J Autom Reasoning (2009) 43:243–261
DOI 10.1007/s10817-009-9145-6

Formalizing an Analytic Proof of the Prime Number Theorem

John Harrison

Received: 13 July 2009 / Accepted: 14 July 2009 / Published online: 19 August 2009

Abstract We describe the computer formalization of a complex-analytic proof of the Prime Number Theorem (PNT), a classic result from number theory characterizing the asymptotic density of the primes. The formalization, conducted using the HOL Light theorem prover, proceeds from the most basic axioms for mathematics yet builds from that foundation to develop the necessary analytic machinery including Cauchy's integral formula, so that we are able to formalize a direct, modern and elegant proof instead of the more involved 'elementary' Erdös-Selberg argument. As well as setting the work in context and describing the highlights of the formalization, we analyze the relationship between the formal proof and its informal counterpart and so attempt to derive some general lessons about the formalization of mathematics.

Keywords Analytic proof • Prime number theorem • Computer formalization

1 Formalizing Mathematics: Pure and Applied

I've always been interested in using theorem provers both for "practical" applications in formally verifying computer systems, and for the "pure" formalization of traditional mathematical proofs. I particularly like situations where there is an interplay between the two. For example, in my PhD thesis [12], written under Mike Gordon's supervision, I developed a formalization of some elementary real analysis. This was subsequently used in very practical verification applications [13], where in fact I even needed to formalize *more* pure mathematics, such as power series for the cotangent function and basic theorems about diophantine approximation.

Dedicated to Mike Gordon on the occasion of his 60th birthday.

J. Harrison (✉)
Intel Corporation, JF1-13 2111 NE 25th Avenue, Hillsboro, OR 97124, USA
e-mail: johnh@ichips.intel.com

I first joined Mike Gordon's HVG (Hardware Verification Group) to work on an embedding in HOL of the hardware description language ELLA. Mike had already directed several similar research projects, and one concept first clearly articulated as a result of these activities was the now-standard distinction between 'deep' and 'shallow' embeddings of languages [5]. Since I was interested in formalizing real analysis, Mike encouraged me to direct my attention to case studies involving arithmetic, and this was the starting-point for my subsequent research. Right from the beginning, Mike was very enthusiastic about my formalization of the reals from first principles using Dedekind cuts. Mike had been involved in Robin Milner's group developing the original Edinburgh LCF [10], a central feature of which was the idea of extending the logical basis with derived inference rules to preserve soundness. Mike had emphasized that the LCF approach is equally applicable to a wide variety of logics [8], and he applied it to higher-order logic to produce the HOL system [9]. Since the HOL logic is suitable as a general foundation for mathematics, it was possible to extend this idea and even develop mathematical concepts themselves in a 'correct by construction' way using definitions. So a definitional construction of the reals fitted in very well with the ideals Mike had for the HOL project, an interest in applications combined with an emphasis on careful foundations that has now become commonplace.

In this paper I will describe a formalization that was undertaken purely for fun, involving complex analysis and culminating in a proof of the Prime Number Theorem. Nevertheless, it doesn't seem entirely far-fetched to imagine some "practical" applications of this result in the future. For example a weak form of the PNT is implicitly used to justify the termination of the breakthrough AKS primality test [1], and some simpler properties of prime numbers have been used in the verification of arithmetical algorithms [15]. But I certainly don't need to give any such justification, because Mike Gordon, as well as introducing me to the fascinating world of theorem proving, has always placed a welcome emphasis on doing "research that's fun".

2 The Prime Number Theorem

Let us write $\pi(x)$ for the number of prime numbers no greater than x, i.e.

$$\pi(x) = |\{p \mid \mathrm{prime(p)} \wedge 0 \leq p \leq x\}|$$

We can think of π as a function out of the set of natural numbers, or as a step function out of the set of real numbers that only changes for integral values. The Prime Number Theorem asserts that

$$\pi(x) \sim \frac{x}{\log(x)}$$

or in other words $\frac{\pi(x)}{x/\log(x)} \to 1$ as $x \to \infty$, where log denotes the natural (base e) logarithm function. It is not hard to show that $x/\log(x) \sim \mathrm{li}(x)$ where $\mathrm{li}(x) = \int_0^x dt/\log t$, so one can interpret the PNT as saying roughly that 'the probability that n is a prime is about $1/\log(n)$'. Table 1 below gives an idea of the rate of convergence for a moderate range of numbers; in fact $\mathrm{li}(x)$ is an even better fit to $\pi(x)$ at the low end.

Table 1 Numerical illustration of the PNT

x	$\pi(x)$	$\frac{x}{\log(x)}$	Ratio
10^1	4	4.34	0.9217
10^2	25	21.71	1.1515
10^3	168	144.76	1.1605
10^4	1229	1085.74	1.1319
10^5	9592	8685.89	1.1043
10^6	78498	72382.41	1.0845
10^7	664579	620420.69	1.0712
10^8	5761455	5428681.02	1.0613
10^9	50847534	48254942.43	1.0537
10^{10}	455052511	434294481.90	1.0478
10^{11}	4118054813	3948131653.67	1.0430
10^{12}	37607912018	36191206825.27	1.0391
10^{13}	346065536839	334072678387.12	1.0359
10^{14}	3204941750802	3102103442166.08	1.0331
10^{15}	29844570422669	28952965460216.79	1.0308
10^{16}	279238341033925	271434051189532.38	1.0288
10^{17}	2623557157654233	2554673422960304.87	1.0270
10^{18}	24739954287740860	24127471216847323.76	1.0254
10^{19}	234057667276344607	228576043106974646.13	1.0240
10^{20}	2220819602560918840	2171472409516259138.26	1.0227
10^{21}	21127269486018731928	20680689614440563221.48	1.0216
10^{22}	201467286689315906290	197406582683296285295.97	1.0206

The PNT was conjectured around 1800 by several mathematicians including Gauss, but no real progress was made towards a proof. The first breakthrough came when Chebyshev proved in 1847 that $\frac{\pi(x)}{x/\log(x)}$ is bounded quite close to 1 asymptotically, and that *if* it tends to a limit, this limit must be 1. While this is tantalizingly close to the Prime Number Theorem, the additional step of proving that the limit does indeed exist resisted attack for some time afterwards.

In 1859 Riemann started a new chapter when he pointed out the deep relationship between the distribution of primes and the complex zeta function, which is the analytic continuation of the function defined for $\Re z > 1$ (we use $\Re z$ and $\Im z$ denote the real and imaginary components of a complex number z) by the series

$$\zeta(z) = \sum_{n=1}^{\infty} 1/n^z$$

The key to the relationship between the zeta function and the distribution of primes is the Euler product formula, again valid for $\Re z > 1$, where the product is taken over all primes p:

$$\zeta(z) = \prod_{p} 1/(1 - p^{-z})$$

In 1894, von Mangoldt was able to reduce the PNT to the conjecture that the zeta function has no zeros on the line $\Re z = 1$. This was finally proved two years later, independently, by both Hadamard and de la Valée Poussin, hence completing the proof of the PNT.

All these proofs made extensive use of contour integrals and other ideas from complex analysis. Since the PNT itself doesn't talk about complex numbers at all, just making basic use of real logarithms, the question naturally arose whether there might be a proof not relying on complex analysis. For some time, there was scepticism about whether any such proof could be found that wasn't just an artificial reformulation of one of the established analytic proofs. Nevertheless, just such a proof was found in 1948 by Selberg and Erdös. On the other hand, this "elementary" proof is only elementary in the sense of avoiding analytic machinery. It is markedly more complicated and intricate than the modern versions of the analytic proof, indicating that there is a price to be paid for avoiding analysis.

3 Mathematical Machinery versus Brute Force

Typically, the direct formalization of a completely explicit argument at the level of a paper or textbook takes a fairly predictable amount of time and the result has a more or less predictable length. An emerging rule of thumb is that formalizing one page of a textbook takes about a work-week [26]. The *de Bruijn factor*, an adjusted ratio between sizes of the formalized and informal text defined in more detail below, is quite often found [25] to be about 4. But it is common for proofs in mathematics to call on an infrastructure of established, standard, results. If these results already exist in the library of the proof assistant, these references to the literature simply become corresponding references to the prover's library of theorems. But if not, then we need to deal with these results from scratch, so all bets are off: we may end up doing an arbitrarily large amount of work formalizing what the informal source regards as standard background. In such cases, when the de Bruijn factor for the direct proof underestimates the actual work, a potentially easier alternative is to adopt a different proof, one that may in itself be significantly longer or more difficult, but which at least does not rely on such a quantity of existing 'machinery' and is hence more practical to formalize.

There isn't a perfect match between what humans find easy and what proof assistants find easy [7, 22, 27], so it's possible that a proof that a human would regard as more difficult is actually easier to formalize, even setting aside a comparison of the amount of knowledge assumed. For example, proofs involving extensive 'brute-force' case analysis can be almost trivial to formalize with a programmable proof assistant, perhaps obviating the need for more sophisticated symmetry arguments. An instance is the proof of the Kochen-Specker paradox from quantum mechanics. A typical human-oriented proof [6] would make use of symmetry to reduce case analysis without loss of generality: 'the three coordinate axes are all orthogonal, so we have a combination 0, 1, 1 in some order, so let's assume the zero is in the x direction...'. By contrast, in a computer formalization it's even easier just to program the computer to run through all the cases [17].

An interesting example of the tension between using mathematical machinery and computational power is proving the associativity of the chord-and-tangent addition operation on elliptic curves. This has been done formally in Coq by Laurent Théry and Guillaume Hanrot [24], with some of the key parts using enormous algebraic computations that were on the edge of feasibility; indeed similar computational issues have obstructed a related project by Joe Hurd. On the practical difficulties caused by this example, Dan Grayson suggested, in an email to the present author, a more 'human-oriented' proof:

> But why not enter one of the usual human-understandable proofs that + is associative? Too many prerequisites from algebraic geometry? [...] The proof I like most is to use the Riemann-Roch theorem to set up a bijection between the rational points of an elliptic curve and the elements of the group of isomorphism classes of invertible sheaves of degree 0. That's a lot of background theory, probably too much for this stage of development, but then the "real" reason for associativity is that tensor product of R-modules is an associative operation up to isomorphism.

It seems likely that formalizing mathematical machinery on that level is still many years away. In this case, therefore, the 'elegant' proof is not a practical option for formalization, and so we need to make the best of it by exploiting the prover's ability to do fast computation, even if that brings problems of its own. It is slightly depressing to reflect that we may be forced to formalize unnatural or 'hacky' proofs because of a lack of general mathematical machinery.

Returning to the Prime Number Theorem, it seems that the elementary proof offers no such advantages in itself from the point of view of formalization: the various manipulations don't offer any great scope for automation. Yet the lack of analytic machinery seems to make it a more attractive target, and in fact the elementary proof was formalized several years ago in a landmark achievement by a team led by Jeremy Avigad [3]. As those authors note, it was hardly practical to contemplate formalizing the analytic proof:

> Since the libraries we had to work with had only a minimal theory of the complex numbers and a limited real analysis library, we chose to formalize the Selberg proof.

Even within the established framework of the elementary proof, some parts were done in a different and arguably less natural way, again to provide a better match to the available libraries [3]:

> But similar issues arose even with respect to the mild uses of analysis required by the Selberg proof. Isabelle's real library gave us a good theory of limits, series, derivatives and basic transcendental functions, but it had almost no theory of integration to speak of. Rather than develop such a theory, we found that we were able to work around the mild uses of integration needed in the Selberg proof. Often, we also had to search for quick patches to other gaps in the underlying library.

On the other hand, with respect to some aspects of the mathematical machinery, Avigad et al. were motivated to develop general libraries that could be of use in other contexts. In particular, their paper [2] presents a general formalization of 'big O' asymptotic expressions.

In our own theorem prover, HOL Light, we now have a solid library of results about analysis in Euclidean space [16] and, based on that, of elementary complex analysis including versions of the Cauchy integral formula for topologically simple regions [18]. Thus, it seemed that it ought to be possible to formalize an analytic proof of the PNT in a fairly direct fashion. Indeed, formalizing such a proof was proposed by Robert Solovay as a good challenge in the formalization of mathematics [26], and this challenge was a significant motivation for us in the work described here.

4 Our Mathematical Machinery

Our formal proof of the Prime Number Theorem rests on an extensive body of formalization that has been developed, on and off, over the course of more than a decade. We will not go into great detail here, but will explain where the reader can look for more details while explaining some of the basic definitions and syntactic constructs that will be used when describing parts of the PNT formalization. As usual in logical systems based on lambda-calculus, we will sometimes write function applications $f(x)$ by simple juxtaposition $f\ x$ unless the bracketing is needed to establish precedence.

The first traditionally 'mathematical' step is the construction of the real numbers and the proof that they form a complete ordered field. The first version of this construction in HOL88 [11] used Dedekind cuts, but the version we now use in HOL Light [12] is based on an encoding of Cauchy sequences using nearly-additive functions. In any case, once the 'axioms' for a complete ordered field are derived, the way in which the structure was constructed is no longer used. We will just note a few syntactic points, especially the presence of the injection '&' from natural numbers to reals.

HOL	Standard	Meaning
`&`	(*none*)	Natural map $\mathbb{N} \to \mathbb{R}$
`--x`	$-x$	Unary negation of x
`inv(x)`	x^{-1}	Multiplicative inverse of x
`abs(x)`	$\|x\|$	Absolute value of x
`log(x)`	$\log(x)$	Natural logarithm of x
`x pow n`	x^n	Real x raised to natural power n

The next step is the formalization of analysis in Euclidean space $\mathbb{R}^N$ [16]. The HOL logic does not have dependent types, but that paper describes how to encode N as a *type* (basically, N is the cardinality of the type's universe) so that HOL's polymorphism naturally gives rise to theorems for general N. Since we mainly use $\mathbb{R}^2$ in what follows, the more general features of this theory are not important here, but we note that many of the analytic theorems that get applied to the complex numbers are in fact applicable to broader classes of functions $\mathbb{R}^M \to \mathbb{R}^N$. The norm $\|z\|$ of a vector z is written `norm(z)`; in the case of the complex numbers we have $\|x + iy\| = \sqrt{x^2 + y^2}$, and the distance between two vectors w and z is `dist(w,z)`. We also define the concept of an open ball (actually a disc in the complex case) $B(x, \epsilon)$ in the usual way as $B(x, \epsilon) = \{y \mid \|y - x\| < \epsilon\}$:

```
|- dist(x,y) = norm(x - y)

|- ball(x,e) = { y | dist(x,y) < e}
```

The theory of complex numbers is then built on top, using `complex` ($\mathbb{C}$) as a mere abbreviation for `real^2` ($\mathbb{R}^2$). Most of the very basic setup was taken over with trivial changes from an earlier theory [14] that was constructed more explicitly from pairs of reals. After being transferred to the new formulation, this theory was slightly revised and greatly extended to include analytic properties like complex derivatives and complex line integrals. This new theory [18], described in a Festschrift tribute to Andrzej Trybulec, another great pioneer in interactive theorem proving, is the direct foundation for the PNT proof. Since that paper was written, we have made only a few incompatible changes, e.g. introducing a subtle distinction between `analytic_on` and `holomorphic_on` as described later. But we (and others) have further extended it with new results, and in the process of proving the PNT we still sometimes needed to go back and add some additional general lemmas to this library. In the rest of this section, we will summarize the main points that are necessary to understand the remainder of the paper.

The usual operations like addition, multiplication, and natural number powers are defined on complex numbers, and we overload the same symbols as are used for the reals; for example `z pow n` for z^n. We use `Cx` for the injection $\mathbb{R} \to \mathbb{C}$ and `Re` and `Im`, both with type $\mathbb{C} \to \mathbb{R}$, for the real and imaginary parts of a number. The `cnj` function denotes complex conjugation and `ii` the imaginary unit i. The more general power function with a complex exponent, also written w^z informally, is `w cpow z`, and we define it for $w \neq 0$ by just $e^{z\log(w)}$ using the complex exponential (`cexp`) and logarithm (`clog`). While in general the properties of the complex logarithm and power function can be subtle (for example, $\log(xy) = \log(x) + \log(y)$ and $(xy)^z = x^z y^z$ fail in general) we will only use w^z or $\log(w)$ in cases where w has zero imaginary part and positive real part, in which case they holds no surprises.

We define complex differentiation in the standard way. We assert that f is differentiable at point z by `f complex_differentiable (at z)`, while $f'(z)$, the complex derivative of f at z, is written `complex_derivative f z`. The more refined variant `f complex_differentiable (at z within s)` indicates complex differentiability where limits at z are restricted to the set s. The distinction between `at z` and `at z within s` appears in the same way in several other limit concepts, e.g. in the following:

```
|- f holomorphic_on s ⇔
   ∀z. z IN s ⇒ f complex_differentiable (at z within s)

|- f analytic_on s ⇔
   ∀z. z IN s ⇒ ∃e. &0 < e ∧ f holomorphic_on ball(z,e)
```

When s is an open set (that is, for each $x \in s$ there is an $\epsilon > 0$ with $B(x, \epsilon) \subseteq s$) the concepts '$f$ is holomorphic on s' and 'f is analytic on s' coincide, but in general the latter, asserting differentiability in an open ball around each point, is stronger.[1]

The most interesting results are connected with line integrals of analytic functions along paths. We use a notion of 'valid path' in these theorems, where `valid_path g` means that g is a piecewise differentiable function out of the canonical interval [0, 1], which strictly speaking is an interval not in $\mathbb{R}$ but in the distinct but equivalent type of 1-dimensional vectors $\mathbb{R}^1$. The function `vec n` here

[1] This is a change compared with [18], where `analytic_on` was used for what we now call `holomorphic_on`. We were persuaded by the experience of the proof to be described below that our earlier usage was non-standard.

gives a vector with all components—in the present case of $\mathbb{R}^1$ there is only one component—equal to the natural number `n`.

```
|- valid_path (f:real^1->complex) ⇔
     f piecewise_differentiable_on interval[vec 0,vec 1]
```

We discuss some of the theorems about line integrals below, and establish there the notation that is used in our formalization.

5 Formalizing Newman's Proof

There are several significantly different analytic proofs of the PNT. The best-established methods, clearly presented in Ingham's book [19], calculate residues for a somewhat involved contour that goes to infinity. A more recent approach, due to Newman, relies only on Cauchy's integral formula applied to a bounded, roughly semicircular contour. Jameson's book [20] is an excellent source for the PNT in general, these various proofs, and related topics like Dirichlet's theorem. Jameson notes that Newman's proof is commonly formulated in two different ways, either in terms of summing a series or in terms of a real 'Laplace transform' integral. We took the former approach, following the presentation by Newman himself in his textbook [21]. In fact, it may well have been somewhat easier to use the 'Laplace transform' version instead, e.g. the slick presentation by Zagier [28]. Subjectively, we guessed that it was likely to be harder to formalize a proof involving additional transcendental functions and a mix of real and complex variables, but perhaps this belief was mistaken.

More specifically, from Newman's book [21] we formalized the "second proof" on pp. 72–74 using the analytic lemma on pp. 68–70. Newman's book is a graduate-level textbook, and while it is written in a friendly and accessible style, it sometimes assumes quite a lot of background or leaves some non-trivial steps to the reader. In particular, the PNT proof assumes some major properties of the zeta function without proof. In formalizing these we used a variety of sources, taking the proof that $\zeta(z)$ has no zeros for $\Re z = 1$ directly from the textbook [4] by Bak and Newman (the same Newman). The overall PNT proof naturally splits up into six parts, which are presented by Newman in somewhat distinct styles and with widely varying levels of explicitness.

1. The Newman-Ingham "Tauberian" analytic lemma.[2]
2. Basic properties of the Riemann ζ-function and its derivative, including the Euler product.
3. Chebyshev's elementary proof that $\sum_{p \leq n} \frac{\log p}{p} - \log n$ is bounded.
4. Deduction using analytic lemma that $\sum_n (\sum_{p \leq n} \frac{\log p}{p} - \log n + c)/n$ is summable for some constant c.

[2] According to MathWorld, "A Tauberian theorem is a theorem that deduces the convergence of an infinite series on the basis of the properties of the function it defines and any kind of auxiliary hypothesis which prevents the general term of the series from converging to zero too slowly."

5. Derivation from that summability that $\sum_{p\leq n}\frac{\log p}{p}-\log n$ tends to a limit.
6. Derivation of the PNT from that limit using partial summation.

We have compared the main parts of our formalization against reverse-engineered TeX for corresponding passages in Newman's book (thanks to Freek Wiedijk for composing these!) The *de Bruijn factor* [25], the size ratio of a compressed formal proof text versus the compressed TeX of its informal counterpart (compressed, currently using `gzip`, for a crude approximation to information content), varies widely:

	Part of proof	de Bruijn factor
1	Analytic lemma	8.2
2	ζ-function	81.3
3	Chebyshev bound	28.2
4	Summability	11.0
5	Limit	5.4
6	PNT	30.4

It is commonly found that the de Bruijn factor for formalizations is about 4, so these seem very high. However, the very high figures in parts 2, 3 and 6 arise where Newman is not giving a proof in any explicit sense. The quotations that follow are the sum total of Newman's text for these parts, which take over half of the 4939 lines in the complete HOL Light formalization. In no cases can Newman's passage really be called a proof, so the comparison is hardly fair; two are really assumptions of background knowledge and the third is the merest hint of how a proof might go:

2 Let us begin with the well-known fact about the ζ-function: $(z-1)\zeta(z)$ is analytic and zero free throughout $\Re z\geq 1$.

3 In this section we begin with Tchebyshev's observation that $\sum_{p\leq n}\frac{\log p}{p}-\log n$ is bounded, which he derived in a direct elementary way from the prime factorization on n!

6 The point is that the Prime Number Theorem is easily derived from '$\sum_{p\leq n}\frac{\log p}{p}-\log n$ converges to a limit' by a simple summation by parts which we leave to the reader.

If we restrict ourselves to parts 1, 4 and 5, the de Bruijn factor is about 8, still higher than normal, but not outrageously so. And indeed, although the proof did not present any profound difficulties, we found that it took more time to formalize Newman's text than we have grown to expect for other formalizations. This may indicate that Newman's style is fairly terse and leaves much to the reader (this does seem to be the case), or that in this area, we sometimes have to work hard to prove things that are obvious informally (this is certainly true for the winding number of the contour mentioned later).

It is also instructive to compute a de Bruijn factor for our formal proof that the ζ-function has no zeros on $\Re z\geq 1$ relative to our source for a 'real proof', the text by Bak and Newman [4]. We find in this case that the de Bruijn factor is only 3.1, actually somewhat *lower* than usual. Although we may be extrapolating too much from a few data points, we believe that this difference is natural given that Newman's analytic number theory text is at the *graduate* level whereas Bak and Newman is at the *undergraduate* level. It certainly doesn't seem implausible on general grounds that we should find larger de Bruijn factors at higher levels of mathematics, where more can be assumed of the reader. It would be interesting to see if this kind of relationship does seem to hold in general.

6 The Analytic Lemma

The centrepiece of Newman's approach is the analytic lemma; this is the only part that uses non-trivial facts about the complex numbers and is thus the place where most of the analytic 'machinery' plays a role.

Threorem *Suppose $|a_n| \leq 1$, and form the series $\sum a_n n^{-z}$ which clearly converges to an analytic function $F(z)$ for $\Re z > 1$. If, in fact, $F(z)$ is analytic throughout $\Re z \geq 1$, then $\sum a_n n^{-z}$ converges throughout $\Re z \geq 1$.*

Here is our HOL formalization of the statement of the theorem, which corresponds pretty directly to the informal one, with `(f sums l) (from k)` corresponding to the traditional notation $\sum_{n=k}^{\infty} f(n) = l$:

```
|- ∀f a. (∀n. 1 <= n ⇒ norm(a(n)) <= &1) ∧
        f analytic_on {z | Re(z) >= &1} ∧
        (∀z. Re(z) > &1 ⇒ ((λn. a(n) / Cx(&n) cpow z) sums (f z))(from 1))
        ⇒ ∀z. Re(z) >= &1
                ⇒ ((λn. a(n) / Cx(&n) cpow z) sums (f z))(from 1)
```

We might note in passing, as an indication of how casual Newman is about minor details, that the theorem is later applied to a particular case where $\|a_n\|$ is bounded, but not necessarily by 1. Since we formalized the analytic lemma before looking closely at its application, we took the above statement at face value, and only later discovered that we needed to generalize it to $\|a_n\| \leq B$ for general B. The generalization is indeed rather trivial, just by assuming without loss of generality that $B > 0$ and applying the basic lemma to $a'_n = a_n/B$. Yet we took another 34 lines of proof script to state and prove the more general version in this way, whereas it would probably have been more efficient to prove this generalization from the start.

```
|- ∀f a b. (∀n. 1 <= n ⇒ norm(a(n)) <= b) ∧
          f analytic_on {z | Re(z) >= &1} ∧
          (∀z. Re(z) > &1 ⇒ ((λn. a(n) / Cx(&n) cpow z) sums (f z)) (from 1))
          ⇒ ∀z. Re(z) >= &1
                  ⇒ ((λn. a(n) / Cx(&n) cpow z) sums (f z)) (from 1)
```

Generally speaking, the formalization is a pretty direct translation of Newman's proof. One systematic change is that we state most integration results in a relational form, i.e. rather than $\int_\gamma f(z)\,dz = I$ in the informal text, use the formal rendering (we could eta-reduce $\lambda z.\ f(z)$ to just f, but maintaining the bound variable z keeps us closer to the original)

```
((λz. f(z)) has_path_integral I) gamma
```

This is because our HOL formalizations use straightforward total functions for notions like 'integral'. Thus, the mere assertion that $\int_\gamma f(z)\,dz = I$ does not imply that f is indeed integrable on the contour γ, since $\int_\gamma f(z)\ dz$ is always equal to *something*. In informal mathematics, by contrast, some convention about partial functions is usually assumed when writing $s = t$, typically 'either s and t are both undefined or both are defined and equal' [12]. Nevertheless, it might have been preferable to stay closer to the original by separating out the assertion of integrability:

```
(λz. f(z)) path_integrable_on gamma ∧ path_integral gamma (λz. f(z)) = I
```

The proof of the analytic lemma involves applying Cauchy's integral formula round a contour and then performing some careful estimations of the sizes of the various line integrals involved. The contour we use, traversed counterclockwise, is shown in Fig. 1. This is slightly different from Newman's, though the distinction is hard to see in the diagram: Newman continues the circular arc into the left-hand side of the complex plane whereas we use an exactly semicircular arc connected to the vertical portion by horizontal straight-line segments. (We only diverged from Newman because we found this version easier to understand informally, not because of any problem specific to formalization.) There are only two points where our formal proof really looks complicated relative to Newman's original, both instances where things that are very easy to see informally require some work to formalize.

One is the observation that because $f(z)$ is analytic for $\Re z \geq 1$, for any $R > 0$ there is a $\delta > 0$ such that $f(z)$ is analytic for $\Re z \geq 1 - \delta$ and $|\Im z| \leq R$. Recall that f is analytic on a set iff it is holomorphic on an open neighbourhood of each point. Thus, for any point z with $\Re z = 1$ there is some $\epsilon > 0$ such that f is holomorphic in the open disc $B(z, \epsilon) = \{w \mid \|w - z\| < \epsilon\}$. With a little experience of such concepts, one can immediately see "by compactness" that the required result follows. Yet in the formal proof this intuition has to be replaced by an explicit paving of a suitable compact set

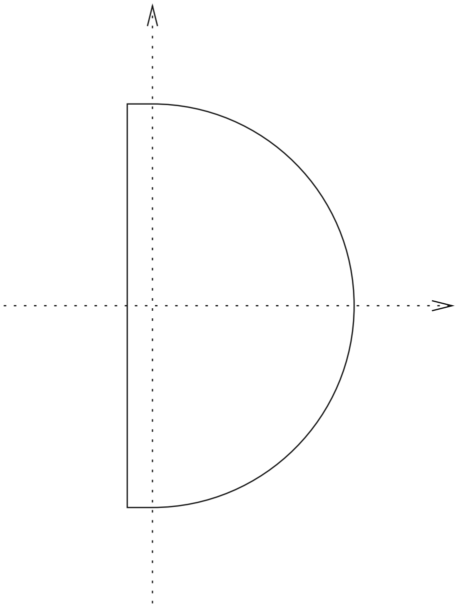

Fig. 1 Contour used in application of Cauchy's integral theorem

with rectangular open sets and the application of the Heine-Borel theorem, requiring 164 lines of proof script.

The other place where formalization presents a striking contrast with informal perception is in the application of Cauchy's integral formula. The formal version we use is quite standard in its statement and hypotheses, though it is limited to cases where the function concerned is analytic on a *convex* set. (A more general version would certainly be appealing, but is not needed for this and many other non-trivial applications.)

```
|- ∀f s g z.
      convex s ∧ f holomorphic_on s ∧
      z IN interior(s) ∧
      valid_path g ∧ (path_image g) SUBSET (s DELETE z) ∧
      pathfinish g = pathstart g
      ⇒ ((λw. f(w) / (w − z)) has_path_integral
        (Cx(&2) ∗ Cx(pi) ∗ ii ∗ winding_number(g,z) ∗ f(z)))g
```

The difficulty comes in verifying the intuitively obvious requirements that arise when applying this result: our contour, composed of three line segments and one semicircular arc, satisfies the conditions (is a valid path whose start is the same as its finish etc.) and has a winding number of 1 about the origin, indicating that it winds exactly once round the origin counterclockwise, where the winding number of γ about z is formally defined as

$$\frac{1}{2\pi i}\int_{\gamma} dz/z$$

These facts are so obvious, based on the picture and an intuitive understanding of what winding numbers mean, that it requires a conscious effort to think about things in a formal way and arrive at the right general lemmas. But when approached systematically, it is not so difficult. The lemmas we use include results like the following, which show how to deduce properties of a 'joined' path `g1 ++ g2` composed of two other paths `g1` and `g2`:

```
|- ∀g1 g2. pathstart(g1 ++ g2) = pathstart g1

|- ∀g1 g2. pathfinish(g1 ++ g2) = pathfinish g2

|- ∀g1 g2.
      pathfinish g1 = pathstart g2
      ⇒ (valid_path(g1 ++ g2) ⇔ valid_path g1 ∧ valid_path g2)

|- ∀g1 g2 z.
      valid_path g1 ∧ valid_path g2 ∧
      ¬(z IN path_image g1) ∧ ¬(z IN path_image g2)
      ⇒ winding_number(g1 ++ g2,z) =
          winding_number(g1,z) + winding_number(g2,z)
```

Proving that the winding number is exactly 1 is a bit more subtle. We take the two-step approach, which perhaps looks a bit strange, of proving that (i) the real part of the winding number is > 0, (ii) it is < 2. (Note that we regard the winding number as a complex number defined for any valid path, and don't restrict ourselves to closed paths.) These results suffice, since we already have a formally proved lemma that the winding number of a closed path is an integer; indeed it is an integer *if and only if* the path is closed:

```
|- ∀g z. valid_path g ∧ ¬(z IN path_image g)
         ⇒ (complex_integer(winding_number(g,z)) ⇔
              pathfinish g = pathstart g)
```

Proving that the winding number is > 0 is again something that we can inherit for a path with multiple components from corresponding results about the components. In fact, we get smoother 'stepping' theorems if we prove such things together with other properties we need anyway:

```
|- ∀g1 g2 z.
        (valid_path g1 ∧ ¬(z IN path_image g1) ∧
         &0 < Re(winding_number(g1,z))) ∧
        (valid_path g2 ∧ ¬(z IN path_image g2) ∧
         &0 < Re(winding_number(g2,z))) ∧
        pathfinish g1 = pathstart g2
        ⇒ valid_path(g1 ++ g2) ∧ ¬(z IN path_image(g1 ++ g2))∧
           &0 < Re(winding_number(g1 ++ g2,z))
```

Straightforward theorems assert that the real part of the winding number is positive for basic path components. For example, for a straight-line path from a to b we have the following sufficient condition for the winding number to be positive around z, based just on components of a, b and z (recall that `cnj` is complex conjugation):

```
|- ∀a b z. &0 < Im((b − a) ∗ cnj(b − z))
           ⇒ &0 < Re(winding_number(linepath(a,b),z))
```

For the dual inequality, proving that the real part of the winding number is < 2, we split the path into two parts, one consisting of the three straight-line segments on the left-hand side of the complex plane, and the other the semicircle on the right. Using the additivity theorem above, it suffices to show that each of these two path components has a winding number with real part < 1. This is justified by the following theorem, asserting that if some semi-infinite ray drawn from z does not meet the curve, then the curve's winding number about z must have real part < 1:

```
|- ∀g w z. valid_path g ∧ ¬(z IN path_image g) ∧ ¬(w = z)∧
           (∀a. &0 < a ⇒ ¬(z + (Cx a ∗ (w − z)) IN path_image g))
           ⇒ Re(winding_number(g,z)) < &1
```

This in turn is derived from the contrapositive observation that if the real part of the winding number is ≥ 1, it must pass through every value between 0 and 1 at some point along the curve. (This depends on the fact that the integral is a continuous function of the parameter, so we can use the Intermediate Value Theorem.) Since the real part of the winding number corresponds to angles as a proportion of a full 2π rotation, the result follows.

7 The Rest of the Proof

The rest of the proof was relatively straightforward to formalize, and we content ourselves with a few brief notes showing some of the main statements in their formal guise.

7.1 Properties of the ζ-Function

We develop the necessary properties of the ζ-function in a fairly ad-hoc way, starting with a definition of an auxiliary function

$$\mathrm{nearzeta}_n(z) = \sum_{m=n}^{\infty} (z-1)/m^z - (1/m^{z-1} - 1/(m+1)^{z-1})$$

and then defining a general family of zeta-like functions with the summation starting at n (the prime denotes the complex derivative):

$$\mathrm{genzeta}_n(z) = \begin{cases} \mathrm{nearzeta}'_n(1) & \text{if } z = 1 \\ (\mathrm{nearzeta}_n(z) + 1/n^{z-1})/(z-1) & \text{otherwise} \end{cases}$$

The usual zeta function is defined as $\zeta = \mathrm{genzeta}_1$, but results about the more general versions are useful to give smoother proofs of some convergence results that crop up now and again. The principal results we derive are the usual summation $\zeta(z) = \sum_{n=1}^{\infty} 1/n^z$ for $\Re z > 1$

```
|- ∀s. &1 < Re s
       ⇒ ((λn. Cx(&1) / Cx(&n) cpow s) sums zeta(s)) (from 1)
```

the fact that the function is analytic on a suitable domain

```
|- zeta holomorphic_on {s | Re(s) > &0 ∧ ¬(s = Cx(&1))}
```

the Euler product formula $\zeta(z) = \prod_p 1/(1 - p^{-z})$, again for $\Re z > 1$, where the limit assertion `(f --> l) sequentially` corresponds to the informal $f(n) \to l$ as $n \to \infty$:

```
|- ∀s. &1 < Re s
       ⇒ ((λn. cproduct {p | prime p ∧ p <= n}
                         (λp. inv(Cx(&1) - inv(Cx(&p) cpow s))))
          -> zeta(s)) sequentially
```

and the key fact that there are no zeros for $\Re z \geq 1$:

```
|- ∀s. &1 <= Re s ⇒ ¬(zeta s = Cx(&0))
```

7.2 The Chebyshev Bound

This part is a formalization of Newman's observation that $\sum_{p\leq n}\frac{\log p}{p}-\log n$ is bounded. Newman attributes this to Chebyshev's original investigations, but it seems to be generally known as Mertens's First Theorem. We formalized a proof from [23]; although we have not computed a de Bruijn factor for this, it might be interesting to do so. Our formalization is as follows, including a loose but explicit bound of 24:

```
|- ∀n. ¬(n = 0)
       ⇒ abs(sum {p | prime p ∧ p <= n} (λp. log(&p) / &p) − log(&n)) <= &24
```

7.3 The Summability Result

Here we use Mertens's theorem to deduce the summability of $\sum_n(\sum_{p\leq n}\frac{\log p}{p}-\log n+c)/n$ for some constant c. To arrive at this result, Newman manipulates various functions that one can see to be analytic modulo a few removable singularities and which have series approximations subject to various constraints. Roughly following his approach, we start by showing that there is a function, which we call `nearnewman`, that is analytic and has a suitable series expansion for $\Re z > 1/2$. (This is a case where we re-use our '`nearzeta`' function, emphasizing that this is not only used to get properties of the zeta function.)

```
|- ∀s. s IN {s | Re(s) > &1 / &2}
       ⇒ ((λp. clog(Cx(&p)) / Cx(&p) * nearzeta p s −
                 clog(Cx(&p)) / (Cx(&p) cpow s * (Cx(&p) cpow s − Cx(&1))))
            sums (nearnewman s)) {p | prime p} ∧
            nearnewman complex_differentiable (at s)
```

We use this to define a function

$$\mathrm{newman}(z) = (\mathrm{nearnewman}(z) - \zeta'(z)/\zeta(z))/(z-1)$$

then apply Mertens's theorem to deduce the following series expansion, where `vsum` gives the sum of a complex or vector-valued function over a set:

```
|- ∀s. &1 < Re s
       ⇒ ((λn. vsum {p | prime p ∧ p <= n} (λp. clog(Cx(&p)) / Cx(&p)) /
                 Cx(&n) cpow s)
            sums (newman s)) (from 1)
```

Newman, who calls our nonce '`newman`' function simply 'f', presents things a little differently, starting by 'defining' $f(z)$ for $\Re z > 1$ by the series expansion formalized in the last box and writing

$$f(z) = \sum_{n=1}^{\infty}\frac{1}{n^z}\Big(\sum_{p\leq n}\frac{\log p}{p}\Big) = \sum_{p}\frac{\log p}{p}\Big[\sum_{n\geq p}\frac{1}{n^z}\Big].$$

(The implied equality between these different orders of summation of the double series, simply stated without comment by Newman, also occurs in our proofs, and we needed a 116-line proof script to justify it. Note that this is an argument about double limits, not just rearranging a finite sum.) Newman then essentially proves that the function f corresponds to our definition on a larger set.

The next step is to analyze the behavior of various functions at 1 and show that for a suitable c, the function $F(z) = \mathrm{newman}(z) + \zeta'(z) + c\zeta(z)$, with a suitable assignment $F(1) = a$, is actually analytic for $\Re z \geq 1$, even though the components have singularities at $z = 1$:

```
|- ∃c a. (λz. if z = Cx(&1) then a
               else newman z + complex_derivative zeta z + c * zeta z)
          analytic_on {s | Re(s) >= &1}
```

Finally we can apply the analytic lemma and deduce the result we want:

```
|- ∃c. summable (from 1)
        (λn. (vsum {p | prime p ∧ p <= n} (λp. clog(Cx(&p)) / Cx(&p)) -
             clog(Cx(&n)) + c) / Cx(&n))
```

7.4 The Limit Result

The next step derives from that summability the fact that $\sum_{p\leq n} \frac{\log p}{p} - \log n$ tends to a limit, or in our formalization

```
|- ∃c. ((λn. sum {p | prime p ∧ p <= n} (λp. log(&p) / &p) - log(&n))
     --> c) sequentially
```

(Here ‘--->’ is a limit over real numbers, whereas ‘-->’ used earlier is over complex numbers or other vectors. It is straightforward to map such theorems about real-valued complex numbers to the actual type of real numbers.) Newman gives an unusually clear and explicit proof of this part, with straightforward manipulations of sums. This must explain the fact observed above that this part of the proof has a relatively low de Bruijn factor.

7.5 The PNT

Here Newman teases the reader a little with his statement that the PNT is now ‘easily derived by a simple summation by parts which we leave to the reader’. We came up with a proof that might plausibly be what Newman had in mind when writing that passage. We start with an artificial-looking expression for the number of primes $\leq n$, where `CARD s` denotes the cardinality $|s|$ of a (finite) set s and `m..n` is the set $\{x \in \mathbb{N} \mid m \leq x \wedge x \leq n\}$:

```
|- &(CARD {p | prime p ∧ p <= n}) =
        sum(1..n)
          (λk. &k / log(&k) *
                (sum {p | prime p ∧ p <= k} (λp. log(&p) / &p) -
                 sum {p | prime p ∧ p <= k - 1} (λp. log(&p) / &p)))
```

The proof then uses the standard partial summation identities, which are sometimes more convenient in 'successor' form:

```
|- sum (m..n) (λk. f(k) * (g(k + 1) - g(k))) =
           if m <= n then f(n + 1) * g(n + 1) - f(m) * g(m) -
                          sum (m..n) (λk. g(k + 1) * (f(k + 1) - f(k)))
           else &0
```

and sometimes in 'predecessor' form:

```
|- sum (m..n) (λk. f(k) * (g(k) - g(k - 1))) =
           if m <= n then f(n + 1) * g(n) - f(m) * g(m - 1) -
                          sum (m..n) (λk. g k * (f(k + 1) - f(k)))
           else &0
```

Following a series of routine algebraic calculations and bounds on the quantities involved, we finally obtain the PNT. The usual informal statement is that $\pi(n) \sim n/\log(n)$. In our formalization we do not use the auxiliary concepts $\pi(x)$ and '$\sim$' (though we easily could), but spell things out:

```
|- ((λn. &(CARD {p | prime p ∧ p <= n}) / (&n / log(&n)))
    --> &1) sequentially
```

8 Conclusions

We have been able to formalize Newman's proof without any undue difficulty. By this we mean that once we understood a passage in the proof informally, its formalization was generally fairly routine. Of course, the formalization presents a few unnatural features, as formalizations often do, such as explicit casts between number systems and relational versions of limits and integrals. But on the whole the structure of the formalization is quite close to that of the underlying text. The most striking exceptions, as we have noted, are those where our geometric understanding of winding numbers and other properties of paths allow us to see at a glance something that is quite hard to formalize. On the whole, we think we actually spent almost as much effort trying to understand Newman's proof and its prerequisites informally as we did typing a formalization into HOL.

It is difficult to compare in any objective way the difficulty of our formalization against the one undertaken by Avigad et al., because the latter involved different people, and a different theorem prover. Our impression is that, as we might expect, the analytic proof in itself was easier to formalize, but that if one includes the analytic preliminaries like Cauchy's theorem, it becomes substantially harder. Of course, the appeal of using this approach is that Cauchy's theorem and similar developments are useful in other contexts and the hard work of formalizing them can be considered as amortized across many future applications. It remains to be seen how this ideal plays out in practice.

We made a conscious effort not to take ad hoc short-cuts and to develop background theory systematically, in an attempt to present a contrasting approach to a heroic assault on the big theorem. Whenever we seemed to be doing something ad hoc, we always tried to go back to our library of complex analysis and prove the most natural, general theorems. For example, we proved at one point that the winding number of a simple closed curve (one that intersects itself only at the single endpoint) is always either -1, 0 or $+1$, because this seemed potentially useful. In full generality, this is a non-trivial result (we used a proof suggested by Tom Hales based on Brouwer's fixed point theorem), and it was never actually used in our PNT proof in the end. Rather to our surprise, simply proving that the 4-component contour we use *is* simple seemed rather laborious, since a priori one needs to consider all the distinct pairs of components and show that they only intersect in the appropriate way. The approach we actually used, as described in Section 6, gives a smoother way of deducing the basic properties of our composite contour in a modular fashion from those of the components.

In any case, we must admit that we have sometimes fallen short of our own ideals. Some estimations of sums were done by comparing them to complex line integrals, simply because we had all the theorems available, when it would have been much more natural to use integrals over $\mathbb{R}$. More significantly, we only proved analyticity of the ζ-function on $\Re z > 0$ and $z \neq 1$, whereas it can actually be analytically continued throughout the complex plane except for the point $z = 1$. It would be more consistent with our approach to develop the general theory of the ζ-function in itself.

Acknowledgement The author would like to thank the anonymous reviewers for their helpful comments, which have significantly improved the final version of the paper.

References

1. Agrawal, M., Kayal, N., Saxena, N.: PRIMES is in P. Ann. Math. **160**, 781–793 (2004)
2. Avigad, J., Donnelly, K.: Formalizing O notation in Isabelle/HOL. In: Basin D., Rusinowitsch M. (eds.) Proceedings of the Second International Joint Conference on Automated Reasoning. Lecture Notes in Computer Science, vol. 3097, pp. 357–371. Springer, Cork (2004)
3. Avigad, J., Donnelly, K., Gray, D., Raff, P.: A formally verified proof of the prime number theorem. Acm Trans. Comput. Log. **9**(1:2), 1–23 (2007)
4. Bak, J. Newman, D.J.: Complex Analysis. Springer, New York (1997)
5. Boulton, R., Gordon, A., Gordon, M., Harrison, J., Herbert, J., Van Tassel, J.: Experience with embedding hardware description languages in HOL. In: Stavridou, V., Melham, T.F., Boute, R.T. (eds.) Proceedings of the IFIP TC10/WG 10.2 International Conference on Theorem Provers in Circuit Design: Theory, Practice and Experience. IFIP Transactions A: Computer Science and Technology, vol. A-10, pp. 129–156. North-Holland, Nijmegen (1993)
6. Conway, J. Kochen, S.: The free will theorem. Found. Phys. **36**, 1441 (2006)
7. Davis, M.: Obvious logical inferences. In: Hayes, P.J. (ed.) Proceedings of the Seventh International Joint Conference on Artificial Intelligence, pp. 530–531. Kaufmann, Ingolstadt (1981)
8. Gordon, M.J.C.: Representing a logic in the LCF metalanguage. In: Néel, D. (ed.) Tools and Notions for Program Construction: An Advanced Course, pp. 163–185. Cambridge University Press, Cambridge (1982)
9. Gordon, M.J.C., Melham, T.F.: Introduction to HOL: A Theorem Proving Environment for Higher Order Logic. Cambridge University Press, Cambridge (1993)
10. Gordon, M.J.C., Milner, R., Wadsworth, C.P.: Edinburgh LCF: A Mechanised Logic of Computation. Lecture Notes in Computer Science, vol. 78. Springer, Cambridge (1979)
11. Harrison, J.: Constructing the real numbers in HOL. In: Claesen, L.J.M., Gordon, M.J.C. (eds.) Proceedings of the IFIP TC10/WG10.2 International Workshop on Higher Order Logic Theorem Proving and its Applications. IFIP Transactions A: Computer Science and Technology, vol. A-20, pp. 145–164. IMEC, Leuven (1992)
12. Harrison, J.: Theorem Proving with the Real Numbers. Springer, New York (1998) (Revised version of author's PhD thesis)

13. Harrison, J.: Formal verification of floating point trigonometric functions. In: Hunt, W.A., Johnson, S.D. (eds.) Formal Methods in Computer-Aided Design: Third International Conference FMCAD 2000. Lecture Notes in Computer Science, vol. 1954, pp. 217–233. Springer, New York (2000)
14. Harrison, J.: Complex quantifier elimination in HOL. In: Boulton, R.J., Jackson, P.B. (eds.) TPHOLs 2001: Supplemental Proceedings, pp. 159–174. Division of Informatics, University of Edinburgh. Published as Informatics Report Series EDI-INF-RR-0046. http://www.informatics.ed.ac.uk/publications/report/0046.html (2001)
15. Harrison, J.: Isolating critical cases for reciprocals using integer factorization. In: Bajard, J.-C., Schulte, M. (eds.) Proceedings, 16th IEEE Symposium on Computer Arithmetic, pp. 148–157. Santiago de Compostela, Spain. IEEE Computer Society, Los Alamitos (2003) (http://www.dec.usc.es/arith16/papers/paper-150.pdf)
16. Harrison, J.: A HOL theory of Euclidean space. In: Hurd, J., Melham, T. (eds.) Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005. Lecture Notes in Computer Science, vol. 3603, pp. 114–129. Springer, Oxford (2005)
17. Harrison, J.: The HOL light tutorial. http://www.cl.cam.ac.uk/~jrh13/hol-light/tutorial.pdf (2006)
18. Harrison, J.: Formalizing basic complex analysis. In: Matuszewski, R., Zalewska, A. (eds.) From Insight to Proof: Festschrift in Honour of Andrzej Trybulec. Studies in Logic, Grammar and Rhetoric, vol. 10(23), pp. 151–165. University of Białystok, Białystok (2007)
19. Ingham, A.E.: The Distribution of Prime Numbers. Cambridge University Press, Cambridge (1932)
20. Jameson, G.J.O.: The Prime Number Theorem. London Mathematical Society Student Texts, vol. 53. Cambridge University Press, Cambridge (2003)
21. Newman, D.J.: Analytic Number Theory. Graduate Texts in Mathematics, vol. 177. Springer, New York (1998)
22. Rudnicki, P.: Obvious inferences. J. Autom. Reason. **3**, 383–393 (1987)
23. Tenenbaum, G., France, M.M.: The Prime Numbers and Their Distribution. Student Mathematical Library, vol. 6. American Mathematical Society, Providence (2000) (Translation by Philip G. Spain, from French original "Nombres premiers", Presses Universitaires de France, 1997.)
24. Théry, L., Hanrot, G.: Primality proving with elliptic curves. In: Schneider, K., Brandt, J. (eds.) Proceedings of the 20th International Conference on Theorem Proving in Higher Order Logics, TPHOLs 2007. Lecture Notes in Computer Science, vol. 4732, pp. 319–333. Springer, Kaiserslautern (2007)
25. Wiedijk, F.: The de Bruijn factor. http://www.cs.ru.nl/~freek/factor/ (2000)
26. Wiedijk, F.: The seventeen provers of the world. In: Lecture Notes in Computer Science, vol. 3600. Springer, New York (2006)
27. Wos, L., Pieper, G.W.: A Fascinating Country in the World of Computing: Your Guide to Automated Reasoning. World Scientific, Singapore (1999)
28. Zagier, D.: Newman's short proof of the prime number theorem. Am. Math. Mon. **104**, 705–708 (1997)