



## Computational Model of Prime Numbers by the Modified Chi-square Function

Daniele Lattanzi<sup>1\*</sup>

<sup>1</sup>Frascati Nuclear Centre, Via E. Fermi, Frascati, Roma, Italy.

### Author's contribution

The sole author designed, analyzed and interpreted and prepared the manuscript.

### Article Information

DOI: 10.9734/BJMCS/2017/31589

#### Editor(s):

(1) Francisco Welington de Sousa Lima, Dietrich Stauffer Laboratory for Computational Physics, Departamento de Física, Universidade Federal do Piauí, Teresina, Brazil.

#### Reviewers:

- (1) G. Y. Sheu, Chang-Jung Christian University, Tainan, Taiwan.
- (2) Luis Angel Gutierrez-Mendez, Emeritus Autonomous University of Puebla, Mexico.
- (3) Hammad Khalil, University of Education, Lahore, Pakistan.
- (4) Bilge Peker, Necmettin Erbakan University, Turkey.

Complete Peer review History: <http://www.sciencedomain.org/review-history/17942>

Received: 15<sup>th</sup> January 2017

Accepted: 15<sup>th</sup> February 2017

Published: 23<sup>rd</sup> February 2017

Original Research Article

## Abstract

An innovative approach that treats prime numbers as raw experimental data and as elements of larger and larger finite sequences  $\{P_m\} \equiv \{P(m_p)\}$  is shown in the present report. The modified chi-square function  $X_k^2(A, m_p/x_0)$  with its three parameters  $A$ ,  $k$  and  $x_0 = x_0(k)$  is the best-fit function of the finite sequences  $\{\rho_m\} \equiv \{\lg P_m / \lg m_p\}$  from the analytical viewpoint thus showing that the property of scale invariance does not hold for the finite sequences of this prime variable and so for primes themselves. In addition an injective map can be set between these  $\{\rho_m\}$  sequences and the  $\{m^a\}$  progressions with domain  $N$  and co-domain  $R^+$  being  $a \in (-1, 0) \subset R^-$  through the parameter  $k = 2 + 2a$  of their common fit function  $X_k^2(A, m_p/x_0)$ . All that leads to induction algorithms and to relationships of the kind  $P_m \approx P(m_p)$ , though within the precisions of the calculations and holding locally.

**Keywords:** Prime number sequences; modified chi-square function; numerical progressions; computational mathematics.

\*Corresponding author: E-mail: [lattanzio.lattanzi@alice.it](mailto:lattanzio.lattanzi@alice.it);

## 1 Introduction

The problem of prime numbers in number theory has always been a challenge to face and still nowadays it remains one of the major open problems notwithstanding the many theoretical successes achieved both historically and recently [1-13]. The main problem concerns the fact that, unlike all the numerical progressions, neither an exact relationship that links the value of a prime  $P_m$  to its counter  $m_p$  i.e.  $P_m = P(m_p)$  has yet been found at present (or simply it does not exist) nor there is an analytical law that links any prime number  $P_m$  to its preceding  $P_{m-1}$ . In other words it is not possible at present to state that the induction principle holds for prime numbers. Moreover, there are strong doubts about the nature itself of prime numbers, whether deterministic or stochastic, in that prime numbers on one hand seem to obey to firm laws on the other they seem to appear in a random way one after another. Thus the present article is aimed at introducing a new viewpoint on prime numbers which makes use of an innovative methodology, never adopted before now, that can describe prime numbers and that, being at its initial stage, can lead, after due refinements, to more precise calculations of their values thus maybe saving computer memory. Besides, this methodology might lead in the future to further unexpected results and findings in addition to that already found of the scale non-invariance of the finite sequences of prime numbers.

In previous articles and reports by the same author [14-17] the finite sequences of prime numbers  $\{P_m\}$  have been examined from both the statistical viewpoint and the analytical one fitting their differential distribution functions and the finite sequences of their frequencies  $\{f_m\} \equiv \{m_p/P_m\}$  by the modified chi-square function  $X_k^2(A, m_p/x_o)$  thus finding remarkable unexpected results among which the scale non-invariance [18,19] of the finite sequences of primes, their scaling laws and their correspondence with the finite progressions  $\{m^\alpha\}$ .

In addition, the same methodology as applied to the truncated progressions  $\{n^{-\alpha}\}$  and their summations  $\{\sum n^{-\alpha}\}$  has led to find an elementary (that is not using the theory of complex functions) and general (that is valid for all the non-trivial zeroes of the  $\zeta(s)$  function up to  $\infty$ ) proof of Riemann's hypothesis [20,21] by means of the modified chi-square function.

In the present report this same innovative approach is suggested again, starting from the computational viewpoint [22-26] and making use of the modified chi-square function with  $k$  degrees of freedom

$$X_k^2[A, m_p/x_o(k)] = A/(2 \cdot \Gamma_{k/2}) \cdot [m_p/2x_o(k)]^{(k/2-1)} \cdot e^{-m_p/2x_o(k)} \quad (1)$$

with  $k \in (1.0, 2.0) \subset \mathbb{R}^+$  as the best-fit function along the whole study to match the finite sequences  $\{\rho_m\} \equiv \{\lg P_m / \lg m_p\}$  from the analytical viewpoint, that is fitting/interpolating the actual data points  $\rho_m = \rho(m_p)$  by the analytic function (1).

The aim is to construct a computational model of the finite sequences  $\{\rho_m\} \equiv \{\rho(m_p)\} \equiv \{\lg P_m / \lg m_p\}$ .

The function (1) is used as the best-fit function along the whole study to match the finite sequences  $\{\rho_m\}$  and the truncated progressions  $\{m^\alpha\}$  having domain  $\mathbb{N}$  and co-domain  $\mathbb{R}^+$  with  $\alpha \in (-1, 0) \subset \mathbb{R}^-$  which can also be written as  $\{m^{-\alpha}\}$  with  $-\alpha \in (-1, 0) \subset \mathbb{R}^-$ . The rationale underlying the entire matter has been to use this function taking advantage of the adjustment of its three parameters  $k$ ,  $A$ , and  $x_o(k)$  which allow to optimize the fits as much as possible up to 99% and even more whenever possible. In other words a *plot&fit* algorithm has been set up.

Furthermore the inverse  $1/X_k^2[A, m_p/x_o(k)]$  of this function (1) can fit both the progressions  $\{m^{+\alpha}\}$  with  $\alpha \in (0, +1) \subset \mathbb{R}^+$  and the prime sequences  $\{\rho_m^{-1}\} \equiv \{\lg m_p / \lg P_m\}$  with the values of its parameters  $A$ ,  $k$  and  $x_o = x_o(k)$  equal to those of the corresponding  $\{\rho_m\}$  sequences and  $\{m^{-\alpha}\}$  progressions fitted by  $X_k^2[A, m_p/x_o(k)]$  what proves all its flexibility and efficacy. This is not a trivial concern, as it might look, but an important topic of the research as shown later on.

The accuracy and precision (i.e. uncertainties), random and systematic errors, error sources, error propagations and reliability of the results have been investigated, being these issues crucial to the whole algorithm as explained in detail in the Methods section.

After all, what has been done is just what is usually done in treating experimental raw data [27-29], a procedure that is common to all the fields of experimental physics. The only difference has been to treat the data points  $\rho_m = \lg P_m / \lg m_p$  just like a set of experimental data, in a broad sense, to which all these concepts and criteria can be applied, with the further undisputable advantage of having zero inaccuracy (i.e. no systematic errors) and zero imprecision (no random errors) on the base data, whereas zero inaccuracy though not zero imprecision (owing to the approximations of the calculations and of the fits) are present in the final results. In such a manner a computer simulation model has been set up and applied to the truncated numerical progressions  $\{m^{-\alpha}\}$  and the finite prime sequences  $\{\rho_m\}$ . Some websites have been used to get the values of the  $\Gamma$  function [30,31].

However, while for any finite progression  $\{m^{\pm\alpha}\}$  the results in no way depend on the number of the terms (apart from the usual improvements of the fit values, an evidence of the improvement of the statistics) what is a consequence of the scale invariance of any single progression, in the case of prime numbers larger and larger finite sets of them have been taken into account one at a time. The reason for such an unconventional choice is a strict consequence of the scale non-invariance of primes and an evidence of the existence of scaling laws holding for them, as shown later on.

## 2 Methods

The modified chi-square function (1) with  $k$  degrees of freedom has been already examined and described in the previous works by the same author with all its features; here it is enough to remark that in the plane  $(x, X^2)$  the coefficient  $A$  (which usually is at one's own choice in all the calculations) shifts rigidly the corresponding curve up or down, the decay parameter  $x_0$  stretches or compresses it along the  $x$  axis and the  $k$  parameter (the number of degrees of freedom) determines the shape of the curve (more or less rounded).

The entire methodology is applied to the analysis of the prime number sequences  $\{\rho_m\}$  and of the finite progressions  $\{m^\alpha\}$  with  $m \in \mathbb{N}$  and  $\alpha \in (-1, 0) \subset \mathbb{R}^-$  getting remarkable results. Larger and larger finite sequences of prime numbers, subsets of their whole infinite sequence, have been examined that is sequences of the kind:  $\{2 \ 3 \ 5 \ 7 \ 11 \ \dots \ P_{h-1} \ P_h\} \equiv \{P_h\} \subset \{P_i\} \subset \{P_j\} \subset \dots \subset \{P_n\} \subset \dots$  being of course  $h < i < j < \dots < n < \dots$  and afterwards the finite sequences  $\{\rho_m\} \equiv \{\lg P_m / \lg m_p\}$  that is  $\{\rho_h\} \subset \{\rho_i\} \subset \{\rho_j\} \subset \dots \subset \{\rho_n\} \subset \dots$  being  $P_m = P(m_p) \approx m_p^{\rho(m_p)}$  have been investigated. Apart from the usual improvements of the statistical values, the reason for such an unconventional choice is a strict consequence of the scale non-invariance of prime finite sequences and of scaling laws holding for them as shown later on. Of course, just few (typically  $n_\Delta = 200$ ) equally spaced data points have been examined for any  $\{\rho_m\}$  sequence, being  $m_p = n_\Delta \cdot \Delta = 200 \cdot \Delta$  and having verified that this choice does not affect the reliability of the results.

Also any  $\{m^\alpha\}$  progression with a finite number of terms has been examined at  $n_\Delta$  (again typically 200) equally spaced data points and it has been fitted, at  $n_\Delta$  data points, by the modified chi-square function with the appropriate values of the parameters  $k$ ,  $A$  and of the decay parameter  $x_0 = x_0(k)$ . Speaking in a rigorous and formal way any  $\{m^\alpha\}$  progression can be analytically continued from the  $\mathbb{N}$  domain to the  $\mathbb{R}$  domain, that is to the function  $f(x) = x^\alpha$  and also to the function  $F(x) = X_k^2[A, m_p/x_0(k)]$ . Both functions are analytic on the whole  $\mathbb{R}^+$  plane.

Maximizing the two statistical markers,  $R$  and  $I$ , means making both of them to approach the value of  $I^-$  as much as possible by adjusting the value of  $x_0$  for any value of  $k$  in order to match the data points  $C = \rho_m$  and the fitting curve  $F = X_k^2[A, m_p/x_0(k)]$  as much as possible, as well as to balance their average values  $\langle C \rangle = \langle \rho_m \rangle = \langle F \rangle = \langle X_k^2[A, m_p/x_0(k)] \rangle$  up to the 12<sup>th</sup> decimal digit, as already said. In addition, even the two standard deviations of the means  $\sigma_{\text{counts}} = \sigma_c$  and  $\sigma_{\text{fit}} = \sigma_F$  have been examined in order to ascertain that each of them would be much lower than its respective mean  $\langle C \rangle$  and  $\langle F \rangle$  and that  $\sigma_c \approx \sigma_F$ . Finally, two further gauges of the fit have been minimizing the values of the Least Square Sum (LSS) according to the principle of maximum likelihood and of the  $X_{\text{test-value}}^2$  in that both of these variables measure the goodness of the fit. Just to summarize, many statistical tools have been used in order to do the best fits at the utmost statistical

reliability. Though complex and time-consuming, this fitting procedure has proven all its effectiveness and reliability in finding out not only the best values of  $k$  and of the decay parameter  $x_0$  for any data set but even the two fundamental relationships  $k=k(\alpha)=k_\alpha$  and  $k=k(m_p)=k_p$  for the progressions  $\{m^\alpha\}$  and the sequences  $\{\rho_m\}$  respectively, as well as  $x_0=x_0(\alpha)$  and  $x_0=x_0(m_p)$ , showing that this basic parameter  $k$  depends only on  $\alpha$  and  $m_p$  respectively while the  $x_0$  decay (or growth in other cases) parameter can be regarded just as a local parameter.

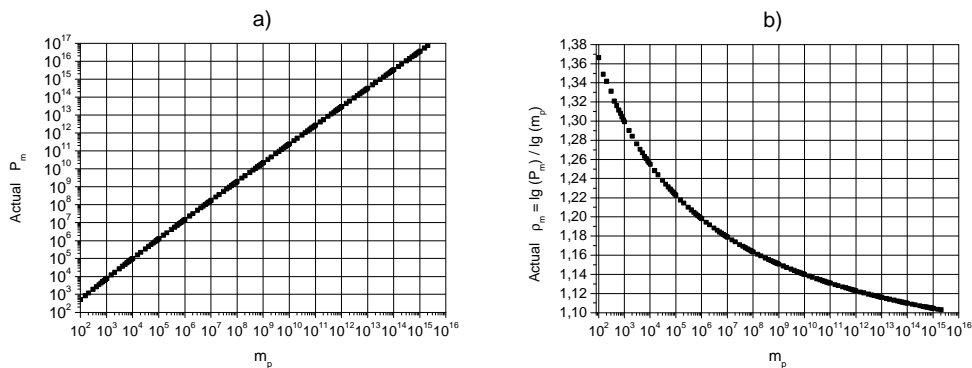
Hence for most cases it is possible, by this methodology, to reduce the problem of prime numbers to a problem of precision of the calculations, though still hard and not yet fully solved. However, in comparison to other approximate methods of prime number evaluation, the present one has the advantage of showing a wide spectrum of possibilities, about six or seven or more, among which it is possible to choose according to the best results. In addition, using a mainframe and acting on the initial fits, there is the possibility to increase the precisions of the final results, i.e. of the  $P_m$  values got in such a manner, as much as possible.

One of the backgrounds of the study has been the use of the principle of extrapolation and interpolation of the data points by means of the fits that has been widely used all over the report whenever required and possible even though not explicitly cited, a principle widely applied in all experimental science.

### 3 Analytical Treatments

The attention has been focused on the prime number finite sequences  $\{\rho_m\} \equiv \{\rho(m_p)\} \equiv \{\lg P_m / \lg m_p\}$  and on the  $\{m^\alpha\}$  truncated progressions having  $m \in \mathbb{N}$  and  $\alpha \in (-1, 0) \subset \mathbb{R}^-$  also written as  $\{m^{-\alpha}\}$  with  $-\alpha \in (-1, 0) \subset \mathbb{R}^-$  with the purpose of fitting them by the modified chi-square function and linking the former to the latter from the analytical viewpoint. The analytical aspects observed in the frame of computational mathematics represent an attempt to reach an algorithm of the kind  $P_m \approx P(m_p)$  for the construction of a prime number starting from its own counter  $m_p$  by the present method, though approximate and valid locally.

The plot of the actual values of  $\{\rho_m\}$  for the first  $m_p \approx 100-200$  primes (not shown) displays wide damped oscillations or fluctuations which tend to disappear completely henceforward that is exhibiting a much more regular trend at higher values as in the two plots of Fig. 1 where the actual  $P_m$  values vs.  $m_p$  (a) and the actual  $\rho_m$  vs.  $m_p$  (b) are shown for all the prime numbers studied from  $m_p=100$  up to  $m_p=2E15$  [32,33]. It is just this regularity that leads to examine the trend of  $\rho_m$  versus  $m_p$  starting from the alleged relationship  $P_m = P(m_p) \approx m_p^\rho$  being  $\rho = \rho_m = \rho(m_p) \approx \lg(P_m) / \lg(m_p)$ .

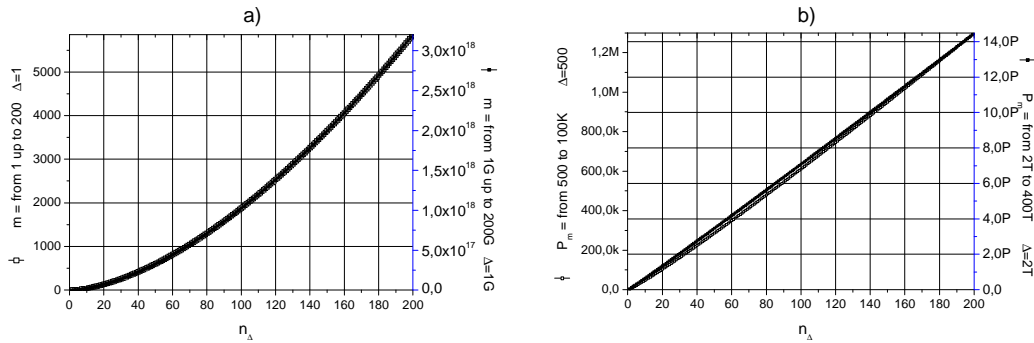


**Fig. 1. The actual values of prime numbers  $P_m$  (a) and of  $\rho_m=\rho(m_p)$  (b) vs.  $m_p$  from 100 up to  $2 \cdot 10^{15}$**

Looking at these two plots of data points it would seem that each of them is fitted by just one single curve (whatever it might be) i.e. that each of these two data sets may be interpolated by just one single analytic

function. It is not so. As a matter of fact one can look at the two plots of the next Fig. 2 showing the huge difference between the two cases of:

- a) The progression  $\{m^\alpha\} \equiv \{m^{1.6357}\}$  ( $\alpha=1.6357$  is a randomly chosen value) with  $n_\Delta=200$  for  $m=500 \rightarrow 100,000$  and for  $m=2T=2E12 \rightarrow 400T=400E12$ ;
- b) The two prime number finite sequences ( $n_\Delta=200$ ) for the sequence  $m_p=500 \rightarrow 100K$  that is  $P_m=3,571 \rightarrow 1,299,709$  and for the sequence  $m_p=2T=2 \cdot 10^{12} \rightarrow 400T=4 \cdot 10^{14}$  that is  $P_m = 61,427,839,512,229 \rightarrow 14,472,680,634,646,931$  [34].



**Fig. 2. The actual values of the progression  $\{m^\alpha\} \equiv \{m^{1.6357}\}$  (a) and of prime numbers  $\{P_m\}$  (b) both for the two cases  $m=500 \rightarrow 100K$  and  $m=2 \cdot 10^{12} \rightarrow 400T$**

While in the first plot (a) there is no difference at all between the two sequences of data points, apart from the scale of course, and their perfect superposition can be easily checked despite the huge difference of their values, in the second plot (b) there is a clear evidence of the difference between the two data point sequences of primes. This is the proof of the scale non-invariance of prime numbers, a property already found in the previous works by the same author. Such a property appears still more evident in examining the finite sequences of the prime variable  $\{\rho_m\} \equiv \{\rho(m_p)\} \equiv \{\lg(P_m)/\lg(m_p)\}$  as shown in the next Fig. 3 which reports some sequences of  $\{\rho(m_p)\} \equiv \{\rho_m\}$  vs. the ordinal number that is the counter of the data points from 1 up to  $n_\Delta=200$  the number of the cells of the spread-sheet used for all the calculations. Thus, otherwise from any progression  $\{m^\alpha\}$  which can be fitted by just one analytic function independently from the number of its terms, that is  $f(x)=x^\alpha$  of course and also  $F(x)=X_k^2[A, x/x_0(k)]$  as shown later on (as a matter of fact  $\lg(m^\alpha)/\lg m = \alpha$ ), in the case of prime numbers any finite sequence  $\{\rho_m\} \equiv \{\lg(P_m)/\lg(m_p)\}$  has to be fitted by a different analytic function, that is  $X_k^2[A, m_p/x_0(k)]$  with the ad-hoc values of  $k$  and  $x_0$  being the value of  $A$  at one's own choice and having chosen  $A=1$  once for all in the present study.

Thus, again the same algorithm as already used in previous cases [14-17] has been applied in the current case, i.e. choosing larger and larger sets of  $\{\rho_h\} \subset \{\rho_i\} \subset \{\rho_j\} \subset \dots \subset \{\rho_1\} \subset \dots \subset \dots \subset \{\rho_m\} \dots \subset \{\rho_{2P}\}$  with  $\{\rho_{2P}\} \equiv \{\rho_1 \rho_2 \rho_3 \dots \rho_h \dots \rho_i \dots \rho_j \dots \rho_l \dots \rho_m \dots \rho_{2P}\}$ , fitting the relationship  $\rho_m = \rho(m_p)$  and the associated plot vs.  $m_p$  for any of them by the modified chi squared function  $X_k^2[A, m_p/x_0(k)]$  where again  $k < 2^-$ . Of course, as already told, for any sequence  $\{\rho_m\} \equiv \{\rho_\Delta \rho_{2\Delta} \rho_{3\Delta} \rho_{4\Delta} \dots \rho_{200\Delta}\}$  being  $m = n_\Delta \cdot \Delta = 200 \cdot \Delta \quad \forall m$ .

In addition, also the finite progressions  $\{m^{-\alpha}\}$  with  $-\alpha \in (-1, 0) \subset \mathbb{R}^- \quad m = m_p \in \mathbb{N} \quad k = 2 + 2 \cdot \alpha$  have been identified as well-fitting the  $\{\rho_m\}$  sequences though with a coefficient  $C_\alpha = C(\alpha)$  so that

$$\{\rho_m\} \approx X_k^2[A, m_p/x_0(k)] \approx C_\alpha \cdot \{m^{-\alpha}\} \equiv \{C_\alpha \cdot m^{-\alpha}\}$$

and the previous scheme of Fig. 4 can be highlighted where  $-\alpha \in (-1, 0) \subset \mathbb{R}^-$   $m = m_p \in \mathbb{N}$   $k = 2 + 2 \cdot \alpha$  and the single arrows reveal the one-way correspondence, that is the injection map, while the double one reveals the one-by-one correspondence, i.e. the bijection map.

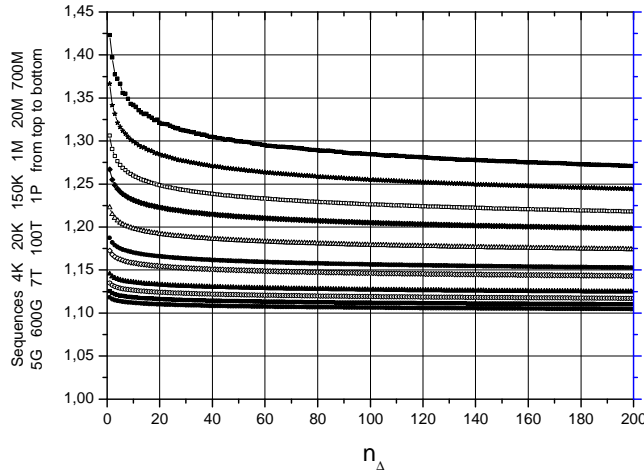


Fig. 3. Some sequences of the actual prime  $p$  values  $\{\rho_m\} \equiv \{\lg(P_m)/\lg(m_p)\}$  vs. the ordinal counter  $n_A$

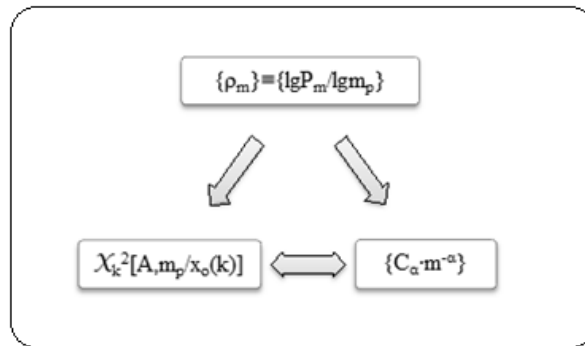
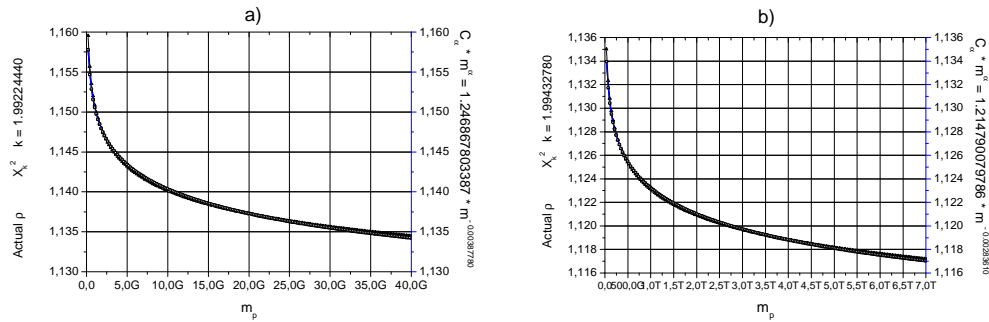


Fig. 4. Fits of the sequences  $\{\rho_m\} \equiv \{\lg(P_m)/\lg(m_p)\}$  with  $k = 2 + 2\alpha$

Thus the aim of the whole research is to fit, that is to interpolate and locally extrapolate, any of the 63 sequences of  $\{\rho_m\}$  examined (this number has been considered adequate to plot final reliable fits) by the analytic functions identified as the modified chi-square function  $X_k^2[A, m_p/x_o(k)]$  and  $x^{-\alpha}$ .

Some examples are reported and discussed in the following pages and in Fig. 5.

The first example concerns the first 40G values of  $\{\rho_m\}$  and it is shown in Fig. 5a where the parameters of the fit by the modified chi-square are  $A=1$   $x_o=1.95028704889E+102$   $k=1.9922444$   $\Gamma_{k/2}=1.00225325280911$   $R=0.998706$   $I=0.997413$   $\langle C \rangle = \langle F \rangle = 1.138583181049570$   $\sigma_c=0.004232$   $\sigma_F=0.004227$   $LSS=0.257949$   $X^2_{\text{test-value}}=8.02996E-6$  while the linked parameters for the fit by the progression  $\{C_\alpha \cdot m^\alpha\} \equiv \{1.246867803387 \cdot m^{-0.0038778}\}$  are:  $R=0.998706$   $I=0.997413$   $LSS=0.257949$   $X^2_{\text{test-value}}=8.024956E-6$ .

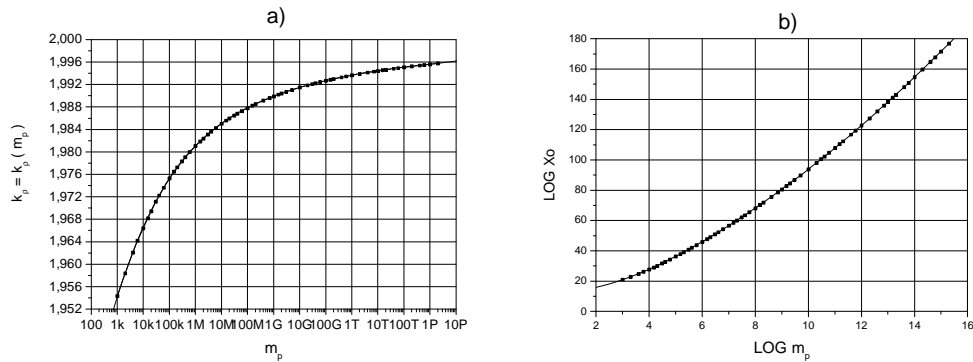


**Fig. 5. Fits of the first  $\{\rho_{40G}\}$  (a) and of the first  $\{\rho_{7T}\}$  (b) by the  $X_k^2[1, m_p/x_0(k)]$  and by  $\{C_a \cdot m^\alpha\}$**

The same Fig. 5b shows the first 7T values of  $\{\rho_m\}$  with the parameters of the fit by the modified chi-square function are  $A=1$   $x_0=7.72278352265E+135$   $k=1.99432780$   $\Gamma_{k/2}=1.0016450175479$   $R=0.999107$   $I=0.998214$   $\langle C \rangle = \langle F \rangle = 1.120166459749330$   $\sigma_c=0.003042$   $\sigma_F=0.003039$   $LSS=0.177957$   $X^2_{test-value}=2.91714E-6$  while the same parameters for the fit by the  $\{C_a \cdot m^\alpha\}$  function are:  $C_a=1.214790079786$   $\alpha= -0.002836108$   $R=0.999107$   $I=0.998214$   $LSS=0.177957$   $X^2_{test-value}=2.916049E-6$ .

It is interesting to remark, just in these few examples, that the fit parameters tend to improve in choosing larger and larger sequences of  $\{\rho_m\}$  what is a clear evidence of the improving of the fits, of course, and of the goodness of the method.

In such a manner many (i.e. 63) sets of values of the parameters  $k$  and  $x_0$  have been gathered, anyone for any set of  $\{\rho(m_p)\}$ , which can be plotted as reported in Fig. 6 showing the trends of  $k=k(m_p)$  (a) and of  $\lg x_0 = \lg x_0(\lg m_p)$  (b) for the 63 sequences of  $\{\rho_m\}$  examined.



**Fig. 6. Data points and fits of  $k$  (a) and of the decay parameter  $x_0$  (b) of  $X_k^2[1, m_p/x_0(k)] \approx \{\rho_m\}$**

It is remarkable to highlight that both trends  $k=k(m_p)$  and  $x_0=x_0(m_p)$  resemble those previously found in other cases [14-17] with the same limit values

$$\begin{aligned} \text{a)} \quad & \lim_{m_p \rightarrow +\infty} k(m_p) = 2^- \\ \text{a)} \quad & \lim_{m_p \rightarrow +\infty} [\Delta k(m_p)/\Delta m_p] = 0^+ \end{aligned} \qquad \begin{aligned} \text{b)} \quad & \lim_{m_p \rightarrow +\infty} x_0(m_p) = \infty \\ \text{b)} \quad & \lim_{m_p \rightarrow +\infty} [\Delta x_0(m_p)/\Delta m_p] = \infty \end{aligned}$$

The best fit of the 63 data points gives the following relationship of  $k$  vs.  $m_p$  (Fig. 9a):

$$k = k(m_p) \approx 1.87715 + 0.03982 \cdot m_p - 0.00598 \cdot m_p^2 + 4.77556E-4 \cdot m_p^3 - 1.95679E-5 \cdot m_p^4 + 3.23842E-7 \cdot m_p^5$$

with  $R^2 = 0.99997$   $\sigma = 6.000E-5$   $N = 63$   $p < 1E-4$

As for the decay parameter  $x_0 = x_0(m_p)$  (Fig. 9b) it is possible to see that the following relationship holds on a log-log scale:

$$\log(x_0) = \log[x_0(m_p)] \approx 14.13218 - 2.49815 \cdot (\log m_p) + 1.99429 \cdot (\log m_p)^2 - 0.15574 \cdot (\log m_p)^3 + 0.00758 \cdot (\log m_p)^4 - 1.47285E-4 \cdot (\log m_p)^5$$

where  $R^2 = 1.000000000000$   $\sigma = 0.06784$   $N = 63$   $p < 1E-4$

where nonetheless it should be kept into account that the six coefficients depend again on  $m_p$  what means that both for  $k$  and for  $x_0$  scaling laws hold. Of course, in both the relationships the coefficients are affected by uncertainties which are approximately between some ‰ and some %.

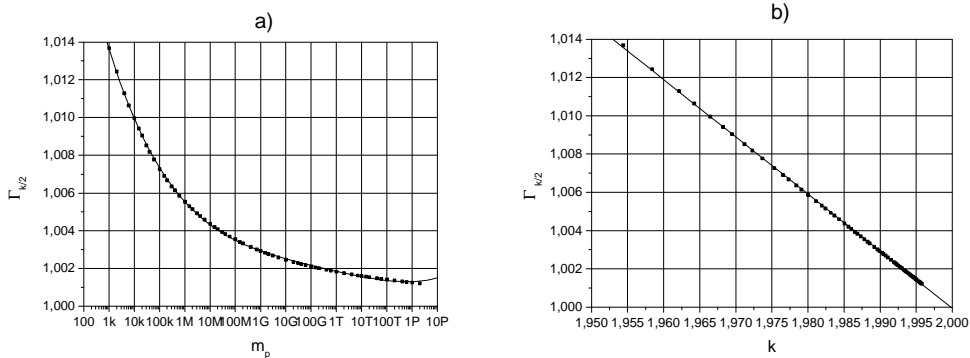
The variable  $\Gamma_{k/2}$  is highly dependent on  $k$ , its law being well-known. However in the present context, being the plots of the data points as reported in the next two graphs of Fig. 7a and 7b, it can be approximated that is fitted by the local (i.e. holding for  $1 < \Gamma_{k/2} < 1.016$ ) relation of  $\Gamma_{k/2}$  vs.  $m_p$ .

$$\Gamma_{k/2}(m_p) \approx 1.03412 - 0.00992 \cdot m_p + 0.0123 \cdot m_p^2 - 7.03545E-5 \cdot m_p^3 + 1.53074E-6 \cdot m_p^4$$

with  $R^2 = 0.998$   $\sigma = 4.53E-5$   $N = 63$   $p < 1E-4$  and the uncertainties on the coefficients are about 1‰ or less. The linear regression of  $\Gamma_{k/2}$  vs.  $k$  (Fig. 10b) in the same range leads to the law

$$\begin{aligned} \Gamma_{k/2}(k) &\approx (1.59804 \pm 0.000728) - (0.29906 \pm 0.0003669) \cdot k \approx \\ &\approx (1.59804 \pm 0.455\%) - (0.29906 \pm 1.22\%) \cdot k \end{aligned}$$

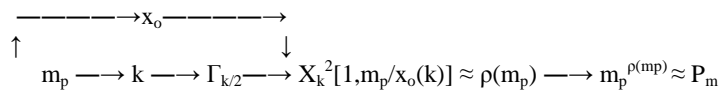
with  $R^2 = 0.99995$   $\sigma = 2.973E-5$   $N = 63$   $p < 1E-4$



**Fig. 7. Data points and fits of  $\Gamma_{k/2}$  of  $X_k^2[1, m_p/x_0(k)]$  for the 63 sequences of  $\{\rho_m\}$  vs.  $m_p$  (a) and vs.  $k$  (b)**

Thus, at least in principle, it would be possible to calculate the value of a prime number  $P_m$  starting from its counter  $m_p$  using the above reported relations and plots and extrapolating and/or interpolating them, apart from precision problems.

What means that the following scheme can be assumed (by means of Figs. 6 and 7):





Using the best fit of the data points to evaluate both  $x_o$  and  $k$  (independently one from each other) from  $m_p$  in order to get the value of the modified chi-square function and thus the value of  $\rho=\rho(m_p)$  and of  $P_m$ . Otherwise, the value of  $k$  can be attained again as  $k=k(m_p)$ , from which the value of  $x_o = x_o(k)$  can be derived, finally getting the value of  $X_k^2[1,m_p/x_o(k)] \approx \rho(m_p)$  leading to  $m_p^{\rho(m_p)} \approx P_m$  according to this alternative scheme

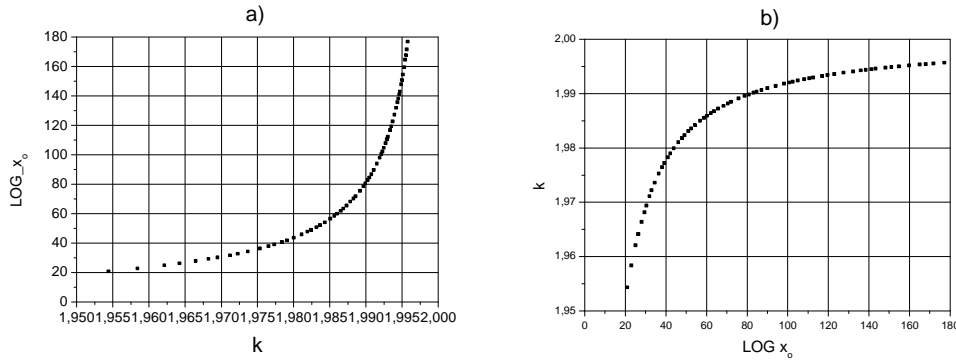
$$m_p \longrightarrow k = k(m_p) \longrightarrow x_o = x_o(k) \longrightarrow X_k^2[1,m_p/x_o(k)] \approx \rho(m_p) \longrightarrow m_p^{\rho(m_p)} \approx P_m$$

or even according to this further scheme (just interchanging the roles of  $k$  and  $x_o$ )

$$m_p \longrightarrow x_o = x_o(m_p) \longrightarrow k = k(x_o) \longrightarrow X_k^2[1,m_p/x_o(k)] \approx \rho(m_p) \longrightarrow m_p^{\rho(m_p)} \approx P_m$$

the choice among these options laying on the best precision attained for the final fits of  $P_m$ .

Another example of the great flexibility and usefulness of the methodology is given in Fig. 8 where the plots of  $\log(x_o)$  vs.  $k$  (a) and of  $k$  vs.  $\log(x_o)$  (b) are shown, once again each with its own polynomial fit and its fit parameters.



**Fig. 8. Plots of  $\log(x_o)$  vs.  $k$  (a) and of  $k$  vs.  $\log(x_o)$  (b)**

For instance, the fit of plot (b) is ( $R^2=0.99902$   $\sigma=3.3517E-4$   $N=63$   $p<1E-4$ )

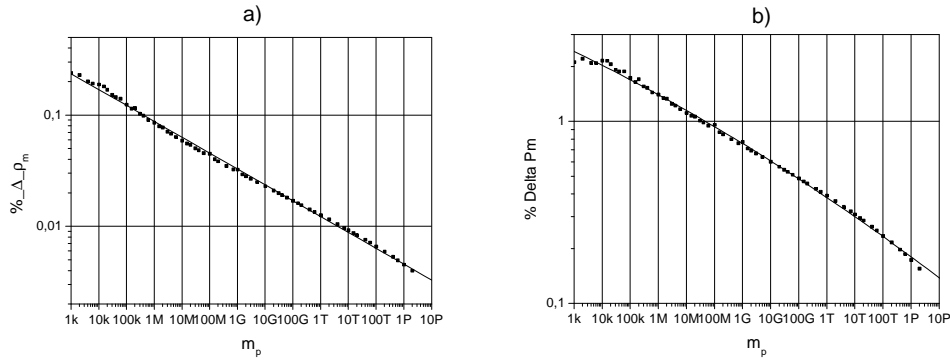
$$k \approx 1.89783 + 0.00403 \cdot \lg x_o - 7.3543E-5 \cdot \lg^2 x_o + 6.8046E-7 \cdot \lg^3 x_o - 3.0908E-9 \cdot \lg^4 x_o + +5.4618E-12 \cdot \lg^5 x_o$$

The two limits  $\lim_{x_o \rightarrow +\infty} k(x_o) = 2^-$   $\lim_{x_o \rightarrow +\infty} [\Delta k(x_o) / \Delta x_o] = 0^+$  are noticeable in both plots, though not in the fit relations. As a matter of fact, just like in the previous plots 6a and 7a, the fitting function, got automatically by the computer code used, is not capable of describing the limit behaviour (for  $m_p \rightarrow \infty$  or  $k \rightarrow 2^-$ ) of the plots. This aspect of the research will be treated in the future with the aim of correcting it.

It has to be remarked that, though at the present time the precisions of the many methods here shown are not very satisfying, nevertheless one of the main advantages of this approach to the problem of primes is given just by the many options by which the final values of  $P_m$  can be obtained. That leads to get also many checks on the final results in addition to the possibility of choosing among the many alternatives of calculating  $P_m$  from  $m_p$ , though approximate.

For instance, the latest method here explained gives the two plots (a) and (b) of Fig. 9 where for the former the difference between the actual values and the calculated values of  $\rho_m$  (a) is shown vs. the prime counter  $m_p$ , while for the latter (b) the difference between the actual values and the calculated values of  $P_m$  is reported versus  $m_p$ .

Nonetheless all these evaluations are still under examination at the present time, in that the error propagation from the values of  $\rho_m$  to the final results  $P_m$  is very high, that is there are high enhancements of all the inaccuracies in passing through the many steps illustrated above. As a matter of fact it is easy to remark that, there is a ten-fold increase at least in the two left scales of Fig. 9a and b, while in other cases the situation is even worse.



**Fig. 9. Error propagation from  $\rho_m$  (a) to  $P_m$  (b)**

Thus another method has been examined. For any sequence  $\{\rho(m_p)\}$  the fit has been made also by the finite progressions  $C_\alpha \cdot \{m^a\}$ , as already said, thus finding further interesting relationships.

Hence, again for any sequence  $\{\rho_m\}$  the corresponding progression  $\{C_\alpha \cdot m^a\} \equiv C_\alpha \cdot \{m^a\}$  has been examined optimizing the fit, again finding the 63 values of both  $\alpha$  and  $C_\alpha$ . The next Fig. 10 plots the related 63 data points and the two fits for the relationships  $\alpha = \alpha(m_p)$  and  $C_\alpha = C_\alpha(m_p)$ . The first set of data points (a) can be fitted by the relationship

$$\alpha = \alpha(m_p) \approx -0.0593 + 0.01807 \cdot m_p - 0.00237 \cdot m_p^2 + 1.33511E-4 \cdot m_p^3 - 2.84515E-7 \cdot m_p^4 - 2.73375E-7 \cdot m_p^5 + 7.95901E-9 \cdot m_p^6$$

where  $R^2 = 0.99997$   $\sigma = 2.8805E-5$   $N = 63$   $p < 1E-4$

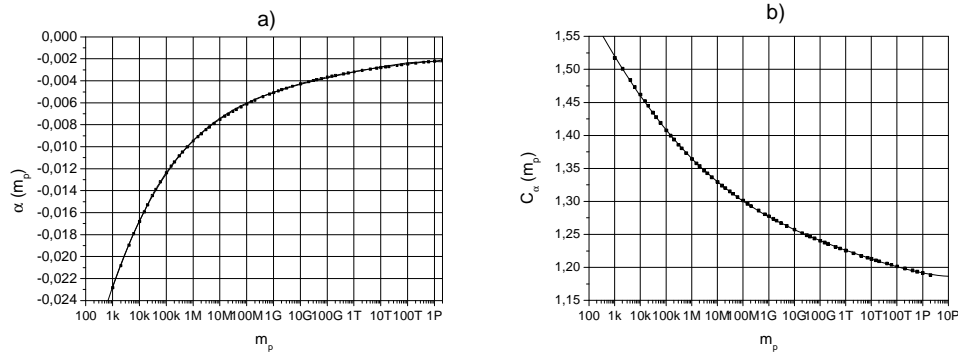
while the analogous fit for the 63 data points  $C_\alpha$  vs.  $m_p$  (Fig.10b) is

$$C_\alpha = C_\alpha(m_p) \approx 1.73777 - 0.0773 \cdot m_p - 2.1417E-4 \cdot m_p^2 + 7.58061E-4 \cdot m_p^3 - 5.82459E-5 \cdot m_p^4 + 1.38533E-6 \cdot m_p^5$$

with  $R^2 = 0.99994$   $\sigma = 7.01773E-4$   $N = 63$   $p < 1E-4$

Once again, the errors (or uncertainties) on the coefficients are between 1‰ and few % and again the limits are

$$\begin{array}{ll} \text{a)} & \lim_{m_p \rightarrow +\infty} \alpha(m_p) = 0^- \\ \text{a)} & \lim_{m_p \rightarrow +\infty} [\Delta\alpha(m_p)/\Delta m_p] = 0^+ \\ \text{b)} & \lim_{m_p \rightarrow +\infty} C_\alpha(m_p) = +1^+ \\ \text{b)} & \lim_{m_p \rightarrow +\infty} [\Delta C_\alpha(m_p)/\Delta m_p] = 0^- \end{array}$$



**Fig. 10. Data points and fits of  $\alpha$  (a) and of the coefficient  $C_\alpha$  (b) both vs.  $m_p$  for the 63 sequences of  $\{C_\alpha \cdot m^\alpha\}$**

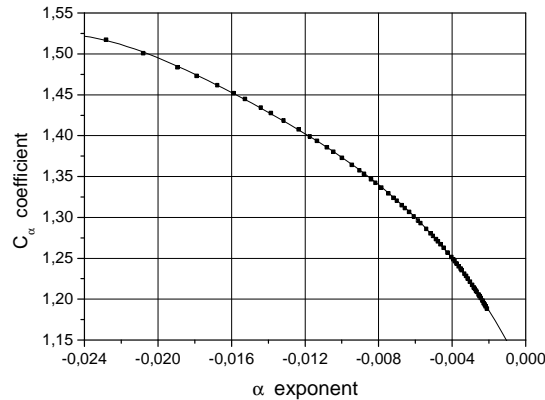
Hence, in this case too, a relationship, though approximate and holding locally, has been found that can be used to calculate the value of a prime number  $P_m$  from its counter  $m_p$ , that is  $P_m \approx m_p^{\rho(m_p)}$  and

$$\rho_m = \rho(m_p) \approx C_\alpha(m_p) \cdot m_p^{\alpha(m_p)} = C_\alpha(m_p) \cdot [m_p^\alpha(m_p)] \approx [1.73777 - 0.0773 \cdot m_p - 2.1417E-4 \cdot m_p^2 + 7.58061E-4 \cdot m_p^3 - 5.82459E-5 \cdot m_p^4 + 1.38533E-6 \cdot m_p^5] \cdot m_p^{[-0.0593 + 0.01807 \cdot m_p - 0.00237 \cdot m_p^2 + 1.33511E-4 \cdot m_p^3 - 2.84515E-7 \cdot m_p^4 - 2.73375E-7 \cdot m_p^5 + 7.95901E-9 \cdot m_p^6]}$$

Also the relation between  $C_\alpha(\alpha)$  and  $\alpha$  can be verified as in the next Fig. 11 leading to the fit

$$C_\alpha(\alpha) = (1.10405 \pm 0.00088) - (46.309 \pm 0.469) \cdot \alpha - (2,920.27 \pm 77.26) \cdot \alpha^2 - (117,419.8 \pm 4,879.6) \cdot \alpha^3 - (1.913 \pm 0.103)E6 \cdot \alpha^4$$

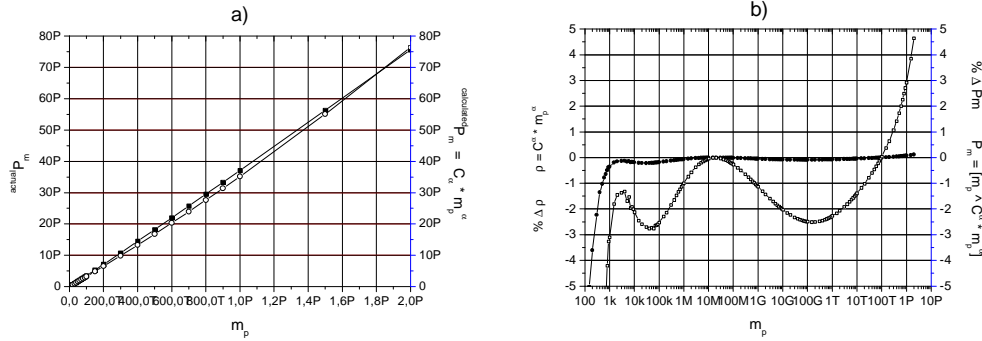
with  $R^2 = 0.99995$   $\sigma = 6.85E-4$   $N = 63$   $p < 1E-4$



**Fig. 11. Data points and fits of the coefficient  $C_\alpha$  vs.  $\alpha$  for the 63 sequences of  $\{C_\alpha \cdot m^\alpha\}$**

The comparison between the factual values and the values calculated by this method  $\rho(m_p) \approx C_\alpha(m_p) \cdot m_p^{\alpha(m_p)}$  and  $P_m \approx m_p^{C_\alpha(m_p)}$  leads to the two plots of Fig. 12 where the former (a) shows the two trends of the actual and the calculated  $P_m$  vs.  $m_p$  on a linear scale (the two plots are undistinguishable on a log-log scale) and the latter (b) displays the relative error between  $\rho_m^{actual}$  and  $\rho_m^{calculated}$  that is  $\% \Delta \rho_m$  and between  $P_m^{actual}$  and  $P_m^{calculated}$  that is  $\% \Delta P_m$  both on the same scale and vs.  $m_p$ , thus displaying the high error propagation from  $\rho_m$  to  $P_m$ . The relative difference between actual and calculated  $P_m$  by this method

remains still too large (some %) as well as the error propagation that ranges from  $\% \Delta \rho_m \in (-1\%, 0.1\%)$  to  $\% \Delta P_m \in (-2.7\%, 4.6\%)$  within the range  $m_p \in (1K, 2P)$  that is neglecting the first data points  $m_p < 1,000$  i.e.  $P_m < 7,919$ .

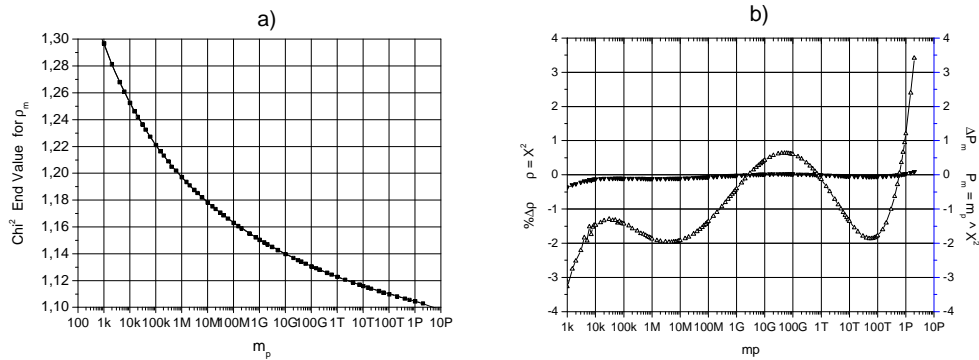


**Fig. 12. Comparison between actual  $P_m$  values and  $\{C_u \cdot m^a\}$  (a) and error propagation (b) from  $\rho_m$  to  $P_m$  vs.  $m_p$**

Thus a further approach has been attempted by examining the trend of the whole function  $X_k^2[C, m_p/x_o(k)] \approx \rho(m_p)$  vs.  $m_p$  with  $C=1$ , where in this case  $m_p$  is the maximum (or end) value of the counter of any prime sequence. The Fig. 13 (plot a) shows the result where again a polynomial regression analysis returns the fit for the data points

$$\begin{aligned} \rho(m_p) &\approx X_k^2[C, m_p/x_o(k)] = X_h^2(m_p) \approx (1.500866 \pm 0.0029) - (0.0956 \pm 0.00155) \cdot \lg(m_p) + \\ &(0.01039 \pm 2.87E-4) \cdot (\lg m_p)^2 - (5.7181 \pm 0.221) \cdot E-4 \cdot (\lg m_p)^3 + (1.22892E-5 \pm 6.036E-7) \cdot (\lg m_p)^4 = \\ &(1.500866 \pm 1.932\%) - (0.0956 \pm 1.621\%) \cdot \lg(m_p) + (0.01039 \pm 2.762\%) \cdot (\lg m_p)^2 - (5.7181 \cdot E- \\ &4 \pm 3.865\%) \cdot (\lg m_p)^3 + (1.22892E-5 \pm 4.911\%) \cdot (\lg m_p)^4 \end{aligned}$$

with the following fit values:  $R^2=0.99989$   $\sigma=5.37937 \cdot E-4$   $p < 1E-4$   $N=63$



**Fig. 13. Data points and fits of  $\rho(m_p) \approx X_k^2[C, m_p/x_o(k)]$  (a) and error propagation from  $\rho_m$  to  $P_m$  (b) both vs.  $m_p$**

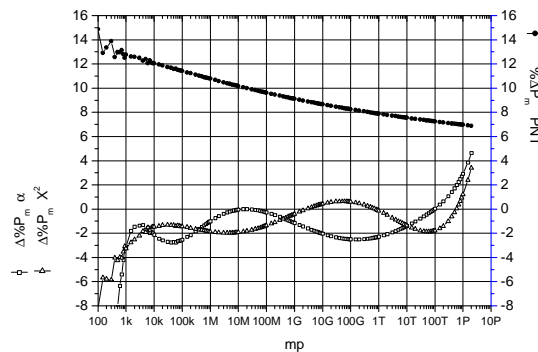
Thus, the value of  $\rho = \rho(m_p) \approx X_k^2[1, m_p/x_o(k)]$  can be directly estimated leading to

$$P_m = P(m_p) \approx m_p^{\rho(m_p)} \approx m_p^{\rho(m_p)} \cdot X_k^2[1, m_p/x_o(k)] \approx m_p^{\rho(m_p)} \cdot [(1.50086 \pm 0.0029) - (0.09506 \pm 0.00155) \cdot \lg(m_p) + (0.01039 \pm 2.87E-4) \cdot (\lg m_p)^2 - (5.7181 \pm 0.221) \cdot E-4 \cdot (\lg m_p)^3 + (1.22892E-5 \pm 6.036E-7) \cdot (\lg m_p)^4]$$

Of course taking into account, here too, both the errors on the fit coefficients and their propagation as well as the scale non-invariance problem. As a matter of fact the second plot (b) of Fig. 16 shows the error propagation from  $\rho_m \approx X_k^2[C, m_p/x_o(k)]$  to  $P_m \approx m_p^{\rho_m}$  vs.  $m_p$  that is the comparison between the percentage difference between the actual  $\rho_m$  values and the calculated  $\rho_m$  values ( $\% \Delta \rho_m$  ranging from  $-3.5\%$  up to  $0.86\%$ ) and the percentage difference between the actual  $P_m$  values and the calculated  $P_m$  values, ranging from  $\% \Delta P_m = -3.27\%$  up to  $\% \Delta P_m = +3.41\%$  thus displaying the downgrading of the precision, approximately a tenfold increase of the uncertainties in passing from  $\rho_m$  to  $P_m$  in the range  $m_p \in (1K, 2P)$ .

Despite the fact that the final precision is not so satisfactory, nonetheless it is better than the difference between the actual  $P_m$  values and the  $P_m$  values calculated by the standard PNT method (that is  $P_m \approx m_p \cdot \ln m_p$ ) which ranges from  $12.77\%$  up to  $6.88\%$  (again from  $m_p = 1K$  up to  $m_p = 2P$ ) as shown in Fig. 14 illustrating the comparisons among the three different methods:  $P_m \approx \rho_m \cdot X_k^2[C, m_p/x_o(k)]$ ,  $P_m \approx \rho_m \cdot C_a \cdot m^\alpha$  and PNT.

Yet, it should be considered that at the present stage the precisions of the results are not so good as those got by other valuable methods [12]. This constraint of the study will be the focus of a further deepening and development later on.



**Fig. 14. Comparisons among the imprecisions of the three methods  $P_m \approx \rho_m \cdot X_k^2$ ,  $P_m \approx \rho_m \cdot C_a \cdot m^\alpha$ ,  $P_m \approx m_p \cdot \ln(m_p)$  vs.  $m_p$**

Thus the relationships got in such a way and reported in the present article should be considered just as preliminary attempts to evaluate a prime number  $P_m$  starting from its counter  $m_p$  using the method of the modified chi-square function.

However, the methodology here shown offers a wide range of different options to calculate the approximate value of a prime number  $P_m$  from its counter  $m_p$  with the additional advantage that the precisions of all the results can be improved by the use of a more powerful computer, that is a mainframe or even a supercomputer, just by acting on both the initial fits of the  $\{\rho_m\}$  sequences and the final fits of the results as well as on the value of  $n_\Delta$ .

Moreover, a whole set of further useful relations are obtained in addition to those reported previously.

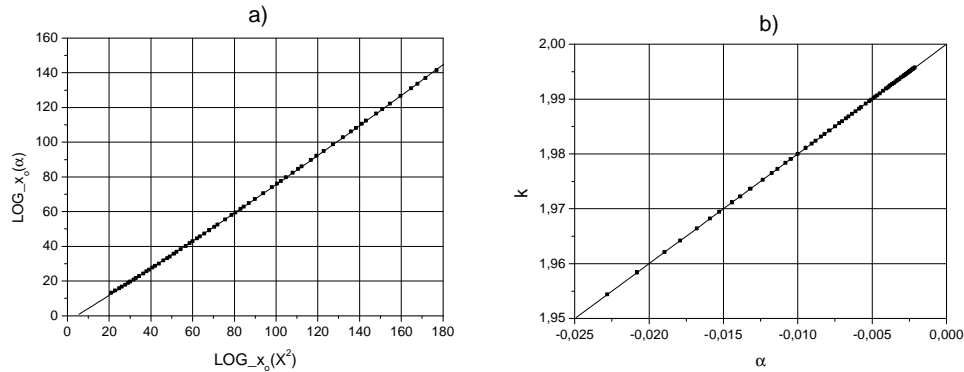
The relationships between the decay parameters  $x_o$  got by the  $C_a \cdot m^\alpha$  method and by the  $X_k^2$  method are shown in the plot (a) of Fig. 15 with the fit given by

$$\lg x_o(\alpha) \approx -3.29733 + 0.74616 \cdot \lg x_o(X^2) + 4.23628E-4 \cdot \lg^2 x_o(X^2)$$

with  $R^2 = 0.99998$ ,  $\sigma = 0.18127$ ,  $N=63$  while the plot (b) of Fig. 15 reports the law  $k(\alpha) = 2 + 2\alpha$  as got by the calculations that is

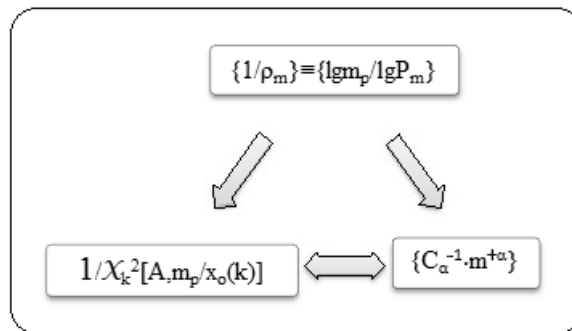
$$k(\alpha) = (2 \pm 2.925E-17) + (2 \pm 2.292E-15) \cdot \alpha$$

with  $R=1.000000000000$  and  $\sigma=1.333E-16$  just a validation of what already told ( $k = 2+2\alpha$ ) with negligible errors.



**Fig. 15. Plot of  $\lg x_0(\alpha)$  vs.  $\lg x_0(X^2)$  (a) and plot of  $k$  vs.  $\alpha$   $k = 2+2\cdot\alpha$  (b) as got by the calculations**

In addition, it can be easily ascertained that also the following scheme of Fig. 16 holds just like the one of Fig. 4 yet with the difference that now  $\alpha \in (0, +1)$  and  $k = 2 - 2\alpha$  and again the single arrows reveal the one-way correspondence (injective map) while the double arrow reflects a one-by-one correspondence (bijective map). That is not a trivial concern as it might look but an important finding of all the research taking into account that both the values of the parameter  $k$  and those of the decay or growth parameter  $x_0$  are equal in the two cases  $\{\rho_m\}$  and  $\{\rho_m^{-1}\}$ . The only difference between the two cases of Fig. 4 and Fig. 16 are that the  $\alpha$  coefficients obey to the two different laws  $k = 2 + 2\alpha$  the former and  $k = 2 - 2\alpha$  the latter.



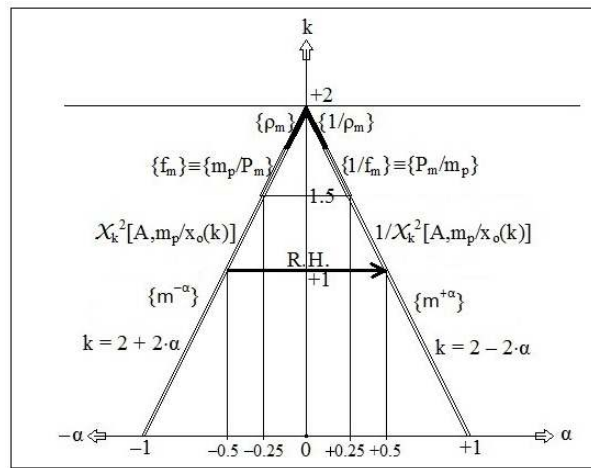
**Fig. 16. Fits of the sequences  $\{1/\rho_m\} \equiv \{\lg(m_P)/\lg(P_m)\}$  with  $k = 2 - 2\alpha$**

Now, just a trivial remark is due to highlight the fact that the fit of any  $\{m^{+\alpha}\}$  progression by the  $(1/\cdot)X_k^2(A, m/x_o)$  function, according to the range, has fit parameters  $R=1=1.0000000000000000$  up to the 12<sup>th</sup> digit i.e. up to the chosen precision, even for so few terms of the progression as 50,000; this is a strict consequence of the features of the modified chi-square function. As a matter of fact choosing in (1)  $A=2\Gamma_{k/2}/(2x_o)^{k/2-1}$  and being always  $x_o \gg m$  it is easy to get  $x^{k/2-1} = x^{+\alpha}$  the natural interpolating function of the  $\{m^{+\alpha}\}$  progressions, as already shown [20,21].

Therefore, passing to the geometric representation of the results in the plane  $(\alpha, k)$  as in the next Fig. 17, with the two half-lines  $k = 2 \pm 2\alpha$  crossing one each other at the point  $(\alpha, k) \equiv (0, 2)$ , at this stage of the current

study it has been ascertained that the left lower neighbourhood of the point  $(\alpha, k) \equiv (0^-, 2^-)$  is a pillar of the whole matter of prime numbers examined in such a way in that capable of directly fitting the finite sequences  $\{\rho_m\}$  by the progressions  $\{m^\alpha\}$  with any number of terms and by the fitting function  $X_k^2[A, m_p/x_o(k)]$ .

Consequently it is possible to summarize the results and the findings of the present article and of the whole research [14-17,20,21] in Fig. 17, where the two half-lines  $k=2 \pm 2 \cdot \alpha$  (according to the ranges  $\alpha < 0$  or  $\alpha > 0$ ) are shown together with all the fits of the  $\{\rho_m\}$  or  $\{1/\rho_m\}$  sequences, that is the  $X_k^2(A, m_p/x_o)$  or the  $1/X_k^2(A, m_p/x_o)$  function as well as the  $\{C_\alpha \cdot m^{-\alpha}\}$  or  $\{C_\alpha^{-1} \cdot m^{+\alpha}\}$  progressions in the left neighbourhood of the point  $(\alpha, k) \equiv (0^-, +2)$  and in the right neighbourhood of the same point  $(\alpha, k) \equiv (0^+, +2^-)$  respectively. In Fig. 17 the neighbourhoods of the point  $(\alpha, k) \equiv (0, 2)$  that is  $(\alpha, k) \equiv (0^\pm, 2^-)$  are enhanced for clearness.



**Fig. 17. The plane  $(\alpha, k)$  in the ranges  $\alpha \in (-1, +1)$  and  $k \in (0, +2)$  with  $k = 2 \pm 2\alpha$**

It is interesting to remark that, while the fits of  $\{\rho_m\}$  or  $\{1/\rho_m\}$  sequences by the  $X_k^2(A, m_p/x_o)$  or the  $1/X_k^2(A, m_p/x_o)$  functions are valid just in the left or right neighbourhoods of the point  $(\alpha, k) \equiv (0^\pm, 2^-)$ , the fits of the finite sequences  $\{m^{-\alpha}\}$  or  $\{m^{+\alpha}\}$  by the functions  $X_k^2(A, m/x_o)$  or  $1/X_k^2(A, m/x_o)$  are valid in the whole ranges  $\alpha \in (-1, 0)$  and  $k \in (0, 2)$  or  $\alpha \in (0, +1)$  and  $k \in (0, 2)$  respectively. This feature has been used as the starting point for an elementary proof of Riemann's hypothesis [20,21] at the point  $(\alpha, k) \equiv (-1/2, +1)$ . Additionally, Fig. 17 is interesting in that already used in the fits of the finite sequences of prime number frequencies  $\{f_m\} \equiv \{m_p/P_m\}$  [16,17] again by the modified chi-square function and by  $\{m^{-\alpha}\}$  progressions in the ranges  $\alpha \in (-0.25, 0) \subset (-1, 0)$  and  $k \in (1.5, +2) \subset (0, +2)$  along the half-line  $k=2+2 \cdot \alpha$  while the examination of the frequency inverses  $\{f_m^{-1}\} \equiv \{P_m/m_p\} \approx 1/X_k^2(A, m_p/x_o) \approx \{m^{+\alpha}\}$  in the ranges  $\alpha \in (0, +0.25) \subset (0, +1)$  and  $k \in (1.5, 2) \subset (0, 2)$  is still in progress with the goal of reaching a more refined version of the PNT.

Lastly, many ways of approximating, that is interpolating and locally extrapolating, the values of prime number finite sequences are possible by the methodology presented here, a methodology which, though still at its early stage, is liable of many interesting future developments and that deserves to be further deepened in the next future.

### 4 Conclusions and Future Developments

The algorithm presented in this report as applied to both the finite sequences of  $\{\rho_m\} \equiv \{\lg P(m_p)/\lg m_p\}$  and the progressions  $\{m^{-\alpha}\}$  by the modified chi-square function  $X_k^2(A, m_p/x_o)$  constitutes an innovative methodology for the former and the latter that with no doubt can be applied to many other cases. In the same manner such an algorithm, that makes a wide use of the modified chi-square function  $X_k^2(A, m/x_o)$  is suitable to fit some basic structures of the finite sequences of prime numbers.

Therefore, the major findings of the present article can be summarized as the following ones also looking at the previous Fig. 17:

- The finite sequences  $\{\rho_m\} \equiv \{\lg P_m / \lg m_p\}$  of prime numbers  $P_m$  do not show the property of scale invariance holding for them the scaling laws given by the modified chi-square function  $+X_k^2(A, m_p/x_0)$  and by the progressions  $\{m^{-\alpha}\}$ ;
- The sequences  $\{\rho_m\} \equiv \{\lg P_m / \lg m_p\}$  are best fitted by the function  $+X_k^2(A, m_p/x_0)$  with  $k=2+2\alpha$  i.e.  $\alpha=k/2-1$  in the lower left neighbourhood of  $(\alpha, k) \equiv (0^-, 2^-)$  and also by the progressions  $\{m^{-\alpha}\}$ ; in other words  $\{\rho_m\} \approx X_k^2(A, m_p/x_0) \approx \{C_\alpha \cdot m^{-\alpha}\}$ ;
- The sequences  $\{\rho_m^{-1}\} \equiv \{\lg m_p / \lg P_m\}$  are best fitted by the modified chi-square function  $+1/X_k^2(A, m_p/x_0) = +1/X_{2-2\alpha}^2(A, m_p/x_0)$  with  $\alpha = 1-k/2$  in the lower right neighbourhood of  $(\alpha, k) \equiv (0^+, 2^-)$  and also by the progressions  $\{m^{+\alpha}\}$ ; in other words  $\{\rho_m\} \approx 1/X_k^2(A, m_p/x_0) \approx \{C_\alpha^{-1} \cdot m^{+\alpha}\}$ ;
- The present methodology allows to get the value of a prime number  $P_m$  from its counter  $m_p$  in an approximate way and valid locally however in many different ways, five or six using the scheme of Fig. 4 and double using also the scheme of Fig. 16;
- One of the main advantages of the methodology is that the final fit uncertainties can be reduced by more precise determinations of all the initial fits and also increasing the value of  $n_\Delta$  thus leading to more precise estimates of the  $P_m$  values.

In addition to these conclusions and final results concerning just the present article, the following ones should be considered as a summary of the main findings of the whole study, at this early stage, having examined the problem of prime numbers from a statistical and analytical viewpoint:

- 1- Prime numbers have not, from either the statistical or the analytical viewpoint, the property of scale invariance holding for them the scaling laws given by  $k=k(m_p)$  of the modified chi-square function  $X_k^2(A, x/x_0)$  [14-17];
- 2- Both the statistics of the progressions  $\{m^\alpha\}$  and the statistics of prime numbers  $\{P_m\}$  are best fitted by the modified chi-square function  $X_k^2(A, x/x_0)$  with the ad-hoc parameters  $A$   $k$   $x_0$  so that the statistics of prime numbers  $\{P_m\}$  are best fitted by the statistics of the progressions  $\{m^\alpha\}$  with  $\alpha \in (1^+, 1.12)$  [14,15];
- 3- The progressions  $\{m^{\pm\alpha}\}$  are analytically best-fitted, according to the range  $\alpha$ , by one of the four forms, either  $(\pm 1 \cdot)$  or  $(\pm 1/)$ , of the modified chi-square function  $X_k^2(A, m/x_0)$  with the appropriate parameters  $A$ ,  $k$  and  $x_0$  holding  $k=2\pm 2\alpha$  according to the range ( $\alpha < 0$   $\alpha > 0$ ) examined [20,21];
- 4- The finite sequences of prime number frequencies  $\{f_m\} \equiv \{m_p/P_m\}$  are analytically best fitted by the modified chi-square function  $X_k^2(A, x/x_0)$  with  $k \in (1.50, 2)$  hence by the progressions  $\{m^\alpha\}$  with  $\alpha \in (-0.25, 0^-)$  holding  $k = 2 + 2\alpha$  [16,17];
- 5- The finite sequences  $\{\rho_m\} \equiv \{\lg(P_m)/\lg(m_p)\}$  are best fitted analytically by the function  $X_k^2(C, m_p/x_0)$  with  $k \in [1.952, 2^-)$  thus also by  $\{m^\alpha\}$  with  $\alpha \in [-0.024, 0^-)$  being  $k=2 + 2\alpha$ ;
- 6- It is possible to find many inductive algorithms which allow to get the value of a prime  $P_m$  starting from its counter  $m_p$  i.e. such that  $P_m \approx P(m_p)$  though with values affected by uncertainties and usable locally;
- 7- As a matter of fact all that holds just taking into account all the approximations adopted, as well as the inaccuracies and propagations of the errors as usually done in experimental physics. However future researches and refinements of the method will be capable to reduce all the uncertainties.
- 8- An elementary (that is not using the theory of complex functions) proof can be given of Riemann's hypothesis implementing this methodology to the progressions  $\{n^{-\alpha}\}$ ,  $\{\sum n^{-\alpha}\}$  and  $\{N^{-\alpha+1}/(\alpha+1)\}$  [20,21].

However much more is still to be done in the matter while in any topic here shown many aspects must be developed and deepened, first of all the accuracy of the calculations. Anyhow it is the author's opinion that the algorithm and the methodology here shown can open a new field of study in number theory and computational mathematics which can reveal all its power more and more in the future.



The use of a much more powerful computer, that is a mainframe, would increase the final precisions a great deal, first of all by allowing to increase the value of the  $n_{\Delta}$  data points making it as high as reasonably attainable (a 1000-fold increase or more). In this frame, further efforts could be focused on the development of an ad-hoc computer code capable of automatically sweeping the plane  $(k, x_0)$  and possibly the 3D space  $(A, k, x_0)$  in order to find, at any point, the parameters of the fit function  $X_k^2(A, x/x_0)$  and to choose among them identifying the best ones.

As for the next future developments, one of the first issues will be the examination of the finite sequences of  $\{P_m\} \equiv B_m \cdot \{m_p^{\rho'(mp)}\}$  or  $\{P_m\} \equiv \{m_p^{Dm \cdot \rho''(mp)}\}$  in order to get more precise evaluations of  $P_m \approx P(m_p)$  and the examination of the finite sequences  $\{1/f_m\} \equiv \{P_m/m_p\} \approx 1/X_k^2(A, m_p/x_0) \approx \{m_p^{+\alpha}\}$   $\alpha \in (0, +1) \subset \mathbb{R}^+$  thus obtaining a more refined version of the PNT. At the present time this work is in progress on behalf of the author.

Further investigations are still in progress on the entire area especially for what concerns the best methodology, among all the described ones and not only, on which to concentrate the efforts to reduce the uncertainties in order to get a highly precise formula of  $P_m \approx P(m_p)$  at the utmost precision achievable. Additional work will be dedicated to the topic of fitting the many decay or growing plots of the final results - see Figs. 6a, 7a, 8a and b, 10a and b, 11, 13a, again by the modified chi-square function in one of its four forms  $\pm(1/\omega)X_k^2(\Omega, x/\omega)$ , instead of a polynomial sum, in that those plots clearly show a finite limit like  $2^-$ ,  $0^-$  and  $1^+$ . This is a very interesting topic in that, if so, the area of prime numbers would present a kind of iteration and maybe even of self-similarity. Moreover, it is highly probable that these polynomial sums themselves may be responsible for the large error propagations and the great uncertainties on the final results.

In addition, many further areas of investigation seem to be disclosed by this advanced and original methodology concerning also the many integer sequences [35] which can be treated in an innovative way, thus maybe revealing undisclosed properties, if any.

## Competing Interests

Author has declared that no competing interests exist.

## References

- [1] Smith DE. A source book in mathematics. Dover Publications, Inc. New York, Unabridged Republication of the 1<sup>st</sup> Edition, Originally Published in 1929 by McGraw-Hill Book Co., Inc; 1959.
- [2] Hardy GH, Littlewood JE. Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes. Acta Mathematica. 1916;41:119-196.  
Available: <http://dx.doi.org/10.1007%2F02422942>
- [3] Babusci D, Dattoli G, Del Franco M. Lectures on mathematical methods for physics RT/2010/58/ENEA, ENEA-Roma-I. Italian.
- [4] Fine B, Rosenberger G. Number theory - An introduction via the distribution of primes ©2007 Birkäuser Boston, USA.
- [5] Du Sautoy M. L'enigma dei numeri primi ©2004 RCS Libri S.p.A. Milano. Italian. Also published as The Music of the Primes ©2003 by Marcus du Sautoy.  
ISBN: 978-8817-05022-7
- [6] Derbyshire J. L'ossessione dei numeri primi – Bernhard Riemann e il principale problema irrisolto della matematica ©2006 Bollati Boringhieri Editore s.r.l. Torino. Italian. Also published as Prime Obsession Bernhard Riemann and the Greatest Unsolved Problem in Mathematics Joseph Henry Press (National Academic Press), Washington D.C; 2003. ©2003 John Derbyshire.

- [7] Languasco A, Zaccagnini A. Alcune proprietà dei numeri primi. I e II, Sito web Bocconi- Pristem; 2005. Italian.  
Available:<http://matematica.uni-bocconi.it/LangZac/home.htm>
- [8] Goldstone DA, Pintz J, Yildirim CY. © 2007 The path to recent progress on small gaps between primes. *Clay Mathematics Proceedings*. 2007;7:125–135.
- [9] Green B. © 2007 Long arithmetic progressions of primes. *Clay Mathematics Proceedings*. 2007;7: 149–167.
- [10] Green B, Tao T. The primes contain arbitrarily long arithmetic progressions. *Annals of Mathematics*. 2008;167(2):481-547.  
Available:<http://www.arXiv:math.NT/0404188>
- [11] Connes A. An essay on the Riemann Hypothesis arXiv:1509.05576v1 [math.NT] 18 Sept.; 2015.
- [12] Cipolla M. La determinazione asintotica del  $n^{\text{mo}}$  numero primo. *Matematiche Napoli*. 1902;3:132-166.
- [13] Sierra G. In collaboration with Latorre J.I., Madrid-Barcelona-Singapore, Primes go Quantum: There is entanglement in the primes, workshop: Entanglement entropy in quantum many-body systems, King's College and City University, London, 2-4 June; 2014. Based on arXiv: 1302.6245 and arXiv: 1403.4765.
- [14] Lattanzi D. Distribution of prime numbers by the modified chi-square function. *Notes on Number Theory and Discrete Mathematics*. 2015;21(1):18-30.  
Available:<http://nntdm.net/volume-21-2015/number-1/18-30/>
- [15] Lattanzi D. Scale laws of prime number distributions by the modified chi-square functions. Oral Presentation at the 3<sup>rd</sup> International Conference on Recent Advances in Pure and Applied Mathematics (ICRAPAM 2016) 19-23 May; 2016. Bodrum, Turkey.
- [16] Lattanzi D. Scale laws of prime number frequencies by the modified chi-square function. *British Journal of Mathematics and Computer Science*. 2016;13(6):1-21. Article no. BJMCS.23200.  
Available:<http://sciencedomain.org/issue/1545>,  
<http://sciencedomain.org/abstract/12890>
- [17] Lattanzi D. Characterization of prime numbers by the modified chi-square function. Oral Presentation at the Global Conference on Applied Physics and Mathematics, Rome, Italy, 25-27 July; 2016.
- [18] Stallinga P. Scalable functions used for empirical forecasting. *British Journal of Mathematics and Computer Science*. 2016;18:2. Article no. BJMCS.28107.
- [19] Ghosh A. Mechanics over micro and nano scales. Chapter 2, Scaling Laws, Editor Chakraborty S. 2011;IX:269.  
ISBN: 978-1-4419-96008
- [20] Lattanzi D. An elementary proof of Riemann's hypothesis by the modified chi-square function. *British Journal of Mathematics and Computer Science*. 2016 ;15(5):1-14. Article no. BJMCS.25419  
Available:<http://sciencedomain.org/issue/1778>,  
<http://sciencedomain.org/abstract/14029>
- [21] Lattanzi D. The modified chi-square function as a tool for Riemann's hypothesis. Oral Presentation at the Global Conference on Applied Physics and Mathematics, Rome, Italy, 25-27 July; 2016.

- [22] Bailey DH, Borwein JM. Exploratory experimentation and computation August 14, 2010 – LBNL Paper LBNL-3313E Lawrence Berkeley National Laboratory; Notices of the AMS. 2011;58(10): 1410-1419.  
Available:<http://escholarship.org/uc/item/26m6noob>
- [23] Andeberhan T, Medina LA, Moll VH. Editors, Contemporary Mathematics - 517 – gems in experimental mathematics. AMS Special Session, Experimental Mathematics, Washington D.C. American Mathematical Society. January 5; 2009.
- [24] Thomas Oliveira e Silva, Dep. De Electronica, Telecomunicações e Informatica, Universidade de Aveiro, Portugal Goldbach Conjecture Verification.  
Available:<http://www.ieeta.pt/~tos/primes.html>
- [25] Sørensen HK. Exploratory experimentation in experimental mathematics: A glimpse at the PSLQ algorithm Institut for Videnskabsstudier, Aarhus Universitet, 8000 Århus C, Denmark.
- [26] Borwein JM. Exploratory experimentation: Digitally-assisted discovery and proof. University of Newcastle, Australia. May 18; 2009.
- [27] Young HD. Carnegie Institute of Technology Statistical Treatment of Experimental Data McGraw-Hill Book Company, Inc. ©1962.
- [28] Morice E. Dizionario di statistica ©1971 by ISEDI. Milano. Italian. Dictionnaire de Statistique ©1967 by Dunod, Paris.
- [29] Mathews J, Walker RL. CalTech mathematical methods of physics ©1964 W.A. Benjamin Inc. New York N.Y.
- [30] Available:<http://www.efunda.com/math/gamma/findgamma>
- [31] Available:<http://keisan.casio.com/exec/system/1180573444>
- [32] Available:<http://primes.utm.edu/lists/small/millions>
- [33] Available:<http://www.bigprimes.net/cruncher/>
- [34] Available:<https://primes.utm.edu/nthprime/index.php>
- [35] Sloane NJA. The On-line Encyclopedia of Integer Sequences<sup>®</sup> (OEIS<sup>®</sup>).  
Available:<https://oeis.org/>

---

© 2017 Lattanzi; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Peer-review history:**

The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)

<http://sciencedomain.org/review-history/17942>