

NOTE ON THE LAW OF FREQUENCY OF  
PRIME NUMBERS.By *J. W. L. Glaisher.**Introduction, §§ 1-4.*

§ 1. It is known that the numbers of primes inferior to any large number  $x$  is approximately equal to  $\text{li}x$ . This formula was discovered by Gauss, but the first satisfactory investigation was given by Tchebycheff. A much more complete and rigorous treatment of the question, by Riemann, showed that the number of primes was more accurately represented by the formula  $\text{li}x - \frac{1}{2} \text{li}x^{\frac{1}{2}} - \frac{1}{3} \text{li}x^{\frac{2}{3}} + \&c.$

§ 2. If the number of primes be represented by  $\text{li}x$ , the average interval between two primes at the point  $x$ , in the ordinal series of numbers, must be  $\log x$ ; and it would seem that it ought to be possible to obtain this result by general reasoning, depending upon the manner in which the composite numbers are formed from the primes.

The method by which Gauss arrived at the conclusion that the average frequency of primes was inversely proportional to the logarithm is not, I believe, known; but a general investigation was given by Hargreave in the *Philosophical Magazine* for 1849, by which he was led independently to the same conclusion, this being in fact the first publication of the law.

§ 3. Although Hargreave's result is the true one, his reasoning is vague and unsatisfactory. Some years ago (in 1880) I devoted much time to the attempt to obtain the known result by more conclusive methods of the same kind. In this I failed, every apparent proof being found to contain serious defects of principle, when critically examined. I then arrived at the conviction that the problem was of such an intricate nature that it would be very difficult to deduce even a moderately satisfactory investigation of the law of frequency by general reasoning from elementary considerations, or by

ordinary algebraical treatment of the formula which expresses accurately the number of primes inferior to  $x$ , viz.

$$x - \Sigma I\left(\frac{x}{p_1}\right) + \Sigma I\left(\frac{x}{p_1 p_2}\right) - \Sigma I\left(\frac{x}{p_1 p_2 p_3}\right) + \dots,$$

where  $p_1, p_2, \dots$ , are the primes inferior to  $x$ , and  $I\left(\frac{x}{p}\right)$  denotes the nearest integer to  $\frac{x}{p}$ , which does not exceed it.

§ 4. It seems to me, however, that it may be interesting to place upon record any investigation founded upon elementary principles, which leads to the law  $\log x$ , or even to a functional equation satisfied by  $\log x$ , and which does not appear to contain any obvious flaw. This is my justification for the present note, in which a functional equation satisfied by  $\log x$  is derived from elementary principles. Judging from experience I think it likely that the investigation would not bear the test of any very careful examination; and not too much importance must be attached to the fact that the result is the true one. The method was suggested by my paper on the series  $\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{11} + \dots$  in vol. XXV. (1891) of the *Quarterly Journal* (pp. 369-375); in fact the connexion is obvious on comparing the first six sections of that paper with the following investigation.

*Investigation of a functional equation satisfied by the function expressing the frequency of prime numbers, §§ 5-15.*

§ 5. If any large number  $x$  be taken at random, the probability that it is not divisible by 2 is  $\frac{1}{2}$ , the probability that it is not divisible by 3 is  $\frac{2}{3}$ , and so on. Thus, regarding these as independent probabilities, the probability that it is not divisible by the primes 2, 3, 5, ...,  $p$  is

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \dots \frac{p-1}{p}.$$

If then we denote by  $P(x)$  the probability that a large number  $x$ , taken at random, is prime; then

$$P(x) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \dots \frac{r-1}{r},$$

where  $r$  is the prime next inferior to the square root of  $x$ .

§ 6. Now

$$\begin{aligned} \log \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots \frac{r-1}{r} \\ = \log \left\{ \left(1 - \frac{1}{2}\right) \left(1 - \frac{1}{3}\right) \left(1 - \frac{1}{5}\right) \left(1 - \frac{1}{7}\right) \cdots \left(1 - \frac{1}{r}\right) \right\} \\ = - \left( \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots + \frac{1}{r} \right) \\ - \frac{1}{2} \left( \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots + \frac{1}{r^2} \right) \\ - \frac{1}{3} \left( \frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{5^3} + \frac{1}{7^3} + \cdots + \frac{1}{r^3} \right), \\ \quad \&c. \qquad \qquad \qquad \&c. \end{aligned}$$

§ 7. Let the series

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots + \frac{1}{x}$$

be denoted by  $\psi(x)$ .

Then, if  $h$  be small compared to  $x$ ,

$$\psi(x+h) - \psi(x) = \frac{N}{x}$$

approximately, where  $n$  is the number of primes between  $x$  and  $x+h$ .

Expanding  $\psi(x+h)$  in ascending powers of  $h$ , this equation may be written

$$h\psi'(x) = \frac{N}{x}.$$

§ 8. Let  $\phi(x)$  denote the number of primes inferior to  $x$ , then

$$N = \phi(x+h) - \phi(x),$$

so that,  $h$  being small compared to  $x$ ,

$$N = h\phi'(x).$$

§ 9. Now the probability that  $x$  is a prime is represented by the fraction  $\frac{N}{h}$ , whence

$$P(x) = \phi'(x),$$

and therefore the product

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \dots \frac{r-1}{r} = \phi'(x),$$

§ 10. From §§ 7 and 8, we have

$$\psi'(x) = \frac{\phi'(x)}{x},$$

so that  $\psi(x) = \text{const.} + \int_1^x \frac{\phi'(x)}{x} dx,$

§ 11. In the same manner, if we denote

$$\frac{1}{2^n} + \frac{1}{3^n} + \frac{1}{5^n} + \dots + \frac{1}{x^n},$$

by  $\psi_n(x)$ , we find

$$\psi_2(x) = \text{const.} + \int_1^x \frac{\phi'(x)}{x^2} dx,$$

$$\psi_3(x) = \text{const.} + \int_1^x \frac{\phi'(x)}{x^3} dx,$$

&c. &c.

§ 12. Substituting in the formula of § 6, we have therefore

$$\begin{aligned} \log \phi'(x) = \text{const.} - \int_1^x \frac{\phi'(r)}{r} dr - \frac{1}{2} \int_1^x \frac{\phi'(r)}{r^2} dr \\ - \frac{1}{3} \int_1^x \frac{\phi'(r)}{r^3} dr - \&c. \end{aligned}$$

Now  $r$  is approximately equal to  $\sqrt{x}$ , so that we may write this equation

$$\log \phi'(x) = \text{const.} - \int_1^{\sqrt{x}} \frac{\phi'(r)}{r} dr - \frac{1}{2} \int_1^{\sqrt{x}} \frac{\phi'(r)}{r^2} dr - \&c.$$

§ 13. Putting  $x^2$  for  $x$ , this equation becomes

$$\log \phi'(x^2) = \text{const.} - \int_1^x \frac{\phi'(r)}{r} dr - \frac{1}{2} \int_1^x \frac{\phi'(r)}{r} dr - \&c.,$$

whence, differentiating with respect to  $x$ ,

$$2x \frac{\phi''(x^2)}{\phi'(x^2)} = - \frac{\phi'(x)}{x} - \frac{1}{2} \frac{\phi'(x)}{x^2} - \frac{1}{3} \frac{\phi'(x)}{x^3} - \&c.$$

§ 14. As we are seeking for a form of  $\phi'(x)$ , which satisfies this equation for large values of  $x$ , we omit all the terms after the first on the right hand side of the equation, which then reduces to

$$2x \frac{\phi''(x^2)}{\phi'(x^2)} = - \frac{\phi'(x)}{x}.$$

Putting  $\chi(x) = \frac{1}{\phi'(x)},$

the equation becomes

$$2x \frac{\chi'(x^2)}{\chi(x^2)} = \frac{1}{x\chi(x)},$$

that is  $\chi(x^2) = 2x^2\chi(x)\chi'(x^2).$

§ 15. Whatever may be the general form of  $\chi(x)$ , which satisfies this functional equation, it is clear that it is satisfied by

$$\chi(x) = \log x,$$

and, taking this value,

$$\phi'(x) = \frac{1}{\log x};$$

whence  $\phi(x) = \text{const.} + \int_1^x \frac{dx}{\log x},$

This formula shows that the number of primes between the numbers  $x$  and  $y$  is  $\text{li } x - \text{li } y$ ; and we may take the number of primes inferior to  $x$  to be  $\text{li } x$ .

*Remarks on the functional equation, §§ 16, 17.*

§ 16. It may be remarked that by putting  $x = e^t$  in the  $\chi$ -equation of § 14, it becomes

$$\chi(e^{2t}) = 2e^{2t}\chi(e^t)\chi'(e^{2t}),$$

that is, putting  $\chi(e^t) = F(t),$

$$F(2t) = 2F(t)F'(2t),$$

giving  $\log F(2t) = \int \frac{dt}{F(t)} + \text{const.}$

§ 17. It may also be noticed that the general equation of § 13 when the terms involving  $x^{-2}$ ,  $x^{-3}$ , &c., are not omitted, may be written

$$2x\phi''(x^2) = \phi'(x) \phi'(x^2) \log\left(1 - \frac{1}{x}\right),$$

or 
$$\chi(x^2) \log\left(1 - \frac{1}{x}\right) = -2x\chi(x) \chi'(x^2).$$

*Second investigation of the functional equation, § 18.*

§ 18. In the following investigation of the functional equation the series  $\psi(x)$  is not introduced.

We have, as in § 5,

$$P(x^2) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \dots \left(1 - \frac{1}{x}\right),$$

where  $P(x^2)$  denotes the probability of a large number near  $x^2$ , taken at random, being prime.

Also,

$$P(x+h)^2 = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \dots \left(1 - \frac{1}{x}\right) \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_n}\right),$$

where  $p_1, p_2, \dots, p_n$  are the primes between  $x$  and  $x+h$ .

Thus

$$\begin{aligned} \frac{P(x+h)^2}{P(x^2)} &= \left(1 - \frac{1}{p_1}\right) \left(1 - \frac{1}{p_2}\right) \dots \left(1 - \frac{1}{p_n}\right) \\ &= \left(1 - \frac{1}{x}\right)^n \end{aligned}$$

approximately.

Since  $n$  is the number of primes between  $x$  and  $x+h$ , we have

$$n = \frac{h}{\chi(x)},$$

where  $\chi(x)$  is the average interval between two primes at the point  $x$ .

Thus

$$\frac{P(x+h)^2}{P(x^2)} = \frac{\chi(x^2)}{\chi(x+h)^2} = \left(1 - \frac{1}{x}\right)^{\frac{h}{\chi(x)}}.$$

This equation gives,  $x$  being very large,

$$\frac{\chi(x^2) + 2hx\chi'(x^2)}{\chi(x^2)} = 1 + \frac{h}{x\chi(x)};$$

whence

$$\chi(x^2) = 2x^2\chi(x)\chi'(x^2),$$

the same functional equation as that found in § 14, and which was satisfied by  $\chi(x) = \log x$ .

*Values of certain constants connected with prime numbers,*  
§§ 19–21.

§ 19. In the paper in the *Quarterly Journal* referred to in § 4, it was shown that

$$\begin{aligned} \log\left(\frac{2}{1} \cdot \frac{3}{2} \cdot \frac{5}{4} \cdot \frac{7}{6} \dots \frac{x}{x-1}\right) &= \frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \dots + \frac{1}{x} \\ &+ \frac{1}{2}\left(\frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \dots + \frac{1}{x^2}\right) \\ &+ \frac{1}{3}\left(\frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{5^3} + \frac{1}{7^3} + \dots + \frac{1}{x^3}\right), \\ &\qquad\qquad\qquad \&c., \qquad\qquad\qquad \&c. \end{aligned}$$

identically, and that, taking Riemann's formula for the number of primes inferior to  $x$ ,

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \dots + \frac{1}{x} = g + \log \log x - \frac{1}{2} \text{li } x^{-\frac{1}{2}} - \&c.,$$

$$\frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{5^2} + \dots + \frac{1}{x^2} = g_2 + \text{li } x^{-1} - \frac{1}{2} \text{li } x^{-\frac{3}{2}} - \&c.,$$

$$\frac{1}{2^3} + \frac{1}{3^3} + \frac{1}{5^3} + \dots + \frac{1}{x^3} = g_3 + \text{li } x^{-2} - \frac{1}{2} \text{li } x^{-\frac{5}{2}} - \&c.,$$

&c.,

&c.,

where  $g$  is an undetermined constant, and  $g_2, g_3, \dots$ , are constants whose numerical values have been calculated to 15 or more decimal places.

It is thus found that

$$\begin{aligned} \log\left(\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \dots \frac{x-1}{x}\right) &= -G - \log \log x + \frac{1}{2} \text{li } x^{-\frac{1}{2}} + \dots \\ &- \frac{1}{2} \text{li } x^{-1} + \frac{1}{4} \text{li } x^{-\frac{3}{2}} + \dots - \frac{1}{3} \text{li } x^{-2} + \frac{1}{6} \text{li } x^{-\frac{5}{2}} + \dots - \&c., \end{aligned}$$

where

$$G = g + \frac{1}{2}g_2 + \frac{1}{3}g_3 + \frac{1}{4}g_4 + \&c.$$

Thus approximately, when  $x$  is very large,

$$\log \left( \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots \frac{x-1}{x} \right) = -G - \log \log x,$$

and 
$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdots \frac{x-1}{x} = \frac{1}{a \log x},$$

where  $G = \log a.$

§ 20. Legendre represented the product

$$\frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots \frac{x-1}{x}$$

by the formula

$$\frac{A}{\log x - 0.08366},$$

and assigned to  $A$  the value 1.104, so that (as was pointed out in § 8 of the paper referred to) since  $a$  and  $A$  are connected by the relation

$$a = \frac{2}{A},$$

Legendre's value of  $A$  gives to  $a$  the value 1.812 and to  $G$  the value of 0.5944.

§ 21. But if the argument in §§ 5, 8, 9 of the present paper is correct, so that

$$P(x^2) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdots \frac{x-1}{x},$$

and 
$$P(x) = \phi'(x),$$

we have 
$$P(x^2) = \frac{1}{\log x^2}.$$

Thus,  $a = 2$  and  $G = \log 2.$

§ 22. If we may attribute this value to  $G$  we are enabled to assign the value of  $g$ , for

$$g = \log 2 - \frac{1}{2}g_2 - \frac{1}{3}g_3 - \frac{1}{4}g_4 - \&c.,$$

Now  $\log 2 = 0.6931\ 4718\ 0559\ 945,$

and  $\frac{1}{2}g_2 + \frac{1}{3}g_3 + \frac{1}{4}g_4 + \&c.$

$$= 0.3157\ 1845\ 2073\ 890.$$

(*Quart. Jour.*, *loc. cit.*, p. 373).

giving  $g = 0.3774\ 2872\ 8486\ 055.$

*Remarks on Legendre's investigation*, §§ 22, 23.

§ 22. Taking Legendre's result that the number of primes between  $x - m$  and  $x + m$  is

$$\frac{2m}{\log x - 1.08366},$$

it follows that

$$P(x^2) = \frac{1}{2 \log x - 1.08366}.$$

Now, if we may put

$$P(x^2) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \dots \left(1 - \frac{1}{x}\right),$$

we have, by equating these results,

$$\begin{aligned} \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \dots \left(1 - \frac{1}{x}\right) &= \frac{1}{2 \log x - 1.08366} \\ &= \frac{1}{2 \log x} \end{aligned}$$

omitting the constant 1.08366 compared to  $\log x$ .

§ 23. Legendre's own investigation, is however, in effect as follows:

$$\text{Let } f(x) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \dots \left(1 - \frac{1}{x}\right).$$

The next prime superior to  $x$  is  $x + \log x - 0.08366$ , or, say,  $x + \alpha$ ; therefore

$$f(x + \alpha) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \dots \left(1 - \frac{1}{x + \alpha}\right);$$

whence  $\frac{f(x + \alpha)}{f(x)} = 1 - \frac{1}{x + \alpha},$

and therefore

$$\frac{f(x+\alpha) - f(x)}{f(x)} = -\frac{1}{x} + \frac{\alpha}{x^2} - \&c.$$

giving 
$$\frac{df(x)}{f(x) dx} = -\frac{1}{\alpha x},$$

or, since 
$$d\alpha = \frac{dx}{x},$$

$$\frac{df(x)}{dx} = -\frac{d\alpha}{\alpha};$$

whence 
$$f(x) = \frac{A}{\alpha} = \frac{A}{\log x - 0.08366},$$

leaving  $A$  to be determined from Legendre's table of the values of  $f(x)$ .

*Remarks on the formula for  $P(x)$ , § 24.*

§ 24. It seems pretty evident that we are justified in regarding the probability that any large number  $x$  is prime as represented by the product

$$\frac{1}{2} \cdot \frac{2}{3} \cdot \frac{4}{5} \cdots \frac{r-1}{r}$$

(§ 5); for, in fact, in endeavouring to determine whether any given large number is or is not prime, we divide it by 7, 11, 13, &c., and the probability of the number not being divisible by these numbers is  $\frac{6}{7}$ ,  $\frac{10}{11}$ ,  $\frac{12}{13}$ , &c.

We may suppose that the large number  $x$  is given by the number of grains in a sack, and the number is prime if the grains do not admit of arrangement in groups of two, three, ..., up to  $r$ .

These arrangements are independent, and the chance that, after dividing the grains into groups of  $p$ , there will be some over is obviously  $\frac{p-1}{p}$ .

*The function  $\log x$ , § 25-27.*

§ 25. The result that at a point  $x$  in the series of ordinal numbers, the average distance between two consecutive primes is  $\log x$ , affords an interesting case of the occurrence of the logarithmic function in mathematics.

Another interesting appearance of the logarithmic function occurs in the formula

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{x} = \gamma + \log x + \frac{1}{2x} - \frac{1}{12x^2} + \&c.$$

Combining the two results we see that the average distance between two primes in the neighbourhood of any large number  $x$  is approximately equal to the sum of the reciprocals of the numbers from unity up to  $x$ .

It was shown by Lejeune-Dirichlet that the average number of the divisors of a large number  $x$  was  $\log x + \gamma$ . Thus, the average distance between two primes is approximately equal to the average number of the divisors of the numbers in the neighbourhood.

§ 26. When the numerals are expressed by a notation depending upon a radix (as in the decimal system),  $\log x$  is roughly proportional to the number of digits in the number  $x$ , diminished by unity. Thus, the intervals between successive primes in the neighbourhoods of different very large numbers are approximately proportional to the numbers of digits in those numbers. In the case of numbers expressed in the decimal scale, this interval is roughly equal to 2.30... times the number of digits in the number.

§ 27. When  $x$  is very large the relative magnitude of  $x$  and  $\log x$  may be represented by the number  $x$  itself, and 2.30... times the number of digits which express it. This gives a good idea of the relative infinities of  $x$  and  $\log x$ , when  $x$  is infinite. And it is interesting to notice that the series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{x}$$

is represented by an infinity of the very small order,  $\log x$ , and that the similar series involving primes only

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \dots + \frac{1}{x}$$

is represented by a correspondingly smaller infinity,  $\log \log x$ .