

The Number of Primes

Pál Dorozslai , Horacio Keller

Independent Researcher, Kékkút, Hungary

Email: paldorozslai@yahoo.com

How to cite this paper: Dorozslai, P. and Keller, H. (2022) The Number of Primes. *Advances in Pure Mathematics*, 12, 81-95. <https://doi.org/10.4236/apm.2022.122008>

Received: November 8, 2021

Accepted: February 20, 2022

Published: February 23, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

It is known that the prime-number-formula at any distance from the origin has a systematic error. It is shown that this error is proportional to the square of the number of primes present up to the square root of the distance. The proposed completion of the prime-number-formula in the present paper eliminates this systematic error. This is achieved by using a quickly converging recursive formula. The remaining error is reduced to a symmetric dispersion of the effective number of primes around the completed prime-number-formula. The standard deviation of the symmetric dispersion at any distance is proportional to the number of primes present up to the square root of the distance. Therefore, the absolute value of the dispersion, relative to the number of primes is approaching zero and the number of primes resulting from the prime-number-formula represents the low limit of the number of primes at any distance.

Keywords

Prime-Number-Formula, Prime Gaps

1. Introduction

Writing the integral of de la Valée Poussin as summing over all integers the local density of primes results in a first approximation for the number of primes. The summing may be written as two steps: First summing up the local density of primes within sections of length (\sqrt{c}) and then summing up to over all the (\sqrt{c}) sections. As the first simplification over each section, the average of the local density is taken. A second simplification is taking instead the average of the local density of primes within each section, the density within the last section at the distance (c) . This simplification results in the well-proved prime-number-formula. It is shown that the difference between the first simplification and the prime-number-formula is proportional to the square of the number of primes present up to the distance (\sqrt{c}) .

Further, it is shown that the error due to the first simplification is proportional to the number of primes present up to the distance (\sqrt{c}) . This proportionality

ty allows for a correction of the first simplification. In fact, this correction may be applied recursively. Avoiding the second simplification, the resulting formula gives without systematic error the number of primes at the distance (c).

2. Evaluation of the Number of Primes

De la Valée Poussin proved 1899, (ref. [1]), that ($\pi(c) \approx Li(c)$), the integral of the local logarithmic density of free positions ($1/\ln(c)$). Approximating the integral by the sum of the logarithmic density over all integers is in the following named as **sum over all integers**:

$$\pi(c) + O(c) \approx \int_2^c \frac{1}{\ln(c)} dc; \quad \pi_{appr}(c) \approx \sum_{n=2}^c \frac{1}{\ln(c)} \tag{1.1}$$

The above sum may be written as summing up first over all integers within the sections of length (\sqrt{c}) and then summing up to over all the (\sqrt{c}) sections of the length (\sqrt{c}). Taking the average value over each section and summing up to over the sections is the first simplification, in the following used as **sum over all sections**:

$$\pi_{appr} \approx \sum_{n=2}^c \frac{1}{\ln(c)} = \sum_{j=2}^{\sqrt{c}} \left[\sum_{n=(j-1)\sqrt{c}}^{j\sqrt{c}} \frac{1}{\ln(c)} \right] \approx \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} \tag{1.2}$$

The well proven prime-number-formula **PNF** results from a second simplification of the above approximation by taking for each of the sections the smallest value of the density at ($j = \sqrt{c}$):

$$\sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} > \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(\sqrt{c}\sqrt{c})} = \frac{c}{\ln(c)} = S(c) \tag{1.3}$$

The difference between the first simplification of the number of primes and the value resulting from the **PNF**, (refs. [2] [3]), is proportional to ($R(c)^2 = S(\sqrt{c})^2$), the square of the number of primes up to the distance (\sqrt{c}):

$$\begin{aligned} & \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(\sqrt{c}\sqrt{c})} \\ &= \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - \frac{\sqrt{c}}{\ln(c)} \sum_{j=1}^{\sqrt{c}} 1 = \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - S(c) = \gamma_c R(c)^2 \end{aligned}$$

because value (γ_c) quickly converges to a constant value, as shown in **Annex A1**:

$$\begin{aligned} \gamma_c &= \frac{\sqrt{c}}{\ln(c)R(c)^2} \sum_{j=1}^{\sqrt{c}} \left(\frac{\ln(c)}{\ln(j\sqrt{c})} - \sqrt{c} \right) \\ &= \frac{1}{2R(c)} \sum_{j=1}^{\sqrt{c}} \left(\frac{\ln(c)}{\ln(j\sqrt{c})} - \sqrt{c} \right) \\ &= \frac{1}{R(c)} \sum_{j=1}^{\sqrt{c}} \left(\frac{\ln(\sqrt{c})}{\ln(j\sqrt{c})} - \frac{\sqrt{c}}{2} \right) \end{aligned} \tag{1.4}$$

Thus, the error of the second simplification resulting in the **PNF** at the distance

(c) is proportional to ($R(c)^2$), the square of the number of primes present at (\sqrt{c}). The factor of proportionality is evaluated in **Annex A1** as ($\gamma_c = 0.2829\dots$).

The relation between the error of the PNF at (c) and the square of the number of primes up to (\sqrt{c}) is invariant. The constant factor of proportionality (γ_c) is an inherent propriety of the number of primes.

The first simplification replacing ($\sum_{n=(j-1)\sqrt{c}}^{j\sqrt{c}} \frac{1}{\ln(n)}$) by ($\frac{\sqrt{c}}{\ln(j\sqrt{c})}$) in (1.2) originates an error ($\Delta\pi_{error}(c)$).

It can be proved, that this error is proportional to ($R(c)$), the number of the series of multiples of primes, which are covering positions at (c). This is because each of the series of multiples of primes contributes its share to the error:

$$\Delta\pi_{error}(c) = \gamma_{sec} \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} \approx \gamma_{sec} \frac{\sqrt{c}}{\ln(\sqrt{c}\sqrt{c})} \sum_{j=1}^{\sqrt{c}} 1 \approx \gamma_{sec} R(c) \quad (1.5)$$

Herewith the error at the distance (c)—meaning the difference between the effective number of primes ($\pi(c)$) and the value resulting from the first simplification ($\pi_{sec_appr_}(c)$), using the average value of the density over the sections in (1.2)—is proportional to the number of primes up to (\sqrt{c}). The factor of proportionality is evaluated in annex 2.2:

$$\Delta\pi_{error}(c) = \pi_{sec_appr_}(c) - \pi(c) = \gamma_{sec} R(c); \quad \gamma_{sec} = -0.03551 \quad (1.6)$$

Because the error of the first simplification resulting from the procedure of taking the average of the number of primes over the sections—evaluated as sum over all sections (1.2)—is proportional to the number of primes present up to the root of the distance, the approximating function (1.2) may be corrected correspondingly. The corrected approximating function, the **corrected sum over all sections** is instead of (1.2):

$$\begin{aligned} \pi_{sec_appr_corr_}(c) &= \pi_{sec_appr_}(c) - \gamma_{sec} \pi_{sec_appr_}(\sqrt{c}) \\ &= \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - \gamma_{sec} \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} \end{aligned} \quad (1.7)$$

But herewith for ($R_{eff}(c)$) and for ($S_{eff}(c)$) it may be written recursively:

$$\begin{aligned} \pi_{sqr_sec_appr_corr_}(c) &= \pi_{sqr_sec_appr_}(c) - \gamma_{sec} \pi_{sqr_sec_appr_}(\sqrt{c}) \\ &= \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - \gamma_{sec} \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} \\ \pi_{sec_appr_corr_}(c) &= \pi_{sec_appr_}(c) - \gamma_{sec} \pi_{sec_appr_}(\sqrt{c}) \\ &= \pi_{sec_appr_}(c) - \gamma_{sec} \left(\pi_{sqr_sec_appr_}(c) - \gamma_{sec} \pi_{sqr_sec_appr_}(\sqrt{c}) \right) \quad (1.8) \\ &= \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - \gamma_{sec} \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} + \gamma_{sec}^2 \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} + \dots \end{aligned}$$

This allows writing for the **complete-prime-number-formula (CPNF)**:

$$\pi_{appr_}(c) = \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - \sum_{m=1}^{\infty} \gamma_{sec}^m \sum_{j=1}^{c^{\left(\frac{1}{2^{m+1}}\right)}} \frac{c^{\left(\frac{1}{2^{m+1}}\right)}}{\ln\left(j \cdot c^{\left(\frac{1}{2^{m+1}}\right)}\right)} \tag{1.9}$$

This formula converge very fast: Two steps with $(m = 2)$ are already sufficient.

The results of the **CPNF** are evaluated in **Annex A3** and compared with the effective number of primes. Compared with another **approximating formula (refs. [4] [5])**, **there is no systematic error. The remaining error is the dispersion around the CPNF.** It is proportional to the number of primes up to $(R(c))$, the series of multiples that are covering positions at (c) . Therefore, the dispersion of the effective number of primes around the **CPNF**, relative to $(R(c))$ has constant boundaries, symmetrical around zero, as demonstrated in **Annex A2** and **Annex A3**:

$$\Delta\pi_{appr_rel_}(c) = \frac{\pi_{sec_appr_}(c) - \pi(c)}{R(c)}; \quad \Delta\pi_{sqr_appr_rel_}(c) = \frac{\pi_{sqr_sec_appr_}(c) - \pi(\sqrt{c})}{R(\sqrt{c})} \tag{1.10}$$

The standard deviation of the relative dispersion is:

$$SD_{\Delta\pi_{appr_rel_}}(c) = \sqrt{\frac{1}{c} \sum_c \left(\frac{\pi_{appr_}(c) - \pi_{eff}(c)}{\pi_{eff}(\sqrt{c})} \right)^2} \tag{1.11}$$

It is evaluated in **Annex A4**. It is found that the standard deviation is constant over the distance. This constant is evaluated in **Annex A4** as $(F_{SD_{\Delta\pi}} = 0.16238)$, it is again an **inherent property of the prime numbers**.

3. Boundaries of the Dispersion of the Number of Primes around the CPNF

Regarding **Figure A3.1**, it is obvious, that the dispersion shows oscillations of different frequencies: There are different oscillations with high, middle, and low frequencies observable. The evaluation of the number of primes with the **CPNF** correct the error of the average on the long range, but at the same time it is the origin of oscillations due to the summation procedures over $(R(c), R(\sqrt{c}), R(\sqrt{\sqrt{c}}), \dots)$ subsections. Each time these values change by unity, there is a discontinuity in the summation procedure. At the distance (c) , with $(R(c))$ being the number of primes present up to (\sqrt{c}) , the summing of the local density values over (j) is influenced only by (\sqrt{c}) . If for the distance succeeding primes were taken, as in the above evaluations, the size of the gaps between primes would account for the long-range fluctuations. If $(R(c))$ changes by unity, there is a jump: At the change of the distance at certain number of primes $(P_{(n)})$ the summation limit changes

$$\begin{aligned} &\text{from } (n = R(c), c = P_{(n)}) \text{ to } (n = R(c)+1, c = P_{(n+1)}), \\ &\text{from } \left(\sqrt{c} \sum_{j=1}^{R(c)} \frac{1}{\ln(j\sqrt{c})}\right) \text{ to } \left(\sqrt{c} \sum_{j=1}^{R(c)+1} \frac{1}{\ln(j\sqrt{c})}\right) \end{aligned} \tag{2.1}$$

The difference makes

$$\left(\sqrt{c} \sum_{j=1}^{R(c)+1} \frac{1}{\ln(j\sqrt{c})} - \sqrt{c} \sum_{j=1}^{R(c)} \frac{1}{\ln(j\sqrt{c})} \approx \frac{\sqrt{c}}{\ln(c)} = \frac{R(c)}{2} \right). \quad (2.2)$$

For illustration, the discontinuity of the approximating function is evaluated and illustrated in **Annex A5**.

Thus, the dispersion of the effective number of the primes around the best estimate value—due to the evaluation procedure of the best estimate value—is proportional to the number of primes, the series of multiples of which primes cover positions at the distance (c). Another source of dispersion is the variable gap between the primes. Concerning the complete dispersion, the following lemma may be formulated:

Lemma 2.1:

The boundaries of the dispersion of the number of primes at any distance (c) are proportional to ($R(c)$), the number of the series of multiples of primes, which are covering positions at (c).

Proof: The cover interval is defined as the interval of the size ($2\sqrt{c}$) with centre at the distance (c). Shifting the centre of the cover interval to the origin, all the series of multiples of the ($R(c)$) primes coincide, and each series covers at least two positions within the cover interval, around the origin. Such a complete coincidence occurs next times at the distance equal to the product of all first ($R(c)$) primes.

In the following, only the series of multiples of the primes in the range ($P_{(n)}$, $R(\sqrt{c}) < n < R(c)$) will be taken into consideration. This does not influence the validity of the proof, since for large distances nearly all primes are in this range.

At any distance ($\sqrt{c} \cdots c$) all of the series of multiples of primes are shifted with reference to (c), except the ones coinciding at (c). All these shifted series liberate by the shifting two positions within the cover interval. The series of multiples of some primes—after the shifting—may cover again two positions, which are liberated by the shifting of the series of multiples of two other primes. This is only possible, if the corresponding prime is—before the shifting—equidistant to two other primes. It can be proved, that the density of the number of such three equidistant primes at the distance (c) is ($\frac{\delta_2 c}{\ln(c)^3}$; $\delta_2 = 1.320323632$), thus ($\frac{\delta_2}{\ln(c)^2}$) times smaller, than the density of primes.

Because of this fact, **the effect of the multiple coverage** after the shifting may be neglected: On the average, each shifted series of multiples of primes liberates two positions within the cover interval and covers only one.

Herewith the shifted series of multiples of the ($R(c) - R(\sqrt{c})$) primes leave on the average ($R(c) - R(\sqrt{c})$) positions free within the cover interval. These free positions are for a large distance about equally distributed between the upper and the lower part of the cover interval. Therefore, within the sections of the length (\sqrt{c}) at the arbitrary distance (c) there are, on the average

$(\frac{1}{2}(R(c) - R(\sqrt{c})) \approx \frac{1}{2}R(c))$ positions left free.

Besides the effect of the multiple coverages, there are some other effects, which may be the reason for divergence from the average value $(R(c)/2)$. All the following effects may be neglected:

The **effect of multiple coincidences** between the series of multiples of primes: The number of coinciding primes $(q = Q(c))$ is limited by the fact, that their product must be smaller than the distance (c) . The largest number of coincidences occurs, if the first (q) primes are coinciding:

$$Per(q) = \prod_{p=1}^q P_{(p)}; \quad Q(c) = Per(q); \quad Per(q) \leq c < Per(q+1) \quad (2.3)$$

The number $(Q(c))$ is growing very slowly with the distance (c) . Its effect may be neglected against $(R(c)/2)$, the average value of free positions remaining.

The **effect of the shifting outside of the cover interval**: Each series of multiples considered may cover two positions within the cover interval after the shifting. By the shifting, the covering capability of some series of multiples of primes may be reduced by shifting outside of the limits of the cover interval one of the covered positions. The number of covered positions within the cover interval may only be reduced by this effect, herewith rising the remaining number of the remaining free positions. The number of the primes, which may participate in this effect is limited, therefore the effect may be neglected against $(\frac{1}{2}R(c))$, the average value of free positions remaining.

Therefore, the total dispersion is proportional to $(\frac{1}{2}R(c))$, as stated in the lemma and **concluding the proof**.

With this lemma the dispersion of the number of primes within each section of the length (\sqrt{c}) at the distance $(j\sqrt{c})$ is certainly smaller than the average number of free positions left $(\frac{1}{2}R(c))$. Therefore, the difference between the effective value of primes up to this distance $(\pi(c))$ and its approximation $(\pi_{appr}(c))$ is limited to the number of the series of multiples of primes, which are covering positions at this distance:

$$\pi_{bounds_eff}(c) = \pm \frac{\sqrt{c}}{2 \ln(c)};$$

$$\pi_{bounds}(c) \pi_{bounds}(c) = \pi_{bounds_appr}(c) + \pi_{bounds_eff}(c) = \pm \frac{\sqrt{c}}{\ln(c)} = \pm R(c) \quad (2.4)$$

Lemma 2.1 is in accordance with the fact—demonstrated in (1.11) and in **Annex A4**—that the standard deviation of the dispersion around the best estimate approximation of the number of primes—resulting from the **CPNF**, and relative to $(R(c))$ —is constant over the distance (c) . **The constancy of the standard deviation of the relative dispersion** around the best estimate approximation of the number of primes—resulting from the **CPNF**—is an inhe-

rent property of the prime numbers.

From Lemma 2.1, the boundary of the dispersion of the effective number of primes around the value resulting from its best estimate expression is proportional to the number of primes, which are covering position at any distance from the origin ($K \cdot R(c)$), with (K) being a constant factor of any size:

$$\lim_{c \rightarrow \infty} \pi_{disp}(c) = \lim_{c \rightarrow \infty} (K \cdot \pi_{bounds_eff}(c)) = \lim_{c \rightarrow \infty} (K \cdot R(c)) = \infty \quad (2.5)$$

With this boundary growing to infinity, an additional proof for the fact, which **the upper limit of the gap between consecutive primes is unlimited**, (see ref. [6]). On the other hand, it follows, that the width of the boundaries of the dispersion of the number of primes, relative to the value resulting from its best estimate expression, is approaching zero with the distance growing without limit:

$$\frac{\pi_{disp}(c)}{\pi(c)} = \frac{K \cdot \pi_{bounds_eff}(c)}{\pi(c)} = \frac{K \cdot R(c)}{S(c)} : \lim_{c \rightarrow \infty} \left(\frac{\frac{K \cdot \sqrt{c}}{\ln(\sqrt{c})}}{c} \right) = \lim_{c \rightarrow \infty} \frac{2 \cdot K}{\sqrt{c}} = 0 \quad (2.6)$$

Therefore, the series of primes approaches a continuum for large distances.

Lemma 2.2:

For any distance from the origin (c) large enough, the difference between the values resulting from the complete-prime-number-formula (CPNF) and from the prime-number-formula (PNF) is larger than the width of the dispersion of the number of primes ($R(c)$) multiplied by any constant factor ($K \geq 1$).

Proof: The difference between the value of the (CPNF) and the (PNF) is with (1.3) growing proportional to ($R(c)^2$), with the factor of proportionality (γ_c). The width of the dispersion of the number of primes grows with Lemma 2.1 only proportional to ($R(c)$).

For any sufficiently large distance ($c_{limit}(\gamma_c) < c$) the difference between the value of the (CPNF) and the (PNF) outgrows ($K \cdot R(c)$) for any (K):

$$\lim_{c \rightarrow \infty} \frac{\pi(c) - S(c)}{K \cdot R(c)} = \frac{\gamma_c}{K} R(c) = \infty$$

as stated in the lemma and **concluding the proof**.

From this Lemma 2.2 follows, that the **PNF** is the low limit of the number of primes. Because the number of primes has a low limit function, growing to infinity, it is infinite itself: An additional proof for the infinity of the number of primes.

4. Conclusions

For large distances, the CPNF gives a result for the number of primes as good as de la Valée Poussin’s formula: There is no systematic error involved. The explanation for this fact is that de la Valée Poussin’s formula uses the integral of the local density given by the inverse of the logarithm, while Equation (1.0) uses the summation of the same values, applied over the sections of the length (\sqrt{c}), with recurring correction. This allows to formulate the following lemma:

Lemma 3.1:

The complete-prime-number-formula (CPNF) gives the number of primes with increasing efficiency for all distances from the origin.

Proof: The approximating formula (1.9) is based on the recurring application of the correction of the replacement of the de la Valée Poussin integral:

$$\sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} - \gamma_{sec} \sum_{j=1}^{\sqrt{\sqrt{c}}} \frac{\sqrt{\sqrt{c}}}{\ln(j\sqrt{\sqrt{c}})} + \gamma_{sec}^2 \sum_{j=1}^{\sqrt{\sqrt{\sqrt{c}}}} \frac{\sqrt{\sqrt{\sqrt{c}}}}{\ln(j\sqrt{\sqrt{\sqrt{c}}})} + \dots$$

For large distances (c) each of the components of the correction goes over into the integral of the inverse of the logarithm of the distance (c):

$$\lim_{c \rightarrow \infty} \left(\sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} \right) = \int_0^c \frac{1}{\ln(c)} dc ; \quad \gamma_{sec} \sum_{j=1}^{\sqrt{\sqrt{c}}} \frac{\sqrt{\sqrt{c}}}{\ln(j\sqrt{\sqrt{c}})} = \gamma_{sec} \int_0^{\sqrt{c}} \frac{1}{\ln(\sqrt{c})} dc$$

For the distance (c) growing to infinity, the contribution of the successive components—relative to the effective number of primes approaches zero. The remaining first component results in the de la Valée Poussin integral as stated in the lemma and **concluding the proof**.

The first evaluation of the factor (γ_c) in (1.6) and (1.7) is somewhat heuristic since the convergence is not strongly proved. But the application of this factor in the CPNF results in a converging approximation of the effective number of primes. This fact justifies the evaluation procedure of the factor (γ_{sec}) and proves the validity of its value. The evaluation of the factor (γ_{sec}) is repeated in **Annex A3** with the results of the CPNF. It gives the exact value.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Titchmarsh, E.C. (1986) *The Theory of the Riemann Zeta-Function*. Oxford University Press, New York.
- [2] Pintz, J. (1980) On the Remainder Term of the Prime Number Formula II. On a Theorem of Ingham. *Acta Arithmetica*, **37**, 209-220.
<https://doi.org/10.4064/aa-37-1-209-220>
- [3] Platt, D.J. and Trudgian, T.S. (2021) The Error Term of the Prime Number Theorem. *Mathematics of Computation*, **90**, 871-881. <https://doi.org/10.1090/mcom/3583>
- [4] Rosser, J.B. and Trudgian, T.S. (1941) Explicit Bounds for Some Functions of Prime Numbers. *American Journal of Mathematics*, **63**, 211-232.
<https://doi.org/10.2307/2371291>
- [5] Rosser, J.B. and Schoenfeld, L. (1962) Approximate Formulas for Some Functions of Primes. *Illinois Journal of Mathematics*, **6**, 64-94.
<https://doi.org/10.1215/ijm/1255631807>
- [6] Green, B. and Tao, T. (2007) The Primes Contain Arbitrary Long Arithmetic Progressions. *Annals of Mathematics*, **167**, 481-547.
<https://doi.org/10.4007/annals.2008.167.481>

Annexes

All above formula are checked in the following annexes with numeric results. The annexes together compose executable files in MATHCAD from PTC. The syntax corresponds to the MATHCAD syntax. Especially the symbols ($ceil(x)$) and ($floor(x)$) stand for rounding up or down a real value to the next larger respectively to the next smaller integer.

Annex 0: General Data, Vectors, and Functions

In the following all formula are checked with numeric results. The set of primes and the listed known formula below are used for this checking. Some vectors of the results of the known formula, which are often used, are evaluated and the results are saved.

The set of primes is read from a file: ($P := READPRN("PRimes_large.prn")$). The number of primes in the set and their numbering are:

$$(N_p := rows(P) - 1 = 5003713, n := 1, 2, \dots, N_p).$$

The number of primes up to (c) is approximated with the prime-number-formula ($S(c)$). At (c) only the multiples of the primes up to (\sqrt{c}) are covering free integral positions ($R(c) = S(\sqrt{c})$). The numbers of these primes are given as well:

$$P_{(S(c))} < c < P_{(S(c)+1)}; \pi_c(c) > S(c) := floor\left(\frac{c}{\ln(c)}\right) \approx \frac{c}{\ln(c)} \tag{A0.1}$$

$$P_{(R(c))} < \sqrt{c} < P_{(R(c)+1)}; \pi_c(\sqrt{c}) > S(\sqrt{c}) = R(c) := floor\left(\frac{\sqrt{c}}{\ln(\sqrt{c})}\right) \approx \frac{\sqrt{c}}{\ln(\sqrt{c})}$$

For the evaluation of the effective number of free positions up to the distance (c) the routine ($n_{next}(c, n_{last})$) resulting the index (n) of the prime next to any integer is needed ($P_{(n)} \leq c < P_{(n+1)}$). The evaluation starts either at the last evaluated index (n_{last}), or at the index resulting from the prime-number-formula. This, to shorten some of the evaluation processes. Further functions are the indexes of the closest primes to any distance, and to the square root of any distance, as well as a short expression for the exponential function:

$$n_{next}(c, n_{last}) = \begin{cases} \text{if } c \neq 0 \\ \left| \begin{array}{l} n \leftarrow S(c) \text{ if } n_{last} = 1 \\ n \leftarrow n_{last} \text{ otherwise} \\ \text{while } P_{(n)} \leq c \\ \quad n \leftarrow n + 1 \\ \text{Res} \leftarrow n - 1 \end{array} \right. & S_{eff}(c) := n_{next}(c, 1); R_{eff}(c) := n_{next}(\sqrt{c}, 1); \exp_{-}(x) := e^x \tag{A0.2} \\ \text{Res} \leftarrow 0 \text{ otherwise} \end{cases}$$

In the following all functions, which are evaluated for illustration, are evaluated at sparse distances, equal to multiples of the square root of the largest distance considered ($c_{sp(k)} = k\sqrt{P_{(N_p)}}$):

$$\Delta c_{sp} := \sqrt{P_{(N_p)}}; k_{limit} := \text{floor} \left[\frac{P_{(N_p)}}{\Delta c_{sp}} \right] - 1 = 9277;$$

$$k := 1, 2, \dots, k_{limit}; c_{sp(k)} := k \cdot \Delta c_{sp} \tag{A0.3}$$

The vectors of the indexes of the primes corresponding to these sparse distances, respectively to the next smaller primes ($P_{[n_{sp(k)}]} < c_{sp(k)} < P_{[n_{sp(k+1)}]}$) are evaluated for $(\pi_{sp}(c))$, $(\pi_{sp}(\sqrt{c}))$, and $(\pi_{sp}(\sqrt{\sqrt{c}}))$. They are evaluated once and written to files. They are read from these files:

$$\pi_{sp(k)} := n_{next} [c_{sp(k)}, 1]; \pi_{sqr_sp(k)} := n_{next} [\sqrt{c_{sp(k)}}, 1];$$

$$\pi_{sqr_sqr_sp(k)} := n_{next} [\sqrt{\sqrt{c_{sp(k)}}}, 1] \tag{A0.4}$$

```
WRITEPRN("index_distance_sp.prn") := \pi_{sp};
\pi_{sp} := READPRN("index_distance_sp.prn"); length(\pi_{sp}) = 9278
WRITEPRN("n_sqr_sp_primes.prn") := \pi_{sqr\_sp};
\pi_{sqr\_sp} := READPRN("n_sqr_sp_primes.prn"); length(\pi_{sqr\_sp}) = 9278
WRITEPRN("n_sqr_sqr_sp_primes.prn") := \pi_{sqr\_sqr\_sp};
\pi_{sqr\_sqr\_sp} := READPRN("n_sqr_sqr_sp_primes.prn")
```

Annex 1. Evaluation of the Number of Primes as Sum over Sections

The number of primes as sum over sections is evaluated with (1.2) as a first simplification:

$$\pi_{sec_appr_}(c) = \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln(j\sqrt{c})} \tag{A1.1}$$

The number of primes resulting from the second simplification (1.3) results the **PNF**. The difference between the first and the second simplification results the error of the PNF. At the distance (c) it is proportional to $(R(c)^2)$, the square of the number of primes covering positions at the distance (\sqrt{c}). The factor of proportionality is evaluated over the distance with (1.4):

$$\gamma_{c_appr_}(k) := \frac{1}{R[c_{sp(k)}]} \left[\frac{\sum_{j=1}^{ceil[\sqrt{c_{sp(k)}}]} \ln[\sqrt{c_{sp(k)}}]}{\ln[j\sqrt{c_{sp(k)}}]} - \frac{1}{2} \sqrt{c_{sp(k)}} \right] \tag{A1.2}$$

The factors are evaluated once and written to a file. They are read from this file:

```
WRITEPRN("gamma_c_appr_sp.prn") := \gamma_{c\_appr\_};
\gamma_{c\_appr\_} := READPRN("gamma_c_appr_sp.prn")
```

The factors are approaching a constant value: **(Figure A1.1)**

$$\gamma_{c_appr} := \gamma_{c_appr_}(k_{limit}) = 0.28312$$

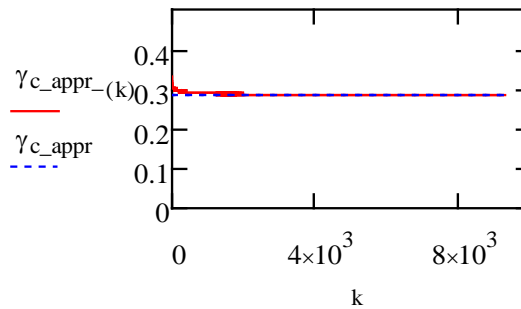


Figure A1.1. Relation of the error of the prime-number-formula to the square of the number of the series of multiples of primes, which are covering positions, over the distance.

The approximating function is evaluated at sparse distances, respectively at the next smaller prime to these distances ($P_{[n_{sp(k)}]} < c_{sp(k)} < P_{[n_{sp(k+1)}]}$) with (1.2).

They are evaluated similarly as well at the square root and at the square root of the square root of these sparse distances. The evaluation at the next smaller prime corresponding to each distance assures that the evaluated numbers of the primes correspond exactly to the distances considered:

$$\pi_{sec_appr(k)} := \pi_{sec_appr-} \left[P_{[n_{sp(k)}]} \right]; \quad \pi_{sqr_sec_appr(k)} := \pi_{sec_appr-} \left[P_{[\pi_{sqr_sp(k)}]} \right] \quad (A1.3)$$

$$\pi_{sqr_sqr_sec_appr(k)} := \pi_{sec_appr-} \left[P_{[\pi_{sqr_sqr_sp(k)}]} \right]$$

They are evaluated once and written to files. They are read from these files:

```
WRITEPRN("pr_sec_appr_sp_t.prn") := pi_sec_appr ;
pi_sec_appr := READPRN("pr_sec_appr_sp_t.prn") ;
WRITEPRN("pr_sqr_sec_appr_sp_t.prn") := pi_sqr_sec_appr ;
pi_sqr_sec_appr := READPRN("pr_sqr_sec_appr_sp_t.prn") ;
WRITEPRN("pr_sqr_sqr_sec_appr_sp_t.prn") := pi_sqr_sqr_sec_appr ;
pi_sqr_sqr_sec_appr := READPRN("pr_sqr_sqr_sec_appr_sp_t.prn")
```

Annex A2. Evaluation of the Factor of Correction of the First Simplification

The result of the first simplification (1.2) giving the sum over the sections of the density of primes has an error. This error is proportional to the number of primes up to (\sqrt{c}) . The error relative to $(\pi(\sqrt{c}))$ results the factor of correction. Assuming the factor of correction (γ_{sec}) is constant over the distance (c), it may be evaluated as relation of the average error to the effective number of primes ($\pi(\sqrt{c}) = \pi_{sp(k)}$). The average error is:

$$\Delta\pi_{sec(k)} := \pi_{sec_appr(k)} - \pi_{sp(k)}; \quad \Delta\pi_{sec_av(k)} := \frac{1}{k} \sum_{j=1}^k [\Delta\pi_{sec(j)}] \quad (A2.1)$$

The value of the factor of correction is herewith:

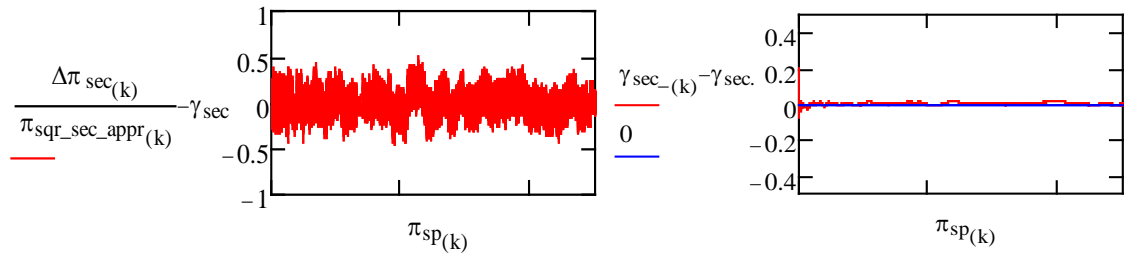


Figure A2.1. Convergence of the relation of the average relative error of the first simplification (1.2) to the final constant value (γ_{sec}).

$$\gamma_{sec}(c) = \frac{\Delta\pi_{sec_{av}}(c)}{\pi(\sqrt{c})}; \quad \gamma_{sec_{(k)}} := \frac{\sqrt{2}\Delta\pi_{sec_{av}(k)}}{\pi_{sqr_{sec_{appr}(k)}}}; \quad \gamma_{sec} := \gamma_{sec_{(k_{limit})}} = -0.035513 \quad (A2.2)$$

Figure A2.1 shows the independence of the factor of correction (γ_{sec}) from the distance. The averaging process (A2.1) to evaluate the factor of correction is therefore justified. **This factor ($\gamma_{sec} = -0.035513$) is invariant, an inherent property of the prime numbers:**

Annex A3. Evaluation of the CPNF

The CPNF (1.9) is evaluated with the following routine:

$$\begin{array}{l}
 \pi_{appr}(c) := \left| \begin{array}{l}
 m \leftarrow 1 \quad * \\
 S \leftarrow \sum_{j=1}^{\text{floor}(\sqrt{c})} \frac{\sqrt{c}}{\ln(j\sqrt{c})} \\
 \text{while } m < \sqrt{c} \\
 \quad \left| \begin{array}{l}
 \Delta_{(m)} \leftarrow (\gamma_{sec})^m \cdot \sum_{j=1}^{\text{floor}(c_{\text{exp}}(c,m))} \text{floor}\left(\frac{c_{\text{exp}}(c,m)}{\ln(j \cdot c_{\text{exp}}(c,m))}\right) \\
 \text{break if } \Delta_{(m)} < 1 \\
 S \leftarrow S - \Delta_{(m)} \\
 m \leftarrow m + 1
 \end{array} \right. \\
 S
 \end{array} \right. \quad (A3.1)
 \end{array}$$

The results of the CPNF are evaluated once at sparse values of the distance ($c_{sp(k)}$), for (c) , (\sqrt{c}) and $(\sqrt{\sqrt{c}})$. They are written to files and are read from these files:

```

WRITEPRN("Pi_appr_sp.prn") :=  $\pi_{appr}$ ,
 $\pi_{appr}$  := READPRN("Pi_appr_sp.prn"),

WRITEPRN("Pi_sqr_appr_sp.prn") :=  $\pi_{sqr_{appr}}$ ,
 $\pi_{sqr_{appr}}$  := READPRN("Pi_sqr_appr_sp.prn")

WRITEPRN("Pi_sqr_sqr_appr_sp.prn") :=  $\pi_{sqr_{sqr_{appr}}}$ ,
 $\pi_{sqr_{sqr_{appr}}}$  := READPRN("Pi_sqr_sqr_appr_sp.prn").
    
```

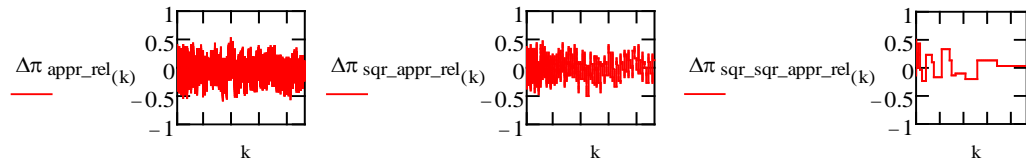


Figure A3.1. The relative dispersion of the difference between the effective number of primes and its value evaluated with the complete-prime-number-formula (CPNF).

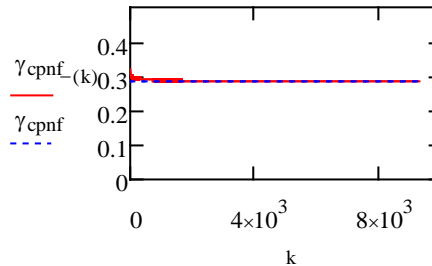


Figure A3.2. Relation of the error of the PNF to $(R(c)^2)$, the square of the number of primes up to (\sqrt{c}) , over the distance.

$$\begin{aligned} \pi_{appr(k)} &:= \pi_{appr_} \left[P_{\left[\pi_{sp(k)} \right]} \right]; \quad \pi_{sqr_appr(k)} := \pi_{appr_} \left[P_{\left[\pi_{sqr_sp(k)} \right]} \right]; \\ \pi_{sqr_sqr_appr(k)} &:= \pi_{appr_} \left[P_{\left[\pi_{sqr_sqr_sp(k)} \right]} \right] \end{aligned} \tag{A3.2}$$

(Figure A3.1) There is no systematic error of the number of primes resulting with the CPNF. The dispersion of the evaluated values relative to the effective number of primes at (\sqrt{c}) is about constant over the distance up to (c) , (\sqrt{c}) and $(\sqrt{\sqrt{c}})$:

$$\begin{aligned} \Delta \pi_{appr_rel(k)} &:= \frac{\pi_{appr(k)} - \pi_{sp(k)}}{R \left[1 \cdot c_{sp(k)} \right]}; \quad \Delta \pi_{sqr_appr_rel(k)} := \frac{\pi_{sqr_appr(k)} - \pi_{sqr_sp(k)}}{R \left[\sqrt{1 \cdot c_{sp(k)}} \right]} \tag{A3.3} \\ \Delta \pi_{sqr_sqr_appr_rel(k)} &:= \frac{\pi_{sqr_sqr_appr(k)} - \pi_{sqr_sqr_sp(k)}}{R \left[\sqrt{\sqrt{1 \cdot c_{sp(k)}}} \right]} \end{aligned}$$

with the results of the CPNF the factor of the proportionality (γ_c) of the error of the PNF relative to the square of the number of primes present up to (\sqrt{c}) evaluated in (A1.2) with (1.4), is reevaluated with the more exact difference as follows: (Figure A3.2)

$$\frac{\pi_{appr_cpnf_}(c)}{R(c)^2} = \gamma_{cpnf_-(k)} := \frac{\pi_{appr(k)} - s \left[c_{sp(k)} \right]}{R \left[c_{sp(k)} \right]^2} \tag{A3.4}$$

Annex A4. Evaluation of the Standard Deviation of the Dispersion of the Effective Number of Primes around the CPNF

The standard deviation SD of the relative dispersion (1.13) is evaluated as fol-

lows:

$$SD_{\Delta\pi_appr_rel(k)} := \sqrt{\frac{1}{k} \sum_{j=1}^k \left[\frac{\pi_{appr(j)} - \pi_{sp(j)}}{R_{eff} [c_{sp(j)}]} \right]^2} \tag{A4.1}$$

The results are evaluated once and written to a file:
 (WRITEPRN("SD_\Delta\prime_sqr_appr_rel.prn") := SD_{\Delta\pi_appr_rel}). They are read from this file: (SD_{\Delta\pi_appr_rel} := READPRN("SD_\Delta\prime_sqr_appr_rel.prn")).

The average of the relation of the standard deviation converges to a final value, to the factor of proportionality (F_{SD_\Delta\pi}). This factor is evaluated as follows:

$$SD_{\Delta\pi_appr_rel_av(k)} := \frac{1}{k} \sum_{j=1}^k SD_{\Delta\pi_appr_rel(j)} \tag{A4.2}$$

The results are evaluated once and written to a file. They are read from this file:

$$\begin{aligned} \text{WRITEPRN("SD_Delta\prime_sqr_appr_rel_av.prn")} &:= SD_{\Delta\pi_appr_rel_av} \\ SD_{\Delta\pi_appr_rel_av} &:= \text{READPRN("SD_Delta\prime_sqr_appr_rel_av.prn")} \end{aligned}$$

The constant factor is equal to the final average value of the standard deviation at large distances. The figure below illustrates that the standard deviation is about constant over the distance. This fact rectifies taking the average over the whole distance for the evaluation:

$$\begin{aligned} F_{SD_\Delta\pi} &:= SD_{\Delta\pi_appr_rel_av(k_{limit-1})} = 0.162389; \\ SD_\pi(c) \cdot SD_{\Delta\pi_sp_rel_av(k)} \cdot R(c) &\approx F_{SD_\Delta\pi} \end{aligned} \tag{A4.3}$$

Figure A4.1 below indicates that the standard deviation of the dispersion of the effective number of primes around its approximation is rising proportionally to (R(c)), the number of the series of multiples of primes, which are covering integer positions at this distance (c). The factor of proportionality (F_{SD_\Delta\pi} = 0.162389) is again an **inherent property of the primes**.

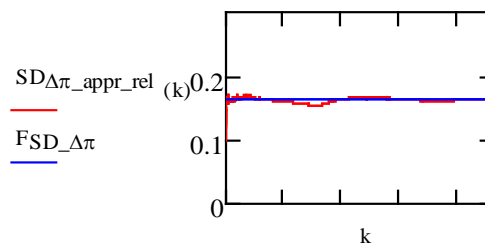


Figure A4.1. Dispersion of the standard deviation of the dispersion of the number of primes around its average, the resulting constant value.

Annex A5. Evaluation of the Boundaries of the Dispersion

At the distance (c), with (R(c)) being the number of the series of multiples of primes covering formerly free positions, the summing of the local density values over (j) is influenced only by (\sqrt{c}). If for the distance succeeding primes were

taken, as in the above evaluations, the size of the gaps between primes would account for the long-range fluctuations. If $(R(c))$ changes by unity, there is a jump: At the change of the distance at certain number of primes $(P_{(n)})$ the summation-limit changes from $(R(c), c = P_{(n)})$ to $(R(c)+1, c = P_{(n+1)})$:

$$\text{from } \left(\sqrt{c} \sum_{j=1}^{R(c)} \frac{1}{\ln(j\sqrt{c})} \right) \text{ to } \left(\sqrt{c} \sum_{j=1}^{R(c)+1} \frac{1}{\ln(j\sqrt{c})} \right) \tag{A5.1}$$

The difference makes

$$\left(\sqrt{c} \sum_{j=1}^{R(c)+1} \frac{1}{\ln(j\sqrt{c})} - \sqrt{c} \sum_{j=1}^{R(c)} \frac{1}{\ln(j\sqrt{c})} \right) \approx \frac{\sqrt{c}}{\ln(c)} = \frac{\sqrt{c}}{2 \cdot \ln(\sqrt{c})} = \frac{R(c)}{2}. \tag{A5.2}$$

For illustration, the discontinuity of the approximating function is shown. For this purpose, the values of the approximating function are evaluated in (addition to its sparse values evaluated in (1.4)) at distances corresponding to each prime $(c = P_{(q)}, \pi_{appr_test(q)} = \pi_{appr_}(q))$ in the range $(q = q_{low} \cdots q_{high})$:

$$\begin{aligned} &\text{from } (q_{low} := 1909500, P(q_{low}) = 30887083) \\ &\text{to } (q_{high} := 1911600, P(q_{high}) = 30923341, q := q_{low} \cdots q_{high}) \\ &\pi_{appr_test(q)} := \pi_{sec_appr_} [P(q)] \end{aligned} \tag{A5.3}$$

The results are written to a file and are read from this file:

```
WRITEPRN("Nprime_appr_c.prn") := pi_appr_test ;
pi_appr_test := READPRN("Nprime_appr_c.prn").
```

The discontinuity of the approximating function is shown in **Figure A5.1**. The size of the jumps (2.2) is the bandwidth of the dispersion due to the approximating function:

$$\pi_{bounds_appr}(c) = \pm \frac{R(c)}{2} \tag{A5.4}$$

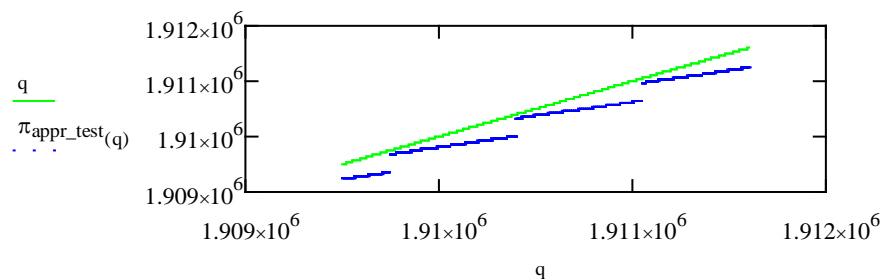


Figure A5.1. Discontinuity due to the approximating function.