Average Running Time of the Fast Fourier Transform

PERSI DIACONIS

*Bell Laboratories, Murray Hill, New Jersey; and Stanford University,
Stanford, California*

Received May 9, 1979; and in revised form October 29, 1979

We compare several algorithms for computing the discrete Fourier transform of n numbers. The number of "operations" of the original Cooley-Tukey algorithm is approximately $2n A(n)$, where $A(n)$ is the sum of the prime divisors of n . We show that the average number of operations satisfies $(1/x) \sum_{n \leq x} 2n A(n) \sim (\pi^2/9)(x^2/\log x)$. The average is not a good indication of the number of operations. For example, it is shown that for about half of the integers n less than x , the number of "operations" is less than $n^{1.61}$. A similar analysis is given for Good's algorithm and for two algorithms that compute the discrete Fourier transform in $O(n \log n)$ operations: the chirp- z transform and the mixed-radix algorithm that computes the transform of a series of prime length p in $O(p \log p)$ operations.

1. INTRODUCTION

The main results of this paper give approximations to the running time of several algorithms for computation of the discrete Fourier transform (DFT) of n numbers. In Section 2 we discuss the need for exact computation of the DFT versus "padding." We also describe the available algorithms for computing the DFT. Direct computation of the DFT is shown to involve approximately $2n^2$ operations—multiplications and additions. If an algorithm is to be used for many different values of n , the average running time is of interest. For direct computation, the average is

$$\frac{1}{x} \sum_{n \leq x} 2n^2 \sim \frac{2}{3} x^2.$$

Several variants of the fast Fourier transform (FFT) involve approximately $2n A(n)$ operations. Here $A(n) = \sum_{p^a | n} p$ is the sum of the prime divisors of n counted with multiplicity (so $A(12) = 2 + 2 + 3 = 7$). In Section 3 we show that the average number of operations satisfies

$$\frac{1}{x} \sum_{n \leq x} 2n A(n) \sim \frac{\pi^2}{9} \frac{x^2}{\log x}.$$

Thus, on the average, these versions of the FFT do not seem to speed things up very much. We will argue that the average is a bad indication of the size of $n A(n)$. Theorem 3 shows that the proportion of integers n less than x such that $n A(n)$ is smaller than $n^{1+\gamma}$ tends to a limit $L(\gamma)$:

$$\frac{1}{x} |\{n \leq x : n A(n) \leq n^{1+\gamma}\}| \sim L(\gamma).$$

The distribution function $L(\gamma)$ is supported on $0 \leq \gamma \leq 1$ and, for example, $L(0.61) = 0.5$. Thus, approximately half of the integers less than or equal to x have $n A(n) \leq n^{1.61}$. The results in Section 3 show that, up to lower-order terms, Good's version of the FFT has the same average case behavior as $n A(n)$.

Section 4 analyzes two algorithms for computing the DFT in $O(n \log n)$ operations. These are the chirp- z algorithm and the mixed-radix algorithm which uses the chirp- z (or number theoretic) transform for series of prime length. Neither approach dominates. Both algorithms have average running time proportional to $x \log_2 x$. For the chirp- z approach, the "constant" of proportionality is a bounded oscillating function of x which oscillates around the constant of proportionality of the mixed-radix approach. For individual n , the better algorithm can speed things up by a factor of $1\frac{1}{2}$ to 2.

Some Notation

Throughout this paper, p is a prime, $\sum_{p|n}$ means a sum over the distinct prime divisors of n , $\sum_{p^a|n}$ means a sum over the prime divisors of n counted with multiplicity. The O, o notation will be used with, for example, O_k meaning that the implied constant depends on k ; $f(x) \sim g(x)$ means $f(x)/g(x) \rightarrow 1$. We write $\lfloor x \rfloor$ for the largest integer less than or equal to x , $\lceil x \rceil$ for the smallest integer greater than or equal to x , and $\{x\}$ for the fractional part of x . The number of elements in a finite set S is denoted $|S|$.

2. THE FAST FOURIER TRANSFORM

The discrete Fourier transform of n real numbers $x_0, x_1, \dots, x_{n-1}$ is the sequence

$$\phi(k) = \sum_{j=0}^{n-1} x_j q_n^{jk}, \quad k = 0, 1, 2, \dots, n-1; \quad q_n = e^{2\pi i/n}. \quad (2.1)$$

The usual assumption is that the numbers q_n^{jk} are stored (or available for

free). Then, for each k , direct computation of $\phi(k)$ involves n multiplications and n additions to good approximation. Computing $\phi(k)$ for $k = 0, 1, 2, \dots, n-1$ involves approximately n^2 multiplications and n^2 additions. We will say that approximately $2n^2$ operations are involved for direct computation.

The FFT is a collection of algorithms for computing the DFT. The basic papers on the FFT are collected together in [14]. A discussion from a modern algorithmic point of view with applications and references is in [1, Chap. 7].

It is useful to divide the ideas behind FFT algorithms into two types. Type 1 concerns methods of "pasting together" transforms of shorter series. Type 2 concerns methods of transforming the sum in (2.1) into a convolution.

First consider the Type 1 ideas. When n is composite, some of the products $x_j q_n^{jk}$ are calculated many times. Suppose $n = pq$. The Cooley-Tukey and Tukey-Sandy algorithms allow computation of the DFT via computation of p transforms of length q and q transforms of length p . If the shorter transforms are computed directly, this leads to $p2q^2 + q2p^2 = 2n(p+q)$ operations approximately. In general, when $n = \prod_{i=1}^r p_i^{a_i}$, the number of operations is

$$2n \sum_{p^a | n} p = 2n A(n). \quad (2.2)$$

We will see later that it is possible to compute the shorter transforms in $O(p \log p)$ operations instead of $O(p^2)$ operations that direct computation entails. Direct computation is suggested by many writers and implemented in published algorithms such as that of Singleton [18].

Another way of linking together shorter transformations, suggested by Good [20], also falls under the Type 1 ideas. Good's algorithm requires that the length of the shorter series be relatively prime. For $n = \prod_{i=1}^r p_i^{a_i}$, the algorithm computes the DFT of series of length $p_i^{a_i}$. If the transforms of length $p_i^{a_i}$ are computed directly using $2p_i^{2a_i}$ operations, then the number of operations is approximately

$$2n \sum_{i=1}^r p_i^{a_i} = 2n G(n). \quad (2.3)$$

The equality in (2.3) defines the function $G(n)$.

Expressions (2.2) and (2.3) may be regarded as the dominant term of the result of a more careful count of operations as presented by Rose [16].

When n is a power of 2, $n = 2^k$, $A(n) = 2k$, and from (2.2), the number of operations is $4n \log_2 n$. This is the oft-quoted result "the FFT allows computation of the DFT in $O(n \log n)$ operations." As we have seen, this statement holds only when n is a power of 2.

When n is not a power of 2, the technique of padding a series by zeros to the next highest power of 2 can be used. If m is the smallest power of 2 larger than n , the DFT of the sequence of length m : $(x_0, x_1, x_2, \dots, x_{n-1}, 0, 0, \dots, 0)$ is computed. This yields $\hat{\phi}(k) = \sum_{j=0}^{n-1} x_j q_m^{jk}$, $k = 0, 1, 2, \dots, m-1$, instead of $\phi(k)$ as defined by (2.1). In many applications, $\hat{\phi}$ can be used as effectively as ϕ .

The difference between ϕ and $\hat{\phi}$ is sometimes important. One example occurs in spectral analysis where one looks at $\phi(k)$ in the hope of detecting periodic oscillations in the sequence x_j . If the period divides n , the difference between ϕ and $\hat{\phi}$ can matter. To give a simple example, let a be a positive integer and for $0 \leq j < n$, define x_j by

$$\begin{aligned} x_j &= 1 && \text{if } j \text{ is a multiple of } a \\ &= 0 && \text{otherwise.} \end{aligned}$$

Then $\phi(k) = \sum_{j=0}^{\lfloor n/a \rfloor - 1} q_n^{ajk}$. If a divides n , then an easy computation shows that

$$\begin{aligned} \phi(k) &= k && \text{if } k = n/a \\ &= 0 && \text{if } k \neq n/a. \end{aligned}$$

Thus, $\phi(k)$ clearly identifies a series of period a . This clear identification is destroyed if n is not a multiple of a , for then

$$\phi(k) = \frac{1 - q_n^{ak \lfloor n/a \rfloor}}{1 - q_n^{ak}},$$

and $\phi(k)$ is never 0. In the language of spectral analysis, there is leakage at all frequencies. Such leakage can cause problems, and n is often chosen as a multiple of a period of interest instead of a convenient power of 2. Further discussion of the need for exact computation of the DFT can be found in [6, 18].

We next turn to Type 2 ideas. These involve transforming the summation index in the sum (2.1) for $\phi(k)$ to convert the sum into a convolution. Convolutions for a series of any length n can be computed exactly by using the FFT on an appropriately extended series.

For example, the chirp- z approach discussed by Rabiner *et al.* [13] and Aho *et al.* [2] makes the change of variables $jk = (j^2 + k^2 - (j - k)^2)/2$. Then

$$\phi(k) = q_n^{k^2/2} \sum_{j=0}^{n-1} (x_j q_n^{j^2/2}) q_n^{-(j-k)^2/2}. \quad (2.4)$$

This is a convolution of the sequence $(x_j q_n^{j^2/2})$ with the sequence $(q_n^{-j^2/2})$,

premultiplied by $q_n^{k^2/2}$. The idea does not depend on n being prime. When n is prime, Rader [15] proposed using a primitive root $g \pmod{n}$ to do the transformation. By definition, g is an integer such that $g^k \not\equiv 1 \pmod{n}$, $k = 1, 2, \dots, n-2$, $g^{n-1} \equiv 1 \pmod{n}$. Make the transformation $j = g^a, k = g^b \pmod{n}$. Then

$$\phi(k) = \phi(g^b) = x_0 + \sum_{a=1}^{n-1} x_{g^{-a}} q_n^{g^{a+b}}. \quad (2.5)$$

The sum is a convolution of the sequence $(x_{g^{-j}})$ with the sequence $(q_n^{g^j})$. It is worth pointing out that an integer n can be factored in less than $O(n^{1/3})$ operations and that primitive roots can be found in less than $O(n^{1/4})$ time. Lehmer [12] gives the number theoretic details.

After transforming to a convolution, the FFT can be used to compute the convolution on an extended series. The extended series must be of length approximately the smallest power of 2 larger than $2n$. For definiteness, we will use the algorithm given by Rader [15]. This requires the use of the FFT on a series of length the smallest power 2 larger than or equal to $2n - 4$. Three FFTs are required to perform convolutions. To get a simple approximation to the number of operations of these algorithms we will neglect lower-order terms.

(2.6) Define $T(n)$ to be the smallest power of 2 larger than $2n - 4$. Let

$$\begin{aligned} C(n) &= 3T(n) \log_2 T(n) && \text{if } n \neq 2^k \\ &= n \log_2 n && \text{if } n = 2^k. \end{aligned}$$

(2.7) ASSUMPTION. The number of operations of the chirp- z transform (or the approach using (2.5)) is well approximated by $\beta C(n)$ for some fixed $\beta > 0$. A reasonable further assumption is to take $\beta = 2$ (for additions and multiplications).

Since $C(n)$ is $O(n \log n)$, when Assumption (2.7) is valid, the DFT of n numbers can be computed in $O(n \log n)$ operations, even when n is prime.

The availability of methods for computing the DFT in $O(n \log n)$ operations immediately suggests a question. Which is more efficient: direct use of the chirp- z idea on a series of length n or splitting a series of length n into pieces of prime length, computation of the shorter transforms by a fast algorithm, and pasting the pieces back together using a mixed-radix algorithm? A detailed analysis of this problem is given in Section 4. To compare the approaches we make a simple approximation to the number of operations required by the mixed-radix approach. The approximation is based on (2.7) and (2.2).

(2.8) ASSUMPTION. The mixed-radix approach which uses a fast algorithm to compute series of prime length uses $\beta F(n) = \beta n \sum_{p|n} (C(p)/p)$ operations, with β as in (2.7).

Some numerical examples of $C(n)$ and $F(n)$ are given in Section 4. Note that in making comparison of the relative size of $\beta C(n)$ and $\beta F(n)$, β cancels out. We also note that none of the asymptotic comparisons depend on the use of $2n - 4$ in (2.6). Using $2n - c$ for any fixed c leads to essentially the same results.

3. AVERAGE RUNNING TIME OF MIXED-RADIX ALGORITHMS THAT COMPUTE THE FFT OF PRIMES IN $O(p^2)$ OPERATIONS

In this section, approximations for the mean, variance, and distribution of the number of operations of some mixed-radix algorithms are derived. As explained in Section 2, $2n A(n)$ and $2n G(n)$ are reasonable approximations to the number of operations used by the Cooley-Tukey and Good algorithms, respectively.

Theorems 1 and 2 provide approximations to the first and second moments of $n A(n)$ and $n G(n)$. Here, if $H(n)$ is any function, the first and second moments of $H(n)$ are

$$\mu_H(x) = \frac{1}{x} \sum_{n \leq x} H(n) \quad \text{and} \quad \mu_{H^2}(x) = \frac{1}{x} \sum_{n \leq x} H^2(n).$$

The variance of $H(n)$ can be computed from the first and second moments via

$$\sigma_H^2(x) = \frac{1}{x} \sum_{n \leq x} (H(n) - \mu_H(x))^2 = \mu_{H^2}(x) - (\mu_H(x))^2.$$

In Theorems 1 and 2 we write $\zeta(s)$ for Riemann's zeta function.

THEOREM 1. *Let $n A(n)$ be defined by (2.2), $n G(n)$ be defined by (2.3). As x tends to infinity, there is a $c > 0$ such that the first moment of $n A(n)$ or $n G(n)$ equals*

$$\int_{1/2}^1 \zeta(s+1) \frac{x^{s+1}}{s+2} ds + O(x^2 \exp(-c(\log x)^{1/10})). \quad (3.1)$$

For any fixed $k \geq 1$,

$$\int_{1/2}^1 \zeta(s+1) \frac{x^{s+1}}{s+2} ds = x^2 \sum_{j=1}^k \frac{a_j}{(\log x)^j} + O_k\left(\frac{x^2}{(\log x)^{k+1}}\right), \quad (3.2)$$

with

$$a_1 = \frac{\zeta(2)}{3}, \quad a_j = (-1)^j \left(\frac{\zeta(s+1)}{s+2} \right)^{(j)} \Big|_{s=1}.$$

THEOREM 2. *As x tends to infinity, there is a $c > 0$ such that the second moment of $n A(n)$ or $n G(n)$ equals*

$$\int_{1/2}^1 \zeta(s+2) \frac{x^{s+3}}{s+4} ds + O(x^4 \exp(-c(\log x)^{1/10})). \quad (3.3)$$

For any fixed $k \geq 1$,

$$\int_{1/2}^1 \zeta(s+2) \frac{x^{s+3}}{s+4} ds = x^4 \sum_{j=1}^k \frac{b_j}{(\log x)^j} + O_k \left(\frac{x^4}{(\log x)^{k+1}} \right), \quad (3.4)$$

with

$$b_1 = \frac{\zeta(3)}{5}, \quad b_j = (-1)^j \left(\frac{\zeta(s+2)}{s+4} \right)^{(j)} \Big|_{s=1}.$$

The first moment and variance are not good indicators of the sizes of $n A(n)$ and $n G(n)$, for the variance is close to the square of the first moment. This suggests that these functions have fluctuations about their mean which are of the same size as their mean.

The next theorem gives the limiting distribution of $n A(n)$ and $n G(n)$. The results show that the proportion of numbers such that $n A(n)$ (or $n G(n)$) is smaller than n^{1+z} tends to a limit.

THEOREM 3. *As $x \rightarrow \infty$, for any fixed z , $0 \leq z \leq 1$,*

$$\frac{1}{x} |\{n \leq x : n A(n) \leq n^{1+z}\}| \sim L(z), \quad (3.5)$$

$$\frac{1}{x} |\{n \leq x : n G(n) \leq n^{1+z}\}| \sim L(z), \quad (3.6)$$

where $L(z)$ is the distribution function of an absolutely continuous measure on $[0, 1]$ with density $L'(z)$. The density satisfies $L'(z) = (1/z)\rho(1/z - 1)$, where $\rho(y) = 1$ for $0 < y \leq 1$ and $\rho(y)$ satisfies the differential difference equation $y\rho'(y) = -\rho(y-1)$.

Remarks. A few values of $L(z)$ are

z	0.33	0.47	0.61	0.78	0.95
$L(z)$	0.05	0.25	0.50	0.75	0.95

The density $L'(y)$ is drawn in Fig. 1.

FIG. 1. Graph of $L'(y)$.

The function $\rho(y)$ was introduced by Dickman [10] in connection with the largest prime divisor of n . It is thoroughly discussed by de Bruijn [8, 9], Billingsley [7], and Knuth and Trabb-Pardo [11]. Bellman and Kotkin [5] and Van de Lune and Wattel [19] give tables of $\rho(y)$ which were used to compute $L(z)$ and $L'(z)$ as given above.

Theorems 1, 2, and 3 are closely related to two other number theoretic functions which we now define:

Let $n = \prod_{i=1}^r p_i^{a_i}$ be the prime decomposition of n . Suppose $p_1 < p_2 < \dots < p_r$.

(3.7) Define $A^*(n) = \sum_{i=1}^r p_i$. We also write $A^*(n) = \sum_{p|n} p$.

(3.8) Define $P_k(n)$ to be the k th largest prime divisor of n . Thus, $P_1(n) = p_r$, $P_2(n) = P_1(n)/P_1(n)$, $P_3(n) = P_1(n)/[P_1(n) \cdot P_2(n)] \dots$ with the convention that $P_i(n) = 1$ if n has fewer than i prime divisors.

The functions $A(n)$ and $A^*(n)$ have been discussed by Alladi and Erdős [3]. They prove the following theorem, which will be needed in the proof of Theorem 3.

THEOREM (Alladi and Erdős). For all $m \geq 1$, as $x \rightarrow \infty$,

$$\sum_{n \leq x} P_m(n) \sim \sum_{n \leq x} \{A(n) - P_1(n) - \dots - P_{m-1}(n)\} \sim k_m \frac{x^{1+(1/m)}}{(\log x)^m},$$

where k_m is a rational multiple of $\zeta(1 + (1/m))$.

When $m = 1$, this gives $\sum_{n \leq x} A(n) \sim k_1(x^2/\log x)$. Alladi and Erdős show that $k_1 = \pi^2/12$. They also show that

$$\sum_{n \leq x} A(n) - A^*(n) = x \log \log x + O(x). \quad (3.9)$$

Saffari [17, p. 213] has given an asymptotic expansion for the mean of $A^*(n)$. The techniques of this paper yield somewhat more precise results:

THEOREM 4. *As x tends to infinity, there is a $c > 0$ such that the first moment of $A(n)$, $A^*(n)$, $G(n)$, or $P_1(n)$ equals*

$$\int_{1/2}^1 \zeta(s+1) \frac{x^s}{s+1} ds + O(x \exp(-c(\log x)^{1/10})) \quad \text{as } x \rightarrow \infty. \quad (3.10)$$

For any fixed $k \geq 1$,

$$\int_{1/2}^1 \zeta(s+1) \frac{x^s}{s+1} ds = x \sum_{j=1}^k \frac{c_j}{(\log x)^j} + O_k\left(\frac{x}{(\log x)^{k+1}}\right), \quad (3.11)$$

with

$$c_1 = \frac{\zeta(2)}{2}, \quad c_j = (-1)^j \left(\frac{\zeta(s+1)}{s+1} \right)^{(j)} \Big|_{s=1}.$$

THEOREM 5. *As x tends to infinity, there is a $c > 0$ such that the second moment of $A(n)$, $A^*(n)$, $G(n)$, or $P_1(n)$ equals*

$$\int_{1/2}^1 \zeta(s+2) \frac{x^{s+1}}{s+2} ds + O(x^2 \exp(-c(\log x)^{1/10})). \quad (3.12)$$

For any fixed $k \geq 1$,

$$\int_{1/2}^1 \zeta(s+2) \frac{x^{s+1}}{s+2} ds = x^2 \sum \frac{d_j}{(\log x)^j} + O_k\left(\frac{x^2}{(\log x)^{k+1}}\right), \quad (3.13)$$

with

$$d_1 = \frac{\zeta(3)}{3}, \quad d_j = (-1)^j \left(\frac{\zeta(s+2)}{s+2} \right)^{(j)} \Big|_{s=1}.$$

Remarks. Theorems 1, 2, 4, and 5 are all proved by using slight modifications of the proof of the prime number theorem. By using the

usual modifications of this proof, it is possible to improve the error terms. Using the Riemann hypothesis, the error terms can be further improved. For example, I believe that, on the Riemann hypothesis, the error term in (3.10) becomes $O(x^{1/2}(\log x)^2)$. The results are given with the error involving $(\log x)^{1/10}$ to allow the proof to rely on the proof of the prime number theorem given by Ayoub [4] without modification.

As with $n A(n)$ and $n G(n)$, the mean and variance are not good indicators of the sizes of $A(n)$, $A^*(n)$, $G(n)$, and $P_1(n)$. Asymptotically, these functions all have the same distribution. The next result implies Theorem 3.

THEOREM 6. *Let $H(n)$ be any one of the functions $A(n)$, $A^*(n)$, $G(n)$, or $P_1(n)$. As x tends to infinity, for any fixed y , $0 \leq y \leq 1$,*

$$\frac{1}{x} |\{n \leq x : H(n) \leq n^y\}| \sim L(y),$$

where $L(y)$ was defined in Theorem 3.

Proof of Theorems 1, 2, 4, and 5. The approach used here is the classical technique using Dirichlet series. The following identities are needed:

LEMMA 7. *Let $A(n)$, $G(n)$, $A^*(n)$, and $P_1(n)$ be as defined in (2.2), (2.3), (3.7), and (3.8), respectively.*

$$\sum_{n=1}^{\infty} \frac{A^*(n)}{n^s} = \zeta(s) \sum_p \frac{1}{p^{s-1}}, \quad \text{re } s > 2, \quad (3.14)$$

$$\sum_{n=1}^{\infty} \frac{(A^*(n))^2}{n^s} = \zeta(s) \sum_p \frac{1}{p^{s-2}} \left(1 - \frac{1}{p^s}\right) + \zeta(s) \left(\sum_p \frac{1}{p^{s-1}}\right)^2, \quad \text{re } s > 3, \quad (3.15)$$

$$\sum_{n=1}^{\infty} \frac{A(n)}{n^s} = \zeta(s) \sum_p \frac{p}{p^s - 1}, \quad \text{re } s > 2, \quad (3.16)$$

$$\sum_{n=1}^{\infty} \frac{A^2(n)}{n^s} = \zeta(s) \sum_p \frac{p^{s+2}}{(p^s - 1)^2} + \zeta(s) \left(\sum_p \frac{p}{p^s - 1}\right)^2, \quad \text{re } s > 3, \quad (3.17)$$

$$\sum_{n=1}^{\infty} \frac{G(n)}{n^s} = \zeta(s) \sum_p \frac{1}{p^{s-1}} \left(1 - \frac{1}{p^s}\right) \left(1 - \frac{1}{p^s - 1}\right)^{-1} \equiv \mu_g(s), \quad \text{re } s > 2, \quad (3.18)$$

$$\sum_{n=1}^{\infty} \frac{G^2(n)}{n^s} = \zeta(s) \sum_p \left\{ \frac{1}{p^{s-2}} \left(1 - \frac{1}{p^s}\right) \left(1 - \frac{1}{p^{s-2}}\right)^{-1} \right. \\ \left. - \frac{1}{p^{2(s-1)}} \left(1 - \frac{1}{p^s}\right)^2 \left(1 - \frac{1}{p^{s-1}}\right)^{-2} \right\} \\ + \zeta(s) (\mu_g(s))^2, \quad \operatorname{re}(s) > 3. \quad (3.19)$$

$$\sum_{n=1}^{\infty} \frac{P_1(n)}{n^s} = \zeta(s) \sum_p \frac{1}{p^{s-1}} \prod_{q>p} \left(1 - \frac{1}{q^s}\right). \quad (3.20)$$

$$\sum_{n=1}^{\infty} \frac{P_1^2(n)}{n^s} = \zeta(s) \sum_p \frac{1}{p^{s-2}} \prod_{q>p} \left(1 - \frac{1}{q^s}\right). \quad (3.21)$$

Lemma 7 can be proved by a number of arguments. The following approach seems to be useful somewhat generally.

Proof of Lemma 7. For fixed real $s > 1$, define a probability measure P_s on the space $\Omega = \{1, 2, 3, \dots\}$ by $P_s(j) = (1/\zeta(s))(1/j^s)$, where $\zeta(s) = \sum_{j=1}^{\infty} 1/j^s$. For each prime number p , let $X_p: \Omega \rightarrow \Omega \cup \{0\}$ be defined by $X_p(n) = p$ if $p|n$, $X_p = 0$ otherwise. Thus $P_s(X_p = p) = 1/p^s$, $P_s(X_p = 0) = 1 - 1/p^s$. The random variables X_p are easily seen to be independent with $E(X_p) = 1/p^{s-1}$, $\operatorname{var}(X_p) = (1/p^{s-2})(1 - 1/p^s)$. Let $A^* = \sum_p X_p$. For $s > 1$, the Borel-Cantelli lemma implies that A^* is almost surely finite. A^* has finite mean if and only if $s > 2$. A^* has finite variance if and only if $s > 3$. For $s > 2$,

$$\frac{1}{\zeta(s)} \sum_{n=1}^{\infty} \frac{A^*(n)}{n^s} = E(A^*) = \sum_p E(X_p) = \sum_p \frac{1}{p^{s-1}}$$

This implies (3.14). To prove (3.15) note that for $s > 3$,

$$\frac{1}{\zeta(s)} \sum \frac{(A^*(n))^2}{n^s} - (E(A^*))^2 = \operatorname{var}(A^*) = \sum_p \operatorname{var}(X_p) \\ = \sum_p \frac{1}{p^{s-2}} \left(1 - \frac{1}{p^s}\right).$$

To prove (3.16) and (3.17) consider the random variables $Y_p: \Omega \rightarrow \Omega \cup \{0\}$, where, if $n = \prod p^{a_p(n)}$, $Y_p(n) = a_p(n)p$. Thus for $a = 0, 1, 2, \dots$; $P_s(Y_p = ap) = (1 - 1/p^s)(1/p^{as})$. It is straightforward to check that $E(Y_p) = p/(p^s - 1)$, $\operatorname{var}(Y_p) = p^{s+2}/(p^s - 1)^2$. To prove (3.18) and (3.19), consider the random variables $Z_p: \Omega \rightarrow \Omega \cup \{0\}$ where if $n = \prod p^{a_p(n)}$, $Z_p(n) = p^{a_p(n)}$. Thus, $P_s(Z_p = 0) = 1 - 1/p^s$ and for $a = 1, 2, \dots$, $P_s(Z_p = p^a)$

$= (1 - 1/p^s)(1/p^{as})$. Again it is easy to check that

$$E(Z_p) = \left(1 - \frac{1}{p^s}\right) \frac{1}{p^{s-1}} \left(1 - \frac{1}{p^{s-1}}\right)^{-1} \quad \text{and}$$

$$E(Z_p^2) = \left(1 - \frac{1}{p^s}\right) \frac{1}{p^{s-2}} \left(1 - \frac{1}{p^{s-2}}\right)^{-1}.$$

The argument used to prove (3.14) and (3.15) now leads to a proof of (3.18) and (3.19). Finally, to prove (3.20) and (3.21), use the fact that for any prime p ,

$$P_s(P_1(n) = p) = \frac{1}{p^s} \prod_{q>p} \left(1 - \frac{1}{q^s}\right),$$

the product being over all primes q larger than p . It follows that

$$E_s(P_1(n)) = \sum_p \frac{1}{p^{s-1}} \prod_{q>p} \left(1 - \frac{1}{q^s}\right), \quad E_s(P_1^2(n)) = \sum_p \frac{1}{p^{s-2}} \prod_{q>p} \left(1 - \frac{1}{q^s}\right).$$

Since also

$$E_s(P_1(n)) = \frac{1}{\zeta(s)} \sum_{n=1}^{\infty} \frac{P_1(n)}{n^s}, \quad E_s(P_1^2(n)) = \frac{1}{\zeta(s)} \sum_{n=1}^{\infty} \frac{P_1^2(n)}{n^s},$$

(3.20) and (3.21) follow.

The arguments prove the identities for all large real s and thus, by analytic continuation, for all s such that the right sides are analytic. The validity for the half planes given follows easily from the known behavior of the function $\sum_p 1/p^s$ (see, for example, [4, Chap. 2, Sect. 4, (16)]). $\square$

Proof of Theorem 1. The argument used here will follow Landau's proof of the prime number theorem as presented by Ayoub [4, Chap. 2]. Since we make constant use of Ayoub's arguments, the reader is advised to follow the present proof with a copy of Ayoub's book in hand.

First consider the function $n A(n)$. The identity (3.16) together with Theorem 3.1 of Ayoub [4] for expressing the sum of the coefficients of a Dirichlet series yields for nonintegral x and any $\alpha > 3$,

$$\sum_{n \leq x} n A(n) = \frac{1}{2\pi i} \int_{\alpha - i\infty}^{\alpha + i\infty} f(s) \frac{x^s}{s} ds \quad (3.22)$$

with

$$f(s) = \zeta(s-1) \sum_p \frac{p}{p^{s-1} - 1}.$$

Changing the variable of integration from s to $s + 2$ in (3.22) gives, for nonintegral x , and any $a > 1$;

$$\sum_{n \leq x} n A(n) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} g(s) \frac{x^{s+2}}{s+2} ds \quad (3.23)$$

with

$$g(s) = \zeta(s+1) \sum_p \frac{p}{p^{s+1}-1}.$$

In what follows the path $\operatorname{re} s = a$ in (3.23) will be deformed so that part of it lies slightly to the left of the line $\operatorname{re} s = 1$. We now show that Ayoub's bounds for $\log \zeta(s)$ apply to $g(s)$. Observe that for $\operatorname{re} s > \frac{1}{2}$, the function $\zeta(s+1)$ is uniformly bounded, in absolute value, by $\zeta(\frac{3}{2})$. Further

$$\sum_p \frac{p}{p^{s+1}-1} = \sum_p \frac{1}{p^s} + \sum_p \frac{1}{p^s(p^{s+1}-1)} = \log \zeta(s) + h(s), \quad (3.24)$$

where

$$h(s) = \sum_p \frac{1}{p^s(p^{s+1}-1)} - \sum_{\substack{k,p \\ k \geq 2}} \frac{1}{kp^{ks}}.$$

Thus $h(s)$ is analytic in the half plane $\operatorname{re} s > \frac{1}{2}$ and uniformly bounded in any half plane $\operatorname{re} s > b$, with $b > \frac{1}{2}$.

Suppose the path of integration is now deformed exactly as in [4, Chap. 2, Sect. 5]. Ayoub's arguments yield bounds for all parts of the path except along the cut running from $b + i\epsilon$ to $b - i\epsilon$, where b is $1 - c \log^{-9} T$ as in [4, Eq. (1), p. 65]. Along the cut, make the substitution $g(s) = \zeta(s+1) \log \zeta(s) + \zeta(s+1)h(s)$ with $h(s)$ defined by (3.24). Since $\zeta(s+1)h(s)$ is analytic and single valued along the cut, the integral along the upper side of the cut cancels the integral along the lower side. From here, the argument in [4, p. 69] yields

$$\begin{aligned} \frac{1}{2\pi i} \int_{\text{cut}} g(s) \frac{x^{s+2}}{s+2} ds &= \int_b^1 \zeta(s+1) \frac{x^{s+2}}{s+2} ds \\ &= \int_{1/2}^1 \zeta(s+1) \frac{x^{s+2}}{s+2} ds + O(x^3 \exp(-c(\log x)^{1/10})). \end{aligned} \quad (3.25)$$

The last equality follows from the choice of b given in [4, p. 70]. This completes the proof of (3.1). Equation (3.2) follows by routine integration by parts.

The argument for $n G(n)$ is virtually the same, and is omitted.

The arguments for Theorems 2, 4, and 5 are also virtually identical to the proof of Theorem 1. In each case, the identity for the Dirichlet series is used together with the inversion theorem (3.1) of Ayoub [4] as in (3.22). Then a change of variables is made to move the path of integration to $re s = a$ with $a > 1$. Again the integrand differs from $\log \zeta(s)$ by a bounded analytic function. Thus, Ayoub's argument can be used to bound the integrals away from $s = 1$. Along the cut the argument given for Theorem 1 holds essentially word for word. Further details are omitted. $\square$

Proof of Theorems 3 and 6. It is useful to have another way to express the limiting relations to be proved.

LEMMA 8. Let $H(n)$ denote one of the functions $A^*(n)$, $A(n)$, $P_1(n)$, or $G(n)$. Then, the following two conditions are equivalent. As $x \rightarrow \infty$,

$$\frac{1}{x} |\{n \leq x : H(n) \leq n^y\}| \rightarrow L(y) \quad \text{for } 0 \leq y \leq 1. \quad (3.26)$$

$$\frac{1}{x} |\{n \leq x : H(n) \leq x^y\}| \rightarrow L(y) \quad \text{for } 0 \leq y \leq 1. \quad (3.27)$$

Proof. Heuristically, Lemma 8 is true because most integers less than x are "large." We argue that (3.27) implies (3.26): Clearly $\{n \leq x : H(n) \leq n^y\} \subset \{n \leq x : H(n) \leq x^y\}$. But,

$$\begin{aligned} \{n \leq x : H(n) \leq x^y\} &= \{n < x/\log x : H(n) \leq x^y\} \\ &\quad \cup \{x/\log x \leq n \leq x : H(n) \leq x^y, H(n) > n^y\} \\ &\quad \cup \left\{ \frac{x}{\log x} \leq n \leq x : H(n) \leq n^y \right\} \\ &= S_1 \cup S_2 \cup S_3. \end{aligned}$$

The set S_1 is negligible and

$$S_2 \subset \left\{ n \leq x : \frac{x^y}{(\log x)^y} \leq H(n) \leq x^y \right\}.$$

If (3.27) holds, this last set, and so S_2 , has density 0. Finally, S_3 differs from $\{n \leq x : H(n) \leq n^y\}$ by a set of density 0. This completes the proof that (3.27) implies (3.26). The proof of the reverse implication is similar and is omitted. $\square$

The results for $A(n)$ and $G(n)$ given in Theorem 6 imply Theorem 3; thus, we need only prove Theorem 6. Theorem 6, for the function $P_1(n)$, was proved by de Bruijn [9]. Nice discussion and simplified proofs are given by Billingsley [7] and Knuth and Trabb-Prado [11]. The idea of the

proof of Theorem 6 is to use the known results for $P_1(n)$ by showing that $A(n)$, $A^*(n)$, and $G(n)$ differ from $P_1(n)$ by a "small" amount.

We now prove Theorem 6 for $A(n)$. Recall that we write $P_i(n)$ for the i th largest prime divisor of n . Let $y \in (0, 1)$ be fixed, and choose an integer m so large that $1/m < y$. Observe that

$$\left\{ n \leq x : \sum_{i=1}^m P_i(n) \leq \frac{x^y}{2} \quad \text{and} \quad \left[A(n) - \sum_{i=1}^m P_i(n) \right] \leq \frac{x^y}{2} \right\} \\ \subset \{ n \leq x : A(n) \leq x^y \} \subset \left\{ n \leq x : \sum_{i=1}^m P_i(n) \leq x^y \right\}. \quad (3.28)$$

Write $Q_x(S)$ for the proportion of integers $n \leq x$ such that $n \in S$. Take S as the smallest set in (3.28)

$$Q_x \left\{ \sum_{i=1}^m P_i(n) \leq \frac{x^y}{2} \quad \text{and} \quad \left[A(n) - \sum_{i=1}^m P_i(n) \right] \leq \frac{x^y}{2} \right\} \\ \geq Q_x \left\{ \sum_{i=1}^m P_i(n) \leq \frac{x^y}{2} \right\} - Q_x \left\{ A(n) - \sum_{i=1}^m P_i(n) > \frac{x^y}{2} \right\}. \quad (3.29)$$

Now, Markov's inequality for positive random variables (for $X > 0$, $P(X > c) \leq E(X)/c$), together with the theorem of Alladi and Erdős quoted above, implies

$$Q_x \left\{ A(n) - \sum_{i=1}^m P_i(n) > \frac{x^y}{2} \right\} \leq Q_x \left\{ A(n) - \sum_{i=1}^m P_i(n) > \frac{x^{1/m}}{2} \right\} \\ = O \left(\frac{x^{-1/m(m+1)}}{(\log x)^{m+1}} \right) = o(1). \quad (3.30)$$

Next we observe that the limiting distribution of $\sum_{i=1}^m P_i(n)$ is the same as the limiting distribution of $P_1(n)$. This follows from the inclusions

$$\{ n \leq x : P_1(n) \leq x^y \} \supset \left\{ n \leq x : \sum_{i=1}^m P_i(n) \leq x^y \right\} \\ \supset \{ n \leq x : mP_1(n) \leq x^y \}$$

along with de Bruijn's result which implies $|\{n \leq x : mP_1(n) \leq x^y\}| \sim xL(y)$ for y and m fixed as $x \rightarrow \infty$. Using the last observation, (3.30) and (3.29) in (3.28), completes the proof for $A(n)$.

The proof of Theorem 6 for $A^*(n)$ follows from the result just proved for $A(n)$ via Markov's inequality together with the estimate (3.9) of Alladi and Erdős for the mean of the difference $A(n) - A^*(n)$.

The last stage in the proof of Theorem 6 is to show that $G(n)$ has the same limiting distribution as $A^*(n)$. The idea of the proof is to show that for almost all integers n , $G(n)$ and $A^*(n)$ differ by at most a bounded amount. Toward this end, define $a_p(n)$ as the largest number such that $p^{a_p(n)}|n$. For integers m and $Z > 0$, let

$$S_{m,Z} = \{n : 0 \leq a_p(n) < m \text{ for } 2 \leq p \leq Z \text{ and } 0 \leq a_p(n) \leq 1 \text{ for } p > Z\}.$$

It is straightforward to show that $S_{m,Z}$ has density

$$\delta_{m,Z} = \left\{ \prod_{p \leq Z} \left(1 - \frac{1}{p^m}\right) \right\} \left\{ \prod_{p > Z} \left(1 - \frac{1}{p^2}\right) \right\}.$$

Since the product $\prod_p (1 - 1/p^2)$ converges, $\delta_{m,Z}$ can be made arbitrarily close to 1 by choosing Z and then m suitably large. Since $G(n) \geq A^*(n)$, $Q_x\{G(n) \leq n^y\} \leq Q_x\{A^*(n) \leq n^y\}$. For the opposite inequality, let y be fixed and note that on $S_{m,Z}$

$$0 \leq G(n) - A^*(n) \leq \sum_{p \leq Z} p^m \equiv c(m, Z).$$

Then,

$$\begin{aligned} Q_x\{G(n) \leq n^y\} &\geq Q_x\{\{S_{m,Z}\} \cap \{G(n) \leq n^y\}\} \\ &\geq Q_x\{\{S_{m,Z}\} \cap \{A^*(n) \leq n^y - c(m, Z)\}\} \\ &\geq Q_x\{A^*(n) \leq n^y - c(m, Z)\} - Q_x\{S_{m,Z}^c\}. \end{aligned}$$

Now, $Q_x\{A^*(n) \leq n^y - c(m, Z)\} \sim L(y)$ as $x \rightarrow \infty$. By choosing m and Z suitably large $Q_x\{S_{m,Z}^c\}$ can be made arbitrarily small. This completes the proof of Theorem 6 for $G(n)$. $\square$

4. TO SPLIT OR NOT TO SPLIT?

The FFT of n numbers can be computed by using the chirp- z transform directly or by splitting n into prime factors, computing the FFT for each factor, and then putting the pieces back together. These two approaches were described in Section 2. Both approaches work in $O(n \log n)$ time. A more careful comparison will now be presented.

As explained in Assumptions (2.7) and (2.8) of Section 2, a reasonable approximation for the number of operations used by the two algorithms is $\beta C(n)$ for the chirp- z transform and $\beta F(n)$ for the full factorization transform. Here $\beta > 0$ is a constant which may be taken as 2 and, if $T(n)$

is the smallest power of 2 larger than $2n - 4$, $C(n)$ and $F(n)$ are defined by

$$C(n) = \begin{cases} 3T(n) \log_2 T(n) & \text{if } n \neq 2^k \\ n \log_2 n & \text{if } n = 2^k, \end{cases} \quad (4.1)$$

$$F(n) = n \sum_{p^a | n} \frac{C(p)}{p}. \quad (4.2)$$

As a numerical example, when $n = 100$,

$$C(100) = 3 \cdot 256 \cdot \log_2 256 = 6144,$$

while

$$\begin{aligned} F(100) &= 100[2(C(2)/2) + 2(C(5)/5)] \\ &= 100[2(2/2) + 2(3 \cdot 8 \cdot 3/5)] = 3080. \end{aligned}$$

Table I gives $F(n)$ for $1000 \leq n \leq 1025$. For each n in Table I (except 1024), $C(n) = 67,584$. The results in Table I suggest that the better algorithm speeds things up by a factor of approximately $\frac{3}{2}$. Neither approach dominates. It is not clear how the approximations used to form Table I compare with actual running times.

Some information about the two approaches can be gleaned from a comparison of averages. Recall that we write $\lceil x \rceil$ for the smallest integer larger than x , $\lfloor x \rfloor$ for the largest integer smaller than x , and $\{x\}$ for the fractional part of x .

THEOREM 7. *Let $C(n)$ and $F(n)$ be defined by (4.1) and (4.2). For $x > 0$ define*

$$w(x) = \begin{cases} \frac{1}{2} & \text{if } \{\log_2(\lfloor x \rfloor - 2)\} = 0 \\ 2^{-(\log_2(\lfloor x \rfloor - 2))} & \text{otherwise.} \end{cases}$$

As x tends to infinity,

$$\frac{1}{x} \sum_{n \leq x} C(n) = 12w(x) \left(1 - \frac{2w(x)}{3}\right) x \log_2 x + O(x), \quad (4.3)$$

$$\frac{1}{x} \sum_{n \leq x} F(n) = \frac{3}{\log 2} x \log_2 x + O(x \log \log x). \quad (4.4)$$

The factor $w(x)$ oscillates boundedly between 1 and $\frac{1}{2}$ so that $12w(x)(1 - \frac{2}{3}w(x))$ oscillates between 4 and 4.5, while $3/\log 2 = 4.33$.

TABLE I
Number of Operations Using Full
Factorization of n With $\beta = 1^a$

n	Factorization	$F(n)$
1000	$2^3 5^3$	46,200
1	$7 \cdot 11 \cdot 13$	108,096
2	$2 \cdot 3 \cdot 167$	85,950
3	$17 \cdot 59$	74,016
4	$2^2 \cdot 251$	57,304
5	$3 \cdot 5 \cdot 67$	108,642
6	$2 \cdot 503$	62,446
7	$19 \cdot 53$	112,128
8	$2^4 \cdot 3^2 \cdot 7$	35,712
9	1009	122,880
10	$2 \cdot 5 \cdot 101$	76,994
11	$3 \cdot 337$	94,182
12	$2^2 \cdot 11 \cdot 23$	96,872
13	1013	122,880
14	$2 \cdot 3 \cdot 13^2$	40,482
15	$5 \cdot 7 \cdot 29$	82,776
16	$2^3 \cdot 127$	52,200
17	$3^2 \cdot 113$	59,364
18	$2 \cdot 509$	62,458
19	1019	122,880
20	$2^2 \cdot 3 \cdot 5 \cdot 17$	47,568
21	1021	122,880
22	$2 \cdot 7 \cdot 73$	115,070
23	$3 \cdot 11 \cdot 31$	84,702
24	2^{10}	10,240
25	$5^2 \cdot 41$	96,720

^aDirect use of the chirp-z idea uses 67,584 operations.

Proof of Theorem 7. To prove (4.3) it is easiest to approximate the distribution of $C(n)$ and then calculate the mean from the distribution. The largest value $C(n)$ can take as n ranges in $1 \leq n \leq x$ is determined by $J = J(x) = \lceil \log_2(2 \lfloor x \rfloor - 4) \rceil$; $C(n)$ takes values $3J2^J, 3(J-1)2^{J-1}, \dots$. When $n = 2^k$, $C(n) = 2^k k$. Since this happens only once for each value of k , these values will not affect the computation. That is, $(1/x) \sum_{k \leq \lfloor \log_2 x \rfloor} k 2^k = O(\log x)$ and the error in (4.3) is $O(x)$. Excepting powers of 2, $C(n) = 3k2^k$ for $2^{k-2} + 2 < n \leq 2^{k-1} + 2$. Write

$$\begin{aligned}
 f(x) &= J(x) - 2 - \log_2 x = -1 && \text{if } x = 2^k + 2 \\
 &= -\{\log_2(\lfloor x \rfloor - 2)\} + O(1/x) && \text{otherwise.}
 \end{aligned}$$

Then

$$\begin{aligned}\frac{1}{x} |(n \leq x : C(n) = 3J2^J)| &= 1 - 2^{(J-2)-\log_2 x} + O\left(\frac{1}{x}\right) \\ &= 1 - 2^{f(x)} + O\left(\frac{1}{x}\right) = 1 - w(x) + O\left(\frac{1}{x}\right)\end{aligned}$$

and similarly, for $1 \leq k \leq J$,

$$\frac{1}{x} |(n \leq x : C(n) = 3(J-k)2^{J-k})| = \frac{w(x)}{2^k} + O\left(\frac{1}{x}\right).$$

The error term is uniform in k . The mean of $C(n)$ is

$$\begin{aligned}&3J2^J \left(1 - w(x) + O\left(\frac{1}{x}\right)\right) + 3(J-1)2^{J-1} \left(\frac{w(x)}{2} + O\left(\frac{1}{x}\right)\right) \\ &\quad + 3(J-2)2^{J-2} \left(\frac{w(x)}{2^2} + O\left(\frac{1}{x}\right)\right) + \cdots + O(\log x) \\ &= 3J2^J \left(1 - \frac{2}{3}w(x)\right) + O(x) \\ &= 3 \cdot 4w(x) \left(1 - \frac{2}{3}w(x)\right)x \log_2 x + O(x).\end{aligned}$$

This proves (4.3).

To prove (4.4), write

$$\begin{aligned}\sum_{n \leq x} F(n) &= \sum_{n \leq x} n \sum_{p^a | n} \frac{C(p)}{p} = \sum_{p^a \leq x} \frac{C(p)}{p} \sum_{\substack{p^a | n \\ n \leq x}} n = \sum_{p^a \leq x} \frac{C(p)}{p} p^a \sum_{j \leq x/p^a} j \\ &= \sum_{p^a \leq x} \frac{C(p)}{p} \frac{p^a}{2} (\lfloor x/p^a \rfloor^2 + \lfloor x/p^a \rfloor).\end{aligned}\quad (4.5)$$

We now argue from (4.5) to the approximation

$$\sum_{n \leq x} F(n) = \frac{x^2}{2} \sum_{p \leq x} \frac{C(p)}{p^2} + O(x^2).\quad (4.6)$$

To derive (4.6), we need the prime number theorem in the crude form $\sum_{p^a \leq x} \log p = O(x)$ (see [4, Theorem 6.2], for example). We also use $C(p) = O(p \log p)$. First, $\lfloor x/p^a \rfloor^2 = (x/p^a)^2 + O(x/p^a)$, so

$$\sum_{p^a \leq x} \frac{C(p)}{p} \frac{p^a}{2} \lfloor x/p^a \rfloor^2 = \frac{x^2}{2} \sum_{p^a \leq x} \frac{C(p)}{p^{a+1}} + O\left(x \sum_{p^a \leq x} \frac{C(p)}{p}\right).\quad (4.7)$$

But

$$\sum_{p^a \leq x} \frac{C(p)}{p} = O\left(\sum_{p^a \leq x} \log p\right) = O(x). \quad (4.8)$$

Using (4.8) in (4.7) and again for the term involving $\lfloor x/p^a \rfloor$ on the right side of (4.5) gives

$$\sum_{n \leq x} F(n) = \frac{x^2}{2} \sum_{p^a \leq x} \frac{C(p)}{p^{a+1}} + O(x^2). \quad (4.9)$$

The proof of (4.6) is completed by the bound:

$$\sum_{\substack{p^a \leq x \\ a \geq 2}} \frac{C(p)}{p^{a+1}} = O\left(\sum_{\substack{p^a \leq x \\ a \geq 2}} \frac{\log p}{p^a}\right) = O\left(\sum_{n=1}^{\infty} \frac{\log n}{n^2}\right) = O(1).$$

Next we need the following bound: For $w < z$, as $w \rightarrow \infty$,

$$\sum_{w < p < z} \frac{1}{p^2} = \left\{ \frac{1}{w \log w} - \frac{1}{z \log z} \right\} + O\left(\frac{1}{w \log^2 w}\right). \quad (4.10)$$

To prove (4.10), write $\pi(t)$ for the number of primes $\leq t$. Then,

$$\sum_{w < p < z} \frac{1}{p^2} = \int_w^z \frac{1}{t^2} d\pi(t) + O(1/w^2) = \frac{\pi(t)}{t^2} \Big|_w^z + 2 \int_w^z \frac{\pi(t)}{t^3} dt + O\left(\frac{1}{w^2}\right).$$

Using the prime number theorem in the form

$$\begin{aligned} \pi(t) &= \frac{t}{\log t} + O\left(\frac{t}{(\log t)^2}\right), \\ \frac{\pi(t)}{t^2} \Big|_w^z &= \frac{1}{z \log z} - \frac{1}{w \log w} + O\left(\frac{1}{w \log^2 w}\right); \end{aligned}$$

also,

$$\begin{aligned} 2 \int_w^z \frac{\pi(t)}{t^3} dt &= 2 \int_w^z \frac{dt}{t^2 \log t} + O\left(\int_w^z \frac{dt}{t^2 \log^2 t}\right) \\ &= 2\left(\frac{1}{w \log w} - \frac{1}{z \log z}\right) + O\left(\frac{1}{w \log^2 w}\right). \end{aligned}$$

This completes the proof of (4.10).

Finally, we complete our argument from (4.6) to (4.4). Write $L = L(x) = \lfloor \log_2 x \rfloor$. We have

$$\begin{aligned} \sum_{p \leq x} \frac{C(p)}{p^2} &= \sum_{k=1}^{L+1} 3k2^k \sum_{2^{k-2}+2 < p \leq 2^{k-1}+2} \frac{1}{p^2} \\ &\quad + 3(L+2)2^{L+2} \sum_{2^L+2 < p \leq x} \frac{1}{p^2} + O(1). \end{aligned} \quad (4.11)$$

The approximation (4.10) gives

$$\sum_{2^L+2 < p \leq x} \frac{1}{p^2} = O\left(\frac{1}{L2^L}\right)$$

and

$$\begin{aligned} \sum_{2^{k-2}+2 < p \leq 2^{k-1}+2} \frac{1}{p^2} &= \left[\frac{1}{2^{k-2} \log 2^{k-2}} - \frac{1}{2^{k-1} \log 2^{k-1}} \right] + O\left(\frac{1}{2^k k^2}\right) \\ &= \frac{1}{2^{k-1}(k-2) \log 2} \left[1 + O\left(\frac{1}{k}\right) \right]. \end{aligned}$$

Using these bounds in (4.11) leads to

$$\begin{aligned} \sum_{p \leq x} \frac{C(p)}{p^2} &= \frac{6}{\log 2} \sum_{k=2}^{L+1} \left(1 + O\left(\frac{1}{k}\right) \right) + O(1) \\ &= \frac{6 \log_2 x}{\log 2} + O(\log \log x). \end{aligned}$$

Using this in (4.6) completes the proof of Theorem 7. $\square$

ACKNOWLEDGMENTS

David Freedman, Andrew Odlyzko, Lawrence Rabiner, Don Rose, Larry Shepp, and Charles Stein all helped at crucial times. I also thank a patient, careful referee.

REFERENCES

1. A. AHO, J. HOPCRAFT, AND J. ULLMAN, "The Design and Analysis of Computer Algorithms," Addison-Wesley, Reading, Mass., 1974.
2. A. AHO, K. STEIGLITZ, AND J. ULLMAN, Evaluating polynomials at fixed sets of points, *SIAM J. Computers* **4** (1975), 533-539.
3. K. ALLADI AND P. ERDŐS, On an additive arithmetic function, *Pacific J. Math.* **71** (1977), 275-294.

4. R. AYOUN, "An Introduction to the Analytic Theory of Numbers," Amer. Math. Soc., Providence, R. I., 1963.
5. R. BELLMAN AND B. KOTKIN, On the numerical solution of a differential-difference equation arising in analytic number. *Theor. Math. Comp.* **16** (1962), 473-475.
6. G. BERGLAND, A guided tour to the fast Fourier transform, *IEEE Trans. Audio Electroacoust.* **AN-16** (1968), 66-76.
7. P. BILLINGSLEY, On the distribution of large prime divisors, *Periodica Math. Hungar.* **2**, (1972), 283-289.
8. N. G. DE BRUIJN, On the number of uncanceled elements in the sieve of Eratosthenes. *Indag. Math.* **12** (1950), 247-256.
9. N. G. DE BRUIJN, On a function occurring in the theory of primes, *J. Indian Math. Soc. A* **15** (1951), 25-32.
10. K. DICKMAN, On the frequency of numbers containing prime factors of a certain relative magnitude, *Ark. Mat., Astronom. Fysik* **22A** **10** (1930), 1-14.
11. D. KNUTH AND L. TRABB-PARDO, Analysis of a simple factorization algorithm, *Theoret. Comp. Sci.* **3** (1976), 321-348.
12. D. LEHMER, Computer technology applied to the theory of numbers, in "Studies in Number Theory," pp. 117-151, Math. Assoc. Amer. (distributed by Prentice-Hall, Englewood Cliffs, N. J.), 1969.
13. L. RABINER, R. SCHAFER, AND C. RADER, The chirp-z transform and its applications. *Bell System Tech. J.* **48** (1969), 1249-1292.
14. L. RABINER AND C. RADER, "Digital Signal Processing," IEEE, New York, 1972.
15. C. RADER, Discrete Fourier transforms when the number of data samples is prime. *IEEE Proc.* **56** (1968), 1107-1108.
16. D. ROSE, "Matrix Identities of the Fast Fourier Transform," *Linear Algebra Appl.*, in press.
17. B. SAFFARI, Sur quelques applications de la "methode de L'Hyperbole" de Dirichlet a la theorie des nombres premiers, *Enseignement Math.* **14** (1969), 205-224.
18. R. SINGLETON, An algorithm for computing the mixed radix fast Fourier transform. *IEEE Trans. Audio Electroacoust.* **AU-17** (1969), 158-161.
19. J. VAN DE LUNE AND E. WATTEL, On the numerical solution of a differential-difference equation arising in analytic number theory. *Math. Comp.* **23** (1969), 417-421.
20. I. J. GOOD, The interaction algorithm and practical Fourier series, *J. Roy. Statist. Soc. Ser. B.* **20** (1958), 361-372.