e., $-1 \in QR_p$ is equivalent with the condition that $4|(p-1)$ and if $p \equiv 1 \pmod{d}$, then $8|(p-1)$. The theorem does prove e.g. that $p = 19$ is not a congruent number, but as $19 \equiv 3 \pmod{d}$ this is known. Yet, the method of this proof seemed interesting enough to me in order to be written down. The method may generalize to other numbers than primes. The primality condition is used only in a few places. The main idea is to exclude branches from a recursion.

Theorem 1. *Let $d > 3$ be a prime such that $-1 \in QNR_d$ and $2 \in QRN_d$. The equation (2) in Lemma 2 does not have solutions $k, j, m, e \in \mathbb{N}$ where $gcd(m, e) = 1$, $gcd(k, j) = 1$, $m > e > 0$.*

Proof. We write (2) with m_1, e_1

$$d = \left(\frac{k}{2j}\right)^2 \frac{m_1^2 - e_1^2}{e_1 m_1} \tag{3}$$

If $d|m_1$ then $d|e_1$ and $gcd(m_1, e_1) \neq 1$, thus $d \not| m_1$ and $d \not| e_1$. If $d|k^2$ then since d is a prime $d|k$. It follows that $k = dk_1$ and as $gcd(k, 2j) = 1$ holds $d \not| 2j$. Thus

$$(2j)^2 m_1 e_1 = dk_1^2(m_1^2 - e_1^2)$$

which is not possible as the left side is not divisible by d. Thus $d \not| k^2$. Therefore $d|m_1^2 - e_1^2$.

If $2 \not| k$ we convert (3) into the form

$$d = \left(\frac{k}{j}\right)^2 \frac{st}{m^2 - e^2} \tag{4}$$

by the substitution $m_1 = m + e$, $e_1 = m - e$, i.e., $2m = m_1 + e_1$, $2e = m_1 - e_1$. As $m_1 e_1 = (m+e)(m-e) = m^2 - e^2$ holds $em = \frac{1}{4}(m_1 + e_1)(m_1 - e_1)$. As $4|(m_1^2 - e_1^2)$ in (3) if $2 \not| k$ it follows that one of $m_1 + e_1$ or $m_1 - e_1$ is even. If so, they are both even and $2|m_1 + e_1$, $2|m_1 - e_1$ and m, e are integers. As $gcd(m_1, e_1) = 1$, $gcd(m_1 + e_1, m_1 - e_1) = 2$. Then $gcd(m, e) = gcd(((m_1 + e_1)/2)((m_1 - e_1)/2)) = 1$. Since $m_1 > e_1 > 0$ holds $m > e > 0$.

If $2|k$ then the substitution is $m = m_1 + e_1$, $e = m_1 - e_1$. Then m, e are integers and $m > e > 0$. In this case $2 \not| j$ gecause $gcd(k, j) = 1$. Therefore $2 \not| (m_1^2 - e_1^2)$.

It follows that $gdc(m, e) = gcd(m_1 + e_1, m_1 - e_1) = 1$. We get the same form (4) since $me = m_1^2 - e_1^2$ and $m^2 - e^2 = 4m_1e_1$.

Then $d|em$ and $j^2|em$. Let us write (4) as

$$j^2(m+e)(m-e)d = k^2me. \tag{5}$$

Since $gcd(m, e) = 1$ it follows that $gcd(m \pm e, m) = 1$. Indeed, if $m \pm e = c_1r$, $m = c_2r$ for some $r, c_1, c_2 \in \mathbb{N}$, then

$$c_1c_2r = c_2m \pm c_2e = c_1m \Rightarrow (c_1 - c_2)m = \pm c_2e$$

$$\Rightarrow m|c_2 \Rightarrow \exists \alpha \in \mathbb{N} \text{ such that } c_2 = \alpha m$$

$$\Rightarrow m = \alpha mr \ \Rightarrow \alpha r = 1 \Rightarrow r = 1.$$

Similarly, $gcd(m \pm e, e) = 1$.

Since $gcd(k, j) = 1$ it follows from (4) that $k^2 = m^2 - e^2$. Therefore (4) implies that $dj^2 = em$. As $dj^2 = em$ and $gcd(e, m) = 1$ there is one of the cases: either $m = ds^2$, $e = t^2$ for some $s, t > 0$ or $m = s^2$, $e = dt^2$.

As $k^2 = (m+e)(m-e)$ and $gcd((m+e)(m-e)) \leq 2$ we have two cases cases: either $m + e = c_1^2$ and $m - e = c_2^2$ for some $c_1, c_2 > 0$ or $m + e = 2c_1^2$ and $m - e = 2c_2^2$.

We have four cases in total.

Case 1. $m = ds^2$, $e = t^2$, $m + e = c_1^2$, $m - e = c_2^2$. Then

$$m - e = s^2d - t^2 = c_2^2.$$

The equation yields $-1 \equiv (c_2t^{-1})^2 \pmod{d}$ which is impossible since $-1 \in QNR_d$.

Case 2. $m = ds^2$, $e = t^2$, $m + e = 2c_1^2$, $m - e = 2c_2^2$. Then

$$s^2 d + t^2 = 2c_1^2 \ , \ s^2 d - t^2 = 2c_2^2.$$

Multiplying the modular equations

$$t^2 \equiv 2c_1^2 \pmod{d} \ , -t^2 \equiv 2c_2^2 \pmod{d}$$

yields $-1 \equiv (2c_1c_2t^{-2})^2 \pmod{d}$ which is impossible since $-1 \in QNR_d$.

Case 3. $m = s^2$, $e = dt^2$, $m + e = c_1^2$, $m - e = c_2^2$. Then

$$s^2 + t^2 d = c_1^2 \ , \ s^2 - t^2 d = c_2^2.$$

Thus

$$2s^2 = c_1^2 + c_2^2 \qquad (6)$$

so

$$4s^2 = c_1^2 + 2c_1c_2 + c_2^2 + c_1^2 - 2c_1c_2 + c_2^2$$

$$(2s)^2 = (c_1 + c_2)^2 + (c_1 - c_2)^2. \qquad (7)$$

It follows from Lemma 1 that $\exists h', e', m' \in \mathbb{N}$, $gcd(m', e') = 1$ such that

$$c_1 + c_2 = h'e'm' \ , c_1 - c_2 = \frac{1}{2}h'(m'^2 - e'^2),$$

$$2s = \frac{1}{2}h'(m'^2 + e'^2).$$

Solving c_1, c_2, s yields

$$c_1 = \frac{1}{4}h'(2e'm' + m'^2 - e'^2),$$

$$c_2 = \frac{1}{4}h'(2e'm' + e'^2 - m'^2),$$

$$s = \frac{1}{4}h'(m'^2 + e'^2).$$

Since

$$2t^2 d = c_1^2 - c_2^2 = (c_1 - c_2)(c_1 + c_2)$$

we get

$$d = \frac{1}{4t^2}h'^2 e'm'(m'^2 - e'^2)$$

i.e.

$$d = \left(\frac{h'e'm'}{2t}\right)^2 \frac{(m'^2 - e'^2)}{e'm'}.$$

Removing the greatest common divisor of $h'e'm'$ and t this equation can be written as

$$d = \left(\frac{k_{i+1}}{2j_{i+1}}\right)^2 \frac{(m_{i+1}^2 - e_{i+1}^2)}{e_{i+1}m_{i+1}}. \tag{8}$$

As $gcd(m', e') = 1$ and we made $gcd(k, j) = 1$, equation (8) is is of the same form as (3)

$$d = \left(\frac{k_i}{2j_i}\right)^2 \frac{(m_i^2 - e_i^2)}{e_i m_i} = \left(\frac{k}{2j}\right)^2 \frac{(m_1^2 - e_1^2)}{e_1 m_1}.$$

We have a recursion that in each step reduces the numbers m_i, e_i to numbers m_{i+1}, e_{i+1} that are of the order of square root of m_i, e_i.

Case 4. $m = s^2$, $e = dt^2$, $m + e = 2c_1^2$, $m - e = 2c_2^2$. We can select $c_1 > c_2 \geq 0$. Then

$$s^2 + t^2 d = 2c_1^2 \ , \ s^2 - t^2 d = 2c_2^2.$$

Thus

$$s^2 = c_1^2 + c_2^2 \ \ dt^2 = c_1^2 - c_2^2 = (c_1 - c_2)(c_1 + c_2). \tag{9}$$

Let us notice that $m + e = 2c_1^2$ and

$$1 = gcd(m + e, e) = gcd(2c_1^2, dt^2) \Rightarrow gcd(c_1, t) = 1, gcd(2, t) = 1$$

$$1 = gcd(m - e, e) = gcd(2c_2^2, dt^2) \Rightarrow gcd(c_2, t) = 1.$$

First we exclude one case in the second equation of (9). If $t > 1$ and $c_1 + c_2 = \alpha_1 t$ and $c_1 - c_2 = \alpha_2 t$ for some $\alpha_1, \alpha_2 \in \mathbb{N}$, then

$$2c_1 = (\alpha_1 + \alpha_2)t \Rightarrow t = 1, 2c_1 = \alpha_1 + \alpha_2,$$

$$2c_2 = (\alpha_1 - \alpha_2)t \Rightarrow t = 1, 2c_1 = \alpha_1 - \alpha_2.$$

Thus, $dt^2 = c_1^2 - c_2^2 = \alpha_1\alpha_2 t^2$. It follows that $d = \alpha_1\alpha_2$ and as d is prime and necessarily $\alpha_1 > \alpha_2$ it follows that $\alpha_1 = d$, $\alpha_2 = 1$. Then $c_1 = d+1$ and $c_2 = d-1$. Consequently $s^2 = c_1^2 + c_2^2 = 2(d^2 - 1)$ is even, so m is even. Since $s^2 + dt^2 = 2c_1^2$ it would follow that t is also even as d is odd, but $t = 1$ in this case. We have a contradiction.

Thus, in (9) must be one of the three cases

$$t^2 | (c_1 + c_2) \Rightarrow (c_1 - c_2) | d \Rightarrow c_1 - c_2 = d \Rightarrow t^2 = c_1 + c_2,$$

or

$$t^2 | (c_1 - c_2) \Rightarrow (c_1 + c_2) | d \Rightarrow c_1 + c_2 = d \Rightarrow t^2 = c_1 - c_2,$$

or

$$t = 1.$$

In the first case

$$2c_1 = t^2 + d \geq 0 \ , \ 2c_2 = t^2 - d \geq 0.$$

In the second case

$$2c_1 = d + t^2 \geq 0 \ , \ 2c_2 = d - t^2 \geq 0.$$

In both of these two cases we can derive in a similar way:

$$s^2 = c_1^2 + c_2^2 \ \Rightarrow (2s)^2 = (2c_1)^2 + (2c_2)^2$$

yields

$$(2s)^2 = (d+t^2)^2 + (d-t^2)^2. \tag{10}$$

By Lemma 2 there exist $h', e', m' \in \mathbb{N}$ such that

$$d + t^2 = h'e'm' \ , \ d - t^2 = \frac{1}{2}h'(m'^2 - e'^2).$$

The first equation implies that $d \not| h'$. Thus

$$4d = h'((m'+e')^2 - 2e'^2)$$

i.e., as $h' \not\equiv 0 \pmod{d}$

$$2 \equiv (m'^2 + e'^2)^2 e'^{-2} \pmod{d} \tag{11}$$

which is a contradiction since $2 \in QNR_d$. There remains the case $t = 1$. Then $2c_1^2 = s^2 + d$, $2c_2^2 = s^2 - d$. Instead of (10) we get

$$(2s)^2 = (d+s^2)^2 + (d-s^2)^2.$$

The contradiction (11) comes in the same way with t replaced by s. This means that Case 4 is not possible.

Because Cases 1, 2 and 4 are not possible, only Case 3 is left. Case 3 gives a recursion formula. The values h', m', e' in Lemma 1 satisfy

$$\frac{e'}{m'} = \frac{a}{b+c} = \frac{c-b}{a}$$

$$h' = gcd(b+c, b-c)$$

giving $a^2 = c^2 - b^2$. The numbers h', m', e' can be chosen to be positive and on the order of a, b, c. Thus, h', m', e' in (8) are of the order c_1, c_2. The numbers c_1, c_2 are of the order $\sqrt{m}$, $\sqrt{e}$. Therefore in each step the numbers m_i, e_i get smaller, they are reduced to the order of their square roots. Consider the problem when the recursion stops.

Let us look at an example of $d = 5$. Then

$$d = 5 = \left(\frac{3}{2 \cdot 2}\right)^2 \frac{9^2 - 1^2}{9 \cdot 1}.$$

We have $m_1 = 9, e_1 = 1, k = 3, j = 2$. We can do the first step and find $m = 5, e = 4$ and

$$d = 5 = \left(\frac{3}{2}\right)^2 \frac{5 \cdot 4}{5^2 - 4^2}.$$

Identifying $k^2 = 3^2 = 5^2 - 4^2 = 9$, $j^2 d = 4 \cdot 5 = 20 = 5 \cdot 4 = me$, $m = ds^2 = 5 \cdot 1^2$, $e = t^2 = 2^2$, $m + e = 5 + 4 = 3^2 = c_1^2$ and $m - e = 5 - 4 = 1^2 = c_2^2$ shows that the logic in the lemma is correct. We have Case 1, but for $d = 5$ the conditions of the lemma are not fulfilled: $-1 \in QNR_5$. This is why Case 1 does not give a contradiction. What happens in Case 1 is that when we remove the term dt^2 in a case resembling (6) we do not get (6) but

$$2t^2 = c_1^2 - c_2^2$$

Therefore we do not get (7) which can be inserted to the equation to Lemma 1 for calculation of the numbers h', m', e'.

Let us look at another example, that of $d = 7$. Here $-1 \in QNR_7$ and the Case is not 1.

$$d = 7 = \left(\frac{24}{2 \cdot 5}\right)^2 \frac{16^2 - 9^2}{16 \cdot 9}.$$

We have $m_1 = 16, e_1 = 9, k = 24, j = 5$. We find $m = 16 + 9 = 25, e = 16 - 9 = 7$. Thus

$$d = 7 = \left(\frac{24}{5}\right)^2 \frac{25 \cdot 7}{25^2 - 7^2}.$$

Here $k^2 = 24^2 = 576 = 25^2 - 7^2 = m^2 - e^2$, $j^2 d = 25 \cdot 7 = 175 = 25 \cdot 7 = me$, $m = s^2 = 5^2$, $e = dt^2 = 7 \cdot 1^2$, $m + e = 25 + 7 = 32 = 2 \cdot 4^2 = 2c_1^2$ and

$m-e=25-7=18=2\cdot 3^2=2c_2^2$. The Case is 4. We notice that $t^2=1$ and $c_1=4, c_2=3$, thus we have the case $t=1$. Then $s^2+d=5^2+7=32=2\cdot 4^2=2c_1^2$ and $s^2-d=5^2-7=18=2\cdot 3^2=2c_2^2$. We get

$$(2s)^2=100=64+36=(2c_1)^2+(2c_2)^2=(5^2+7)^2+(5^2-7)^2$$

and therefore find the numbers h',m',e' for $10^2=8^2+6^2$. The numbers are $h'=gcd(10+6,10-6)=4$, $e'=1$, $m'=2$. Thus

$$d+t^2=h'e'm'=7+1=8\ ,\ d-t^2=\frac{1}{2}h'(m'^2-e'^2)=6$$

are true and

$$4d=h'((m'+e')^2-2e'^2)=28=4\cdot(3^2-2).$$

We get the modular equation $3^2\equiv 2 \mod (7)$, which violates the assumption $2\in QNR_d$, but indeed $2\in QR_7$. Therefore for $d=7$ we do not get a contradiction.

The way the lemma works is that in (2) the numbers m_1 and e_1 must be squares $m_1=s_1^2$, $e_1=t_1^2$ so that k^2 can cancel them. The condition $-1\in QNR_d$ excludes the larger branch $(s_1^2+t_1^2)$ of

$$m_1^2-e_1^2=(s_1^2+t_1^2)(s_1^2-t_1^2)$$

by $(s_1^2+t_1^2)\equiv 0 \pmod d$ being impossible.

Therefore $4d|(m_1^2-e_1^2)$ leads to $4d|(s_1^2-t_1^2)$. The condition $2\in QNR_d$ excludes Case 4 and leaves only Case 3 which gives a recursion. Thus, the numbers m_i,e_i get smaller.

If there is a congruent number d with $-1\in QNR_d$, the recursion must continue until it stops in some way and not to a contradiction, but the recursion does not stop and continues to a contradiction. At each stage $4d|(m_i^2-e_i^2)$ or $d|(m_i^2-e_i^2)$

depending on if k_i is odd or even. The numbers m_i and e_i become smaller on each step. Finally we must have $4d = m_i^2 - e_i^2$ or $d = m_i^2 - e_i^2$.

Changing variables in (2) to $m = (m_i + e_i)/2$, $e = (m_i - e_i)/2$ if k is odd and $m = m_i + e_i$, $e = m_i - e_i$ if k is even we get

$$d = \frac{k^2}{j^2} \frac{me}{m^2 - e^2}. \tag{12}$$

When the recursion has reached $4d = m_i^2 - e_i^2$ or $d = m_i^2 - e_i^2$ the number $j = 1$. In (12) necessarily $k^2 = m_i^2 e_i^2$ and consequently $d = me$. As d is prime either $m = d$, $e = 1$ or $m = 1$, $e = d$. As in Cases 1 and 2 the choice $m = d$ leads to $-1 \in QR_d$ and is impossible. Thus $m = 1$ and $t = d$, but then $m^2 - e^2 < 0$ and $d > 0$ is negative. This is a contradiction. The recursion leads to a contradiction and the claim of the lemma follows. □

There are primes d filling the conditions of the lemma: for $d = 19$ holds $-1 \in QNR_{19}$ and $2 \in QNR_{19}$. We also get a small result:

Corollary 1. *If p is a prime and $p \equiv 7 \pmod 8$, then $2 \in QR_p$.*

Proof. If p is a prime and $p \equiv 7 \pmod 8$, then p is a congruent number. Therefore the conditions of Theorem 1 cannot be fulfilled. The condition $-1 \in QR_p$ is equivalent with $4|(p-1)$. As $p - 1 = 6 + 8k$ for some k, it follows that $4 \not| (p-1)$. Thus $-1 \in QNR_p$. The only other condition in Theorem 1 is that $2 \in QNR_p$. □

While working with the Birch and Swinnerton-Dyer conjecture in 2010 I derived in [1] a theorem of congruent primes. The theorem (Lemma 11) in [1] was never needed for the result it gives to primes, but as an easy case of the proof method that I hoped to generalize to other d. Now I have rewritten the 2010 paper and do not use the method of Lemma 11. Yet, it has some own interest in the proof method. Therefore I moved it into this short paper. There were some typos in [1] in the proof of Lemmas 10 and 11, which are Lemma 2 and Theorem 1 in this paper. Now the errors are fixed. The method of the proof does work. Maybe some application for the method will be found later.

References

1. J. Jormakka, On the rank of elliptic curves, arXiv:0806.4091, first version from 2010.

A lemma and a calculation from the Riemann Hypothesis proofs

12 About this short paper

When I wrote the two proofs of the Riemann Hypothesis [1] and [2] some results were deleted as unnecessary. As they have some interest of their own, I write them down in this short paper.

Lemma 1 was originally in the totally new prood of the Riemann Hypothesis [1] that I wrote in 2020. This lemma requires reading the introduction and Lemma 1 of [1]. I will not repeat the text in this paper.

Lemma 1. *The following two claims hold: (i) The only poles of the sum*

$$\sum_{j=1}^{\infty} h_j(s)$$

that remain after cancellations of poles of $h_j(s)$ by poles of other $h_m(s)$ are poles of $h_j(s)$ at points s_k of the type

$$\frac{r_k}{s-s_k}+f_k(s)$$

where $f_k(s)$ is analytic close to s_k. The number $r \neq 0$ is an integer. Only for one s_k the number r_k is negative and has the value -1. It is the pole of $\zeta(s)$ at $s_k = 1$. For the other poles s_k, $k > 0$, the value r_k is a positive integer. (ii) The function $h_1(s)$ has an infinite number of poles in $Re\{s\} > 0$.

Proof. Claim (i) follows directly. All poles of every $h_j(s)$, $j > 0$, that have a noninteger value of r must be cancelled or partially cancelled by poles of other $h_m(s)$, $m > 0$, because at a pole of $\zeta(s)$ and a zero of $\zeta(s)$ the value r is always an integer. Additionally, there cannot be any other poles of $\zeta(s)$ than the one at $s = 1$. Thus, in the sum of $h_j(s)$ all negative values of r sum to zero or to a positive integer value except for $h_1(s)$ in the pole $s = 1$. There $r = -1$.

For the claim (ii) we give one possible cancellation process and then notice that every cancellation process has the same features leading to (ii).

Let the set of natural numbers $\{1, 2, 3, 4, \dots\}$ be divided into disjoints sets $C_j = \{t_j, 2t_j, \dots, 2^k t_j, \dots\}$, where 2 does not divide t_j and $t_{j+1} > t_j$. Thus, $t_1 = 1$, $t_2 = 3$, $t_3 = 5$, and so on.

Let us take the sum of two pairs of poles of $h_1(s)$

$$\frac{r_k}{s - s_k} + \frac{r_k}{s - s_k^*}$$

$$-\frac{2^{-1} r_k}{s - 2^{-1} s_k} - \frac{2^{-1} r_k}{s - 2^{-1} s_k^*}. \tag{1}$$

Because $h_j(s) = h_1(js)$ (see Lemma 1 in [1]) there must be corresponding poles at $h_m(s)$

$$\frac{\frac{1}{m} r_k}{s - \frac{1}{m} s_k} + \frac{\frac{1}{m} r_k}{s - \frac{1}{m} s_k^*}$$

$$-\frac{\frac{1}{m} 2^{-1} r_k}{s - \frac{1}{m} 2^{-1} s_k} - \frac{\frac{1}{m} 2^{-1} r_k}{s - \frac{1}{m} 2^{-1} s_k^*}.$$

In C_j there are corresponding poles for each $m = 2^n t_j$.

Let us sum these poles over C_j. We see that most terms cancel

$$\frac{\frac{1}{t_j} r_k}{s - \frac{1}{t_j} s_k} + \frac{\frac{1}{t_j} r_k}{s - \frac{1}{t_j} s_k^*} \tag{1}$$

$$-\frac{\frac{1}{t_j} 2^{-1} r_k}{s - \frac{1}{t_j} 2^{-1} s_k} - \frac{\frac{1}{t_j} 2^{-1} r_k}{s - \frac{1}{t_j} 2^{-1} s_k^*}.$$

$$+\frac{\frac{1}{t_j} 2^{-1} r_k}{s - \frac{1}{t_j} 2^{-1} s_k} + \frac{\frac{1}{t_j} 2^{-1} r_k}{s - \frac{1}{t_j} 2^{-1} s_k^*}$$

$$-\frac{\frac{1}{t_j}2^{-2}r_k}{s-\frac{1}{t_j}2^{-2}s_k}-\frac{\frac{1}{t_j}2^{-2}r_k}{s-\frac{1}{t_j}2^{-2}s_k^*}$$

$$\dots$$

$$+\frac{\frac{1}{t_j}2^{-i}r_k}{s-\frac{1}{t_j}2^{-i}s_k}+\frac{\frac{1}{t_j}2^{-i}r_k}{s-\frac{1}{t_j}2^{-i}s_k^*}$$

$$-\frac{\frac{1}{t_j}2^{-i-1}r_k}{s-\frac{1}{t_j}2^{-i-1}s_k}-\frac{\frac{1}{t_j}2^{-i-1}r_k}{s-\frac{1}{t_j}2^{-i-1}s_k^*}.$$

$$+\frac{\frac{1}{t_j}2^{-i-1}r_k}{s-\frac{1}{t_j}2^{-i-1}s_k}+\frac{\frac{1}{t_j}2^{-i-1}r_k}{s-\frac{1}{t_j}2^{-i-1}s_k^*}$$

$$-\frac{\frac{1}{t_j}2^{-i-2}r_k}{s-\frac{1}{t_j}2^{-i-2}s_k}-\frac{\frac{1}{t_j}2^{-i-2}r_k}{s-\frac{1}{t_j}2^{-i-2}s_k^*}$$

$$\dots$$

$$=\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k}+\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k^*}. \tag{2}$$

There is left only one pole pair (2) in each C_j. Especially in C_1 the function $h_1(s)$ has left the poles

$$\frac{r_k}{s-s_k}+\frac{r_k}{s-s_k^*}. \tag{3}$$

Let t_j be a prime larger than 2. We can cancel the pole pair (2) of C_j by addding two pairs of poles to $h_1(s)$

$$-\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k}-\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k^*} \tag{4}$$

$$+\frac{2^{-1}\frac{1}{t_j}r_k}{s-2^{-1}\frac{1}{t_j}s_k}+\frac{2^{-1}\frac{1}{t_j}r_k}{s-2^{-1}\frac{1}{t_j}s_k^*}.$$

If $\frac{1}{t_j}r_k$ is fractional, the pole pair (2) of C_j must be cancelled, but we will show that it is cancelled even if $\frac{1}{t_j}r_k$ is an integer: Because of (5), if we add the poles (4) to $h_1(s)$, then there must be corresponding poles at $h_m(s)$

$$\frac{\frac{1}{mt_j}r_k}{s-\frac{1}{mt_j}s_k}+\frac{\frac{1}{mt_j}r_k}{s-\frac{1}{mt_j}s_k^*} \tag{5}$$

$$-\frac{\frac{1}{mt_j}2^{-1}r_k}{s-\frac{1}{mt_j}2^{-1}s_k}-\frac{\frac{1}{mt_j}2^{-1}r_k}{s-\frac{1}{mt_j}2^{-1}s_k^*}.$$

For a sufficiently large prime m the number $\frac{1}{mt_j}r_k$ is not an integer. Therefore this pole of $h_m(s)$ must be cancelled. It can only be cancelled by a pole of $h_1(s)$ and therefore the poles (4) are necessary. Therefore we must add the first pole pair in (4) even if $\frac{1}{t_j}r_k$ is an integer. The pole pair in C_j is cancelled by adding the pole pairs of (4).

We do not get new poles to each C_m. The new poles (5) are added to C_m only if t_j divides t_m. When we sum these new poles (5) to each such C_m, it is the same calculation as in (1). Most poles cancel and only one pole pair remains for each C_m, namely

$$-\frac{\frac{1}{t_m}r_k}{s-\frac{1}{t_m}s_k}-\frac{\frac{1}{t_m}r_k}{s-\frac{1}{t_m}s_k^*}. \tag{6}$$

Let $j=1$, so $t_j=3$. We add the poles (4) to $h_1(s)$ and there remains the poles (6). Adding (6) to (2) the new poles of $h_1(s)$

$$-\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k}-\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k^*}$$

cancels the remaining pole pair in C_j

$$\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k}+\frac{\frac{1}{t_j}r_k}{s-\frac{1}{t_j}s_k^*}.$$

More generally, we notice that for each C_m such that 3 divides t_m the pole pair (2) has been cancelled. Especially, the remaining pole at C_2 has been cancelled.

But now comes a complication. We continue the process by adding to $h_1(s)$ a pole that cancels the remaining pole at C_3 where $t_3 = 5$, the smallest prime larger than t_2. Thus, we add the poles (4) for $j = 3$. The remaining pole pair in (2) in the set C_3 is cancelled, but in each C_m where $15 = 3 \cdot 5$ divides t_m we have a new pole pair. That is, the original pole pair (2) in C_m was cancelled when we added (4) with $t_j = t_2 = 3$ and in this process added the poles also to C_m where t_m is divisible by $15 = 3 \cdot 5$, but now we add $t_j = t_3 = 5$ and again have a pole pair at C_m since t_m is divisible by $5 \cdot 3$. The remaining pole at C_3 was cancelled, but we made new poles to every C_m where t_m has the factor 15.

Continuing this process by adding to $h_1(s)$ poles where $t_j = t_4 = 7$ as in (4) we cancel the remaining pole at C_4 in (2), but make new poles to C_m where t_m has the factors $4 \cdot 3$ or $4 \cdot 5$. Continuing the process by adding poles (4) to $h_1(s)$ for each prime number t_j in the increasing order we cancel the remaining pole in each C_j where t_j is a prime. At the same time we are creating new poles to each C_m where t_m has two prime factors larger than two. This is the first step of the pole cancellation process.

In the second step we add poles to $h_1(s)$ that have t_j a product of two primes larger than two and select r-values that cancel the remaining poles of C_j for every t_j that is a product of two primes larger than two. Again we add new poles to C_m where t_m has more than two prime factors that are larger than two.

In the nth step we add to $h_1(s)$ poles which are products of n primes larger than two, and select r-values that cancel the poles of C_j for each t_j that is a product of exactly n primes larger than two.

This process continues and on each step we must add poles to $h_m(s)$ where m is a so large number that the resulting r value for $h_m(s)$ is fractional. Such a pole must be cancelled and it can only be cancelled by making a next step (4) by adding poles to $h_1(s)$. Thus, this process cannot stop. It does not stop even if the r value of a pole in C_j that we want to cancel is a positive integer. There is always a large $m = t_j m_1$ that will also be created because of adding the poles to $h_1(s)$ in (4). This C_m has a fractional r value and must be cancelled. Therefore the remaining pole of C_j is always cancelled by a new pole of $h_1(s)$.

The new poles that we are adding in each step to $h_1(s)$ are on each step closer and closer to $s = 0$. The r-values of the poles that we have to add to $h_1(s)$ become very large in absolute value when the number n of steps grows and the new poles added to $h_1(s)$ approach $s = 0$. The absolute value of r grows because the numbers of the remaining C_m on each step have many factors and we add as many poles as there are factors to those poles that do not get cancelled in each step. We cannot reach $s = 0$, but every C_j will have the poles completely cancelled at some step.

As a conclusion, $h_1(s)$ must have an infinite number of poles because this pole cancellation process cannot stop. Only some poles of $h_1(s)$ remain uncancelled in the sum $\sum_{j=1}^{\infty} h_j(s)$.

The cancellation of poles on the x-axis is the same, only there is one pole and not a pole pair. The procedure is obtained by setting $Im\{s_k\} = 0$ in the described process and removing the part with s_k^*. There is a pole of $h_1(s)$ on the x-axis at $s = \frac{1}{2}$ with the r-value $\frac{1}{2}$. This is so because $h_1(s)$ has a pole at $s = 1$ with the r-value -1. Consequently $h_2(s)$ has a pole with r-value $-\frac{1}{2}$ at $s = \frac{1}{2}$. This pole must be cancelled and can only be cancelled by $h_1(s)$. Therefore there must be the pole of $h_1(s)$ at $s = \frac{1}{2}$, but it is cancelled and the only pole remaining on the x-axis from this sequence is the pole of $h_1(s)$ at $s = 1$. However, $h_1(s)$ may have other uncancelled poles with a positive integer r-value on the x-axis. If there are such poles of $h_1(s)$, then they are not cancelled by this process.

We will not prove that the described pole cancellation process is the only possible process, though this claim may be true. However, every possible pole cancellation process has the same feature: every pole at $h_1(s)$ requires a corresponding pole at each $h_m(s)$. If m is sufficiently large, such a pole has a noninteger r-value and must be cancelled. Cancelling such a pole by adding (i.e., noticing that the pole exists) a pole to any $h_j(s)$ always implies adding a new pole to $h_1(s)$ and this again requires new poles to all $h_m(s)$. Again some of these new poles have noninteger r-values and must be cancelled. This process cannot stop, thus $h_1(s)$ must have an infinite number of poles. Only a subset of the poles of $\sum_{j=1}^{\infty} h_j(s)$ remain uncancelled. □

Lemma 2 was included in one version of the old proof from 2008 [2]. There was an error in Lemma 2 of the original paper and I had some hard time correcting the error. In the present form [2] has a very short Lemma 2, but I tried all kinds of things before finding the short proof. Among other things I made the following calculation.

Assume that $f(s)$ has a first order zero at s_0. Then $f(s)$ has the Taylor series

$$f(s) = C(s - s_0) + (s - s_0)^{r+1}A(s) \tag{7}$$

where $A(s)$ has only nonzero powers of $(s - s_0)$ and $A(s_0) \neq 0$. Define

$$h(s) = f'(s)f(s)^{-1} = \frac{d}{ds}\ln f(s).$$

Then

$$h(s) = \frac{1}{s - s_0} + f_1 \tag{8}$$

where $f_1(s)$ is analytic in a small environment of s_0. Rewriting (7) and (8) gives

$$A(s) = (s - s_0)^{-2}f - C(s - s_0)^{-1}. \tag{9}$$

and

$$f_1 = f'f^{-1} - (s - s_0)^{-1}. \tag{10}$$

Let us define

$$B(s) = Cf_1$$

$$w = A' - C^{-1}A^2$$

$$v = B - A. \tag{11}$$

Inserting (9) and (10) to $B - A$ shows that

$$B - A = Cf^{-1}(f' - \frac{1}{C}\frac{1}{(s-s_0)^2}f^2).$$

and

$$A' - C^{-1}A^2 = \frac{1}{(s-s_0)^2}(f' - \frac{1}{C}\frac{1}{(s-s_0)^2}f^2).$$

Thus

$$v = B - A = (s-s_0)(c(s-s_0)f^{-1})(A' - \frac{1}{C}A^2)$$

and

$$w = (s-s_0)^{-1}(1 + AC^{-1}(s-s_0))v \tag{12}$$

$$w' = (s-s_0)^{-1}(1 + AC^{-1}(s-s_0))v' - (s-s_0)^{-2}v + A'C^{-1}v. \tag{13}$$

From (11) follows

$$B' - C^{-1}B^2 = A' - C^{-1}A^2 + v' - C^{-1}v^2 - 2AC^{-1}v$$

$$= w + v' - C^{-1}v^2 - 2AC^{-1}v. \tag{14}$$

Inserting (12) and (13) to (14) gives

$$(s-s_0)^{-2}C^{-2}f^2(B'-C^{-1}B^2) = 2w-(A'+C^{-1}A^2)C^{-1}(s-s_0)^2w+C^{-1}fw'-C^{-1}(s-s_0)^2w^2. \tag{15}$$

A calculation shows that

$$C^{-1}(B' - C^{-1}B^2) = f''f^{-1} - 2(f'f^{-1})^2 + 2f'f^{-1}\frac{1}{s-s_0}$$

and inserting (7) yields

$$f''f - 2(f')^2 + 2(2-s_0)^{-1}f'f$$

$$= (s-s_0)^2 2C(A' - C^{-1}A^2) + (s-s_0)^3 C(A'' - 2C^{-1}A'A)$$

$$+(s-s_0)^4(A''A - 2(A')^2).$$

Inserting

$$w' = A'' - 2C^{-1}A'A$$

and

$$A''A - 2(A')^2 = A^3\frac{d}{ds}(A'A^{-2}) = A^3\frac{d}{ds}(C^{-1} + wA^{-1})$$

$$= A^3w'A^{-2} + A^3w\frac{d}{ds}A^{-2} = Aw' - 2A'wA^3A^{-3} = -2A'w + Aw'$$

gives

$$f''f - 2(f')^2 + 2(2-s_0)^{-1}f'f$$

$$= 2Cw(s-s_0)^2 + C(s-s_0)^3w' + (s-s_0)^4(-2A'w + Aw').$$

and so

$$(s-s_0)^{-2}C^{-2}(B' - C^{-1}B^2)f^2 =$$

$$= 2w + (s-s_0)w' + (s-s_0)^2C^{-1}(-2A'w + Aw'). \tag{16}$$

Combining (15) and (16) gives an equation

$$2w - (A' + C^{-1}A^2)C^{-1}(s-s_0)^2w + C^{-1}fw' - C^{-1}(s-s_0)^2w^2$$

$$= 2w + (s-s_0)w' + (s-s_0)^2C^{-1}(-2A'w + Aw').$$

Simplified this is

$$w = A' - C^{-1}A^2. \tag{17}$$

So, no new equation: it is an identify, but this identity is true for any $f(s)$ that has a simple pole and this fact I found quite interesting. Naturally you notice that if $w = 0$ in (17), the solution for A has a pole, but not necessarily at s_0. I think it is worth to save this calculation. Maybe in one hundred years someone will find some use for it.

References

1. J. Jormakka, On the zeros of the Riemann zeta function, 2020.
2. J. Jormakka, On the zeros of the Riemann zeta function, the old proof. 2020.

A Proof of the Poincaré Conjecture

Abstract. This version is very similar to the original proof of the Poincaré Conjecture that I submitted in 1987 to the Annals of Mathematics and sent to Professor Siebenman. Editor Raoul Bott rejected the paper and warned me of submitting any such papers ever in the future as it will seriously affect my career. Professor Siebenman wrote that my English is so poor that he cannot understand what the author tries to say. My English then was as now. I tried to get this proof reviewed for thirteen years without any better luck. The last time I tried was in 2001 when I sent a (different) version to John Milnor. He did not comment it.

Key words: Homotopy spheres, 3-manifold topology, Morse theory.

13 Introduction

A manifold is closed if it is compact and has no boundary. In an orientable 3-manifold every 2-dimensional embedded submanifold has two sides. Embedding g of a submanifold N to a manifold M means, that there is a homeomorphism between N and $g(N) \subset M$. If g mapping a submanifold N to M is not an embedding then $g(N)$ has self-intersections.

An embedded circle S^1 is called a loop. If l is a loop a homomorphism is a continuous mapping $h : I \times I \to M$ such that $l = \{h(0,x)|x \in I\}$. If l' is another loop in M and there is a homomorphism h such that $l' = \{h(0,x)|x \in I\}$, then l is homomorphic to l', we write it $l \simeq l'$. Homotopic images of S^1 belong to the same homotopy class and the set of homotopy classes is a group, the fundamental group $\pi_1(M)$ of the manifold M. As a special case, l' does not need to be an embedded loop but a point. If it is a point, then l is said to be contractible. If every loop in the manifold is contractible, $\pi_1(M) = 1$ and M is said to be simply-connected.

The Poincaré Conjecture states that every simply-connected closed 3-manifold is homeomorphic to the 3-dimensional sphere S^3.

A Morse function is a function $f : M \rightarrow \mathbb{R}$ such that in all but isolated points there is a diffeomorphism g from a neighborhood V of a point $p \in M$ to $\mathbb{R}^3$ so, that if $g(p) = (x, y, z)$, $f(p) = z + c$ where c is a constant.

In a closed 3-manifold there are finitely many points where there is no such homeomorphism. The points are called critical points of f and their structure is well known: they are classified by the index $i(p)$. Critical points of index 0 are points where there is a homeomorphism $g : V \rightarrow \mathbb{R}^3$ from a local neighborhood of a critical point p such that if $g(p) = (r, \theta, \phi)$, then $f(p) = r + c$ where c is a constant. A critical point of index 3 for f is a critical point of index 0 for $-f$, so the Morse levels $f^{-1}(x)$ are spheres and the level grows towards the center. A critical point of index 1 is a point $p \in M$ where f has a saddlepoint. There is a local homeomorphism $g : V \rightarrow \mathbb{R}^3$ from a neighborhood V of p which maps the Morse levels $f^{-1}(x)$ to surfaces shown in Figure 1. The level x grows to the direction of the arrows in Figure 1. A critical point of index 2 for f is a critical point of index 1 for $-f$.

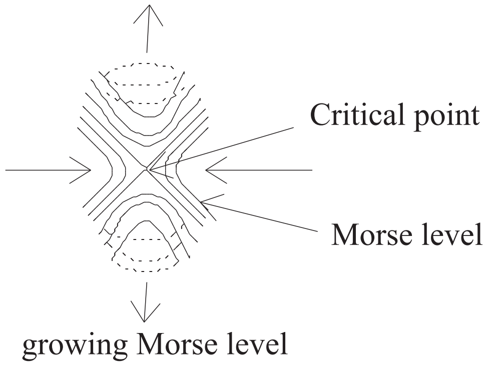

Figure 1.

We will assume that M is a closed and differentiable 3-manifold.

The following notations will be used: I = a closed interval; S^n = a closed n-sphere; D^2 = a closed disc; T^2 = a torus; B^3 = a closed 3-ball. $D^2_r = \{(x, y) | x^2 +$

$y^2 \leq r\}$ = a closed disc of radius r. $U^3 = \{(x,y,z)|x^2+y^2 \leq 1, z \in [0,1], (x,y,1) = (x,y,0)\}$ = a filled torus. A loop is an image of an embedding $g : S^1 \to M$. A path is an image of an embedding $g : I \to M$. A disc is an image of an embedding $g : D^2 \to M$. A collar is an image of an embedding $g : D^2_1 \backslash D^2_{\frac{1}{2}} \to M$.

A closed tubular neighborhood V of a loop l is an image of an embedding $h : U^3 \to M$ such that l is the image of the points $x = y = 0$ in U^3. A closed tubular neighborhood V of a path l is an image of a homeomorphism $h : D^2 \times I \to M$ such that l is the image of the points $x = y = 0$ in D^2. A closed cylinder neighborhood of a disc D is an embedding of $D^2 \times I$ to M such that D is the image of the points $(r, \theta, 1/2)$. When necessary, we assume that these embeddings are smooth.

Let $X \approx Y$ mean that the manifold X (possibly with a boundary) is homeomorphic to Y. Usually we will assume, that the homeomorphism is a diffeomorphism. Let $\gamma \simeq \gamma'$ mean that the loop γ is homotopic to γ'.

We will assume, that the Morse functions are smooth on all points except for a finite set of critical points. Smooth means here that f is sufficiently many times (say, 3 to be sure) continuously differentiable in local coordinates. Let us write $M^a_f = \{x \in M | f(x) \leq a\}$, $\partial M^a_f = \{x \in M | f(x) = a\}$ and if no confusion can arise we write $M^a = M^a_f$. ∂M^a is here called a Morse level. We will write $M_a = M \backslash int(M^a)$. Let a_{max} and a_{min} denote the highest and lowest levels a respectively such that $f^{-1}(a)$ is not empty.

A Heegaard split is a general way of expressing 3-manifolds by glueing two handlebodies at their boundaries. The genus g of the split is the number of handles in the handlebodies.

Definition 1. *A standard handlebody H_g of genus g is the following subset of* $\mathbb{R}^3$

$H_g = \{(x,y,z) \in \mathbb{R}^3 | x = 4j - 2 + (2+r)\cos(\theta), y = (2+r)\sin(\theta), |z| \leq \sqrt{1-r^2}, |r| \leq 1, 0 \leq \theta < 2\pi, j = 1, ..., g\}$

That is, H_g is obtained as a union of g filled toruses which have center at $(4j-2,0,0)$, $j=1,...,g$, laying on the (x,y,-1)-plane and the filled tube of the filled torus has radius 1.

There are homeomorphisms g_1 and g_2 such that $g_1(M^a)$ and $g_2(M_a)$ are handlebodies embedded in $\mathbb{R}^3$ and a boundary homeomorphism $\psi : g_1(M^a) \to g_2(M_a)$. We can take g_1 and g_2 differentiable if needed. Standard noncontractible generators $x_{a,1},...,x_{a,g}$ and contractible generators $y_{a,1},...,y_{a,g}$ are defined for $\partial g_1(M^a)$. Similarily, the loops $x_{b,1},...,x_{b,g}$, $y_{a,1},...,y_{a,g}$ are standard generators for $g_2(M_a)$.

Definition 2. *A contractible standard generator of a standard handlebody H_g of genus g is a loop:*

$$y_j = \{(x,y,z) \in \mathbb{R}^3 | x = 4j-2+(2+r)\cos(\theta), y = (2+r)\sin(\theta), |z| \leq \sqrt{1-r^2}, |r|=1, \theta=\pi/2\}$$

That is, y_j is the loop on the surface of the jth filled torus of H_g, has the center point $(4j-2,2,0)$ and radius 1 and is in the (y,z)-plane.

Definition 3. *A noncontractible standard generator of a standard handlebody H_g of genus g is a loop:*

$$x_j = \{(x,y,1) \in \mathbb{R}^3 | x = 4j-2+2\cos(\theta), y = 2\sin(\theta), 0 \leq \theta < 2\pi\}$$

That is, x_j is the loop on the surface of the jth filled torus of H_g, has the center point $(4j-2,0,0)$ and radius 2 and is in the (z,y)-plane.

Figure 2 shows a standard handlebody with standard generator loops.

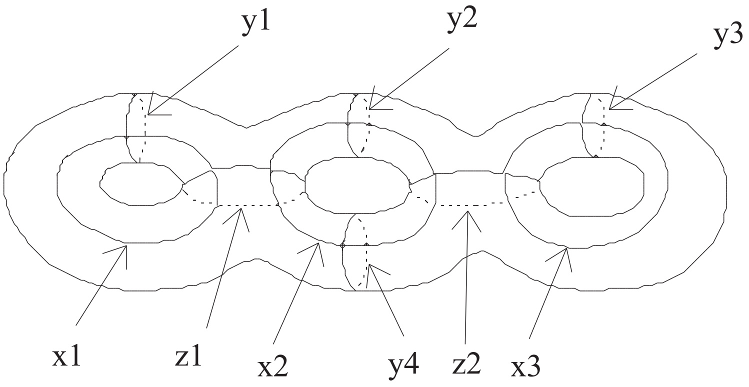

Figure 2.

Definition 4. *Let M be a closed 3-manifold. The 3-manifold M has a Heegaard split M_1, M_2 if there there are homeomorphisms g_1, g_2 of M_1 and M_2 to standard handlebodies $g_1(M_1)$ and $g_2(M_2)$ in $\mathbb{R}^3$ and a boundary identification homeomorphism ψ such that the following diagram commutes*

$$\begin{tikzcd}
M_1 \arrow[r, "g_1"] & g_1(M_1) \arrow[r, hook] & \mathbb{R}^3 \\
\partial M_1 \arrow[u] \arrow[d, "\approx"'] & \partial g_1(M_1) \arrow[u] \arrow[d, "\psi"'] & \\
\partial M_2 \arrow[d] & \partial g_2(M_2) \arrow[d] & \\
M_2 \arrow[r, "g_2"] & g_2(M_2) \arrow[r, hook] & \mathbb{R}^3
\end{tikzcd}$$

The boundary homeomorphism $\psi : g_1(M^a) \to g_2(M_a)$ induced by a Morse function has the property, that ψ and ψ^{-1} map loops to loops.

14 The simple proof

We will show that every differentiable simply-connected closed 3-manifold is homeomorphic to the 3-dimensional sphere S^3 as proving the claim for differentiable

3-manifolds proves it for all. The manifold M is assumed simply-connected and orientable.

We do not change the manifold M in the proof but construct different Morse functions on it. First we recall from a theorem of Smale that it is possible to find a Morse function which has only one critical point of index 0 and of index 3.

Lemma 1. *Let M be a closed simply-connected differentiable 3-manifold. There exists a Morse function $f : M \to \mathbb{R}$ such that f has one critical point of index 0 only, one critical point of index 3 only and on each critical level there is one critical point only.*

Proof. By Theorem of [2] there is a Morse function f such that f has one critical point of index 0 only and one critical point of index 3 only, and the number of critical points is finite. By small modification to the Morse function we can construct a Morse function which has the same critical points (and of same indices) as f but such that all critical points are on distinct levels. □

Another simple lemma is needed.

Lemma 2. *Let M be a closed simply-connected differentiable 3-manifold. If a Morse function $f : M \to \mathbb{R}$ has one critical point of index 0 only and one critical point of index 3 only, then for each level a ∂M^a is connected.*

Proof. If ∂M^a is not connected, then there exists a noncontractible loop passing the critical point of index 0, the critical point of index 3 and two points in two components of ∂M^a. □

We can change the Morse function in a tubular neighborhood of a path so, that in the new Morse function critical point of index 2 are on a higher level. We say, that a 2-handle is moved up by changing the Morse function. In a similar way we can construct a new Morse function where a selected critical point of index 1 is on a lower level. We say, that a 1-handle is moved down by changing the Morse function. The modification to the Morse function for moving a 2-handle up or a

1-handle down is local. We pull Morse levels through the handle as is shown in Figure 3.

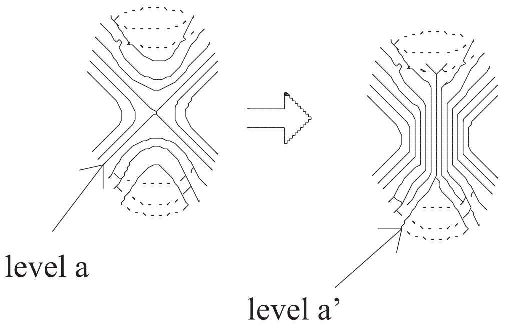

Figure 3.

Let us move one 1-handle to a so low level b that below it is only the 0-handle. Let $a = b + \epsilon$. The manifold $M_1 = M^a$ is a filled torus. The boundary of the manifold $M_a = M \int (M^a)$ is a torus. Inside the manifold M_a has an embedded exotic ball. We replace this embedded exotic ball by corresponding inside of a normal ball. Then M_a is changed to M_2, which is a filled torus. The identification map of the boundary ∂M^a and ∂M_a defines a glueing map g_1 from the boundary of the filled torus M_1 to the boundary of the filled toros M_2. That is, M_1, M_2 is a Heegaard split of genus 1 of $M_1 \cup M_2$.

The manifold $M_1 \cup M_2$ is a simply-connected closed 3-manifold. It is known that if a genus 1 Heegaard split gives a simply-connected closed manifold, then the glueing map is trivial, that is, it maps a standard noncontractible generator of the torus M_1 to a standard contractible generator of M_2. It means that we can find a 2-disc $D \in M_2$ bounding the contractible generator in M_2, and it is a standard noncontractible generator in M_1, a loop around the 1-handle.

We can remove this 1-handle of M_1 by adding the thickened 2-disc (i.e., a closed cylindar neighborhood of the disc) to M_1 and removing its interior from M_2. This operation changes the Morse level a so that on a the manifold M^a now contains the 2-disc and there is no critical point on the level b. The 1-handle disappears, and as there always must be equally many 1-handles and 2-handles, one 2-handle also disappears. That is, adding the thickened 2-disc to M^a adds the thickened 2-disc to every Morse level $a' > a$, and there will be no need to one 2-handle.

Performing this operation of moving a 1-handle down to a level where it is the lowest above the 0-handle, finding the 2-disc in M_2 which bounds a non-contractible standard generator in M_1, and adding the thickened 2-disc to M^a, we remove one pair (1-handle, 2-handle) every time. As there are only finitely many 1-handles, we end up with a Morse function that has no 1-handles and no 2-handles, yet the manifold M is closed and simply-connected. This implies that the manifold M must be a real sphere, not an exotic sphere. Thus, assuming we can do this operation, the Poincaré Conjecture is proven.

There is one problem left. The manifold M_a is not M_2. The manifold M_a has an embedded exotic homotopy ball. If the 2-disc in M_2 does not intersect the exotic homotopy ball, then we can do as described above, but let us assume the 2-disc cuts through the embedded exotic homotopy ball. We have to move the homotopy ball away from the 2-disc in some manner.

Notice that a 2-handle is a 1-handle for the Morse function $-f$. That is, in the direction of $-f$ the manifold M_2 has a 1-handle and the 2-disc of M_2 cuts this 1-handle of the Morse function $-f$. The embedded exotic ball goes through this 1-handle of M_a. Figure 3 illistrates this:

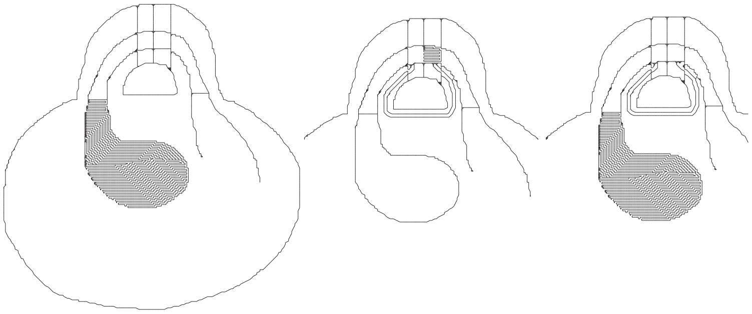

Figure 3.

The exotic homotopy ball must be moved away from the 1-handle of M_2 that is cut by the 2-disc. Then we can remove the thickened 2-disc from M_2 and the proof of the Conjecture is complete.

Figure 3 shows what is done in Lemma 3. The blackened part is a part of an exotic homotopy ball that goes through a 1-handle. We connect the parts of the embedded exotic homotopy ball by a (by-pass) tube that does not go through the 1-handle. The moved part of the embedded exotic homotopy ball is inside this tube. This is a form of Kirby surgery. We show in Lemma 3 that the embedded exotic ball is not essentially changed by this by-pass surgery.

There are wild embeddings and one may wonder if an exotic homotopy ball might be wildly embedded and we could not find a tube like in Figure 3. This is not a problem because the level a can be changed. The embedding of the exotic homotopy ball does not depend on a and we can select a slightly smaller a. For this new a there is area close to the boundary of M_a where there is no part of the exotic homotopy ball and we can find the tube, but as the following lemma says, the exotic homotopy ball cannot be wildly embedded as it is of the dimension 3 in a 3-dimensional manifold.

The lemma has an unnecessarily complicated way of moving the exotic homotopy ball. One tube is enough, but as this paper is only of historical interest and not to be published, I keep the lemma as it was in one of the versions of this paper.

Lemma 3. *Assume that a 3-manifold M_1 (with boundary) is obtained from a 3-manifold M_2 by addition of a 1-handle $H \approx D^2 \times I$, $M_1 = M_2 \cup H$. Assume, that M_1 and M_2 are orientable. If M_1 contains an exotic homotopy ball B_e^3. Then M_2 contains an embedded exotic homotopy ball ${B'}_e^3$.*

Proof. The embedding of B_e^3 to M_1 is not wild as both manifolds M_1 and B_e^3 have the same dimension 3. Wild embeddings happen only with submanifolds of lower dimension. As B^3_e is not wildly embedded, $\partial B^3_e \approx S^2$ is not a wildly embedded sphere.

Let F be a smoothly embedded disc cutting the 1-handle H. F can be assumed to be in a general position so, that $F \cap \partial B^3_e$ is a set of disjoint circles $A_0 = \{C_i | i \in I_0\}$, $C_i \approx S^1$, i_0 finite.

The set A_0 contains a nonempty subset $B_0 = \{S_j | j \in I_1\} \subset A_0\}$ where each $S_j \subset \partial B^3_e$ is a boundary of a disc $D_j \subset \partial B^3_e$ satisfying $int(D_j) \cap A_0\} = \emptyset$. In order to see that B_0 is nonempty, notice, that each C_i separates $\partial B^3_e \approx S^2$. There must be such circles on ∂B^3_e that one of the separated sides does not contain smaller circles. Such a circle C_i is a circle S_j.

Take one S_j and a circle $S'_j \subset F$ which is sligthly bigger than S_j. There is a disc $D'_j \subset F$ such that $\partial D'_j = S'_j$. There may be other circles C_i than S_j inside D'_j but that does not matter here. Replace the disc D'_j by a disc D''_j which is close to D_j. Then D_j, D''_j and the annulus $F_0 \subset F$ between the circles S_j and S'_j separates M_1 and one side is a 3-ball B_0, it is the side which does not contain points of $int(B^3_e)$. The disc D''_j separates M_2 into two components. Let U be the component which contain points of $int(B_0)$.

Let V be a closed neighborhood of F in H and V_1 a smaller closed neighborhood of F in V. Then F separates V and V_1 and ∂V is the union of a collar, which is subset of ∂H, and two discs D_a, D_b which have boundary at ∂H.

By construction D''_j and D_j insersect with F only at S'_j and S_j. Therefore we can select V so small, that the component in $V \cap B^3_e$ which contains D_j is a collar $C = c(S^1 \times I)$, for an embedding $c : S^1 \times I \to H$ such that $C \subset V$, $c(S^1 \times \{0\}) \subset D_a$, $c(S^1 \times \{1\}) \subset D_b$, $C \cap F = S_j$, $C \subset \partial B^3_e$.

Similarily we can find a collar $C' = c'(S^1 \times I)$, for an embedding $c' : S^1 \times I \to H$ such that $C' \subset V$, $c'(S^1 \times \{0\}) \subset D_a$, $c'(S^1 \times \{1\}) \subset D_b$, $C' \cap F = S'_j$, $C' \cap \partial B^3_e = \emptyset$ and $D''_j \cap V \subset C'$. $D''_j \cap V$ is on one side of F in V, let that be the side where D_a is.

Let us define a homeomorphism $g' : M_1 \to M_1$ so, that g' keeps the disc D''_j fixed, expands the 3-ball B_0 to B'_0so, that the disc D_j is pushed to the component K_b of $V \int V_1$ which contains D_b. The mapping g' pushes a part A of the homotopy ball B_e^3 which was in M_2 to K_b.

We must move this part of B_e^3 back to M_a but not through F. Let l be a path from a point in $V_1 \cap K_b$ to a point in $V_1 \cap K_a$ which does not go through F. We can select l so that it does not intersect with B_e^3 as $\partial B_e^3 \approx S^2$. Let V_2 be a small closed tubular neighborhood of l in M_1. We connect the tube V_2 to A and move A through the tube V_2 to the other component K_a of $V \int V_1$ by expanding a 3-ball into K_b. Finally we restore A back to M_2 by decreasing the 3-ball in U. Now the tube V_2 is filled with a 3-ball, M_2 is restored and the handle H is again a handle. Let us call this homeomorphism of M_1 to M_1 by g''. See Figure 3 where the mapping is shown.

Define g as the combined homeomorphism $g : M_1 \to M_1$, $g(p) = g''(g'(p))$. Then $g(M_2) = M_2$.

Let us rename the exotic homotopy ball $g'(B_e^3)$ as B_e^3.

There are now less circles in the intersection set $A_1 = F \cup \partial B_e^3$. If A_1 is nonempty, let us return to the step of selecting one S_j, repeat the procedure and get a smaller set A_2.

We repeat the procedure as long as $F \cup \partial B_e^3 \neq \emptyset$. When the set is empty, we have obtained an exotic homotopy ball which is contained in $M_2 \cup K_1 \cup K_2 \approx M_2$. So, there is an exotic homotopy ball B'^3_e in M_2. □

The Poincaré Conjecture is proven.

Theorem 1. *Every simply connected closed 3-manifold is homeomorphic to the 3-sphere.* □

The first generally accepted proof of Theorem 1 was given by Grigory Perelman in the year 2002. This my simple proof from 1987 was not accepted, but for thirty three years I have considered it as a correct proof. Some proofs are checked, some are not checked. That is simply so.

References

1. J. Hempel, 3-manifolds, Princeton University Press, Princeton, N,J.,1976
2. S. Smale, Generalized Poincaré's Conjecture in dimensions greater than 4, Annals of Math.,74 (1961), 391-406

On the existence of polynomial-time algorithms to the subset sum problem

Abstract. This paper proves that there does not exist a polynomial-time algorithm to the the subset sum problem. As this problem is in NP, the result implies that the class P of problems admitting polynomial-time algorithms does not equal the class NP of problems admitting nondeterministic polynomial-time algorithms.

Key words: computational complexity, polynomial-time, algorithm, knapsack problem.

15 Introduction

Let $\mathbb{N}$ and $\mathbb{R}$ indicate natural and real numbers respectively.

Definition 1. *A knapsack is a pair of the form $(j, (d_1, \ldots, d_n))$ where $j, n \in \mathbb{N}$, $j, n > 0$ and $d_k \in \mathbb{N}$, $d_k > 0$ for $1 \leq k \leq n$.*

The knapsack problem means the following: given a knapsack $(j, (d_1, \ldots, d_n))$ determine if there exist binary numbers $c_k \in \{0, 1\}$, $1 \leq k \leq n$, such that

$$j = \sum_{k=1}^{n} c_k d_k.$$

Let $B, \alpha \in \mathbb{R}$, $B \geq 1$, $\alpha \geq 0$ be fixed numbers. An algorithm A is called polynomial-time algorithm to the knapsack problem if there exist numbers $C, \beta \in \mathbb{R}$ that depend on B and α but not on n such that the following condition is true: For any sequence of knapsacks of the form

$$((j_n, (d_{1,n}, \ldots, d_{n,n})))_{n \geq 1}$$

satisfying

$$\log_2 j_n < Bn^{\alpha},\ \log_2 d_{k,n} < Bn^{\alpha},\ (1 \leq k \leq n), (n \geq 1) \tag{1.1}$$

the number N_n of elementary operations that the algorithm A needs to produce an answer *yes* or *no* to the question if there exists binary numbers $c_{k,n} \in \{0,1\}$, $1 \leq k \leq n$, such that

$$j_n = \sum_{k=1}^{n} c_{k,n} d_{k,n} \tag{1.2}$$

satisfies $N_n < Cn^{\beta}$ for all $n \geq 1$.

The problem that has been described is used in the Merkle-Hellman knapsack cryptosystem and today it is commonly known as the knapsack problem. The name Subset sum problem is used for it in [2] p. 301, while the name Knapsack problem is reserved for a more general problem involving selecting objects with weights and profits. The name *knapsack* is more convenient than *subset sum* and it is ofen used in this paper.

In the definition of a polynomial-time algorithm for the knapsack problem we have included an upper bound on j_n and on each $d_{k,n}$, $1 \leq k \leq n$. Such bounds are necessary for the following two reasons (i) and (ii).

(i) The number m of bits in the binary representation of j_n satisfies $m \leq \log_2 j_n < m+1$. Thus, if $\log_2 j_n$ grows faster than any polynomial as a function of n then so does the length of j_n in the binary representation. It is necessary to verify that (1.2) is satisfied. It requires making some operations (like compare, copy, read, add, subtract, multiply, divide, modulus) that act on a representation of j_n on some base number. We may assume that the number base is 2 as changing a number base does not change the character of the algorithm from polynomial-time to non-polynomial-time. Any operations that require all bits of j_n must require more than a polynomial number of elementary operations from any algorithm A if the number of bits in j_n grows faster than any polynomial. Similar comments apply to $d_{k,n}$.

(ii) If j_n has an upper bound independent of n, then there exist a polynomial-time algorithm solving the knapsack problem. The Annex gives one such algorithm in Lemma A2. The algorithm in Lemma A2 calculates an exponentially growing number of combinations of c_k in the same polynomial time run.

Because of (i) and (ii) j_n must grow polynomially with n. We can select j_n as growing linearly as in (1.1). It gives an NP-complete knapsack problem.

Remark 1. Lemma A2 in the Annex solves all possible values of $j_n < Bn^\alpha$ with the same polynomial time run of Algorithm A0 because j_n is not used in A0 before checking the final result $b_{n,k}$. Let us consider the case when j_n is not limited from above by a polynomial of n. Lemma A1 runs in polynomial time even if the upper bound for j_n grows faster than a polynomial of n but it does not produce results that can tell if there exists a solution for a particular value j_n. A polynomial-time test, such as taking a modulus in (A1), maps the superpolynomial set of possible values of $j = \sum_{k=1}^{n} c_{k,n} d_{k,n}$ into a polynomial number of classes. In (A1) the classes are all sums j with the same moduli by r_n. At least one such a class corresponds to an superpolynomial number of values j. In order to check if any value j in the class equals j_n the algorithm should in some way check all of the values j in the class, but if the algorithm at the same run checks all values of j then it should in some way loop over a superpolynomial set which is not possible for a polynomial time algorithm. In general, we can say that a single polynomial time run of an algorithm cannot solve all values of j_n that are below a superpolynomial upper bound because the algorithm can only produce a polynomial number of results and there exist a superpolynomial number of possible values j_n. A polynomial time algorithm that solves the subset sum problem for any value j_n below a superpolynomial upper bound must limit search and there must be values j_n that are solved with different runs of the algorithm.

Remark 2. An algorithm is a finite set of rules that at every step tell what to do next. We can implement an algorithm as a computer program in a second generation language on a von Neumann machine and a polynomial time algorithm can be implemented in this way so that it requires time and memory that grow polynomially with respect to the problem dimension. In the case when the smallest upper bound of j_n in Remark 1 grows exponentially a program in a second generation computer language implementing a polynomial time algorithm needs to limit search by branching instructions, or by acting differently on different data (like in add, subtract and compare instructions). Thus, we can find values of j_n such that the algorithm uses different branches, or acts differently on data, in solving the subset sum problem.

16 The inequality (2.6) means non-polynomial time

It is not possible to select a fixed sequence of specific subset sum problems and show that no algorithm can solve this specific sequence of problems in polynomial time. This is so because we can create an algorithm that treats these specific problems in a particular way and can solve that specific sequence of problem in a fast way. Instead, we must first select the algorithm and pose that selected algorithm a sequence of subset sum problems that are particularily hard for that specific algorithm. As the algorithm can be any possible algorithm, the sequence of problems can only be defined by using some suitable definition of a difficult problem to the selected algorithm and we cannot give any numerical values for all of the numbers $c_{k,n}$ in (1.2). The selection will be done by using the following definition of the computation time of a subset sum problem.

For convenience, let us select n to be of the form $n = 2^{i+2}$ for some $i > 0$. This simplifies expressions since it is not necessary to truncate numbers to integers.

Definition 2.

We define a function $f(n)$ that describes (in a certain sense) the worst computation time for a selected algorithm.

Let the worst in the median n-tuple as be defined as follows. Let

$$h(d_{1,n},\ldots,d_{n,n},j_n)$$

be the computation time for deciding if the knapsack

$$(j_n,(d_{1,n},\ldots,d_{n,n}))$$

has a solution or not. Let

$$Median_{j_n}\ h(d_{1,n},\ldots,d_{n,n},j_n) \tag{2.1}$$

be the median computation time where j_n ranges over numbers

$$j_n \in \{C+1,\ldots,2^{n+1}-1\} \tag{2.2}$$

satisfying the two conditions

$$j_{n,l} = j_n - C\left\lfloor \frac{j_n}{C} \right\rfloor > 2^{\frac{n}{4}+2} \tag{2.3}$$

where $C = 2^{\frac{n}{2}+1}$, and that there is no solution to the knapsack $(j_n,(d_{1,n},\ldots,d_{n,n}))$. That is, $j_{n,l}$ are the lower half bits if j_n. The values of j_n are computed separately in calculation of the median, i.e., no partial results from previously computed values of j_n are used.

Let $(d_{1,n},\ldots,d_{n,n})$ range over all knapsack sequences with

$$\lceil \log_2 \sum_{k=1}^{n} d_{k,n} \rceil = n$$

and $d_{k,n} \le \frac{2^n-1}{n}$. Because of this requirement at most every second value of j_n in (2.2) is a solution to the knapsack, i.e., there are 2^n combinations of $(c_{1,n},\ldots,c_{n,n})$ mapped to numbers from zero to $2^{n+1}-1$. The worst in the median tuple for n

is an n-tuple $(d_{1,n}, \ldots, d_{n,n})$ (possibly not unique) that maximizes the median computation time (2.1).

Let this maximal median computation time be denoted by $f(n)$. Thus

$$f(n) = \max_{d_{1,n},\ldots,d_{n,n}} Median_{j_n}\ h(d_{1,n}, \ldots, d_{n,n}, j_n). \tag{2.4}$$

We use the median in Definition 2 instead of the worst case or the worst in the average case because we need $\frac{n}{2}$ almost as long computations as the worst in (2.6). In the worst and in the worst in the average, a very slow computation of one value j_n can be the reason for the long computation time. By using the median we can find many values j_n giving almost the median computational time because the distribution of the computational time for j_n becomes almost normally distributed when n grows due to the law of large numbers. We include only unsuccessful cases of j_n in the computation of the median because this choice implies that a more complicated knapsack problem (i.e., more cases to check) gives a longer computation time. If there are more cases to choose, there are more successful cases. Therefore the time for finding a solution decreases if there are more cases to check.

Lemma 1. *Let m be fixed and n be a power of m. If $f(n)$ satisfies the inequality*

$$\frac{n}{m} f\left(\frac{n}{m}\right) < f(n) \tag{2.5}$$

then $f(n)$ does not grow polynomially with n.

Proof. Iterating we get

$$\frac{n}{m}\frac{n}{m^2} f\left(\frac{n}{m^2}\right) < f(n)$$

and iterating up to k yields

$$\frac{n^k}{m^{\sum_{i=1}^{k} i}} f\left(\frac{n}{m^k}\right) < f(n)$$

i.e.,

$$e^{k\ln n-\frac{1}{2}k^2\ln m-\frac{k}{2}\ln m}f\left(\frac{n}{m^k}\right)<f(n).$$

Setting $k=\frac{\ln n}{\ln m}$ gives

$$\left(n^{\ln n}\right)^{\frac{1}{2\ln m}}n^{-\frac{1}{2}}f(1)<f(n).$$

If m is any fixed number we see that $f(n)$ satisfying (2.5) is not bounded by a polynomial function of n. □

Lemma 2. *Let n be a power of 2. If $f(n)=f_1(n)+f_2(n)$ where $f_1(n)$ is a polynomial function of n and $f_2(n)$ satisfies the inequality*

$$\frac{n}{2}f_2\left(\frac{n}{2}\right)<f_2(n) \qquad (2.6)$$

then $f(n)$ does not grow polynomially with n.

Proof. If $f(n)$ is a polynomial function of n and since $f_1(n)$ is a polynomial function of n by assumption, it follows that $f_2(n)$ must also be a polynomial function of n. By Lemma 1, $f_2(n)$ is not a polynomial function of n, thus neither is $f(n)$. □

17 Construction of a special subset sum problem

In this section we will define a special subset sum problem K_{1,j_n} in Definition 3 and show that it can only be solved by solving $n_1=n/2$ subknapsacks $(j_i', (d_{1,n},\dots,d_{n_1,n}))$ with different values of j_i'. We will use the denotation $n_1=n/2$ throughout this article for brevity.

Definition 3. Construction of K_{1,j_n}. We first make a knapsack where the only solutions must satisfy the condition that exactly one c_k must be 1 and the others must be zero for $k=n_1+1$ to $k=n$. Let us construct the values $d_{k,n}$,

$k = n_1 + 1, \ldots, n$ of K_{1,j_n} for a given j_n. Let $C = 2^{\frac{n}{2}+1}$ and

$$j_{n,h} = C \left\lfloor \frac{j_n}{C} \right\rfloor \ , \ j_{n,l} = j_n - j_{n,h} \tag{3.1}$$

be the high and low bit parts of j_n. Because of (2.2), $j_{n,h} \neq 0$. Let

$$d_{n_1+k,n} = j_{n,h} + a_k \tag{3.2}$$

where $0 < a_k < \min\{j_{n,l}, \frac{2^{n_1}-1}{n_1}\}$ are distinct integers and there exists no solution to the knapsack problem for the knapsack

$$(j'_i, (d_{1,n}, \ldots, d_{n_1,n}))$$

where

$$j'_i = j_{n,l} - a_i. \tag{3.3}$$

Let us also require that the computation time for j'_i is at least as long as the median computation time $f(n_1)$ for $(j, (d_{1,n}, \ldots, d_{n_1,n}))$. We can select j'_i filling this condition because half of the values j are above the median. Notice that we compute the median only over values j that do not give a solution to the knapsack. We will also assume that the j'_i are in the set corresponding to (2.2)-(2.3) for $f(n_1)$, i.e.,

$$j'_i \in \{C' + 1, \ldots, 2^{n_1+1} - 1\} \tag{3.4}$$

satisfying the condition

$$j'_i - C' \left\lfloor \frac{j'_i}{C'} \right\rfloor > 2^{\frac{n}{8}+2} \tag{3.5}$$

where $C' = 2^{\frac{n_1}{2}+1}$. We may assume so because there are enough values from which to choose j'_i.

In (3.2) we select the numbers a_k in such a way that the $d_{n_1+k,n}$ satisfy the size condition $d_{n_1+k,n} \leq \frac{2^n-1}{n}$. Because of the bound (2.3) we have an exponential number of choices for a_i. It is possible to find numbers j_i' such that there is no solution since only for about half of the values of j there exists a solution for $(j,(d_{1,n},\dots,d_{n_1,n}))$. If $j_{n,l}$ is too small and we cannot find values j_i', we take a carry from $j_{n,h}$ in (3.3) and reselect a_k. Because of the lower bound on j in (2.2), $j_{n,h}$ is not zero and we can take the carry. Then $j_{n,h}$ is decreased by the carry.

Exactly one c_k must be 1 and the others must be zero for $k=n_1+1$ to $k=n$. There cannot be more values $c_k=1$ for $k>n_1$ because then the higher bits of j_n are not matched. The unknown algorithm can try also other combinations but these are the only possible combinations and the algorithm must also try them (i.e., check these cases in some way unknown to us). The sum of the numbers $d_{k,n}$, $k\le\frac{n}{2}$ is less than $2^{\frac{n}{2}+1}-1$. Adding one c_k can give a carry and there may not be a solution to the knapsack because the high bits of j_n do not match but this is not an issue since we do not want solutions. We select the n-tuple so that there are no solutions to the knapsack already because the lower bits do not match.

Lemma 3. *The algorithm cannot stop to finding a solution because for every j_n none of the $\frac{n}{2}$ values of j_i' solve the knapsack problem. Every value j_i' gives at least as long computation as the median computation time $f(n_1)$.*

Proof. We have selected K_{1,j_n} such that $(j_i',(d_{1,n},\dots,d_{n_1,n}))$ has no solution for any j_i'. Thus the algorithm cannot stop because it finds a solution. By construction the values j_i' give at least as long computation time as the median for the tuple at $k=1,\dots,n_1$. Since that tuple is the worst in the median tuple for n_1, the computation time for each j_i' is at least $f(n_1)$. □

Lemma 4. *There is no way to discard any values j_i' without checking if they solve the subknapsack from $k=1$ to $k=n_1$. Any case of using the values of $d_{k,n}$ in order to get the result is considered checking.*

Proof. We can select any a_k in such a way that there either exists a solution or does not exist. Knowledge from other $c_{i,n}$ $(i\neq n_1+k)$ cannot give any information on how this a_k was selected. Thus, the existence of a solution must be checked using the value $d_{n_1+k,n}$. □

Lemma 5. *Several values of j_i' cannot be evaluated on the same run. The median computation time of K_{1,j_n} is at least*

$$f_1(n_1) + n_1 f_2(n_1)$$

where $f(n) = f_1(n) + f_2(n)$ is a lower bound for the computation time of one j_i' and $f_1(n)$ is a polynomial function of n, the shared part of the computation time of all j_i'.

Proof. As explained in Remark 1, a polynomial time algorithm cannot solve all values of j_i' at the same run because it would require an exponential amount of memory. As explained in Remark 2, we can assume that the algorithm is implemented in a second generation computer language on a von Neumann machine and its code has branching instructions, or it acts differently on different data in an instruction (like add depends on the data), which has the same effect as a branching instruction: for a different j_n there is needed a different run. These branching instructions define a branching tree describing the execution of the algorithm for any input data. The tree is fixed when the algorithm is selected. At each branching point the input data is divided into a finite number of classes. Because this division is fixed, we can always find two values j_i' which are not executed by the same polynomial time run. After finding two, we can continue to find three values j_i' which all are executed by different polynomial time runs of the algorithm. This can be extended to $\frac{n}{2}$ values j_i': we can select j_i' in such a way that no two values j_i' are computed in the same run. The runs for different values j_i' can have parts that are shared, as long as the shared parts are computed in polynomial time. This is necessarily the case for practical algorithms: the runs must share at least the beginning of the code before branch instructions are reached and this shared part must take only polynomial time for the algorithm to make any sense. The shared part of the computation time can be described by a polynomial function $f_1(n)$ and a lower bound for the nonshared computation time can be denoted by a function $f_2(n)$. □

18 Proving the inequality (2.6)

Let the algorithm be chosen. We selected a tuple K_{1,j_n} for a chosen j_n and showed in Lemma 5 that the computation time for the set of K_{1,j_n} for the single value j_n is at least as high as the left hand side of (2.6). We have obtained the left side of the inequality (2.6) for an arbitrarily chosen algorithm solving the knapsack problem. However, the set of K_{1,j_n} is a (reasonably) hard problem only for the chosen value j_n. Let us call this j_n with the name j_{n0}. In the right side of (2.6) the number j_n must range over all values and we calculate the median computation time over those values of j_n where there is no solution. In $K_{1,j_{n0}}$ it is very fast to conclude that most values for j_n do not have a solution: it is usually enough just to check the bits of j_n in the most significant half of the number. If they do not match the most significant bits of j_{n0}, then there is no solution.

We want to change the knapsack problem $K_{1,j_{n0}}$ to another knapsack problem K_2 (the problem K_2 will be defined later in Definition 5) where j_n can range over all numbers and for many values of j_n there is no solution and the knapsack problem is difficult. The knapsack problem K_2 has at most as long median computation time as the worst in the median tuple for n because the worst is the worst.

We will do the change in two steps. First we change $K_{1,j_{n0}}$ to $K_{3,j_{n0}}$ where the bits in the lower half of j_n can vary. In the second step we change $K_{3,j_{n0}}$ to K_2 where also the upper half bits of j_n can vary. What we have to show is that the computaton time of the set $K_{1,j_{n0}}$ with a single $j_n = j_{n0}$ is not larger than the median computation time for $K_{3,j_{n0}}$ when j_n can have any lower half bits. In $K_{3,j_{n0}}$ only one $d_{j,n}$, the one with $j = n$, has the most significant bits of j_{n0}. Therefore $c_{n,n}$ must be one in order to have a possibility of finding a solution for j_n

that has the high bits of j_{n0}. We put some numbers to $d_{j,n}$ for $j = n_1+1, \ldots, n-1$. These numbers have zero high bits. There are more combinations that can give a solution in $K_{3,j_{n0}}$ than in $K_{1,j_{n0}}$, thus it is easier (and faster) to find a solution, provided that there is a solution for a chosen j_n. The trick here is that in the calculation of the median computation time we take only those j_n where there is no solution. Then the fact that there are more possible combinations only makes it harder to conclude that there is no solution. We conclude in Lemma 6 that the median computation time for $K_{3,j_{n0}}$ when the lower half bits of j_n vary is larger than the computation time of $K_{1,j_{n0}}$.

Next we have to show that K_2 gives a larger median computation time when j_n varies over all numbers than $K_{3,j_{n0}}$ when the bits of the lower half of j_n vary. It is a similar situation here: there are more combinations in K_2 that can give a solution for a given j_n, but only those j_n that give no solution are counted in the median computation time. Therefore adding complexity makes the median computation time longer. In K_2 we replace $d_{n,n}$ of $K_{3,j_{n0}}$ by a difficult knapsack problem in the upper half bits. As this difficult knapsack problem in the upper half has n numbers $d_{j,n}$ and the bit length of each $d_{j,n}$ is only $n/2$, there usually always are solutions to the upper half knapsack problem. Looking at the upper half knapsack problem does not help in finding values j_n that give no solution to the knapsack problem K_2. Because of this, the knapsack problem K_2 is not any easier than the knapsack problem $K_{3,j_{n0}}$.

Figure 1 shows the main idea.

In Figure 1 the set $K_{1,j_{n0}}$ has the worst in the median n_1-tuple in the left side and the right side has numbers from which it is necessary to select exactly one in order to satisfy the high bits of j_{n0}. This yields n_1 separate subset sum problems and we get the computation time corresponding to the left side of (2.6). The set $K_{3,j_{n0}}$ has only one element which has high order bits and it must always be selected in order to satisfy the high bits of j_n. Here the bits of the upper half of j_n are the same as in j_{n0}. There is the same worst in the median n_1-tuple and

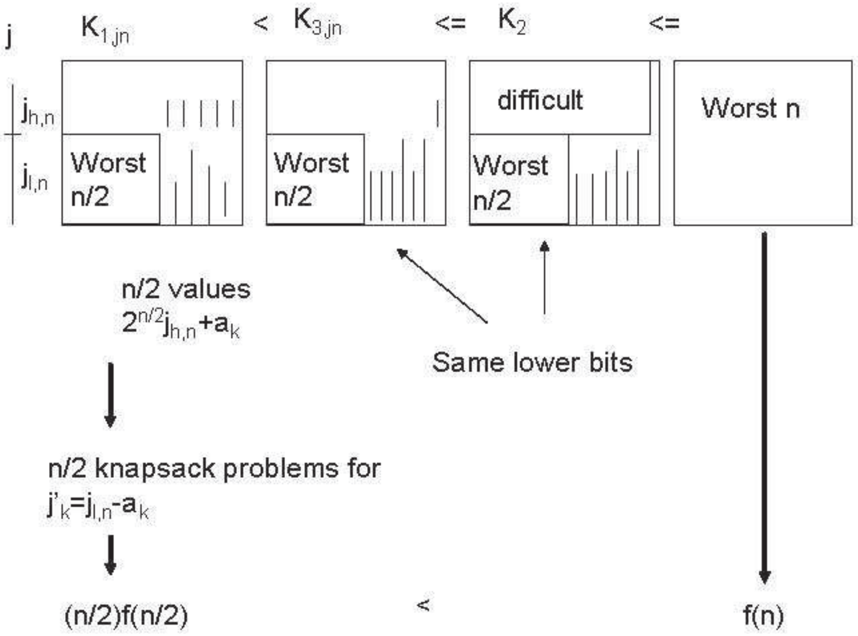

Fig. 1. The idea of the proof.

the remaining $n_1 - 1$ elements can be assigned in any way yielding of the order n^2 knapsack problems. It is easier to find a solution than in $K_{1,j_{n0}}$, but it is harder to conclude that there are no solutions. Lemma 6 shows that the time of solving K_{1,j_n} is not higher than the median computation time for K_{3,j_n} for almost any j_n that does not yield a solution.

The n-tuple K_2 has some difficult upper half knapsack problem which has to be satistifed with the same values c_k as the lower half knapsack. It is not of any use to check if the upper half knapsack half has a solution when trying to show that there is no solution to the whole knapsack since there almost always are many solutions to the upper half knapsack problem. The algorithm must look at all bits. As finding a solution in K_2 requires looking at both the upper and lower half bits, it should be more difficult to conclude that there are no solutions. We will show that at least it is not faster. Finally, the inequality from K_2 to the worst in the median n-tuple is obtained directly by the definition of what the worst means.

Definition 4. Construction of $K_{3,j_{n0}}$. Let j_n be given and let us define a n-tuple $K_{3,j_{n0}}$ as an n-tuple with elements $(d_{1,n,3},\dots,d_{n,n,3})$ by specifying the elements

$$\begin{aligned} d_{k,n,3} &= d_{k,n} \quad (k=1,\dots,\frac{n}{2}) \\ d_{k,n,3} &= e_1 \quad (k=\frac{n}{2}+1,\dots,\frac{3n}{4}) \\ d_{k,n,3} &= e_2 \quad (k=\frac{3n}{4}+1,\dots,n-1) \\ d_{n,n,3} &= j_{n0,h}. \end{aligned} \tag{4.1}$$

We select two nonnegative integers $e_i \le \frac{2^{n_1}-1}{n_1}$, $i=1,2$. The selected e_1 and e_2 are so small that if $c_n=0$ the higher bits of j_n are not matched because there is no carry. That is, the worst in the median knapsack for $n_1=n/2$ is still in the left side. The high bits of j_{n0} are in $d_{n,n,3}$. We choose some numbers to the elements $d_{k,n,3}$ for $k=n_1+1,\dots,n-1$.

This n-tuple has a simple upper half tuple. The sum of the numbers $d_{k,n,3}$, $k\le\frac{n}{2}$ is less than $2^{\frac{n}{2}+1}-1$. It is always necessary to set $c_n=1$ and this satisfies the upper half bits of j_n when j_n ranges over numbers that have the same upper half bits as j_{n0}.

Definition 5. Construction of K_2. We will define K_2 as an n-tuple with elements $(d_{1,n,2},\dots,d_{n,n,2})$. Let us remember that the n-tuple $(d_{1,n},\dots,d_{\frac{n}{2},n})$ is the worst in the median tuple for $\frac{n}{2}$. Let $(d_{0,1},\dots,d_{0,n})$ be an n-tuple where each $d_{0,k}\le\frac{2^{n_1}-1}{n_1}$. We define

$$d_{k,n,2} = Cd_{k,n,2} + d_{k,n} \tag{4.2}$$

for $k = 1, \ldots, n_1$. The numbers e_1 and e_2 are as in $K_{3,j_{n0}}$ and we define the elements of K_2 for $k = n_1 + 1$ to $k = n$ as

$$d'_k = Cd_{0,k} + e_1 \quad (k = \frac{n}{2} + 1, \ldots, \frac{3n}{4})$$

$$d'_k = Cd_{0,k} + e_2 \quad (k = \frac{3n}{4} + 1, \ldots, n-1) \tag{4.3}$$

$$d'_n = Cd_{0,n.}$$

Thus, K_2 has the same lower half tuple elements as K_{3,j_n} and in the upper half there is the n-tuple $(d_{0,1}, \ldots, d_{0,n})$. In this definition we do not specify the n-tuple $(d_{0,1}, \ldots, d_{0,n})$, but it will be chosen as a sufficietly difficult n-tuple.

In K_{3,j_n} our chosen algorithm may fast find a solution and stop for any j_n, but we are only interested at such j_n that give no solution. The tuple K_2 can be split into two n-tuples: the lower half tuple with elements smaller than C and the upper half tuple that has the higher bit parts. In K_2 the algorithm usually does not stop to a solution of the lower half tuple since the upper half tuple is usually not satisfied by c_k that satisfy the lower half knapsack.

Lemma 6. *The time for the chosen algorithm to solve $K_{1,j_{n0}}$ is not larger than the median computation time for the algorithm for solving $K_{3,j_{n0}}$ when j_n ranges over all values where $j_{n,h} = j_{n0,h}$.*

Proof. In $K_{3,j_{n0}}$ the indices $k > n_1$ give $\frac{(n+4)n}{16}$ values of j for a knapsack problem in the indices $k \leq n$. Let us name these values j'_i where i, $i = 1, \ldots, \frac{(n+4)n}{16}$.

In the indices $k = 1, \ldots, n_1$ there is the worst in the median n_1-tuple. The values j'_i that we get are a sample of all possible values j_{n_1} for the knapsack problem for this worst in the median n_1-tuple.

Half of all possible values of j_{n_1} yield a longer computation time than $f(n_1)$ in the worst in the median knapsack problem for n_1 because $f(n_1)$ is the median computation time. If the values of j'_i that we get are a representative sample of all

j_{n_1}, then about half of the values of j_i' that do not give a solution yield a longer computation time than $f(n_1)$.

We can select e_1 and e_2 from an exponential set of numbers. Therefore we can assume that the numbers j_i' are sufficiently well randomly distributed over the possible range of the numbers j_{n_1} for the knapsack problem for $n_1 = n/2$ and they are a representative sample of all numbers j_{n_1}.

Also, because the numbers j_i' are sufficiently randomly distributed over all possible values of j_{n_1} we may assume that about half of the values j_i' are on the range (3.4).

There are more values j_i' to check in $K_{3,j_{n0}}$ than the $n/2$ in $K_{1,j_{n0}}$. If there is no solution for some j_n, then it is necessary to check all j_i' before the algorithm can conclude that there are no solutions. Therefore the computation time of the chose algorithm to solve $K_{1,j_{n0}}$ is not longer than the median computation time for the algorithm to compute $K_{3,j_{n0}}$ when j_n ranges over all numbers that have $j_{n,h} = j_{n0,h}$. □

The median computation time in (2.1) is calculated over the *no* instances only. Thus, *yes* instances are ignored. It is sufficient that there are at least some *no* instances so that (2.1) can be calculated. We give an argument that estimates the number of solutions to the knapsack problem (j_n, K_2). The argument makes use of averages but it is quite sufficient for showing that there are some *no* instances for computation of (2.1) if the upper bits of K_2 are selected in a suitable way, indeed a random selection of these bits is likely to yield many *no* instances.

Lemma 7. *There are in average $2^{\frac{n}{2}}$ solutions possible choices of $(c_1, \ldots, c_n)$ that give the same sum $\sum_{k=1}^{n} c_k d_{o,k}$.*

Proof. The number of combinations of c_k is 2^n and the sum $\sum_{k=1}^{n} d_{o,k}$ is at most $2^{\frac{n}{2}}$. There are fewer combinations that yield very small or large sums and most sums are in the middle ranges. □

Lemma 11. *We can select the numbers $d_{o,k}$ in such a way that there are in average about $2^{\frac{n}{4}}$ solutions possible choices of $(c_1, \ldots, c_{n_1})$ that give the same sum $\sum_{k=1}^{n_1} c_k d_{o,k}$.*

Proof. Most random selections of the numbers $d_{o,k}$ give this result. There are fewer combinations that yield very small or large sums and most sums are in the middle ranges. □

Lemma 12. *The lower half tuple in the indices $k = n_1 + 1, \ldots, n$ has only $\frac{n+4}{4}\frac{n}{4}$ possible values j.*

Proof. These numbers are

$$j = \sum_{k=n_1+1}^{n} c_k(d_{k,n,2} - Cd_{0,k}) = k_1e_1 + k_2e_2 \tag{4.4}$$

where $0 \leq k_1 \leq \frac{n}{4}$ and $0 \leq k_2 \leq \frac{n}{4} - 1$. □

The elements in the worst in the median tuple for n_1 satisfy $d_{k,n} \leq \frac{2^{n_1}-1}{n_1}$ because we only consider such values of $d_{k,n}$ when finding the worst in the median tuple for n_1. Also $e_i \leq \frac{2^{n_1}-1}{n_1}$. Thus, there is no carry from the lower half tuple to the upper half tuple.

Lemma 13. *It is possible to compute the median (2.1) for K_2.*

Proof. Let us assume that the values c_k are fixed for the indices $k > n_1 + 1$. This fixes some value j that must be obtained from the knapsack in the indices $k = 1, \ldots, n_1$ as the subset sum. By Lemma 12 there are only $\frac{n+4}{4}\frac{n}{4}$ possible values j. The upper half tuple yields about $2^{\frac{n}{4}}$ possible solutions for a given j in the indices $k = 1, \ldots, n_1$ by Lemma 11. The worst in the median tuple in the lower half tuple has $\frac{n}{2}$ elements, thus $2^{\frac{n}{2}}$ possible numbers can be constructed as sums $\sum_{k=1}^{n_1} c_k d'_k$ in the lower half tuple. The set of the about $2^{\frac{n}{4}}$ possible solutions of the upper half tuple for a randomly selected j is a small subset of all possible combinations of c_k in the lower half tuple in the indices $k = 1, \ldots, n_1$.

The probability that any of the possible solutions from the upper half tuple is a solution of the lower half tuple is only on the range of $\frac{(n+4)n}{16}2^{-\frac{n}{4}}$. The events of selecting the upper half tuple, the lower half tuple, and the value j can all be considered independent events. There are only a polynomial number of sums (4.4), thus when j_n is selected, there are only a polynomial number of possible values for the lower half of j in $(j, (d'_1, \ldots, d'_{n_1}))$. For a randomly selected j_n there are then only a polynomial number of c_k, $k \leq n_1$, that satisfy the lower half bits of j_n. The choice of c_k, $k \leq n_1$, fixes the upper half of j. We are left with an upper half knapsack problem for the indices $k = n_1 + 1, \ldots, n$. In this knapsack problem the elements have the size about 2^{n_1} and there are n_1 elements. Thus, for a randomly selected j_n we expect about one solution. The solution is constrained by the demand that the lower half bits give j, i.e., not all combinations are possible. We conclude that we get at least some *no* instances for computation of (2.1) for some choice of $(d_{0,1}, \ldots, d_{0,n})$. □

Lemma 14. *The time for the chosen algorithm to solve $K_{3,j_{n0}}$ when j_n ranges over numbers satisfying $j_{n,j} = j_{n0,h}$ is not larger than the median computation time for the algorithm for solving K_2 when j_n ranges over all values of j_n.*

Proof. In $K_{3,j_{n0}}$ the upper bits are easily satisfied by selecting $c_{n,n} = 1$. In order to find a solution to the subset sum problem for K_2 the algorithm must find a common solution to two knapsacks, i.e., both the upper bits and the lower bits knapsacks in K_2 must be solved with the same numbers $(c_1, \ldots, c_n)$. We may choose any difficult knapsack $(d_{0,1}, \ldots, d_{0,n})$ to the upper bits of K_2.

The algorithm cannot conclude that there are no solutions to the whole knapsack problem because there are no solutions to the upper half knapsack problem. This is so since there almost always are many solutions to the upper half knapsack problem for any value of j: the upper half knapsack problem has n elements of the bit length at most $n/2$. This means that there are 2^n possible combinations of c_k and they are mapped to $2^{n/2}$ different numbers j. Each number j is likely to come from many combinations of c_k since in average $2^{n/2}$ combinations give the same j.

It is also not possible to the algorithm to check that none of the solutions to the upper half knapsack problem give a solution to the lower half knapsack problem. This is so because there are exponentially many (i.e., $2^{n/2}$) solutions to the upper half knapsack problem. They cannot be checked in a polynomial time.

Because of these two reasons the median computation time of $K_{3,j_{n0}}$ when j_n ranges over all j_n that has the same high bits as j_{n0} cannot be higher than the median computation time for K_2 where j_n ranges over all numbers. In the computation of the median time we only take cases of j_n where there is no solution and a more complicated n-tuple must give a longer time for concluding that there are no solutions. □

Lemma 15. *The inequality (2.6) holds for the chosen algorithm.*

Proof. By Lemma 6 the median computation time for $K_{3,j_{n0}}$ when the median is taken over the set of j_n having $j_{n,h} = j_{n0,h}$ is at least as high as the time to solve $K_{1,j_{n0}}$. By Lemma 13 we can calculate the median of computation times over cases when there is no solution for K_2. By Lemma 14 the median computation time for K_2 when j_n ranges over all values is not smaller than the median computation time for $K_{3,j_{n0}}$ when the median is computed over the set j_n where $j_{n,h} = j_{n0,h}$. As K_2 is a fixed n-tuple it follows from the definition of the worst in the median tuple that K_2 has at most as long median computation time as the worst in the median tuple for n, i.e., $f(n)$. Thus the inequality (2.6) holds. □

Theorem 1. *Let an algorithm for the knapsack problem be selected. There exist numbers $B, \alpha \in \mathbb{R}$, $B \geq 1$, $\alpha \geq 0$ and a sequence*

$$((j_n, (d_{1,n}, \ldots, d_{n,n})))_{n\geq 1}$$

of knapsacks satisfying

$$\log_2 j_n < Bn^\alpha,\ \log_2 d_{k,n} < Bn^\alpha, (1 \le k \le n),\ (n \ge 1)$$

such that the algorithm cannot determine in polynomial time if there exist binary numbers $c_{k,n}$, $1 \le k \le n$, *satisfying*

$$j_n = \sum_{k=1}^{n} c_{k,n} d_{k,n}.$$

Proof. The idea of this proof is to compare the computation time of the worst (in some sense) knapsack of size n to the computation time of (in the same sense) worst knapsack of $\frac{n}{2}$. The computation time was defined in (2.4) and denoted by $f(n)$. By Lemma 15 the inequality (2.6) holds for an arbitrary chosen algorithm. By Lemma 2 the arbitrarily chosen algorithm is not a polynomial time algorithm. □

Theorem 2. **P** *does not equal* **NP**.

Proof. The knapsack problem is well known to be in **NP**. □

References

1. S. Cook, The P versus NP problem. *available on-line at* www.claymath.org.
2. D. L. Kreher and D. R. Stinson, Combinatorial algorithms, generation, enumeration, and search, CRC Press, Boca Raton, 1999.

19 Annex

Lemma A1. *Let* $B \ge 1$, $\alpha \ge 0$ *and* $\gamma \ge 0$ *be selected. Let* $r_n > 0$ *and* j_n *be integers satisfying*

$$r_n < n^\gamma, \qquad \log_2 j_n < Bn^\alpha \qquad (n \ge 1).$$

There exist numbers $C, \beta \in \mathbb{R}$,$C \geq 1$, $\beta \geq 0$ and an algorithm that given any sequence of knapsacks

$$((j_n, (d_{1,n}, \dots, d_{n,n})))_{n \geq 1}$$

can determine for each n if there exist binary numbers $c_{k,n}$, $1 \leq k \leq n$, such that

$$j_n \equiv \sum_{k=1}^{n} c_{k,n} d_{k,n} \pmod{r_n}. \tag{A1}$$

The number N_n of elementary operations needed by the algorithm satisfies $N_n < Cn^\beta$ for every $n > 1$.

Proof. The bound on the logarithm of j_n guarantees that modular arithmetic operations on $d_{k,n}$ can be made in polynomial time since we can assume that $d_{k,n} \leq j_n$. We can find the numbers $c_{k,n}$ by computing numbers $s_{k,j,n}$ from the recursion equations for k

$$s_{k,j,n} = s_{k-1,j,n} + s_{k-1,(j-d_{k,n}) (\bmod r_n),n} \tag{A2}$$

$$s_{0,j,n} = \delta_{j=0},$$

where the index j ranges from 0 to $r_n - 1$ and is calculated modulo r_n. The index n is fixed and only indicates that the numbers are for the n^{th} knapsack. Here δ_x is an indicator function: $\delta_x = 1$ if the statement x (i.e., j equals 0 in (A2)) is true and $\delta_x = 0$ if x is false. Let

$$G_{k,n}(x) = \sum_{j=0}^{r_n - 1} s_{k,j,n} x^j,$$

where $|x| < 1$. From (A2) follows

$$\sum_{j=0}^{r_n-1} s_{k,j,n} x^j = \sum_{j=0}^{r_n-1} s_{k-1,j,n} x^j + \sum_{j=0}^{r_n-1} s_{k-1,(j-d_{k,n}) (\bmod r_n),n} x^j.$$

Changing summation to $j' = j - d_{k,n}$ yields

$$G_{k,n}(x) = G_{k-1,n}(x) + \sum_{j'=-d_{k,n}}^{r_n-1-d_{k,n}} s_{k-1,j'(\mod r_n),n} x^{j'+d_{k,n}}.$$

Changing the order of summation of j' shows that

$$G_{k,n}(x) = G_{k-1,n}(x) + x^{d_{k,n}} \sum_{j'=0}^{r_n-1} s_{k-1,j',n} x^{j'}. \qquad (A3)$$

Simplifying (A3) gives

$$G_{k,n}(x) = G_{k-1,n}(x) + x^{d_{k,n}} G_{k-1,n}(x).$$

As $G_{0,n}(x) = s_{0,0,n} = 1$, we get

$$G_{n,n}(x) = \prod_{k=1}^{n} (1 + x^{d_{k,n}}).$$

Expanding the product shows that $s_{k,j,n} \neq 0$ if and only if there exist binary numbers c_m, $c_m \in \{0,1\}$, $1 \leq m \leq n$, satisfying

$$j \equiv \sum_{m=1}^{n} c_m d_{m,n} \qquad (\mod r_n).$$

For $j = j_n$ and $k = n$ we get the knapsack problem. This means that we can solve the knapsack problem by computing all $s_{k,j,n}$ form (A2). We do not actually need the numbers $s_{k,j,n}$ but only the information if $s_{k,j,n} \neq 0$. Therefore we will not compute the terms $s_{k,j,n}$ directly but calculate binary numbers $b_{j,k} \in \{0,1\}$ by Algorithm A0 below. The number $b_{k,j}$ calculated by A0 is zero if and only if the number $s_{k,j,n} = 0$ is zero.

```
Algorithm A0:
  Loop from $k = 0$ to $k = n$ with the step $k := k + 1$ do {
    Loop from $j = 0$ to $j = r_n - 1$ with the step $j := j + 1$ do
    $b_{j,k} := 0$
}
$b_{0,0} := 1$
Loop from $k = 1$ to $k = n$ with the step $k := k + 1$ do {
  $M := \min\{r_n - 1, \sum_{m=1}^{k} d_{m,n}\}$
  Loop from $j = 0$ to $j = M$ with the step $j := j + 1$ do {
    If $(b_{k-1,j} = 0$ and $b_{k-1,(j-d_{k,n})(\mod r_n)} = 0)$ do $b_{j,k} := 0$
    else do $b_{j,k} := 1$
  }
}
If $b_{n,j_n} = 1$ do $result := TRUE$ else do $result := FALSE$
```

Algorithm A0 loops from $k = 0$ to $k = n$ and from $j = 0$ to $j = r_n - 1 < n^\gamma$. Thus A0 needs a polynomial number of elementary operations as a function of n in order to give the result $TRUE$ or $FALSE$ to the existence of a solution to (A1). □

Lemma A2. *Let $B, \alpha \in \mathbb{R}$, $B \geq 1$, $\alpha \geq 0$ be fixed. There exist numbers $C, \beta \in \mathbb{R}$, $C \geq 1$, $\beta \geq 0$ and an algorithm that for any sequence*

$$((j_n, (d_{1,n}, \ldots, d_{n,n})))_{n \geq 1}$$

of knapsacks satisfying

$$j_n \leq Bn^\alpha, \qquad d_{k,n} \leq j_n \qquad (1 \leq k \leq n),$$

can determine if there exist binary numbers $c_{k,n}$, $1 \leq k \leq n$, such that

$$j_n = \sum_{k=1}^{n} c_{k,n} d_{k,n}.$$

The number N_n of elementary operations needed by the algorithm satisfies $N_n < Cn^{\beta}$ for every $n > 1$.

Proof. The result follows directly from Lemma A1 by selecting $r_n = \sum_{k=1}^{n} d_{k,n} \leq nj_n$. □