The number of primes

Pal Doroszlai, Horacio Keller

Abstract

The prime-number-formula at any distance from the origin has a systematic error, proportional to the square of the number of primes up to the square root of the distance. The proposed completion in the present paper eliminates by a quickly converging recursive formula the systematic error. The remaining error is reduced to a symmetric dispersion, with standard deviation proportional to the number of primes at the square root of the distance.

1: Evaluation of the number of primes

The total number of the primes is the integral of the local logarithmic density of free positions, evaluated by Riemann. The first approximation of the integral is the sum of the logarithmic density over all integers, in the following used as **sum over all integers**:

$$\pi(c) = \int_2^c \frac{1}{\ln(c)}\,dc \qquad \pi_{\ln\_appr}(c) \approx \sum_{n=2}^{c} \frac{1}{\ln(n)} \tag{1.1}$$

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

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

The well proven prime-number-formula **PNF** results as simplification of the above approximation taking for all sections the largest value ($j = \sqrt{c}$):

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

The difference between the approximation of the number of primes and the value resulting from the **PNF** is proportional to the square of the number of primes present up to the distance ($\sqrt{c}$):

$$\sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\sqrt{c}}{\ln(j\cdot\sqrt{c})} - \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\sqrt{c}}{\ln(\sqrt{c}\cdot\sqrt{c})} = \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \left(\frac{\sqrt{c}}{\ln(j\cdot\sqrt{c})} - \frac{\sqrt{c}}{\ln(\sqrt{c}\cdot\sqrt{c})} \right) = \sqrt{c}\cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(\sqrt{c}\cdot\sqrt{c}) - \ln(j\cdot\sqrt{c})}{\ln(j\cdot\sqrt{c})\cdot\ln(\sqrt{c}\cdot\sqrt{c})} =$$

$$= \frac{\sqrt{c}}{\ln(c)} \cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(c) - \ln(j\cdot\sqrt{c})}{\ln(j\cdot\sqrt{c})} = \frac{\sqrt{c}}{\ln(c)} \cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(c)}{\ln(j\cdot\sqrt{c})} - \frac{\sqrt{c}}{\ln(c)} \cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} 1 = \sqrt{c}\cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{1}{\ln(j\cdot\sqrt{c})} - S(c) = \gamma_c\cdot R(c)^2$$

$$\gamma_c = \frac{\sqrt{c}}{\ln(c)\cdot R(c)^2} \cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(c) - \ln(j\cdot\sqrt{c})}{\ln(j\cdot\sqrt{c})} = \frac{\sqrt{c}}{2\cdot\ln(\sqrt{c})\cdot R(c)^2} \cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(c) - \ln(j\cdot\sqrt{c})}{\ln(j\cdot\sqrt{c})} = \tag{1.4}$$

$$= \frac{1}{2\cdot R(c)} \cdot \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(c) - \ln(j\cdot\sqrt{c})}{\ln(j\cdot\sqrt{c})} = \frac{1}{2\cdot R(c)} \cdot \left(\sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(c)}{\ln(j\cdot\sqrt{c})} - \sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} 1 \right) =$$

$$= \frac{1}{2\cdot R(c)} \cdot \left(\sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{\ln(\sqrt{c})}{\ln(\sqrt{j\cdot\sqrt{c}})} - \sqrt{c} \right) = \frac{1}{R(c)} \cdot \left(\sum_{j=1}^{\mathrm{ceil}(\sqrt{c})} \frac{2\cdot\ln(\sqrt{\sqrt{c}})}{\ln(\sqrt{j\cdot\sqrt{c}})} - \frac{1}{2}\cdot\sqrt{c} \right)$$

The error of this approximation, of the **PNF**, at the distance (c) 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 in annex 1. It is ($\gamma_c := 0.28298$):

The relation between the error of the PNF at (c) and 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.

Replacing ($\sum_{n=\text{ceil}[(j-1)\cdot\sqrt{c}]}^{\text{floor}(j\cdot\sqrt{c})} \frac{1}{\ln(n)}$) by ($\frac{\sqrt{c}}{\ln(j\cdot\sqrt{c})}$) in (1.2) originates an error ($\Delta\pi_{error}(c)$).

It can be proved, that his error is proportional to the number of primes at ($\sqrt{c}$), the number of the series of multiples of primes, which are covering positions at (c), because each of the series of multiples of primes contributes its share to the error:

$$\Delta\pi_{error\_}(c) = \gamma_{sec}\cdot\sum_{j=1}^{\text{ceil}\left(\sqrt{\sqrt{c}}\right)} \frac{\sqrt{\sqrt{c}}}{\ln\left(j\cdot\sqrt{\sqrt{c}}\right)} \approx \gamma_{sec}\cdot\frac{\sqrt{\sqrt{c}}}{\ln\left(\sqrt{\sqrt{c}}\cdot\sqrt{\sqrt{c}}\right)}\cdot\sum_{j=1}^{\text{ceil}\left(\sqrt{\sqrt{c}}\right)} 1 \approx \gamma_{sec}\cdot R_{eff}(c) \tag{1.5}$$

Herewith the error at the distance (c), meaning the difference between the effective number of primes ($\pi(c) = \pi_{sp.sp_{(k)}}$) and the value resulting from the approximation ($\pi_{sec\_appr\_}(c)$), disregarding, respectively eliminating the local dispersion by using the average value over the sections in (1.2) . The factor of proportionality is evaluated in annex 2.2:

$$\Delta\pi_{sec\_}(c) = \pi_{sec\_appr\_}(c) - \pi(c) = \gamma_{sec}\cdot R_{eff}(c) \;\; ; \; \gamma_{sec} := -0.035513 \tag{1.6}$$

Because the error of the approximation of the number of primes over the distance - evaluated as sum over all sections (1.2) - is proportional to the number of primes at the square 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):

$$\pi_{sec\_appr\_corr\_}(c) = \pi_{sec\_appr\_}(c) - \gamma_{sec}\cdot\pi_{sec\_appr\_}\left(\sqrt{c}\right) = \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln\left(j\cdot\sqrt{c}\right)} - \gamma_{sec}\cdot\sum_{j=1}^{\sqrt{\sqrt{c}}} \frac{\sqrt{\sqrt{c}}}{\ln\left(j\cdot\sqrt{\sqrt{c}}\right)} \tag{1.7}$$

But herewith for ($R_{eff}\left[c_{sp_{(k)}}\right]$) and for ($S_{eff}\left[c_{sp_{(k)}}\right]$) it may be written recursively:

$$\pi_{sqr\_sec\_appr\_corr\_}(c) = \pi_{sqr\_sec\_appr\_}(c) - \gamma_{sec}\cdot\pi_{sqr\_sec\_appr\_}\left(\sqrt{c}\right) = \sum_{j=1}^{\sqrt{\sqrt{c}}} \frac{\sqrt{\sqrt{c}}}{\ln\left(j\cdot\sqrt{\sqrt{c}}\right)} - \gamma_{sec}\cdot\sum_{j=1}^{\sqrt{\sqrt{\sqrt{c}}}} \frac{\sqrt{\sqrt{\sqrt{c}}}}{\ln\left(j\cdot\sqrt{\sqrt{c}}\right)}$$

$$\pi_{sec\_appr\_corr\_}(c) = \pi_{sec\_appr\_}(c) - \gamma_{sec}\cdot\pi_{sec\_appr\_}\left(\sqrt{c}\right) =$$

$$= \pi_{sec\_appr\_}(c) - \gamma_{sec}\cdot\left(\pi_{sqr\_sec\_appr\_}(c) - \gamma_{sec}\cdot\pi_{sqr\_sec\_appr\_}\left(\sqrt{c}\right)\right) =$$

$$= \sum_{j=1}^{\sqrt{c}} \frac{\sqrt{c}}{\ln\left(j\cdot\sqrt{c}\right)} - \gamma_{sec}\cdot\sum_{j=1}^{\sqrt{\sqrt{c}}} \frac{\sqrt{\sqrt{c}}}{\ln\left(j\cdot\sqrt{\sqrt{c}}\right)} - \left(\gamma_{sec}\right)^2\cdot\sum_{j=1}^{\sqrt{\sqrt{\sqrt{c}}}} \frac{\sqrt{\sqrt{\sqrt{c}}}}{\ln\left(j\cdot\sqrt{\sqrt{c}}\right)} + \blacksquare\ldots \tag{1.8}$$

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

$$\pi_{appr\_}(c) = \sum_{j=1}^{\text{floor}\left(\sqrt{c}\right)} \frac{\sqrt{c}}{\ln\left(j\cdot\sqrt{c}\right)} - \sum_{m=1}^{\text{floor}\left(\sqrt{c}\right)} \left[\left(\gamma_{sec}\right)^m\cdot\sum_{j=1}^{\text{floor}\left(c^{\frac{1}{2^{m+1}}}\right)} \frac{c^{\frac{1}{2^{m+1}}}}{\ln\left(j\cdot c^{\frac{1}{2^{m+1}}}\right)}\right] \tag{1.9}$$

The results of the **CPNF** are evaluated in annex 3 and compared with the effective number of primes. There is no systematic error. The remaining error, the dispersion around the **CPNF** is proportional to the number of primes ($R(c)$), the multiples of which are covering positions at (c).

Therefore the dispersion relative to the **CPNF** has a constant width and symmetrical around zero:

$$\Delta\pi_{appr\_rel\_}(c) = \frac{\pi_{sec\_appr\_}(c) - \pi(c)}{R(c)} \;\; ; \;\; \Delta\pi_{sqr\_appr\_rel\_}(c) = \frac{\pi_{sqr\_sec\_appr\_}(c) - \pi\left(\sqrt{c}\right)}{R\left(\sqrt{c}\right)} \tag{1.10}$$

The standard deviation of the relative dispersion is:

$$SD_{\Delta\pi\_appr\_rel}(c) = \sqrt{\frac{1}{c}\cdot\sum_{c}\left(\frac{\pi_{appr}(c) - \pi_{eff}(c)}{\pi_{eff\_}\left(\sqrt{c}\right)}\right)^2} \tag{1.11}$$

It is evaluated in annex 4. It is found, that the standard deviation is constant over the distance. This constant is evaluated in annex 4 as ($F_{SD\_\Delta\pi} := 0.162389$), is again an **inherent property of the prime numbers.**.

2: Boundaries of the dispersion of the number of primes around the CPNF

Regarding figure 1.4, it is obvious, that the dispersion shows oscillations of different frequencies: there are oscillations of relatively low frequencies and oscillations with short and even shorter 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 the oscillations due to the summing procedures over ($R(c),\ R(\sqrt{c}),\ R\left(\sqrt{\sqrt{c}}\right)$ ...) subsections. Each time ($R(c)$) changes by unity, there is a discontinuity in the summation procedure. The same is happening by each of the following summing, but with diminishing amplitude and higher frequency. For the first and largest effect this is illustrated below for the original **PNF**:

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 ($R(c)+1,\ c = P_{(n+1)}$):

$$\text{from } \left(\sqrt{c}\cdot\sum_{j=1}^{R(c)} \frac{1}{\ln\left(j\cdot\sqrt{c}\right)}\right) \text{ to } \left(\sqrt{c}\cdot\sum_{j=1}^{R(c)+1} \frac{1}{\ln\left(j\cdot\sqrt{c}\right)}\right) \tag{2.1}$$

$$\text{The difference makes } \left(\sqrt{c}\cdot\sum_{j=1}^{R(c)+1} \frac{1}{\ln\left(j\cdot\sqrt{c}\right)} - \sqrt{c}\cdot\sum_{j=1}^{R(c)} \frac{1}{\ln\left(j\cdot\sqrt{c}\right)} \approx \frac{\sqrt{c}}{\ln(c)} = \frac{\sqrt{c}}{2\cdot\ln\left(\sqrt{c}\right)} = \frac{R(c)}{2}\right). \tag{2.2}$$

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 ($c = P_{(q)},\ \pi_{appr\_test(q)} = \pi_{appr\_}\left[P_{(q)}\right]$) in the range , ($q = q_{low} .. q_{high}$):

from ($q_{low} := 1909500$) to ($q_{high} := 1911600\ ,\ q := q_{low} .. q_{high}$)

$$\pi_{appr\_test(q)} := \pi_{sec\_appr\_}\left[P_{(q)}\right] \tag{2.3}$$

The results are written to a file: ($\mathrm{WRITEPRN}(\text{"Nprime\_appr\_c.prn"}) := \pi_{appr\_test}$). They are read from these files each time the present paper is evaluated: ($\pi_{appr\_test} := \mathrm{READPRN}(\text{"Nprime\_appr\_c.prn"})$). The discontinuity of the approximating function is shown in the next figure.

The size of the jumps (2.2) is the bandwidth of the dispersion due to the approximating function:

$$\pi_{bounds\_appr}(c) = \pm\frac{R(c)}{2}). \tag{2.4}$$

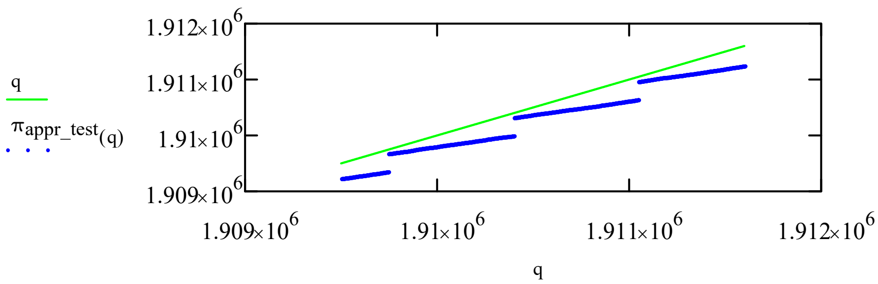

Figure 2.1: Discontinuity of the approximating function

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 covers positions at the distance (c). 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) is 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 ($\sqrt{c}$) both side of the point at the distance (c). At the origin all the series of multiples of the ($R(c)$) primes coincide and each series covers at least two positions around the origin, within the cover interval. 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 nearly all primes are in this range.

At any distance (c), up to (c^2) most of the series of multiples of primes are shifted with reference to (c). All these shifted series liberate by the 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 such three equidistant primes at the distance (c) is ($\frac{\delta_2 \cdot c}{\ln(c)^3}$; $\delta_2 = 1.3203236325$), thus ($\frac{\delta_2}{\ln(c)^2}$) times smaller, then 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 leaves on the average ($R(c) - R(\sqrt{c})$) positions free within the cover interval. These free positions are for 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 ($\frac{1}{2}\cdot(R(c) - R(\sqrt{c})) \approx \frac{1}{2}\cdot R(c)$) positions left

Besides the effect of the multiple coverage there are some other effects, which may be the reason for divergence from the average value ($\frac{1}{2}\cdot R(c)$). It can be shown, that all other 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 has to be smaller, then the distance (c). The largest number of the series of multiples of primes are coinciding, if the the first (q) primes are coinciding:

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

The number ($Q(c)$) is growing very slowly with the distance (c). Its effect may be neglected against ($\frac{1}{2}\cdot R(c)$), 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 of the series of multiples of primes may be reduced because 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, rising the remaining free positions. The number of the primes, which may participate in this effect is limited and the effect may be neglected against ($\frac{1}{2}\cdot R(c)$), the average value of free positions remaining.

Therefore the total dispersion is proportional to ($R(c)$), as stated in the lemma and **concluding the proof.**

With this lemma the divergence of the number of primes within each section of the length ($\sqrt{c}$) at the distance ($j\cdot\sqrt{c}$) is certainly smaller, then the average number of free positions left ($\frac{1}{2}\cdot R(c)$). 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 covering positions at this distance:

$$\pi_{bounds\_eff}(c) = \pm\frac{\sqrt{c}}{2\cdot\ln(c)} \;;\quad \pi_{bounds}(c) = \pi_{bounds\_appr}(c) + \pi_{bounds\_eff}(c) \quad = \pm\frac{\sqrt{c}}{\ln(\sqrt{c})} = \pm R(c) \qquad (2.5)$$

Lemma 2.1 is in accordance with the fact - demonstrated in (1.13) - that the standard deviation of the dispersion around the best estimate approximation of the number of primes - resulting from the **CPNF** - is constant over the distance to the origin. **The constancy of the standard deviation of the dispersion** around the best estimate approximation of the number of primes - resulting from the **CPNF - is an inherent property of the prime numbers.**

From this lemma 2.1 follows, that 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 any constant factor):

$$\lim_{c\to\infty} \pi_{disp}(c) = \lim_{c\to\infty} \left(K\cdot\pi_{bounds\_eff}(c)\right) = \lim_{c\to\infty} (K\cdot R(c)) = \infty \qquad (2.6)$$

With this boundary growing to infinity, **the upper limit of the gap between consecutive primes is unlimited.** 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:

$$\lim_{c \to \infty} \frac{\pi_{disp}(c)}{\pi(c)} = \lim_{c \to \infty} \left(\frac{K \cdot \pi_{bounds\_eff}(c)}{\pi(c)} \right) = \frac{K \cdot R(c)}{\pi_{appr}(c)} = \lim_{c \to \infty} \frac{K \cdot \frac{\sqrt{c}}{\ln(\sqrt{c})}}{\frac{c}{\ln(c)}} = \lim_{c \to \infty} \frac{K}{R(c)} = 0 \qquad (2.7)$$

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 then 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 ($\gamma_c = 0.28298$). The width of the dispersion of the number of primes is growing 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):

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

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

From this lemma 2.2 follows, that the **CPNF** 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.

3: Conclusions

For large distances the CPNF gives a result for the number of primes as good as Riemann's formula: There is no systematic error involved. The explication for this fact is, that Riemann's equation 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 (1.10) is based on the recurring application of the correction of the replacement of the Riemann integral:

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

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):

$$\lim_{c \to \infty} \left(\sum_{j=1}^{ceil(\sqrt{c})} \frac{\sqrt{c}}{\ln(j \cdot \sqrt{c})} \right) = \int_0^c \frac{1}{\ln(c)}\, dc \;;\; \gamma_{sec} \cdot \sum_{j=1}^{\sqrt{\sqrt{c}}} \frac{\sqrt{\sqrt{c}}}{\ln\left(j \cdot \sqrt{\sqrt{c}}\right)} = \gamma_{sec} \cdot \int_0^{\sqrt{c}} \frac{1}{\ln(c)}\, dc \ldots$$

For the distance (c) growing to infinity, the contribution of the successive components - relative to the effective number of primes approaches zero, leaving the first component, which results the Riemann integral as stated in the lemma, and **concluding the proof.**

It has to be noted, that the evaluation of the factor (γ_{sec}) 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 a converging approximation of the effective number of primes and this fact justifies the evaluation procedure of the factor (γ_{sec}) and proofs the validity of its value.

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 is 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: ($P := \mathrm{READPRN}(\text{"Primes\_large.prn"})$). The number of primes in the set and their numbering are: ($N_P := \mathrm{rows}(P) - 1 = 5003713$, $n := 1, 2 .. N_P$).

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

$$P_{(S(c))} < c < P_{(S(c)+1)} \quad ; \quad P_{(R(c))} < \sqrt{c} < P_{(R(c)+1)} \tag{A0.1}$$

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

For the evaluation of the 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 ($P_{(n)} \le c < P_{(n+1)}$). The evaluation starts either at the last evaluated index (n_{last}), or at the index resulting from the prime-number-formula. This, in order 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:

$$n_{next}(c, n_{last}) := \left| \begin{array}{l} \text{if } c \ne 0 \\ \quad \left| \begin{array}{l} n \leftarrow S(c) \ \text{ if } n_{last} = 1 \\ n \leftarrow n_{last} \ \text{ otherwise} \\ \text{while } P_{(n)} \le c \\ \quad n \leftarrow n + 1 \\ \mathrm{Res} \leftarrow n - 1 \end{array} \right. \\ \mathrm{Res} \leftarrow 0 \ \text{ otherwise} \end{array} \right. \qquad R_{eff}(c) := n_{next}(\sqrt{c}, 1) \qquad S_{eff}(c) := n_{next}(c, 1) \tag{A0.2}$$

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 ($c_{sp(k)} = k \cdot \sqrt{P_{(N_P)}}$):

$$\Delta c_{sp} := \sqrt{P_{(N_P)}} \quad ; \quad k_{limit} := \mathrm{floor}\left[\frac{P_{(N_P)}}{\Delta c_{sp}}\right] - 1 = 9277 \quad ; \quad k := 1, 2 .. k_{limit} \quad ; \quad c_{sp(k)} := k \cdot \Delta c_{sp} \tag{A0.3}$$

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

$$\pi_{sp(k)} := n_{next}\left[c_{sp(k)}, 1\right] \quad ; \quad \pi_{sqr\_sp(k)} := n_{next}\left[\sqrt{c_{sp(k)}}, 1\right] \tag{A0.4}$$

$$\pi_{sqr\_sqr\_sp(k)} := n_{next}\left[\sqrt{\sqrt{c_{sp(k)}}}, 1\right] \quad ; \quad \pi_{sqr\_sqr\_sqr\_sp(k)} := n_{next}\left[\sqrt{\sqrt{\sqrt{c_{sp(k)}}}}, 1\right]$$

$$\mathrm{WRITEPRN}(\text{"index\_distance\_sp.prn"}) := \pi_{.sp} \quad ; \quad \pi_{sp} := \mathrm{READPRN}(\text{"index\_distance\_sp.prn"}) \quad ; \quad \mathrm{length}(\pi_{sp}) = 9278$$

$$\mathrm{WRITEPRN}(\text{"n\_sqr\_sp\_primes.prn"}) := \pi_{sqr\_sp} \quad ; \quad \pi_{sqr\_sp} := \mathrm{READPRN}(\text{"n\_sqr\_sp\_primes.prn"}) \quad ; \quad \mathrm{length}(\pi_{sqr\_sp}) = 9278$$

$$\mathrm{WRITEPRN}(\text{"n\_sqr\_sqr\_sp\_primes.prn"}) := \pi_{sqr\_sqr\_sp} \quad ; \quad \pi_{sqr\_sqr\_sp} := \mathrm{READPRN}(\text{"n\_sqr\_sqr\_sp\_primes.prn"})$$

$$\mathrm{WRITEPRN}(\text{"n\_sqr\_sqr\_sqr\_sp\_primes.prn"}) := \pi_{sqr\_sqr\_sqr\_sp} \quad ;$$

$$\pi_{sqr\_sqr\_sqr\_sp} := \mathrm{READPRN}(\text{"n\_sqr\_sqr\_sqr\_sp\_primes.prn"})$$

ANNEX 1: Evaluation of the approximation

The approximation of the number of primes is with (1.2):

$$\pi_{sec\_appr\_}(c) := \sum_{i=1}^{floor(\sqrt{c})} \frac{\sqrt{c}}{\ln(j\cdot\sqrt{c})} * \qquad (A1.1)$$

With (1.3) the error of this approximation, of the **PNF**, at the distance (c) 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 as follows:

$$\pi_{sec\_appr\_}(c) - S(c) = ; \quad \gamma_{c\_appr\_}(c) = \frac{\pi_{sec\_appr\_}(c) - S(c)}{R(c)^2} \qquad (A1.2)$$

$$\gamma_{c\_appr\_(k)} := \frac{1}{R\left[c_{sp(k)}\right]}\cdot\left[\sum_{j=1}^{ceil\left[\sqrt{c_{sp(k)}}\right]} \frac{\ln\left[\sqrt{\sqrt{c_{sp(k)}}}\right]}{\ln\left[\sqrt{j\cdot\sqrt{c_{sp(k)}}}\right]} - \frac{1}{2}\cdot\sqrt{c_{sp(k)}}\right]$$

They are evaluated once and written to files. They are read from these files:
(WRITEPRN("gamma_c_appr_sp.prn") := $\gamma_{c\_appr\_}$; $\gamma_{c\_appr\_}$:= READPRN("gamma_c_appr_sp.prn")

$$\gamma_{c\_appr} := \gamma_{c\_appr\_}(k_{limit}) = 0.28312$$

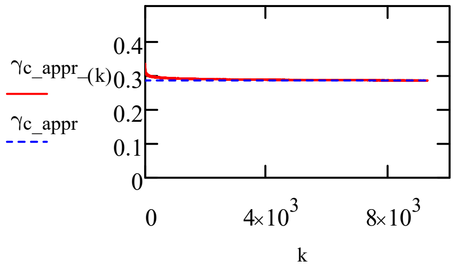

Figure A1.1: Relation of the error of the prime-number-formula to the square of the number of series of multiples of primes, which are covering positions, over the distance

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

$$\pi_{sec\_appr(k)} := \pi_{sec\_appr\_}\left[P_{\left[\pi_{sp(k)}\right]}\right] \qquad \pi_{sqr\_sec\_appr(k)} := \pi_{sec\_appr\_}\left[P_{\left[\pi_{sqr\_sp(k)}\right]}\right] \qquad (A1.3)$$

$$\pi_{sqr\_sqr\_sec\_appr(k)} := \pi_{sec\_appr\_}\left[P_{\left[\pi_{sqr\_sqr\_sp(k)}\right]}\right]$$

They are evaluated once and written to files. They are read from these files:
(WRITEPRN("pr_sec_appr_sp_t.prn") := π_{sec_appr} ; π_{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 approximating formula

Assuming the factor of correction (γ_{sec}) is constant over the distance (c), it may be evaluated as relation of the average of the dispersion of the effective number of primes ($\pi(c) = \pi_{sp(k)}$) around the value of the approximation (1.2), meaning the sum over all sections ($\pi_{sec\_appr\_}(c)$). The average is:

$$\Delta\pi_{sec(k)} := \pi_{sec\_appr(k)} - \pi_{sp(k)} \; ; \; \Delta\pi_{sec\_av(k)} := \frac{1}{k}\cdot\sum_{j=1}^{k}\left[\Delta\pi_{sec(j)}\right] \qquad (A2.1)$$

The value of the factor is herewith:

$$\gamma_{sec}(c) = \frac{\Delta\pi_{sec\_av}(c)}{\pi\left(\sqrt{c}\right)} \quad ; \quad \gamma_{sec\_(k)} := \frac{\sqrt{2}\cdot\Delta\pi_{sec\_av(k)}}{\pi_{sqr\_sec\_appr(k)}} \quad ; \quad \gamma_{sec} := \gamma_{sec\_}\left(k_{limit}\right) = -0.035513 \qquad \text{(A2.2)}$$

The figure below shows the independence of the factor (γ_{sec}) from the distance. The averaging process to evaluate the correction is therefore justified. **This factor (γ_{sec}) is invariant, an inherent property of the prime numbers:**

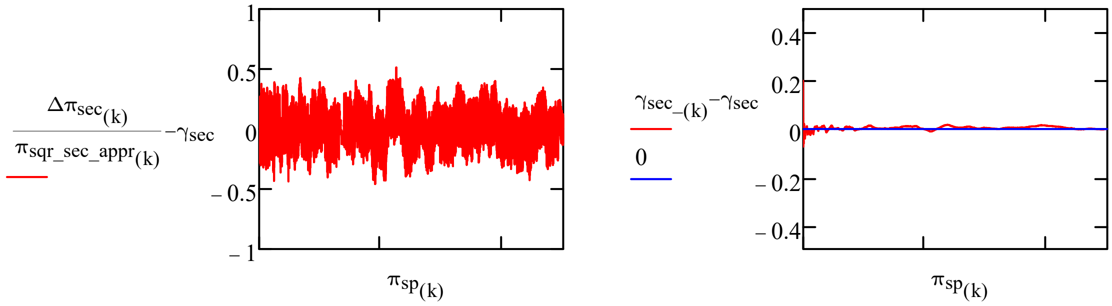

Figure A2.1: Convergence of the relation of the average number of the dispersion of primes to the number of primes at the square root of the distance, to the constant (γ_{sec})

ANNEX A3: Evaluation of the CPNF

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

$$c\_exp(c,m) := c^{\frac{1}{2^{m+1}}} \qquad \pi_{appr\_}(c) := \left| \begin{array}{l} m \leftarrow 1 \\ S \leftarrow \displaystyle\sum_{j=1}^{floor\left(\sqrt{c}\right)} \frac{\sqrt{c}}{\ln\left(j\cdot\sqrt{c}\right)} \\ \text{while } m < \sqrt{c} \\ \quad \left| \begin{array}{l} \Delta_{(m)} \leftarrow \left(\gamma_{sec}\right)^{m} \cdot \displaystyle\sum_{j=1}^{floor(c\_exp(c,m))} \frac{c\_exp(c,m)}{\ln(j\cdot c\_exp(c,m))} \\ \text{break if } \Delta_{(m)} < 2 \\ S \leftarrow S - (-1)^{m}\cdot floor\left[\Delta_{(m)}\right] \\ m \leftarrow m + 1 \end{array} \right. \\ S \end{array} \right. \qquad \text{(A3.1)}$$

The results of the **CPNF** are evaluated once at sparse values of the distance ($c_{sp(k)}$), for (c), ($\sqrt{c}$) and ($\sqrt{\sqrt{c}}$).

The results are written to files. They are read from these files:

$$\pi_{appr(k)} := \pi_{appr\_}\left[P\left[\pi_{sp(k)}\right]\right] \qquad \pi_{sqr\_appr(k)} := \pi_{appr\_}\left[P\left[\pi_{sqr\_sp(k)}\right]\right] \qquad \pi_{sqr\_sqr\_appr(k)} := \pi_{appr\_}\left[P\left[\pi_{sqr\_sqr\_sp(k)}\right]\right] \qquad \text{(A3.2)}$$

(WRITEPRN("Pi_appr_sp.prn") := π_{appr}), (π_{appr} := READPRN("Pi_appr_sp.prn")),
(WRITEPRN("Pi_sqr_appr_sp.prn") := π_{sqr_appr}), (π_{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")).

There is no systematic error of the approximation of the number of primes with the CPNF. The relative error, the relative dispersion of the evaluated value around the effective number of primes at ($\sqrt{c}$) is constant:

$$\Delta\pi_{appr\_rel(k)} := \frac{\pi_{appr(k)} - \pi_{sp(k)}}{R\left[1\cdot c_{sp(k)}\right]} \quad ; \quad \Delta\pi_{sqr\_appr\_rel(k)} := \frac{\pi_{sqr\_appr(k)} - \pi_{sqr\_sp(k)}}{R\left[\sqrt{1\cdot c_{sp(k)}}\right]} \qquad \text{(A3.3)}$$

$$\Delta\pi_{sqr\_sqr\_appr\_rel(k)} := \frac{\pi_{sqr\_sqr\_appr(k)} - \pi_{sqr\_sqr\_sp(k)}}{R\left[\sqrt{\sqrt{1\cdot c_{sp(k)}}}\right]}$$

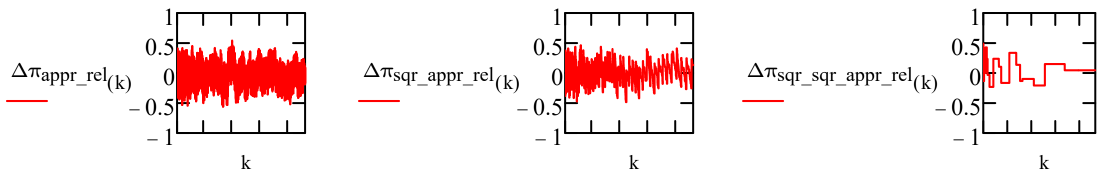

Figure A3.1: The relative dispersion of the difference between the effective number of primes and its value evaluated with the complete-prime-number-formula (CPNF)

With the results of the **CPNF** the factor of the proportionality (γ_{c_appr}) between the error of the **PNF** and 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:

$$\pi_{appr\_cpnf\_}(c) - S(c) = ; \quad \gamma_{cpnf\_}(k) := \frac{\pi_{appr}(k) - S\left[c_{sp}(k)\right]}{R\left[c_{sp}(k)\right]^2} \qquad \text{(A3.4)}$$

$\gamma_{cpnf} := \gamma_{cpnf\_}(k_{limit}) = 0.28279$ $\qquad$ $\gamma_c = 0.28298$

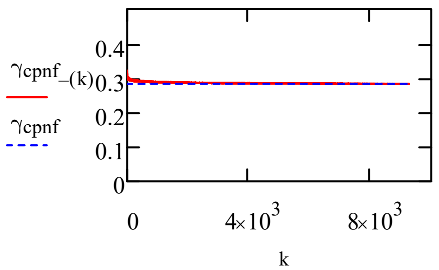

Figure A3.2: Relation of the error of the PNF to the square of the number of series of multiples of primes ($R(c)^2$), which are covering positions, over the distance

ANNEX A4: Evaluation of the of the standard deviation of the dispersion of the number of primes around the CPNF

The standard deviation SD of the relative dispersion (1.13) is evaluated as follows:

$$SD\Delta\pi\_appr\_rel(k) := \sqrt{\frac{1}{k} \cdot \sum_{j=1}^{k} \frac{\left[\pi_{appr}(j) - \pi_{sp}(j)\right]^2}{R_{eff}\left[c_{sp}(j)\right]^2}} \qquad \text{(A4.1)}$$

The results are evaluated once and written to a file: (WRITEPRN("SD_Δprime_sqr_appr_rel.prn") := $SD\Delta\pi\_appr\_rel$).
They are read from this file: ($SD\Delta\pi\_appr\_rel$:= READPRN("SD_Δprime_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:

$$SD\Delta\pi\_appr\_rel\_av(k) := \frac{1}{k} \cdot \sum_{j=1}^{k} SD\Delta\pi\_appr\_rel(j) \qquad \text{(A4.2)}$$

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

WRITEPRN("SD_Δprime_sqr_appr_rel_av.prn") := $SD\Delta\pi\_appr\_rel\_av$

$SD\Delta\pi\_appr\_rel\_av$:= READPRN("SD_Δprime_sqr_appr_rel_av.prn")

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:

$$F_{SD\_\Delta\pi} := SD\Delta\pi\_appr\_rel\_av(k_{limit-1}) = 0.162389 \qquad SD\pi(c) = SD\Delta\pi\_sp\_rel\_av(k) \cdot R(c) \approx F_{SD\_\Delta\pi} \qquad \text{(A4.2)}$$

The figures below indicate, 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 prime numbers.**

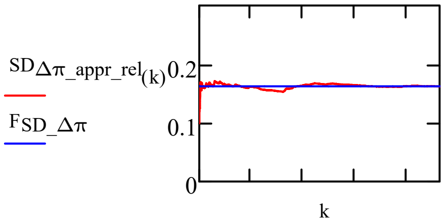

Figure A4.1: Dispersion of the standard deviation of the dispersion of the number of primes around its average, a constant value ($F_{SD\_\Delta\pi\_sec}$)

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 ($R(c) + 1$, $c = P_{(n+1)}$):

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

The difference makes ($\sqrt{c} \cdot \sum_{j=1}^{R(c)+1} \frac{1}{\ln(j \cdot \sqrt{c})} - \sqrt{c} \cdot \sum_{j=1}^{R(c)} \frac{1}{\ln(j \cdot \sqrt{c})} \approx \frac{\sqrt{c}}{\ln(c)} = \frac{\sqrt{c}}{2 \cdot \ln(\sqrt{c})} = \frac{R(c)}{2}$). (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 ($c = P_{(q)}$, $\pi_{appr\_test(q)} = \pi_{appr\_}\left[P_{(q)}\right]$) in the range , ($q = q_{low} .. q_{high}$):

from ($q_{low} := 1909500$, $P_{(q_{low})} = 30887083$) to ($q_{high} := 1911600$, $P_{(q_{high})} = 30923341$, $q := q_{low} .. q_{high}$)

$$\pi_{appr\_test(q)} := \pi_{sec\_appr\_}\left[P_{(q)}\right] \qquad \text{(A5.3)}$$

The results are written to a file: (WRITEPRN("Nprime_appr_c.prn") := π_{appr_test}). They are read from these files each time the present paper is evaluated: (π_{appr_test} := READPRN("Nprime_appr_c.prn"). The discontinuity of the approximating function is shown in the next figure.

The size of the jumps (2.2) is the bandwidth of the dispersion due to the approximating function:

$$\pi_{bounds\_appr}(c) = \pm \frac{R(c)}{2}. \qquad \text{(A5.4)}$$

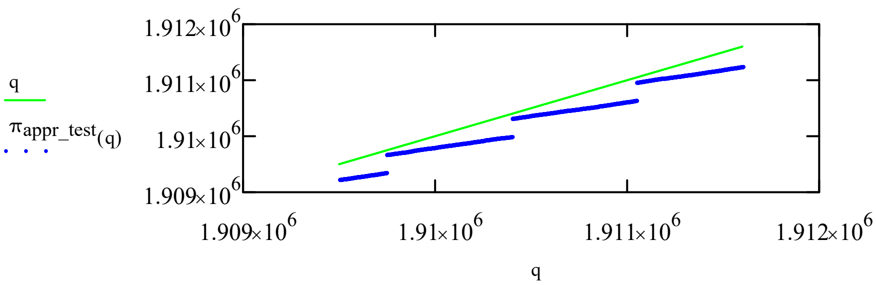

Figure A5.1: Discontinuity due to the approximating function