A formalized proof of Dirichlet's theorem on primes in arithmetic progression

JOHN HARRISON
Intel Corporation JF1-13, 2111 NE 25th Ave, Hillsboro OR 97124, USA
johnh@ichips.intel.com

We describe the formalization using the HOL Light theorem prover of Dirichlet's theorem on primes in arithmetic progression. The proof turned out to be more straightforward than expected, but this depended on a careful choice of an informal proof to use as a starting-point. The goal of this paper is twofold. First we describe a simple and efficient proof of the theorem informally, which is otherwise difficult to find in one self-contained place at an elementary level. We also describe its, largely routine, HOL Light formalization, a task that took only a few days.

1. INTRODUCTION

Dirichlet's theorem asserts that for all pairs of positive integers k and d that are coprime (have no common integer factor besides 1), there are infinitely many primes p such that $p \equiv k \pmod{d}$, i.e. that the infinite arithmetic progression $k, k+d, k+2d, k+3d, \ldots$ contains infinitely many primes. (The coprimality condition is easily seen to be necessary, for any common divisor of k and d would divide all members of this progression.) This result was first conjectured by Euler in the case $k = 1$, and by Legendre in full generality. It was first proved in 1837 by Dirichlet [Dir37], who in the process introduced *L-functions* and, indeed, more or less began the subject of analytic number theory in its modern form.

In this paper, we will present an elementary self-contained proof of Dirichlet's theorem culled from various sources, and describe its complete formalization in the HOL Light theorem prover [Har96], the culmination of which is the following formal statement:

```
|- ∀d k. 1 <= d ∧ coprime(k,d)
        ⇒ INFINITE {p | prime p ∧ (p == k) (mod d)}
```

We have extensive experience of formalizing mathematics, both for its use in applications and just for general intellectual interest. Among the theorems that have been formalized by others [ADGR07] and by ourselves [Har09] is the Prime Number Theorem. And yet for a long time, we had been reluctant to embark on Dirichlet's theorem because it seemed intimidatingly difficult. For example, Hardy and Wright's famous textbook [HW79] gives a detailed elementary proof of the Prime Number Theorem yet stops short of proving Dirichlet's theorem. We were finally persuaded that perhaps it wasn't as difficult as we thought by skimming through the proof given by Gelfond and Linnik [GL65], which was only about 5 pages of fairly elementary-looking manipulations of sums (section 3.2, pp. 47–52). This does assume some background concerning Dirichlet characters, but this was mostly easy to put together for ourselves after a quick glance at a Wikipedia page. Moreover, while browsing the Web we discovered a paper by Monsky [Mon90]

that gives an even more straightforward proof of the most complicated step, the nonvanishing of the L-series associated with a real nonprincipal character. Thus our final proof is a synthesis of three parts:

—A 'roll your own' development of basic properties of Dirichlet characters.
—Monsky's proof that an L-series for a real nonprincipal character does not vanish.
—The remainder of the proof following Gelfond and Linnik

Although it took some time to browse these sources and understand the basic outline of the argument, the actual translation to a formal HOL Light proof was unproblematic. The majority of the formalization was completed over a Thanksgiving long weekend, representing about 3.5 days of full-time work (half a day off for celebrations). But including additional time completing a few details, the entire proof from start to finish probably represents about 5 days of full-time work. This represents at least one page of informal text per day. Though those unfamiliar with formalization may find that grindingly slow, it is actually much better than the commonly quoted estimate of a page a week [Wie06].

2. GROUNDWORK

In this section we cover the basic mathematical infrastructure used in the proof. Subsection 2.1 covers established parts of the HOL Light library, largely to be explicit about notation and prerequisites. In 2.2 we show how some previous results from the Prime Number Theorem proof were taken over and adapted. Finally in 2.3 we show how new background material to support the Dirichlet proof was formalized.

2.1 Number theory, real and complex numbers

We start by noting the formal HOL notations we use for notions from number theory. These are defined in established HOL Light library files (all these are in one of `int.ml`, `Examples/prime.ml` and `Examples/pocklington.ml`). Note that all the properties here are of natural numbers $\mathbb{N}$, and the types of variables are assumed to be $\mathbb{N}$. (Analogous notions are defined for integers $\mathbb{Z}$ but we make little or no use of them.) The following means that 'a divides b' (i.e. b is exactly divisible by a) usually written $a|b$:

```
|- a divides b ⇔ ∃x. b = a * x
```

Congruence of x and y modulo n, written $x \equiv y \pmod{n}$, is in fact defined in terms of its integer counterpart using the injection '`&`' from $\mathbb{N}$ to $\mathbb{Z}$:

```
|- (x == y) (mod n) ⇔ (&x == &y) (mod &n)
```

while this is itself defined in the natural way as $x \equiv y \pmod{n} \Leftrightarrow \exists q.\, x - y = q \cdot n$. (Note that this doesn't work directly over $\mathbb{N}$ because subtraction is defined so that $x - y = 0$ for $x < y$.)

We say two natural numbers are *coprime* or *relatively prime* if they have no common (integer) factor besides 1. Informally this is often written just $(a, b) = 1$ where in this context (a, b) is an abbreviation for the greatest common divisor of a and b. We could express it in that way too using `gcd(a,b) = 1`, but instead we directly use the following binary predicate:

```
|- coprime(a,b) ⇔ ∀d. d divides a ∧ d divides b ⇒ d = 1
```

We already have a theorem relating this to the characterization using GCDs:

```
|- ∀a b. coprime(a,b) ⇔ gcd(a,b) = 1
```

Finally, a natural number p is prime if it is not 1 and has no factors besides itself and 1:

```
|- prime(p) ⇔ ¬(p = 1) ∧ ∀x. x divides p ⇒ x = 1 ∨ x = p
```

We often use sums over sets of natural numbers. While these sums are defined over general sets, we often use the special notation $m..n$ for set of natural numbers between m and n (inclusive):

```
|- m..n = {x:num | m <= x ∧ x <= n}
```

Then $\sum_{i=m}^{n} f[i]$ can be expressed using `sum(m..n) (λi. f[i])`, and a more general sum over a set $\sum_{i \in s} f[i]$ as `sum s (λi. f[i])`. More precisely there are three different notions of summation, with `nsum` for sums of natural numbers, `sum` for real numbers and `vsum` for complex numbers or vectors in $\mathbb{R}^N$. In each case, the indexing set over which the sum is defined can be of any type, though very often it is a number segment $m..n$. In the present work, sums are only taken over finite sets, though the notion also works as expected for infinite sets where the function being summed is nonzero only finitely often (e.g. $\sum_{n \in \mathbb{N}} \lfloor x^2/n \rfloor$). We also have products $\Pi_{i \in s} f[i]$ for natural numbers (`nproduct`), real numbers (`product`) and complex numbers (`cproduct`), which are likewise defined over arbitrary sets where the function is 1 for all but finitely many elements of the set.

As well as a collection of useful theorems, the HOL Light system contains some convenient automated procedures for proving routine facts of algebra, arithmetic and number theory. One of the more interesting [Har07a] can prove many otherwise tedious lemmas about divisibility automatically, for example

```
|- ∀d a b. d divides (a * b) ∧ coprime(d,a) ⇒ d divides b
```

We also use the Euler totient function $\phi(n)$, which is the number of natural numbers $0 < m \leqslant n$ and coprime to n. (Note that n can only be coprime to itself or to zero in the case $n = 1$, so for $n \neq 1$ it's immaterial whether one has $0 \leqslant m$ or $0 < m$, and whether one has $m < n$ or $m \leqslant n$. But we want to ensure $\phi(1) = 1$.)

```
|- phi(n) = CARD { m | 0 < m ∧ m <= n ∧ coprime(m,n) }
```

The HOL types `real` and `complex`, the real and complex numbers, play an important role in our proof. In this section we just mention some points of notation, and refer the reader elsewhere for details about how these types are constructed [Har98, Har01, Har07b]. Note that the HOL Light numeric types are completely distinct, so explicit injections are used to cast between them. The ones that will appear a lot in what follows are '`&`', which is overloaded to casts from the natural numbers `num` to several other number systems including `real`, and `Cx`, which is the cast from real to complex numbers. For example, the complex number 0 has the rather verbose representation '`Cx(&0)`'. Generally speaking, we use the standard operator names like '`+`', which are overloaded to various number systems. In both

real and complex numbers *unary* negation $-x$ is denoted by '`--x`'. Powers of real or complex numbers x^n are written '`x pow n`', while those of natural numbers m^n are written '`m EXP n`'. (Note that in either case the exponent n is a natural number.) The imaginary unit i is denoted by `ii` and the complex conjugation operation is `cnj`. The absolute value $|x|$ of a real number is denoted by '`abs(x)`' and the norm $\|z\|$ of a complex number by '`norm(z)`'. In fact complex numbers are synonymous with the type $\mathbb{R}^2$, and some concepts like norms are defined for general vectors in $\mathbb{R}^N$ of which the complex-number versions are just a special case.

2.2 von Mangoldt function and Mertens's estimates

The present proof re-uses some definitions and results that were proved in our formalization of the Prime Number Theorem [Har09]. (Note that this is also assumed without proof in our main source text [GL65], so it is fair to re-use it without counting it as a direct part of the formalization when comparing lengths.) The re-used material includes a definition of the von Mangoldt function $\Lambda(n)$, which is defined as $\Lambda(p^k) = \log p$ for a prime power p^k and $\Lambda(n) = 0$ otherwise. The formal definition uses the Hilbert choice operator ε, with $\varepsilon x.\ P[x]$ to be read as 'some x such that $P[x]$' [Lei69]. The characterizing axiom of this operator in HOL Light is essentially $(\exists x.\ P[x]) \Rightarrow P[\varepsilon x.\ P[x]]$; if there is no x such that $P[x]$ then $\varepsilon x.\ P[x]$ is just an element (of the appropriate type) about which little is known. The predicate $P[x]$ can involve additional variables, and since from $\forall x.\ (\exists y.\ P[x, y]) \Rightarrow P[x, \varepsilon y.\ P[x, y]]$ we can deduce $(\forall x.\ \exists y.\ P[x, y]) \Rightarrow (\forall x.\ P[x, \varepsilon y.\ P[x, y]])$, this builds in the Axiom of Choice.

```
|- mangoldt n = if ∃p k. 1 <= k ∧ prime p ∧ n = p EXP k
                 then log(&(εp. prime p ∧ p divides n))
                 else &0
```

We use some key properties of this function directly, notably the following lemma `LOG_MANGOLDT_SUM` expressing $\log n$ as a sum over divisors of n of the von Mangoldt function:

```
|- ∀n. ¬(n = 0)
       ⇒ log(&n) = sum {d | d divides n} (λd. mangoldt(d))
```

More significantly, we use some non-trivial estimates about the density of primes that were derived via properties of Λ. The principal result used as a starting point for the Prime Number Theorem is Mertens's theorem, called `MERTENS`. (Note that the explicit bounds like 24 here are seldom sharp, nor are they intended to be. It's just slightly more convenient to have a concrete number than an existentially quantified variable, but not worth the trouble of making the bound near-optimal.)

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

For the present application, we split the formerly monolithic proof of this into two components and slightly generalized one of them so that we could use two of the intermediate results, `MERTENS_LEMMA`

```
|- ∀n. ¬(n = 0)
       ⇒ abs(sum(1..n) (λd. mangoldt(d) / &d) - log(&n)) <= &21
```

and MERTENS_MANGOLDT_VERSUS_LOG, which was formerly proved only for the special case where s is the full number segment $1..n$:

```
|- ∀n s.
       s SUBSET 1..n
       ⇒ abs(sum s (λd. mangoldt d / &d) -
             sum {p | prime p ∧ p ∈ s} (λp. log (&p) / &p)) <= &3
```

2.3 New formalized background

We had to formalize two general pieces of "background" theory that we hadn't already dealt with but which were assumed in our source text [GL65]. The first is *Möbius inversion*, which took several hours just in itself. (Note that there is no novelty involved in formalizing this: it has been done before on more than one occasion [ADGR07, AA08].) This uses the Möbius function, where $\mu(n)$ is defined to be zero if n has a squared (prime) factor and otherwise $(-1)^k$ where k is the number of distinct prime factors of n. The definition that follows is as a function $\mathbb{N} \to \mathbb{R}$, i.e. of HOL type '`num->real`', but it can be defined as a mapping into any other ring in the same way.

```
|- mobius(n) = if ∃p. prime p ∧ (p EXP 2) divides n then &0
               else --(&1) pow CARD {p | prime p ∧ p divides n}
```

The crucial property of this function that we use is the Möbius inversion formula, which allows one to invert the definition of a function defined over divisors: if $g(n) = \sum_{d|n} f(d)$ for all $n \geqslant 1$ then conversely $f(n) = \sum_{d|n} \mu(d)g(\lfloor n/d \rfloor)$, or in HOL Light:

```
|- ∀f g. (∀n. 1 <= n ⇒ g(n) = sum {d | d divides n} f)
         ⇒ ∀n. 1 <= n
                ⇒ f(n) = sum {d | d divides n}
                             (λd. mobius(d) * g(n DIV d))
```

For example, combining Möbius inversion and the theorem LOG_MANGOLDT_SUM we easily obtain the following theorem MANGOLDT_LOG_SUM:

```
|- ∀n. 1 <= n
       ⇒ mangoldt(n) = --(sum {d | d divides n}
                             (λd. mobius(d) * log(&d)))
```

The Möbius function is an example of a *multiplicative* function, meaning that $\mu(mn) = \mu(m)\mu(n)$ whenever m and n are coprime. (If this is true for arbitrary m and n, we say a function is *completely multiplicative*.) We define this notion for functions $f : \mathbb{N} \to \mathbb{R}$ as follows:

```
|- real_multiplicative f ⇔
     f(1) = &1 ∧ ∀m n. coprime(m,n) ⇒ f(m * n) = f(m) * f(n)
```

An important lemma we use later, which was already employed in the derivation of Möbius inversion, is that sums of multiplicative functions over divisors are also

multiplicative, i.e. if f is multiplicative then so is $g(n) = \sum_{d|n} f(d)$. The proof is straightforward based on the observation that if m and n are coprime, each $d|mn$ factors uniquely into $d = d'd''$ where $d'|m$ and $d''|n$.

```
|- ∀f. real_multiplicative f
      ⇒ real_multiplicative (λn. sum {d | d divides n} f)
```

The other new background lemmas we had to formalize involve the Dirichlet convergence test for series. We started with `SERIES_DIRICHLET_COMPLEX`, the typical high-level statement, which essentially asserts that if the partial sums $\sum_{n=a}^{b} f_n$ of a complex series $\sum_n f_n$ are bounded and g_n is a sequence of reals that decreases monotonically to zero, then the product series $\sum_n f_n g_n$ is convergent.

```
|- ∀f g N k m.
       bounded {vsum(m..n) f | n ∈ ℕ} ∧
       (∀n. real(g n)) ∧
       (∀n. N <= n ⇒ Re(g(n + 1)) <= Re(g n)) ∧
       (g --> Cx(&0)) sequentially
       ⇒ summable (from k) (λn. f(n) * g(n))
```

The proof is not difficult using partial summation, i.e. the transformation of the sum $\sum_{k=m}^{n} f(k)(g(k) - g(k-1))$ into $f(n+1)g(n) - f(m)g(m-1) - \sum_{k=m}^{n}(f(k+1) - f(k))g(k)$ etc. This is the version of partial summation we use, for a general bilinear function over pairs of vectors:

```
|- ∀f g h m n.
       bilinear h
       ⇒ vsum(m..n) (λk. h (f k) (g k - g (k - 1))) =
           (if m <= n
            then h (f (n + 1)) (g n) -
                 h (f m) (g (m - 1)) -
                 vsum(m..n) (λk. h (f (k + 1) - f k) (g k))
            else vec 0)
```

We started out with even more abstract and general forms of the Dirichlet convergence test than one typically sees in books, where the series $\sum_n f_n$ is in $\mathbb{R}^n$. However, when we actually came to apply the Dirichlet test in our formalization, it turned out that the proof uses more than just the top-level statement, and we needed to unpack the proof to make some bounds on the convergence more explicit, resulting in `SERIES_DIRICHLET_COMPLEX_EXPLICIT`:

```
|- ∀f g p q. 1 <= p ∧
            bounded {vsum(q..n) f | n ∈ ℕ} ∧
            (∀n. p <= n ⇒ real(g n) ∧ &0 <= Re(g n)) ∧
            (∀n. p <= n ⇒ Re(g(n + 1)) <= Re(g n))
            ⇒ ∃B. &0 < B ∧
                    ∀m n. p <= m
                          ⇒ norm(vsum(m..n) (λk. f k * g k))
                                  <= B * norm(g m)
```

and later an even more explicit form SERIES_DIRICHLET_COMPLEX_VERY_EXPLICIT of which this is an easy corollary:

```
|- ∀f g B p. &0 < B ∧ 1 <= p ∧
            (∀m n. p <= m ⇒ norm(vsum(m..n) f) <= B) ∧
            (∀n. p <= n ⇒ real(g n) ∧ &0 <= Re(g n)) ∧
            (∀n. p <= n ⇒ Re(g(n + 1)) <= Re(g n))
            ⇒ ∀m n. p <= m
                      ⇒ norm(vsum(m..n) (λk. f k * g k))
                          <= &2 * B * norm(g m)
```

These versions are used to place explicit bounds on the partial sums, which appear in several later arguments. We do still use the 'top-level' statement to verify the basic fact that L-function series converge (see section 4), but in several other places use explicit bound information.

3. DIRICHLET CHARACTERS

Roughly speaking, a Dirichlet character modulo d is a multiplicative homomorphism $\chi : \mathbb{N} \to \mathbb{C}$ from the integers modulo d to the complex numbers. We define the concept as follows:

```
|- dirichlet_character d c ⇔
        (∀n. c(n + d) = c(n)) ∧
        (∀n. c(n) = Cx(&0) ⇔ ¬coprime(n,d)) ∧
        (∀m n. c(m * n) = c(m) * c(n))
```

That is, a character χ is periodic with period d, $\chi(n) \neq 0$ iff n and d are coprime, and χ is completely multiplicative. Although this formulation is most convenient for us, we can equivalently consider a Dirichlet character simply as a homomorphism from the multiplicative group $\{n \mid 0 \leqslant n < d \wedge \mathrm{coprime}(n, d)\}$ that we extend to $\mathbb{N}$ or $\mathbb{Z}$ by periodicity and set to zero elsewhere. The advantage of this formulation is that the concept can be related to more general group characters in deriving useful theorems. However, we were not directly interested in developing this group-theoretic machinery, and with one exception described below it was easy to verify the properties we need directly. First of all, it follows from multiplicativity that $\chi(1) \cdot \chi(1) = \chi(1 \cdot 1) = \chi(1)$ and so, since 1 is coprime to any d and therefore $\chi(1) \neq 0$, that $\chi(1) = 1$:

```
|- ∀d c. dirichlet_character d c ⇒ c(1) = Cx(&1)
```

Using multiplicativity again with this as the base case, we can prove by induction that $\chi(m^n) = \chi(m)^n$:

```
|- ∀d c m n. dirichlet_character d c ⇒ c(m EXP n) = c(m) pow n
```

and we can derive similarly easy consequences of periodicity such as the following:

```
|- ∀d c m n.
        dirichlet_character d c ∧ (m == n) (mod d) ⇒ c(m) = c(n)
```

By Euler's generalization of Fermat's Little Theorem, if n is coprime to d we have $n^{\phi(d)} \equiv 1 \pmod{d}$. Using the power property we deduce that $\chi(n)^{\phi(d)} = 1$:

```
|- ∀d c n. dirichlet_character d c ∧ coprime(d,n)
          ⇒ c(n) pow phi(d) = Cx(&1)
```

It is an immediate consequence that the values $\chi(n)$ are unimodular complex numbers, or zero if n is not coprime to d:

```
|- ∀d c n. dirichlet_character d c
          ⇒ norm(c n) = if coprime(d,n) then &1 else &0
```

An important character in what follows is χ_0, the *principal character* mod d, which is defined to be 1 for all n coprime to d:

```
|- chi_0 d n = if coprime(n,d) then Cx(&1) else Cx(&0)
```

We note that for $d < 2$ there is no Dirichlet character mod d other than the principal character. This sometimes allows us to omit trivial $2 \leqslant d$ assumptions on theorems because they become degenerately true anyway.

An important lemma for us is that $\sum_{n=1}^{d} \chi(n) = 0$ for any $\chi \neq \chi_0$. To prove this, note that since χ is nonprincipal, there must be some m coprime to d such that $\chi(m) \neq 1$. It therefore suffices to show that $\chi(m) \cdot \sum_{n=1}^{d} \chi(n) = \sum_{n=1}^{d} \chi(n)$. But this follows since $\chi(m) \cdot \sum_{n=1}^{d} \chi(n) = \sum_{n=1}^{d} \chi(mn) = \sum_{n=1}^{d} \chi(n)$ since multiplication by m of the numbers $1 \leqslant n \leqslant d$ coprime to d just permutes the same set, modulo d. Thus we conclude:

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d)
         ⇒ vsum(1..d) c = Cx(&0)
```

It follows that we can reduce sums of Dirichlet characters over segments of the natural numbers by collapsing such blocks of d integers to zero:

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d)
         ⇒ vsum(1..n) c = vsum(1..(n MOD d)) c
```

On the other hand, summing the *principal* character over the integers $1..d$ will give $\phi(d)$, so we have:

```
|- ∀d c. dirichlet_character d c
         ⇒ vsum(1..d) c =
             if c = chi_0 d then Cx(&(phi d)) else Cx(&0)
```

For the proofs that follow, it's convenient to know that the set of Dirichlet characters mod d is always finite. Later, we will prove that there are exactly $\phi(d)$ of them, but in order to derive such results we first need to justify taking sums over the set of Dirichlet characters. Our finiteness result is simply derived by observing that a Dirichlet character is, by periodicity, determined by its effect on the numbers $1..d$, and for each of those values it is either zero or a $\phi(d)^{th}$ root of unity, of which there are finitely many.

```
|- ∀d. FINITE {c | dirichlet_character d c}
```

That concludes the most basic properties of Dirichlet characters. The subsequent results start to exploit algebraic structure on the set of characters mod d. If χ is a Dirichlet character mod d, then so is the *conjugate character* $\overline{\chi}$ defined by pointwise complex conjugation:

```
|- ∀d c. dirichlet_character d c
        ⇒ dirichlet_character d (λn. cnj(c n))
```

The set of Dirichlet characters mod d is also closed under multiplication defined pointwise:

```
|- ∀d c1 c2. dirichlet_character d c1 ∧ dirichlet_character d c2
             ⇒ dirichlet_character d (λn. c1(n) * c2(n))
```

It is clear that this operation is associative and commutative. Moreover, because the nonzero values of a Dirichlet character are unimodular, we have $\overline{\chi(n)}\chi(n) = \|\chi(n)\|^2 = 1^2 = 1$ whenever n and d are coprime, so the conjugation gives an inverse with respect to our multiplication with the principal character as the identity:

```
|- ∀d c. dirichlet_character d c ⇒ (λn. cnj(c n) * c n) = chi_0 d
```

In later informal text, we will occasionally write this multiplication of characters χ and χ' simply as $\chi\chi'$, though in the HOL formalization, this is not defined as an operator and we use the explicit lambda-term '$\lambda n.\ \chi(n)\chi'(n)$' for this product as we did in the inverse law above.

By virtue of the finiteness result above, we can consider sums of the form $\sum_\chi \chi(n)$ where the sum is taken over all Dirichlet characters mod d. The interesting case is when n and d are coprime, since otherwise every Dirichlet character mod d has $\chi(n) = 0$ so we also have $\sum_\chi \chi(n) = 0$. Note that for any specific Dirichlet character χ' we have $\chi'(n)\sum_\chi \chi(n) = \sum_\chi \chi'(n)\chi(n) = \sum_\chi (\chi'\chi)(n) = \sum_\chi \chi(n)$, because by the group properties multiplication by χ' merely permutes the set of characters. We therefore conclude that either $\chi'(n) = 1$ or $\sum_\chi \chi(n) = 0$. So *either* the sum is zero or n and d are coprime and $\chi(n) = 1$ for *every* Dirichlet character.

```
|- ∀d n. vsum {c | dirichlet_character d c} (λx. x n) = Cx(&0) ∨
        coprime(n,d) ∧
        ∀c. dirichlet_character d c ⇒ c(n) = Cx(&1)
```

We now seek to clarify just when this special situation, $\chi(n) = 1$ for all Dirichlet characters, occurs. It is clear that it must happen when $n \equiv 1 \pmod{d}$, because then by the elementary properties noted above we have $\chi(n) = \chi(1) = 1$. In fact it turns out that this is the *only* case where the situation arises, i.e. that if $n \not\equiv 1 \pmod{d}$ then there exists a Dirichlet character with $\chi(n) \neq 1$. This result is used essentially in the Dirichlet proof, and so we needed to formalize it.

Our first thought was a fairly straightforward proof, entirely number-theoretic in character, based on primitive roots modulo prime power factors $p^k|d$. If $n \not\equiv 1 \pmod{d}$, then we must have $n \not\equiv 1 \pmod{p^k}$ for some such factor $p^k|d$. Now if r is a primitive root modulo p^k, then we can define a suitable character by $\chi(a) = e^{2\pi i b/\phi(p^k)}$ where $a \equiv r^b \pmod{p^k}$. While this can be made to work [FR07], and a similar approach is used in Dirichlet's original paper, it needs some additional arguments to cope with the cases $p = 2$, $k \geqslant 3$ when primitive roots modulo p^k do not exist. Besides, we didn't have a ready formalization of the existence of primitive roots modulo powers of odd primes, so all in all, quite a lot of work would have been involved.

Instead we proved a lemma that has a natural and obvious group-theoretic intuition and was inspired by results of that nature in [Jam03]. The intuition is

simply that if we have a "partial" Dirichlet character mod d defined on a subgroup H of the multiplicative group of integers $G = \{n \mid 0 \leqslant n < d \wedge \mathrm{coprime}(n, d)\}$, and $a \in G - H$, then we can extend the definition of that character χ to a larger subgroup H' with $H \cup \{a\} \subseteq H' \subseteq G$ such that $\chi(a) \neq 1$. While that statement seems quite natural, it becomes rather messy when stated in a precise way, purely for the natural numbers:

```
|- ∀f h a d.
       h SUBSET {x | x < d ∧ coprime(x,d)} ∧
       1 ∈ h ∧
       (∀x y. x ∈ h ∧ y ∈ h ⇒ ((x * y) MOD d) ∈ h) ∧
       (∀x. x ∈ h ⇒ ∃y. y ∈ h ∧ (x * y == 1) (mod d)) ∧
       (∀x. x ∈ h ⇒ ¬(f x = Cx(&0))) ∧
       (∀x y. x ∈ h ∧ y ∈ h ⇒ f((x * y) MOD d) = f(x) * f(y)) ∧
       a ∈ {x | x < d ∧ coprime(x,d)} DIFF h
       ⇒ ∃f' h'. (a INSERT h) SUBSET h' ∧
                 h' SUBSET {x | x < d ∧ coprime(x,d)} ∧
                 (∀x. x ∈ h ⇒ f'(x) = f(x)) ∧
                 ¬(f' a = Cx(&1)) ∧
                 1 ∈ h' ∧
                 (∀x y. x ∈ h' ∧ y ∈ h' ⇒ ((x * y) MOD d) ∈ h') ∧
                 (∀x. x ∈ h' ⇒ ∃y. y ∈ h' ∧ (x * y == 1) (mod d)) ∧
                 (∀x. x ∈ h' ⇒ ¬(f' x = Cx(&0))) ∧
                 (∀x y. x ∈ h' ∧ y ∈ h'
                        ⇒ f'((x * y) MOD d) = f'(x) * f'(y))
```

The intuition behind the proof is similarly straightforward inside the group G, but again becomes messy in the detailed realization in terms of integers. Consider the least positive integer m such that $a^m \mod d \in H$. There must indeed be such an m since at worst $a^{\phi(d)} \equiv 1 \pmod{d}$, and $1 \in H$ as it is a subgroup. We cannot have $m = 1$ for that would contradict $a \notin H$, so we can assume $m \geqslant 2$.

Now suppose that for some other m' we have $a^{m'} \in H$. Let us write q and r for the quotient and remainder on dividing m' by m, so $m' = qm + r$ with $0 \leqslant r < m$, and $a^{m'} = (a^m)^q a^r$. Since $a^m \in H$, it has an inverse $b \in H$, and therefore $a^r \mod d = (b^q a^{m'}) \mod d \in H$. But since m was the minimal positive integer with $a^m \mod d \in H$, we must have $r = 0$, i.e. $m | m'$.

Our larger group H' is going to be the set $(xa^k) \mod d$ for $x \in H$. Note that each member of H' can be expressed uniquely in this form with $k < m$ because if $xa^k \equiv ya^l \pmod{d}$, then assuming without loss of generality $l \leqslant k$ and letting $y' \in H$ be a multiplicative inverse for y modulo d we have $a^{l-k} \equiv (xy') \pmod{d}$, which means $m | l - k$, and since $l, k < m$ we have $l = k$ and then by cancellation $x = y$.

Choose $z \in \mathbb{C}$ such that $z^m = f(a^m)$ and $z \neq 1$. Note that this is possible even if $f(a^m) = 1$ since $m \geqslant 2$ and therefore there are at least two distinct m^{th} roots. Because of the uniqueness noted in the previous paragraph, we can define $f' : H' \to \mathbb{C}$ by $f'((xa^k) \mod d) = f(x)z^k$. It is now fairly straightforward to establish all the properties we need, though again they become somewhat obscured by the explicit 'modulo d' operation appearing in many places.

Having established this lemma, it is now easy to use it to get the main result we need, that if $n \not\equiv 1 \pmod{d}$, there is a Dirichlet character χ with $\chi(n) \neq 1$:

```
|- ∀d n. 1 < d ∧ ¬((n == 1) (mod d))
        ⇒ ∃c. dirichlet_character d c ∧ ¬(c n = Cx(&1))
```

We apply the lemma once with H being the trivial subgroup 1, f being the constant function with value 1, and a being $n \mod d$. This establishes a new subgroup H' with $a \in H'$ and an extension f' with $f'(a) \neq 1$. Now it follows by induction on the size of $G - H'$ where again $G = \{n \mid 0 \leqslant n < d \wedge \text{coprime}(n, d)\}$ that we can extend any such mapping f' from a subgroup H' to the whole of G: the induction step just uses the same lemma once again (not taking advantage of the special properties of the mapping, but just the fact that we get at least one more element in the domain). Now given the final f' defined on G we extend it to a character on $\mathbb{N}$ by $\chi(x) = f'(x \mod d)$, and all the properties are assured. In particular we have $\chi(n) = f'(n \mod d) = f'(a) \neq 1$ as required.

Using this we can finally strengthen the result above about sums over the set of all Dirichlet characters to the following so-called 'orthogonality relation':

```
|- ∀d n. vsum {c | dirichlet_character d c} (λc. c n) =
           if (n == 1) (mod d)
           then Cx(&(CARD {c | dirichlet_character d c}))
           else Cx(&0)
```

We can now also deduce that in fact the set of Dirichlet characters mod d has size $\phi(d)$, just by evaluating the double sum $\sum_\chi \sum_{n=1}^{d} \chi(n)$ in both orders. Using our earlier theorem about summing $\sum_{n=1}^{d} \chi(n)$ we obtain

$$\sum_{\chi} \sum_{n=1}^{d} \chi(n) = \sum_{\chi} (\text{if } \chi = \chi_0 \text{ then } \phi(d) \text{ else } 0) = \phi(d)$$

while using the above result about summing over the set of characters we get

$$\sum_{n=1}^{d} \sum_{\chi} \chi(n) = \sum_{n=1}^{d} (\text{if } n \equiv 1 \pmod{d} \text{ then } |X| \text{ else } 0) = |X|$$

where X is the set of all characters mod d. Hence we can get the orthogonality relation in the somewhat nicer form:

```
|- ∀d n. 1 <= d
        ⇒ vsum {c | dirichlet_character d c} (λc. c(n)) =
               if (n == 1) (mod d) then Cx(&(phi d)) else Cx(&0)
```

4. L-FUNCTIONS

The proof of Dirichlet's theorem uses the notion of an *L-function* corresponding to a Dirichlet character. Given a character χ, the corresponding L-function $L_\chi : \mathbb{C} \to \mathbb{C}$

is defined by[1]

$$L_\chi(z) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^z}$$

This series is analogous to $\sum_{n=1}^{\infty} 1/n^z$ for the ζ-function, analytic properties of which play a central role in the Prime Number Theorem. And indeed, sharper extensions of Dirichlet's theorem rely on similar analytic properties [Jam03, New98]. However, for the basic Dirichlet theorem, we only need relatively few properties, and in fact we only need to consider the L-functions at the point $z = 1$, i.e. the values $L_\chi(1)$. Accordingly, we hardwire this into our HOL definition, and hereafter sometimes follow our source text [GL65] in just writing $L(\chi)$ instead of $L_\chi(1)$. Here the HOL Light construct `infsum s x` denotes the limit as $N \to \infty$ of $\sum_{n \in s \wedge n \leqslant N} x_n$, while `from k` is the set of natural numbers $\{n \in \mathbb{N} \mid n \geqslant k\}$, and so this corresponds to the infinite sum in the standard definition.

```
|- Lfunction c = infsum (from 1) (λn. c(n) / Cx(&n))
```

To get any useful results, we first need to establish that for any character $\chi \neq \chi_0$, this series converges. (Note that the series for χ_0, like the series for the ζ-function, does *not* converge for $z = 1$, which indeed is a pole even for their analytic continuations that are holomorphic everywhere else.) This is a direct consequence of the Dirichlet convergence test since for a nonprincipal character the partial sums of the character are bounded; this follows from periodicity and the earlier result $\sum_{n=1}^{d} \chi(n) = 0$ for $\chi \neq \chi_0$. Thus the sum of the series indeed converges to the limit $L_\chi(1)$:

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d)
        ⇒ ((λn. c(n) / Cx(&n)) sums (Lfunction c)) (from 1)
```

In fact, we will later need explicit bounds on the error incurred in truncating the series at the n^{th} term, which is a similar direct application of the theorem `SERIES_DIRICHLET_COMPLEX_EXPLICIT`:

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d)
        ⇒ ∃B. &0 < B ∧
                ∀n. norm(Lfunction c - vsum(1..n) (λn. c(n) / Cx(&n)))
                     <= B / (&n + &1)
```

Few other general properties of the L-functions are used, besides elementary results like the following, which states that the L-function for a conjugate character $\overline{\chi}$ is the conjugate of the L-function for χ:

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d)
        ⇒ Lfunction (λn. cnj(c n)) = cnj(Lfunction c)
```

The key results that follow concern the nonvanishing of the L-functions, i.e. showing that for $\chi \neq \chi_0$ we have $L_\chi(1) \neq 0$. We first consider the case of a (nonprincipal) *real* character, i.e. one where all the values $\chi(n)$ are real numbers, and then turn to other characters.

[1] These are more usually written either $L(\chi, z)$ or $L(z, \chi)$.

5. NONVANISHING OF L-SERIES FOR REAL NONPRINCIPAL CHARACTER

As mentioned earlier, we here follow the article by Monsky [Mon90]. Fix a real Dirichlet character χ, and for any n consider the sum of χ over all divisors of n, i.e. $\sum_{m|n} \chi(m)$. In the case where n is a prime power p^k where $p|d$, we have $\sum_{m|p^k} \chi(m) = 1$, because $\chi(1) = 1$ while $\chi(m) = 0$ whenever m and d are not coprime:

```
|- ∀d c p k. dirichlet_character d c ∧ prime p ∧ p divides d
             ⇒ vsum {m | m divides (p EXP k)} c = Cx(&1)
```

Now we use the fact that χ is a *real* character. For any prime p, regardless of whether it divides d, we have $\sum_{m|p^k} \chi(m) = \sum_{i=0}^{k} \chi(p^i) = \sum_{i=0}^{k} \chi(p)^i$. Since $\chi(p)$ is zero or unimodular and real, we have either $\chi(p) = -1$, $\chi(p) = 0$ or $\chi(p) = 1$, and in each case we see that $0 \leqslant \sum_{m|p^k} \chi(m)$:

```
|- ∀d c p k. dirichlet_character d c ∧ (∀n. real(c n)) ∧ prime p
             ⇒ &0 <= Re(vsum {m | m divides (p EXP k)} c)
```

Now, because Dirichlet characters are (completely) multiplicative and sums over divisors preserve multiplicativity, it follows that the same is true for arbitrary nonzero n, not just prime powers:

```
|- ∀d c n. dirichlet_character d c ∧ (∀n. real(c n)) ∧ ¬(n = 0)
           ⇒ &0 <= Re(vsum {m | m divides n} c)
```

Since $0 \leqslant \sum_{m|n} \chi(m)$ always holds and the sum is 1 for p^k whenever $p|d$, it follows that $\sum_{n=1}^{N} \sum_{m|n} \chi(m) \to \infty$ as $N \to \infty$, which underlies the argument to follow.

We consider the series $f(z) = \sum_{n=1}^{\infty} \chi(n) z^n/(1 - z^n)$ for our given real character χ. By the comparison test, it is easy to see that this series is convergent for $\|z\| < 1$. Moreover, by expanding $1/(1 - z^n) = 1 + z^n + z^{2n} + \cdots$ we have $f(z) = \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \chi(n) z^{kn}$. By a rearrangement of the double infinite series (this is assumed without comment by Monsky, but took us more than 70 lines to formalize), we obtain $f(z) = \sum_{n=1}^{\infty} (\sum_{m|n} \chi(m)) z^n$. Since we noted above that $\sum_{n=1}^{N} \sum_{m|n} \chi(m) \to \infty$, it follows that f is unbounded as $z \to 1$ from below.

Suppose now that the L-function is in fact zero, i.e. that $\sum_{n=1}^{\infty} \chi(n)/n = 0$, to obtain a contradiction. In that case, setting $b_n(z) = \frac{1}{n(1-z)} - \frac{z^n}{(1-z^n)}$ we also have, for $\|z\| < 1$, that

$$\begin{aligned} \sum_{n=1}^{\infty} \chi(n) b_n(z) &= \sum_{n=1}^{\infty} \frac{\chi(n)}{n(1-z)} - \sum_{n=1}^{\infty} \frac{\chi(n) z^n}{(1-z^n)} \\ &= L_\chi(z)/(1-z) - f(z) \\ &= 0/(1-z) - f(z) = -f(z) \end{aligned}$$

We will therefore obtain a contradiction if we can show that $\sum_{n=1}^{\infty} \chi(n) b_n(z)$ is bounded in the half-open interval $[0, 1)$. We know the partial sums of any nonprincipal character are bounded

```
|- ∀d c.
       dirichlet_character d c ∧ ¬(c = chi_0 d)
       ⇒ bounded {vsum(1..n) c | n ∈ ℕ}
```

so by our Dirichlet convergence test with explicit bounds, it would be sufficient to prove that the series b_k is decreasing, i.e. $b_1 \geqslant b_2 \geqslant b_3 \geqslant \cdots$. Note that

$$\begin{aligned}(1-t)(b_n - b_{n+1}) &= \frac{1}{n} - \frac{1}{n+1} - \frac{t^n}{1+t+\cdots+t^{n-1}} + \frac{t^{n+1}}{1+t+\cdots+t^n} \\ &= \frac{1}{n(n+1)} - \frac{t^n}{(1+t+\cdots+t^{n-1})(1+t+\cdots+t^n)}.\end{aligned}$$

But by the arithmetic-geometric mean inequality, we have $(1+t+\cdots+t^{n-1}) \geqslant nt^{(n-1)/2} \geqslant nt^{n/2}$ and $(1+t+\cdots+t^n) \geqslant (n+1)t^{n/2}$, and since $0 < 1-t$ for $t \in [0,1)$, the result follows:

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d) ∧ (∀n. real(c n))
        ⇒ ¬(Lfunction c = Cx(&0))
```

6. THE MAIN PROOF

The main proof uses a sort of "bootstrapping" technique, first considering the order of magnitude w.r.t. x of the sum $\sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n}$ in three hypothetical cases:

—χ is a nonprincipal character where $L(\chi) \neq 0$
—χ is a nonprincipal character where $L(\chi) = 0$
—χ is the principal character χ_0.

This intermediate result will already allow us to rule out the second hypothetical possibility, and hence 'strengthen itself', at which point it is a short step to Dirichlet's theorem. Several of the proofs below use the rearrangement of a double sum

$$\sum_{1 \leqslant n \leqslant x} \sum_{m|n} f_{n,m} = \sum_{1 \leqslant n \leqslant x} \sum_{k \leqslant x/n} f_{kn,n}$$

or in HOL:

```
|- ∀f x. vsum(1..x) (λn. vsum {d | d divides n} (f n)) =
         vsum(1..x) (λn. vsum(1..(x DIV n)) (λk. f (k * n) n))
```

6.1 The nonzero case

Let us assume that χ is a nonprincipal character where $L(\chi) \neq 0$. The key observation is that the following difference of sums up to x is bounded for all $x \in \mathbb{N}$:

$$\begin{aligned}
& L(\chi) \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} - \sum_{1 \leqslant n \leqslant x} \frac{\chi(n) \log n}{n} \\
= & \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} \sum_{k} \frac{\chi(k)}{k} - \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \Lambda(m) \\
= & \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} \sum_{k > x/n} \frac{\chi(k)}{k} + \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} \sum_{k \leqslant x/n} \frac{\chi(k)}{k} - \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \Lambda(m) \\
= & \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} \sum_{k > x/n} \frac{\chi(k)}{k} + \sum_{1 \leqslant n \leqslant x} \Lambda(n) \sum_{k \leqslant x/n} \frac{\chi(nk)}{nk} - \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \Lambda(m) \\
= & \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} \sum_{k > x/n} \frac{\chi(k)}{k} + \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \Lambda(m) - \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \Lambda(m) \\
= & \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} \sum_{k > x/n} \frac{\chi(k)}{k} \\
= & \sum_{1 \leqslant n \leqslant x} \frac{\Lambda(n)}{n} O\left(\frac{n}{x}\right) = O\left[\frac{1}{x} \sum_{1 \leqslant n \leqslant x} \Lambda(n)\right] = O(1)
\end{aligned}$$

So we conclude after 50 lines of HOL proof script:

```
|- ∀d c.
    dirichlet_character d c ∧ ¬(c = chi_0 d)
    ⇒ bounded {Lfunction c * vsum(1..x) (λn. c n * Cx(mangoldt n / &n)) -
                vsum(1..x) (λn. c n * Cx(log(&n) / &n)) | x ∈ ℕ}
```

Moreover, using the Dirichlet test again we see that the second series of our difference is convergent, and therefore its partial sums are bounded:

```
|- ∀c d. dirichlet_character d c ∧ ¬(c = chi_0 d)
         ⇒ summable (from 1) (λn. c(n) * Cx(log(&n) / &n))
```

We deduce that $L(\chi) \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n}$ is bounded, and since we assumed $L(\chi) \neq 0$, it follows that $\sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n}$ is itself bounded.

```
|- ∀d c.
    dirichlet_character d c ∧ ¬(c = chi_0 d) ∧ ¬(Lfunction c = Cx(&0))
    ⇒ bounded { vsum(1..x) (λn. c n * Cx(mangoldt n / &n)) | x ∈ ℕ}
```

6.2 The zero case

Now suppose (although we will later rule out the possibility) that χ is a nonprincipal character where $L(\chi) = 0$. Recalling that $\Lambda(n) = -\sum_{m|n} \mu(m) \log(m)$ (`MANGOLDT_LOG_SUM` above) and using the degenerate case of Möbius inversion:

```
|- ∀n. 1 <= n ⇒ sum {d | d divides n} (λd. mobius d) =
                  if n = 1 then &1 else &0
```

we have:

$$\begin{aligned}
&\sum_{1\leqslant n\leqslant x} \frac{\chi(n)\Lambda(n)}{n} + \log x \\
&= \log x - \sum_{1\leqslant n\leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \mu(m)\log(m) \\
&= \log x \sum_{1\leqslant n\leqslant x} \frac{\chi(n)}{n}(\text{if } n = 1 \text{ then } 1 \text{ else } 0) - \sum_{1\leqslant n\leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \mu(m)\log(m) \\
&= \log x \sum_{1\leqslant n\leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \mu(m) - \sum_{1\leqslant n\leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \mu(m)\log(m) \\
&= \sum_{1\leqslant n\leqslant x} \frac{\chi(n)}{n} \sum_{m|n} \mu(m)\log\frac{x}{m} \\
&= \sum_{1\leqslant n\leqslant x} \frac{\chi(n)\mu(n)}{n} \log\frac{x}{n} \sum_{k\leqslant x/n} \frac{\chi(k)}{k} \\
&= L(\chi) \sum_{1\leqslant n\leqslant x} \frac{\chi(n)\mu(n)}{n} \log\frac{x}{n} - \sum_{1\leqslant n\leqslant x} \frac{\chi(n)\mu(n)}{n} \log\frac{x}{n} \sum_{k>x/n} \frac{\chi(k)}{k} \\
&= 0 - \sum_{1\leqslant n\leqslant x} \frac{\chi(n)\mu(n)}{n} \log\frac{x}{n} \sum_{k>x/n} \frac{\chi(k)}{k} \\
&= O\left(\sum_{1\leqslant n\leqslant x} \frac{1}{x}\log\frac{x}{n}\right) = O(1).
\end{aligned}$$

or in HOL:

```
|- ∀d c.
     dirichlet_character d c ∧ ¬(c = chi_0 d) ∧
     Lfunction c = Cx(&0)
     ⇒ bounded { vsum(1..x) (λn. c n * Cx(mangoldt n / &n)) +
                   Cx(log(&x)) | x ∈ ℕ}
```

6.3 The principal case

Now consider the principal character $\chi = \chi_0$. In this case we have

$$\sum_{1\leqslant n\leqslant x} \frac{\chi_0(n)\Lambda(n)}{n} = \sum_{1\leqslant n\leqslant x} \frac{\Lambda(n)}{n} + O(1) = \log x + O(1)$$

This simple equation took some work to formalize, whereas it is just stated as obvious in the source text. The involved part is a bound on $\sum_{1\leqslant n\leqslant x} \frac{\chi_0(n)\Lambda(n)}{n} - \sum_{1\leqslant n\leqslant x} \frac{\Lambda(n)}{n}$:

```
|- ∀d. 1 <= d
       ⇒ norm(vsum(1..x) (λn. (chi_0 d n - Cx(&1)) *
                              Cx(mangoldt n / &n)))
           <= sum {p | prime p ∧ p divides d} (λp. log(&p))
```

For a fixed d, the upper bound is itself obviously bounded, and from `MERTENS_LEMMA`, we know that the difference $\sum_{1 \leqslant n \leqslant x} \frac{\Lambda(n)}{n} - \log x$ is also bounded, which yields:

```
|- ∀d. 1 <= d
       ⇒ bounded { vsum(1..x) (λn. chi_0 d n * Cx(mangoldt n / &n)) -
                   Cx(log(&x)) | x ∈ ℕ}
```

6.4 Combining the cases

Summarizing the results we have obtained, we can write

$$\sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} = \delta(\chi) \log x + O(1)$$

where

$$\delta(\chi) = \begin{cases} 1, & \text{if } \chi = \chi_0 \\ -1, & \text{if } \chi \neq \chi_0 \text{ and } L(\chi) = 0, \\ 0, & \text{if } \chi \neq \chi_0 \text{ and } L(\chi) \neq 0, \end{cases}$$

If we sum this over all characters χ, we obtain

$$\sum_{\chi} \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} = \sum_{1 \leqslant n \leqslant x} \frac{\Lambda(n)}{n} \sum_{\chi} \chi(n) = \left[\sum_{\chi} \delta(\chi) \right] \log x + O(1)$$

The middle expression makes clear that this is real and $\geqslant 0$, because $\sum_{\chi} \chi(n)$ is either $\phi(d)$ or 0 and $\Lambda(n)$ is either 0 or $\log(p)$ for $n = p^k$. Suppose now that indeed we have $L(\chi) = 0$ for some $\chi \neq \chi_0$. Since we know $L(\chi) \neq 0$ for a *real* character, χ cannot be real so the conjugate character $\overline{\chi}$ must be distinct, and of course we also have $L(\overline{\chi}) = 0$. Therefore $\delta(\overline{\chi}) = \delta(\chi) = -1$, and the sum over all characters above is $\leqslant (1-2) \log x + O(1) = -\log x + O(1)$, contradicting the fact that it is $\geqslant 0$. So we obtain

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d)
        ⇒ ¬(Lfunction c = Cx(&0))
```

We can now conclude that for any $\chi \neq \chi_0$ the sums $\sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n}$ are bounded for all x

```
|- ∀d c. dirichlet_character d c ∧ ¬(c = chi_0 d)
        ⇒ bounded { vsum(1..x) (λn. c n * Cx(mangoldt n / &n))
                  | x ∈ ℕ}
```

6.5 The finale

It is now a fairly short step to Dirichlet's theorem. Let l be coprime to our modulus d and consider the sum

$$\sum_{\chi} \chi(l) \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n}$$

For $\chi \neq \chi_0$, each summand $\chi(l) \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n}$ is bounded, while for $\chi = \chi_0$,

$$\chi_0(l) \sum_{1 \leqslant n \leqslant x} \frac{\chi_0(n)\Lambda(n)}{n} = \sum_{1 \leqslant n \leqslant x} \frac{\chi_0(n)\Lambda(n)}{n} = \log x + O(1)$$

We therefore deduce that the difference between the sum and $\log x$ is bounded:

```
|- ∀d l.
      1 <= d ∧ coprime(l,d)
       ⇒ bounded
          { vsum {c | dirichlet_character d c}
                 (λc. c(l) * vsum(1..x) (λn. c n * Cx(mangoldt n / &n))) -
            Cx(log(&x)) | x ∈ ℕ}
```

However, if this l is chosen so that $kl \equiv 1 \pmod{d}$, we can also write this sum as

$$\begin{aligned}
&\sum_{\chi} \chi(l) \sum_{1 \leqslant n \leqslant x} \frac{\chi(n)\Lambda(n)}{n} \\
&= \sum_{\chi} \sum_{1 \leqslant n \leqslant x} \frac{\chi(ln)\Lambda(n)}{n} \\
&= \sum_{1 \leqslant n \leqslant x} \frac{\Lambda(n)}{n} \sum_{\chi} \chi(ln) \\
&= \sum_{1 \leqslant n \leqslant x} \frac{\Lambda(n)}{n} (\text{if } ln \equiv 1 \pmod{d} \text{ then } \phi(d) \text{ else } 0) \\
&= \phi(d) \sum_{\substack{1 \leqslant n \leqslant x \\ ln \equiv 1 \pmod{d}}} \frac{\Lambda(n)}{n} = \phi(d) \sum_{\substack{1 \leqslant n \leqslant x \\ n \equiv k \pmod{d}}} \frac{\Lambda(n)}{n}
\end{aligned}$$

Thus we conclude:

```
|- ∀d k. 1 <= d ∧ coprime(k,d)
         ⇒ bounded { Cx(&(phi d)) *
                     vsum {n | n ∈ 1..x ∧ (n == k) (mod d)}
                          (λn. Cx(mangoldt n / &n)) -
                     Cx(log(&x)) | x ∈ ℕ}
```

All the terms involved here are real, so transferring back to the reals, expanding the definition of 'bounded' and dividing the inequality by $\phi(d)$, we can write this more straightforwardly as

```
|- ∀d k. 1 <= d ∧ coprime(k,d)
         ⇒ ∃B. &0 < B ∧
                 ∀x. abs(sum {n | n ∈ 1..x ∧ (n == k) (mod d)}
                             (λn. mangoldt n / &n) -
                         log(&x) / &(phi d)) <= B
```

Calling on the lemma `MERTENS_MANGOLDT_VERSUS_LOG`, we can replace $\sum_n \Lambda(n)/n$ by $\sum_p \log p/p$:

```
|- ∀d k. 1 <= d ∧ coprime(k,d)
        ⇒ ∃B. &0 < B ∧
                ∀x. abs(sum {p | p ∈ 1..x ∧ prime p ∧ (p == k) (mod d)}
                            (λp. log(&p) / &p) -
                        log(&x) / &(phi d)) <= B
```

This is actually a somewhat stronger result than Dirichlet's theorems, giving the order of magnitude with respect to x of the series

$$\sum_{\substack{p \leqslant x \\ p \equiv k \pmod d}} \log p/p$$

as $\log(x)/\phi(d)$, which in particular implies that the sum

$$\sum_{p \equiv k \pmod d} \log p/p$$

diverges, and therefore that the set of primes p with $p \equiv k \pmod d$ is infinite

```
|- ∀d k. 1 <= d ∧ coprime(k,d)
        ⇒ INFINITE {p | prime p ∧ (p == k) (mod d)}
```

as required.

7. DE BRUIJN FACTOR COMPUTATION

We have compared the main parts of our formalization against reverse-engineered TeX for corresponding passages in the original sources. The *de Bruijn factor* [Wie00] is the size ratio of a gzipped formal proof text versus the gzipped TeX of its informal counterpart. For our proof, the de Bruijn factor is 4.66, which is about in line with a fair number of other formalization case studies [Wie00] and markedly better than the value of at least 8 we noted in our proof of the Prime Number Theorem [Har09]. However, if we analyze parts of the proof separately, we find quite a wide variation, as the following table shows.

Portion	Lines (HOL/TeX)	Bytes (HOL/TeX)	Gzipped (HOL/TeX)
Monsky	444/34 = 13.06	23322/1938 = 12.03	5380/874 = 6.16
Convergence	82/15 = 5.47	4182/471 = 8.88	1304/287 = 4.54
Nonzero case	106/24 = 4.42	5692/949 = 6.00	1628/375 = 4.34
Zero case	124/15 = 8.27	6720/714 = 9.41	1944/289 = 6.73
Principal	120/6 = 20.00	6610/223 = 29.64	1826/186 = 9.82
Delta sum	105/17 = 6.18	4908/765 = 6.42	1618/436 = 3.71
Finale	117/30 = 3.90	5845/1153 = 5.07	1645/544 = 3.02
TOTAL	1183/192 = 6.16	61636/7823 = 7.88	11762/2524 = 4.66

The worst part is the estimation of the sum $\sum_{1 \leqslant n \leqslant x} \frac{\chi_0(n)\Lambda(n)}{n}$ for the principal character. This only occupies a couple of lines in the source text, but the computations on sums turned out to be quite lengthy when formalized. Monsky's proof of nonvanishing for the L-series corresponding to a real nonprincipal character also has a slightly higher de Bruijn factor than the average for the formalization as a

whole. On general grounds, it's not surprising that a different author with a different expository style should have different de Bruijn factor characteristics when formalized. Indeed, it is perhaps a reflection of the fact that when one sets out to write an avowedly short proof, the temptation is stronger to write in a more condensed style.

8. CONCLUSIONS

In one sense, this formalization effort is not very interesting, since it does not involve any great difficulties, surprises or subtleties. On the other hand, it is pleasant to reflect that the formalization of a fairly interesting theorem was essentially routine, perhaps indicating the increasing maturity of proof assistants. We have made some use of automation for basic algebraic and arithmetic reasoning, but the most important property of HOL Light has probably been the availability of a solid library of elementary lemmas that we could call upon. We have tried to present the proof here in a fairly explicit and self-contained way for readers interested in the proof per se or in trying its formalization in other systems. The actual formal proof is available in recent HOL Light snapshots as `100/dirichlet.ml`.

We hope that it will be useful to have a formalized version of this theorem available for use as a lemma when working on other results in number theory. But it might also be interesting to revisit some parts of the present formalization and try to rewrite them in a more systematic way. In particular, it would be nice to formalize enough of the general theory of groups and group characters so that some arguments involving Dirichlet characters could be developed in a more elegant way. It would also be natural to extend the formalization to more general properties of L-functions. This could lead on to stronger theorems about the distribution of primes in arithmetic progression, and to the formalization of many other parts of modern mathematics where L-functions play an important role.

Acknowledgements

The author would like to thank Rob Arthan and the anonymous JFR referees, whose comments have significantly improved the final version of this paper.

References

[AA08] Andrea Asperti and Cristian Armentano. A page in number theory. *Journal of Formalized Reasoning*, 1:1–23, 2008.

[ADGR07] Jeremy Avigad, Kevin Donnelly, David Gray, and Paul Raff. A formally verified proof of the prime number theorem. *ACM Transactions on Computational Logic*, 9(1:2):1–23, 2007.

[Dir37] Gustav Lejeune Dirichlet. Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält. *Abhandlungen der Königlichen Preußischen Akademie der Wissenschaften zu Berlin*, pages 45–81, 1837. English translation "There are infinitely many prime numbers in all arithmetic progressions with first term and difference coprime" by Ralf Stefan available from `http://arxiv.org/abs/0808.1408`.

[FR07] Benjamin Fine and Gerhard Rosenberger. *Number theory: An Introduction via the Distribution of Primes*. Birkhäuser, 2007.

[GL65] A. E. Gelfond and U. V. Linnik. *Elementary methods in analytic number theory*. Rand McNally, 1965. Translated by L. J. Mordell.

[Har96] John Harrison. HOL Light: A tutorial introduction. In Mandayam Srivas and Albert Camilleri, editors, *Proceedings of the First International Conference on Formal Methods in Computer-Aided Design (FMCAD'96)*, volume 1166 of *Lecture Notes in Computer Science*, pages 265–269. Springer-Verlag, 1996.

[Har98] John Harrison. *Theorem Proving with the Real Numbers*. Springer-Verlag, 1998. Revised version of author's PhD thesis.

[Har01] John Harrison. Complex quantifier elimination in HOL. In Richard J. Boulton and Paul B. Jackson, editors, *TPHOLs 2001: Supplemental Proceedings*, pages 159–174. Division of Informatics, University of Edinburgh, 2001. Published as Informatics Report Series EDI-INF-RR-0046. Available on the Web at `http://www.informatics.ed.ac.uk/publications/report/0046.html`.

[Har07a] John Harrison. Automating elementary number-theoretic proofs using Gröbner bases. In Frank Pfenning, editor, *Proceedings of the 21st International Conference on Automated Deduction, CADE 21*, volume 4603 of *Lecture Notes in Computer Science*, pages 51–66, Bremen, Germany, 2007. Springer-Verlag.

[Har07b] John Harrison. Formalizing basic complex analysis. In R. Matuszewski and A. Zalewska, editors, *From Insight to Proof: Festschrift in Honour of Andrzej Trybulec*, volume 10(23) of *Studies in Logic, Grammar and Rhetoric*, pages 151–165. University of Białystok, 2007.

[Har09] John Harrison. Formalizing an analytic proof of the Prime Number Theorem (dedicated to Mike Gordon on the occasion of his 60th birthday). *Journal of Automated Reasoning*, 43:243–261, 2009.

[HW79] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Clarendon Press, 5th edition, 1979.

[Jam03] G. J. O. Jameson. *The Prime Number Theorem*, volume 53 of *London Mathematical Society Student Texts*. Cambridge University Press, 2003.

[Lei69] A. C. Leisenring. *Mathematical logic and Hilbert's ϵ-symbol*. Macdonald, 1969.

[Mon90] Paul Monsky. Simplifying the proof of Dirichlet's theorem. *The American Mathematical Monthly*, 100:861–2, 1990.

[New98] Donald J. Newman. *Analytic Number Theory*, volume 177 of *Graduate Texts in Mathematics*. Springer-Verlag, 1998.

[Wie00] Freek Wiedijk. The de Bruijn factor. See `http://www.cs.ru.nl/~freek/factor/`, 2000.

[Wie06] Freek Wiedijk. *The Seventeen Provers of the World*, volume 3600 of *Lecture Notes in Computer Science*. Springer-Verlag, 2006.