SOME PROBLEMS OF DIOPHANTINE APPROXIMATION.

BY

G. H. HARDY and J. E. LITTLEWOOD,

TRINITY COLLEGE, CAMBRIDGE.

I.

The fractional part of $n^k\theta$.

1.0 — Introduction.

1.00. Let us denote by $[x]$ and (x) the integral and fractional parts of x, so that

$$(x) = x - [x], \qquad 0 \leq (x) < 1.$$

Let θ be an irrational number, and α any number such that $0 \leq \alpha < 1$. Then it is well known that it is possible to find a sequence of positive integers $n_1, n_2, n_3, \cdots$ such that

$$(1.001) \qquad (n_r\theta) \rightarrow \alpha$$

as $r \rightarrow \infty$.

It is necessary to insert a few words of explanation as to the meaning to be attributed to relations such as (1.001), here and elsewhere in the paper, in the particular case in which $\alpha = 0$. The formula (1.001), when $\alpha > 0$, asserts that, given any positive number ε, we can find r_0 so that

$$-\varepsilon < (n_r\theta) - \alpha < \varepsilon \qquad (r > r_0).$$

The points $(n_r\theta)$ may lie *on either side of* α. But $(n_r\theta)$ is never negative, and so, in the particular case in which $\alpha = 0$, the formula, if interpreted in the obvious manner, asserts *more* than this, viz. that

$$0 \leq (n_r\theta) < \varepsilon \qquad (r > r_0).$$

The obvious interpretation therefore gives rise to a distinction between the value $\alpha = 0$ and other values of α which would be exceedingly inconvenient in our subsequent analysis.

These difficulties may be avoided by agreeing that, when $\alpha = 0$, the formula (1.001) is to be interpreted as meaning '*the set of points* $(n_r\theta)$ *has, as its sole limiting point or points, one or both of the points 1 and 0*', that is to say as impying that, for any r greater than r_0, *one or other* of the inequalities

$$0 \leq (n_r\theta) < \varepsilon, \quad 1 - \varepsilon < (n_r\theta) < 1$$

is satisfied. In the particular case alluded to above, this question of interpretation happens to be of no importance: our assertion is true on either interpretation. But in some of our later theorems the distinction is of vital importance.

Now let $f(n)$ denote a positive increasing function of n, integral when n is integral, such as

$$n, n^2, n^3, \cdots 2^n, 3^n, \cdots, n!, 2^{n^2}, \cdots 2^{2^n}, \cdots.$$

The result stated at the beginning suggests the following question, which seems to be of considerable interest: — *For what forms of $f(n)$ is it true that, for any irrational θ, and any value of α such that $0 \leq \alpha < 1$, a sequence (n_r) can be found such that*

$$(1.002) \qquad (f(n_r)\theta) \to \alpha\,?$$

It is easy to see that when the increase of $f(n)$ is sufficiently rapid the result suggested will not always be true. Thus if $f(n) = 2^n$ and θ is a number which, expressed in the binary scale, shows at least k 0's following upon every 1, it is plain that

$$(2^n\theta) < \frac{1}{2} + \lambda_k,$$

when λ_k is a number which can be made as small as we please by increasing k sufficiently. There is thus an »excluded interval» of values of α, the length of which can be made as near to $\frac{1}{2}$ as we please. If $f(n) = 3^n$ we can obtain an excluded interval whose length is as near to $\frac{2}{3}$ as we please, and so on; while if $f(n) = n!$ it is (as is well known) possible to choose θ so that $(n!\theta)$ tends to a unique limit. Thus $(n!e) \to 0$.

At the end of the paper we shall return to the general problem. The immediate object with which this paper was begun, however, was to determine whether the relation (1.002) always holds (if θ is irrational) when $f(n)$ is a power of n, and we shall be for the most part concerned with this special form of $f(n)$.

1.01. The following generalisation of the theorem expressed by (1.001) was first proved by KRONECKER.[1]

Theorem 1.01. *If $\theta_1, \theta_2, \cdots \theta_m$ are linearly independent irrationals (i. e. if no relation of the type*

$$a_1\theta_1 + a_2\theta_2 + \cdots + a_m\theta_m + a_{m+1} = 0,$$

where $a_1, a_2, \cdots a_{m+1}$ are integers, not all zero, holds between $\theta_1, \theta_2, \cdots \theta_m$), and α_1, $\alpha_2, \cdots \alpha_m$ are numbers such that $0 \leq \alpha_p < 1$, then a sequence (n_r) can be found such that

$$(n_r\theta_1) \to \alpha_1, (n_r\theta_2) \to \alpha_2, \cdots, (n_r\theta_m) \to \alpha_m,$$

as $r \to \infty$. Further, in the special case when all the α's are zero, it is unnecessary to make any restrictive hypothesis concerning the θ's, or even to suppose them irrational.

This theorem at once suggests that the solution of the problem stated at the end of 1.00 may be generalised as follows.

Theorem 1.011. *If $\theta_1, \theta_2, \cdots \theta_m$ are linearly independent irrationals, and the α's are any numbers such that $0 \leq \alpha < 1$, then a sequence (n_r) can be found such that*

$$(1.011) \qquad \begin{cases} (n_r\theta_1) \to \alpha_{11}, (n_r\theta_2) \to \alpha_{12}, \cdots, (n_r\theta_m) \to \alpha_{1m}; \\ (n_r^2\theta_1) \to \alpha_{21}, (n_r^2\theta_2) \to \alpha_{22}, \cdots, (n_r^2\theta_m) \to \alpha_{2m}; \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ (n_r^k\theta_1) \to \alpha_{k1}, (n_r^k\theta_2) \to \alpha_{k2}, \cdots, (n_r^k\theta_m) \to \alpha_{km}. \end{cases}$$

---

[1] KRONECKER, *Berliner Sitzungsberichte,* 11 Dec. 1884; *Werke,* vol. 3, p. 49.

A number of special cases of the theorem were known before. That in which all the α's are zero was given by DIRICHLET (*Berliner Sitzungsberichte,* 14 April 1842, *Werke,* vol. 1, p. 635). Who first stated explicitly the special theorems in which $m = 1$ we have been unable to discover. DIRICHLET (l. c.) refers to the simplest as »längst bekannt»: it is of course an immediate consequence of the elementary theory of simple continued fractions. See also MINKOWSKI, »*Diophantische Approximation*», pp. 2, 7. KRONECKER's general theorem has been rediscovered independently by several writers. See *e. g.* BOREL, *Leçons sur les séries divergentes,* p. 135; F. RIESZ, *Comptes Rendus,* 29 Aug. 1904. Some of the ideas of which we make most use are very similar to those of the latter paper. It should be added that DIRICHLET's and KRONECKER's theorems are presented by them merely as particular cases of more general theorems, which however represent extensions of the theory in a direction different from that with which we are concerned.

A number of very beautiful applications of KRONECKER's theorem to the theory of the RIEMANN ζ-function have been made by H. BOHR.

Further, if the α's are all zero, it is unnecessary to suppose the θ's restricted in any way.

1.02. This theorem is the principal result of the paper: it is proved in section 1.2. The remainder of the paper falls into three parts. The first of these (section 1.1) consists of a discussion and proof of KRONECKER's theorem. We have thought it worth while to devote some space to this for two reasons. In the first place our proof of theorem **1.011** proceeds by induction from k to $k+1$, and it seems desirable for the sake of completeness to give some account of the methods by which the theorem is established in the case $k=1$. In the second place the theorem for this case possesses an interest and importance sufficient to justify any attempt to throw new light upon it; and the ideas involved in the various proofs which we shall discuss are such as are important in the further developments of the theory. We believe, moreover, that the proof we give is considerably simpler than any hitherto published.

The second of the remaining parts of the paper (section 1.3) is devoted to the question of the *rapidity* with which the numbers $(n^{\varkappa}\theta_p)$ in the scheme (1.011) tend to their respective limits. Our discussion of the problems of this section is very tentative, and the results very incomplete;[1] and something of the same kind may be felt about the paper as a whole. We have not solved the problems which we attack in this paper with anything like the definiteness with which we solve those to which our second paper is devoted. The fact is, however, that the first paper deals with questions which, in spite of their more elementary appearance, are in reality far more difficult than those of the second. Finally, the last section (1.4) contains some results the investigation of which was suggested to us by an interesting theorem proved by F. BERNSTEIN.[2] The distinguishing features of these results are that they are concerned with a single irrational θ and with sequences which are not of the form $(n^k\theta)$, and that they hold for *almost all* values of θ, *i. e.* for all values except those which belong to an exceptional and unspecified set of measure zero.

1.1 — Kronecker's Theorem.

1.10. KRONECKER's theorem falls naturally into two cases, according as to whether or not all the α's are zero. We begin by considering the simpler case,

[1] Some of the results that we do obtain, however, are important from the point of view of applications to the theory of the series $\Sigma e^{n^k\theta i}$ and that of the RIEMANN ζ-function. It was in part the possibility of these applications that led us to the researches whose results are given in the present paper. The applications themselves will, we hope, be given in a later paper.

[2] *Math. Annalen*, vol. 71, p. 421.

when all the α's are zero. Unlike most of the theorems with which we are concerned, this is not proved by induction, and there is practically no difference between the cases of one and of several variables. The proof given is DIRICHLET'S.

Let $\bar{x}$ denote the number which differs from x by an integer and which is such that $-\frac{1}{2} < \bar{x} \leq \frac{1}{2}$. Then the theorem to be proved is equivalent to the theorem that, given any integers q and N, we can find an n not less than N and such that

$$|\overline{n\theta_1}| \leq 1/q,\ |\overline{n\theta_2}| \leq 1/q,\ \cdots,\ |\overline{n\theta_m}| \leq 1/q.$$

Let us first suppose that $N = 1$. Let R be the region in m-dimensional space for which each coordinate ranges from 0 to 1. Let the range of each coordinate be divided into q equal parts: R is then divided into q^m parts. Consider now the $q^m + 1$ points

$$(\nu\theta_1),\ (\nu\theta_2),\ \cdots,\ (\nu\theta_m);\ (\nu = 0, 1, 2, \cdots q^m).$$

There must be one part of R which contains two points; let the corresponding values of ν be ν_1 and ν_2. Then clearly

$$|\overline{(\nu_1 - \nu_2)\theta_1}| \leq 1/q,\ |\overline{(\nu_1 - \nu_2)\theta_2}| \leq 1/q,\ \cdots,|\overline{(\nu_1 - \nu_2)\theta_m}| \leq 1/q,$$

and
$$|\nu_1 - \nu_2| \geq 1.$$

We have therefor only to take $n = |\nu_1 - \nu_2|$. We observe that we have also

$$n \leq q^m,$$

a result to which we shall have occasion to return in section 1.3.

If $N > 1$ we have only to consider the points $(\nu N\theta_1), (\nu N\theta_2), \cdots$ instead of the points $(\nu\theta_1), (\nu\theta_2), \cdots$.

1.11. We turn now to the case when the α's are not all necessarily zero. In this case the necessity of the hypothesis that the θ's are linearly independent is obvious, for the existence of a linear relation between the θ's would plainly involve that of a corresponding relation between the α's; naturally, also, the added restriction makes the theorem much more difficult than the one just proved.

Our proof proceeds by induction from m to $m + 1$; it is therefore important to discuss the case $m = 1$. The result for this case may be proved in a variety of ways, of which we select four which seem to us to be worthy of separate discussion. These proofs are all simple, and each has special advantages of its own. It is important for us to consider very carefully the ideas involved in them with a view to selecting those which lend themselves most readily to generalisation. For example, it is essential that our proof should make no appeal to the theory of continued fractions.

(a). The first proof is due to KRONECKER. It follows from the result of 1.10, with $m=1$, or from the theory of continued fractions, that we can find an arbitrarily large q such that

$$\theta=\frac{p}{q}+\frac{\delta}{q^2},$$

and so

$$q\theta-p=\delta/q. \tag{1.111}$$

where

$$|\delta|<1.$$

It is possible to express any integer, and in particular the integer $\{q\alpha\}$ nearest to $q\alpha$, in the form

$$q n_1+p n$$

where n and n_1 are integers, and $|n|\leq q/2$. From the two equations

$$q\theta-p=\delta/q,\ q n_1+p n=\{q\alpha\}$$

we obtain

$$q(n\theta+n_1)=\frac{n\delta}{q}+q\alpha+\frac{1}{2}\delta_1, \qquad |\delta_1|<1,$$

and so

$$-1<q(n\theta+n_1-\alpha)<1,$$

or

$$|(n\theta)-\alpha|<1/q.$$

If we write $\nu=n+q$ and use (1.111), we see that

$$|(\nu\theta)-\alpha|<2/q, \qquad q/2<\nu<3q/2;$$

so that

$$|(\nu\theta)-\alpha|<3/\nu$$

for some value of ν between $q/2$ and $3q/2$. This evidently establishes the truth of the theorem.

If we attempt to extend this proof to the case of several variables we find nothing to correspond to the equation

$$\{q\alpha\} = q n_1 + p n.$$

But KRONECKER'S proof has, as against the proofs we shall now discuss, the very important advantage of furnishing a definite result as to the order of the approximation, a point to which we shall return in 1.3.

(*b*). Let ε be an arbitrary positive constant. By the result of 1.10, we can find an n such that $0 < \theta_1 < \varepsilon$ or $1 - \varepsilon < \theta_1 < 1$, where $\theta_1 = (n\theta)$. Since θ is irrational, θ_1 is not zero. Let us suppose that $0 < \theta_1 < \varepsilon$; the argument is substantially the same in the other case. We can find an m such that

$$m\theta_1 \leq \alpha < (m+1)\theta_1,$$

$$|m\theta_1 - \alpha| < \theta_1;$$

and so

$$|(nm\theta) - \alpha| < \varepsilon,$$

which proves the theorem.

(*c*).[1] Let S denote the set of points $(n\theta)$. S', its first derived set, is closed. It is moreover plain that, if α is not a point of S', then neither is $(\alpha + n\theta)$ nor $(\alpha - n\theta)$.

The theorem to be proved is clearly equivalent to the theorem that S' consists of the continuum $(0,1)$. Suppose that this last theorem is false. Then there is a point α which is not a point of S', and therefore an interval containing α and containing[2] no point of S'. Consider I, the greatest possible such interval containing α.[3] The interval obtained by translating I through a distance θ, any number of times in either direction,[4] must, by what was said above, also contain no point of S'. But the interval thus obtained cannot overlap with I, for then I would not be the »greatest possible» interval of its kind.

[1] This proof was discovered independently by F. RIESZ, but, so far as we know, has not been published.

[2] In its interior, in the strict sense.

[3] The existence of such a »greatest possible» interval is easily established by the classical argument of DEDEKIND.

[4] Taking the congruent interval in $(0, 1)$. This interval may possibly consist of two separate portions $(0, \xi_1)$, and $(\xi_2, 1)$.

Hence, if we consider a series of $[1/\delta]$ translations, where δ is the length of I, it is clear that two of the corresponding $[1/\delta]+1$ intervals must coincide. Clearly this can only happen if θ is rational, which is contrary to our hypothesis.

(*d*). We argue as before that, if the theorem is false, there is an interval I, of length 2ε and middle point α, containing no point of S'. By the result of 1.10 we can find n so that, if $\theta_1 = (n\theta)$, then $0 < \theta_1 < \varepsilon$ or $1 - \varepsilon < \theta_1 < 1$.

By the reasoning used in (*c*) it appears that the interval obtained by translating I through a distance θ_1, any number of times in either direction, must contain no point of S'. But since each new interval overlaps with the preceding one it is clear that after a certain number of translations we shall have covered the whole interval 0 to 1 by intervals containing no point of S', and shall thus have arrived at a contradiction.

1.12. Let us compare the three last proofs. It is clear that (*b*) is considerably the simplest, and that (*d*) appears to contain the essential idea of (*b*) together with added difficulties of its own. It appears also that, in point of simplicity, there is not very much to choose between (*c*) and (*d*), and that (*c*) has a theoretical advantage over (*d*) in that it dispenses the assumption of the theorem for the case $\alpha = 0$, an assumption which is made not only in (*b*) and (*d*), but also in (*a*). When, however, we consider the theorem for several variables, it seems that (*b*) does not lend itself to direct extension at all, that the complexity of the region corresponding to I in (*c*) leads to serious difficulties, and that (*d*) provides the simplest line of argument. It is accordingly this line of argument which we shall follow in our discussion of the general case of KRONECKER's theorem.

1.13. We pass now to the general case of KRONECKER's theorem. We shall give a proof by induction. For the sake of simplicity of exposition we shall deduce the theorems for three independent irrationals θ, φ, ψ, from that for two. It will be obvious that the same proof gives the general induction from n to $n+1$ irrationals.

We wish to show that *if we form the set S of* points within the cube $0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 1$, which are congruent with

$$(\theta, \varphi, \psi),\ (2\theta, 2\varphi, 2\psi), \cdots\cdots (n\theta, n\varphi, n\psi), \cdots$$

then every point of the cube is a point of the first derived set S'. It is plain that, if (α, β, γ) is not a point of S, then neither is $((\alpha + n\theta), (\beta + n\varphi), (\gamma + n\psi))$ nor $((\alpha - n\theta), (\beta - n\varphi), (\gamma - n\psi))$. If now our theorem is not true, there must exist a sphere, of centre (α, β, γ) and radius ϱ, which contains[1] no point of S'. By

[1] Within or upon the boundary.

the result of 1.10, there is an n such that the distance δ of $((n\theta), (n\varphi), (n\psi))$ or $(\theta_1, \varphi_1, \psi_1)$ from one of the vertices of the cube is less than $\varrho/\sqrt{2}$. Let us suppose, for example, that the vertex in question is the point (0, 0, 0). Consider the straight line

$$(1.131) \qquad \frac{x-\alpha}{\theta_1} = \frac{y-\beta}{\varphi_1} = \frac{z-\gamma}{\psi_1},$$

and the infinite cylinder of radius δ with this line as axis. It is clear that the finite cylinder C obtained by taking a length δ on either side of (α, β, γ) is entirely contained in the sphere and therefore contains no point of S'. Hence the cylinder obtained by translating C through $(\theta_1, \varphi_1, \psi_1)$, any number of times in either direction, also contains no point of S', so that, since each new position of C overlaps with the preceding, the whole of the infinite cylinder, or rather of the congruent portions of the cube, is free from points of S'.

Let us now consider the intersections of the totality of straight lines in the cube, which are congruent with portions of the axis of the cylinder, with an arbitrary plane $x = x_0$. We shall show that they are everywhere dense in the square in which the plane cuts the cube, whence clearly follows that no point of the cube is a point of S', and so a contradiction which establishes the theorem.

The intersections (y, z) are congruent with the intersections of the axis (1.131) with

$$x = x_0 + \nu, \quad (\nu = \cdots, -2, -1, 0, 1, 2, \cdots),$$

and so they are the points congruent with

$$\beta + \frac{(x_0-\alpha)\varphi_1}{\theta_1} + \frac{\nu\varphi_1}{\theta_1}, \quad \gamma + \frac{(x_0-\alpha)\psi_1}{\theta_1} + \frac{\nu\psi_1}{\theta_1}.$$

But, under our hypothesis, φ_1/θ_1 and ψ_1/θ_1 are linearly independent irrationals, and so, by the theorem for two irrationals, this set of points is everywhere dense in the square. The proof is thus completed.

1.14. We add two further remarks on the subject of KRONECKER's theorem, in which, for the sake of simplicity of statement, we confine ourselves to the case of two linearly independent irrationals θ, φ.

(*a*) Suppose that $0 < \alpha < 1$, $0 < \beta < 1$. KRONECKER's theorem asserts the existence of a sequence (n_s) such that $(n_s\theta) \to \alpha$, $(n_s\varphi) \to \beta$. Let us choose a sequence of points

$$(\alpha_\mu, \beta_\mu), \quad (\mu = 1, 2, 3, \cdots),$$

such that

$$\alpha_\mu > \alpha, \ \beta_\mu > \beta, \quad \alpha_\mu \to \alpha, \ \beta_\mu \to \beta.$$

There is, for any value of μ, a sequence $(n_{s\mu})$ such that

$$(n_{s\mu}\theta) \to \alpha_\mu, \ (n_{s\mu}\varphi) \to \beta_\mu,$$

as $s \to \infty$. From this it is easy to deduce the existence of a sequence (n_r) for which $(n_r\theta)$ and $(n_r\varphi)$ tend to the limits α and β and are always greater than those limits, so that the direction of approach to the limit is in each case from the right hand side.[1] Similarly, of course, we can establish the existence of a sequence giving, for either θ or φ, either a right-handed or a left-handed approach to the limit.

If we apply similar reasoning to the case in which α or β or both are zero we see that, when θ and φ are linearly independent irrationals, we may abandon the convention with respect to the particular value 0 which was adopted in 1 . 00, and assert that there is a sequence for which $(n\theta) \to \alpha$ and $(n\varphi) \to \beta$, α and β having any values between 0 and 1, both values included, and the formulae having the ordinary interpretation. This result is to be carefully distinguished from that of 1 . 10. The latter is, the former is *not*, true without restriction on the θ's, as may be seen at once by considering the case in which $\varphi = -\theta$.

(b). It is easy to deduce from KRONECKER's theorem a further theorem, which may be stated as follows:[2] *if we take any portion γ of the square $0 \leq x \leq 1$, $0 \leq y \leq 1$, bounded by a finite number of regular curves, and of area δ; and if we denote by $N_\gamma(n)$ the number of the points*

$$((\nu\theta), (\nu\varphi)), \qquad (\nu = 1, 2, \cdots n),$$

which fall inside γ; then

$$N_\gamma(n) \sim \delta n$$

as $n \to \infty$.

This result, when compared with the various theorems of this paper, suggests a whole series of further theorems. The proofs of these appear likely to be very difficult, and we have, up to the present, considered only the case of a single irrational θ. We have proved that, if $N_\gamma(n)$ denotes the number of the points

$$(\nu^\varkappa\theta), \qquad (\nu = 1, 2, \cdots n),$$

[1] The reasoning by which this is established is essentially the same as that of 1 . 20.

[2] This is a known theorem. For a proof and references see the tract 'The Riemann Zeta-function and the Theory of Prime Numbers', by H. BOHR and J. E. LITTLEWOOD, shortly to be published in the *Cambridge Tracts in Mathematics and Mathematical Physics*.

which fall inside a segment γ of (0, 1), of length δ, then $N_\gamma(n) \sim \delta n$. This result may be compared with that of Theorem **1.483** at the end of the paper. But results of this character will find a more natural place among our later investigations than among those of which we are now giving an account.

1.2. — The generalisation of Kronecker's theorem.

1.20. We proceed now to the proof of theorem **1.011**. Our argument is based on the following general principle, which results from the work of PRINGSHEIM and LONDON on double sequences and series.[1]

1.20. *If*

$$\lim_{r_1 \to \infty} \lim_{r_2 \to \infty} \cdots \lim_{r_k \to \infty} f_p(r_1, r_2, \cdots r_k) = A_p, \; (p = 1, 2, \cdots m),$$

then we can find a sequence of sets $(r_{1n}, r_{2n}, \cdots r_{kn})$ *such that, as* $n \to \infty$,

$$r_{qn} \to \infty, \; (q = 1, 2, \cdots k),$$

and

$$f_p(r_{1n}, r_{2n}, \cdots r_{kn}) \to A_p, \; (p = 1, 2, \cdots m).$$

We shall show that, if this principle is true for all values of m and a particular k, then it is true for $k + 1$. As it is plainly true for $k = 1$, we shall thus have proved it generally.

We shall abbreviate '$\lim\limits_{r_1 \to \infty} \lim\limits_{r_2 \to \infty} \cdots \lim\limits_{r_k \to \infty}$' into '$\lim\limits_{r_1, r_2, \cdots r_k}$', or, when there is no danger of confusion, into 'lim'.

Let

$$\lim_{r_1, r_2, \cdots r_k} f_p(r_1, r_2, \cdots r_{k+1}) = f_p(r_{k+1}).$$

Then by hypothesis

$$f_p(r_{k+1}) \to A_p$$

as $r_{k+1} \to \infty$. Let us choose an integer $r_{k+1, n}$, greater than 2^n, for which

$$|f_p(r_{k+1, n}) - A_p| < 2^{-n-1}, \quad (p = 1, 2, \cdots m).$$

By the principle for k variables, we can find $r_{1n}, r_{2n}, \cdots r_{kn}$, all greater than 2^n, and such that

[1] PRINGSHEIM, *Münchener Sitzungsberichte*, vol. 27, p. 101, and *Math. Annalen*, vol. 53, p. 289; LONDON, *Math. Annalen*, vol. 53, p. 322.

$$|f_p(r_{1n}, r_{2n}, \cdots r_{kn}, r_{k+1,n}) - f_p(r_{k+1,n})| < 2^{-n-1}, \ (p = 1, 2, \cdots m).$$

We thus obtain a sequence of sets $(r_{1n}, r_{2n}, \cdots r_{k+1,n})$, such that every member of the n^{th} set is greater than 2^n and

$$|f_p(r_{1n}, r_{2n}, \cdots r_{k+1,n}) - A_p| < 2^{-n}, \quad (p = 1, 2, \cdots m).$$

This sequence evidently gives us what we want.

An important special case of the principle is the following:

1.201. *If for all values of t we can find a sequence $n_{1t}, n_{2t}, \cdots, n_{rt}, \cdots$ such that*

$$f_p(n_{rt}) \to A_{pt}, \quad (p = 1, 2, \cdots m),$$

as $r \to \infty$, and if

$$A_{pt} \to A_p, \quad (p = 1, 2, \cdots m),$$

as $t \to \infty$, then there is a sequence (n_s) such that

$$f_p(n_s) \to A_p, \quad (p = 1, 2, \cdots m),$$

as $s \to \infty$.

This is in reality merely a case of the principle that a limiting-point of limiting-points is a limiting-point.

1.21. We consider first the case in which all the α's are zero, and the θ's are unrestricted. In this case the proof is comparatively simple.

Theorem 1.21. *There is a sequence (n_r) such that, as $r \to \infty$*

$$(n_r^\varkappa \theta_p) \to 0, \quad (\varkappa = 1, 2, \cdots k; \ p = 1, 2, \cdots m).$$

We prove this theorem by induction from k to $k + 1$: we have seen that it is true when $k = 1$. We suppose then that there is a sequence (μ_s) such that

$$(\mu_s^\varkappa \theta_p) \to 0, \quad (\varkappa = 1, 2, \cdots k; \ p = 1, 2, \cdots m). \tag{1.211}$$

The sequence

$$(\mu_s^{k+1}\theta_1), \ (\mu_s^{k+1}\theta_2), \cdots (\mu_s^{k+1}\theta_m), \quad (s = 1, 2, \cdots),$$

has at least one limiting point $\varphi_1, \varphi_2, \cdots \varphi_m$; hence, by restricting ourselves to a subsequence selected from the sequence (μ_s), we can obtain a sequence (ν_s) such that, as $s \to \infty$,

$$(\nu_s^\varkappa \theta_p) \to 0, \ (\varkappa \leq k); \ (\nu_s^{k+1}\theta_p) \to \varphi_p; \ (p = 1, 2, \cdots m).$$

We then have, for $\varkappa \leq k + 1$,

$$\lim_{s_1, s_2, \cdots s_\lambda} ((\nu_{s_1} + \nu_{s_2} + \cdots + \nu_{s_\lambda})^\varkappa \theta_p) = \sum_{q=1}^{\lambda} \lim (\nu_{s_q}^\varkappa \theta_p) + \sum C \lim (\nu_{s_1}^{\varkappa_1} \nu_{s_2}^{\varkappa_2} \cdots \nu_{s_\lambda}^{\varkappa_\lambda} \theta_p),$$

where the C's are constants, $\varkappa_1 + \varkappa_2 + \cdots + \varkappa_\lambda = \varkappa$, and $\varkappa_q \leq k$. In virtue of (1.211) we can evaluate at once every repeated limit on the right hand side, and it is clear that we obtain $\lambda \varphi_p$ or 0 according as $\varkappa = k + 1$ or $\varkappa \leq k$. It follows from the general principle **1.20** that we can find a sequence $(n_{r\lambda})$, $(r = 1, 2, \cdots)$, such that, as $r \to \infty$,

$$(n_{r\lambda}^\varkappa \theta_p) \to 0, \ (\varkappa \leq k); \ (n_{r\lambda}^{k+1} \theta_p) \to \lambda \varphi_p; \ (p = 1, 2, \cdots m).$$

But, by theorem **1.01**, we can find a sequence (λ_s) such that

$$(\lambda_s \varphi_p) \to 0, \quad (p = 1, 2, \cdots m);$$

and we have only to apply the principle **1.201** to obtain the theorem for $k + 1$.

1.22. We pass now to the general case when the α's are not all zero. We have to prove that *if $\theta_1, \theta_2, \cdots \theta_m$ are linearly independent irrationals, there is a sequence (n_r) such that, as $r \to \infty$,*

$$(n_r^\varkappa \theta_p) \to \alpha_{\varkappa p}, \ (\varkappa = 1, 2, \cdots k; \ p = 1, 2, \cdots m).$$

We shall prove this by an induction from k to $k + 1$ which proceeds by two steps.

(i). We assume the existence, for a particular k, any number m of θ's, and any corresponding system of α's, of a sequence giving the scheme of limits

$$\begin{array}{c|ccccccc}
 & \theta_1 & \theta_2 & . & . & . & . & \theta_m \\
\hline
n & \alpha_{11} & \alpha_{12} & . & . & . & . & \alpha_{1m} \\
n^2 & \alpha_{21} & \alpha_{22} & . & . & . & . & \alpha_{2m} \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . \\
n^k & \alpha_{k1} & \alpha_{k2} & . & . & . & . & \alpha_{km}
\end{array} \quad ;$$

and we prove the existence, for any number m of θ's, and any corresponding system of α's, of a sequence giving the scheme

$$
\begin{array}{c|cccc}
 & \theta_1 & \theta_2 & \cdots & \theta_m \\
\hline
n & \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1m} \\
n^2 & \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2m} \\
\cdot & \cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdot & \cdots & \cdot \\
n^k & \alpha_{k1} & \alpha_{k2} & \cdots & \alpha_{km} \\
n^{k+1} & 0 & 0 & \cdots & 0 .
\end{array}
$$

It will be understood that neither m, nor the θ's, nor the α's are necessarily the same in these two schemes, all of them being arbitrary.

(ii). We then show that we can pass from the last written scheme of limits to the general scheme in which the elements of the last row also are arbitrary.

1.23. *Proof of the first step.* To fix our ideas we shall show that we can pass from a sequence (n_r) giving[1]

$$n_r\theta \to \alpha_1,\ n_r\varphi \to \beta_1,\ n_r\psi \to \gamma_1,\ n_r\chi \to \delta_1,\ n_r\omega \to \eta_1,\ n_r\tau \to \zeta_1,$$

$$n_r^2\theta \to \alpha_2,\ n_r^2\varphi \to \beta_2,\ n_r^2\psi \to \gamma_2,\ n_r^2\chi \to \delta_2,\ n_r^2\omega \to \eta_2,\ n_r^2\tau \to \zeta_2,$$

to a sequence (m_r) giving

$$m_r\theta \to \alpha_1,\ m_r\varphi \to \beta_1,$$

$$m_r^2\theta \to \alpha_2,\ m_r^2\varphi \to \beta_2,$$

$$m_r^3\theta \to 0,\ m_r^3\varphi \to 0.$$

It will be clear that the argument is in reality of a perfectly general type.

Suppose we are given $\alpha'_1, \alpha'_2, \beta'_1, \beta'_2$, and that $\theta, \varphi, \alpha'_1, \alpha'_2, \beta'_1, \beta'_2$, are linearly independent irrationals. Then by hypothesis we can find a sequence giving the scheme

$$n_r\theta \to \alpha'_1,\ n_r\varphi \to \beta'_1,\ n_r\alpha'_1 \to 0,\ n_r\beta'_1 \to 0,\ n_r\alpha'_2 \to 0,\ n_r\beta'_2 \to 0,$$

$$n_r^2\theta \to \alpha'_2,\ n_r^2\varphi \to \beta'_2,\ n_r^2\alpha'_1 \to 0,\ n_r^2\beta'_1 \to 0.$$

Further, the set of points $(n_r^3\theta, n_r^3\varphi)$ has at least one limiting-point (λ, μ), and, by restricting ourselves to a subsequence of (n_r), we may suppose that we have also

$$n_r^3\theta \to \lambda,\ n_r^3\varphi \to \mu.$$

[1] In what follows we shall omit the brackets in $(n\theta, \ldots$; it is of course to be understood that integers are to be ignored.

We express all this by saying that we can find a sequence (n_r) giving the scheme

$$
(1.221) \qquad \begin{cases} \alpha'_1, \ \beta'_1, \ 0, \ 0, \ 0, \ 0, \\ \alpha'_2, \ \beta'_2, \ 0, \ 0, \\ \lambda, \ \ \mu. \end{cases}
$$

The sequence (k_r), where $k_r = 2 n_r$, gives us the scheme

$$
(1.222) \qquad \begin{cases} 2\alpha'_1, \ 2\beta'_1, \ 0, \ 0, \ 0, \ 0, \\ 4\alpha'_2, \ 4\beta'_2, \ 0, \ 0, \\ 8\lambda, \ \ 8\mu. \end{cases}
$$

By the general principle **1.20**, we can find a sequence (l_r) giving the scheme

$$
\begin{cases} \lim (n_{r_1} + n_{r_2} + \cdots + n_{r_s})\, \theta, \ \lim (n_{r_1} + n_{r_2} + \cdots + n_{r_s})\, \varphi, \ \ldots . \\ \lim (n_{r_1} + n_{r_2} + \cdots + n_{r_s})^2 \theta, \qquad \ldots . \qquad , \ \ldots . \\ \lim (n_{r_1} + n_{r_2} + \cdots + n_{r_s})^3 \theta, \qquad \ldots . \qquad , \end{cases}
$$

where 'lim' stands for $\lim\limits_{r_1, r_2, \cdots r_s}$.

Consider the repeated limit

$$
\lim_{r_1, r_2, \cdots r_s} (n_{r_1} + n_{r_2} + \cdots + n_{r_s})^3 \theta,
$$

which is easily evaluated with the aid of the table (1.221). The limit of a term $n_{r_i}^3 \theta$ is λ: that of a 'cross-term'

$$
n_{r_i}^a n_{r_j}^b n_{r_k}^c \theta \quad (a + b + c = 3;\ a, b, c < 3;\ i < j < k)
$$

is zero, since $n_{r_k}^c \theta$ tends to an α' or a β', and $n_{r_j}^b \alpha'$ and $n_{r_j}^b \beta'$ tend to zero. Thus we obtain the repeated limit 8λ. In all the other repeated limits the cross-terms give zero in the same way, and we see that the sequence (l_r) gives the scheme

$$
\begin{array}{l} 8\alpha'_1, \ 8\beta'_1, \ 0, \ 0, \ 0, \ 0, \\ 8\alpha'_2, \ 8\beta'_2, \ 0, \ 0, \\ 8\lambda \ , \ 8\mu. \end{array}
$$

Consider now the repeated limits

$$\lim_{r, r_1, \cdots r_m} (l_r + k_{r_1} + k_{r_2} + \cdots + k_{r_m})^\varkappa \chi$$

where

$$\chi = \theta, \varphi, \alpha'_1, \beta'_1, \alpha'_2, \beta'_2; \ \varkappa = 1, 2, 3.$$

All the cross-terms contribute zero as before, and we obtain the scheme

$$\begin{array}{l} (8 + 2m)\alpha'_1, \ (8 + 2m)\beta'_1, \ 0, 0, 0, 0, \\ (8 + 4m)\alpha'_2, \ (8 + 4m)\beta'_2, \ 0, 0, \\ (8 + 8m)\lambda \ \ , \ (8 + 8m)\mu, \end{array}$$

or

$$\begin{array}{l} 6\alpha'_1 + (m + 1)2\alpha'_1, \ 6\beta'_1 + (m + 1)2\beta'_1, \ 0, 0, 0, 0, \\ 4\alpha'_2 + (m + 1)4\alpha'_2, \ 4\beta'_2 + (m + 1)4\beta'_2, \ 0, 0, \\ \qquad (m + 1)8\lambda \ \ , \qquad\quad (m + 1)8\mu. \end{array}$$

It is possible, then, to find a sequence giving this scheme. But now, since it is possible to find a sequence of m's such that

$$(m + 1)\psi \to 0, \ (\psi = 2\alpha'_1, 2\beta'_1, 4\alpha'_2, 4\beta'_2, 8\lambda, 8\mu),$$

it follows (in virtue of the principle **1.20**) that we can find a sequence giving the scheme

$$\begin{array}{l} 6\alpha'_1, 6\beta'_1, \ 0, 0, 0, 0, \\ 4\alpha'_2, 4\beta'_2, \ 0, 0, \\ \ \ 0 \ , \ \ 0. \end{array}$$

This gives us what we want (and something more) provided it is possible to choose

$$\alpha'_1 = \frac{1}{6}\alpha_1, \ \beta'_1 = \frac{1}{6}\beta_1, \ \alpha'_2 = \frac{1}{4}\alpha_2, \ \beta'_2 = \frac{1}{4}\beta_2.$$

This is the case provided $\theta, \varphi, \alpha_1, \beta_1, \alpha_2, \beta_2$ are linearly independent irrationals: it remains only to show that this restriction on $\alpha_1, \beta_1, \alpha_2, \beta_2$ may be removed. It is obvious, in virtue of the principle **1.20**, that this may be done provided we can find a sequence $(\alpha_{1n}, \beta_{1n}, \alpha_{2n}, \beta_{2n})$ such that, for each n, $\theta, \varphi, \alpha_{1n}, \beta_{1n}, \alpha_{2n}, \beta_{2n}$ are linearly independent irrationals, and such that

$$\alpha_{1n} \to \alpha_1, \ \beta_{1n} \to \beta_1, \ \alpha_{2n} \to \alpha_2, \ \beta_{2n} \to \beta_2.$$

Now it is easy to see that there must be points $(\alpha_{1n}, \beta_{1n}, \alpha_{2n}, \beta_{2n})$ interior to the 'cube' with $(\alpha_1, \beta_1, \alpha_2, \beta_2)$ as centre and of side 2^{-n}, and exterior to that with the same centre and of side 2^{-n-1}, and such that $\theta, \varphi, \alpha_{1n}, \beta_{1n}, \alpha_{2n}, \beta_{2n}$ are linearly independent irrationals. By selecting one such point corresponding to each value of n we obtain a sequence of the kind desired.[1]

1.24. *Proof of the second step.* Here also we shall consider a special case for simplicity: the argument is really general. We shall show that we can pass from a sequence giving the scheme

$$\begin{array}{c|cccc} & \theta & \varphi & \psi & \chi \\ \hline n & \alpha_1 & \beta_1 & \gamma_1 & \delta_1 \\ n^2 & \alpha_2 & \beta_2 & \gamma_2 & \delta_2 \\ n^3 & 0 & 0 & & \end{array}$$

to one giving

$$\begin{array}{c|cc} & \theta & \varphi \\ \hline n & \alpha_1 & \beta_1 \\ n^2 & \alpha_2 & \beta_2 \\ n^3 & \alpha_3 & \beta_3. \end{array}$$

As in 1.23, we may suppose, without real loss of generality, that $\theta, \varphi, \alpha_1, \beta_1$ are linearly independent irrationals. Let (n_r) be a sequence giving

$$\begin{array}{c|cccc} & \theta & \varphi & \alpha_1 & \beta_1 \\ \hline n & \frac{1}{2}\alpha_1 & \frac{1}{2}\beta_1 & \alpha_2 & \beta_2 \\ n^2 & 0 & 0 & \frac{2}{3}\alpha_3 & \frac{2}{3}\beta_3 \\ n^3 & 0 & 0 & & \end{array}$$

---

[1] This argument depends ostensibly on ZERMELO's 'Auswahlsprinzip' (or WHITEHEAD and RUSSELL's 'Multiplicative Axiom'). This difficulty can however be surmounted with a little trouble. It should perhaps be observed that we have ignored several similar points early in the paper: in all of these the difficulty is comparatively trivial, and we have only called attention to it in the present instance because it occurs in a more serious form than is usual in constructive mathematics.

An alternative line of argument from that in the text proceeds as follows. It is easy to show that if at most a finite number of primes are omitted, any four of the sequence $\log 2, \log 3, \log 5, \log 7, \log 11, \ldots$, together with θ and φ, form a set of six linearly independent irrationals. Moreover it can be deduced from known results concerning the distribution of the primes that we can find a sequence $(\log p_n, \log q_n, \log r_n, \log s_n)$, where p_n, q_n, r_n, and s_n are primes, such that

$$(\log p_n) \longrightarrow \alpha_1,\ (\log q_n) \longrightarrow \beta_1,\ (\log r_n) \longrightarrow \alpha_2,\ (\log s_n) \longrightarrow \beta_2.$$

Then

$$\lim_{r,s}(n_r + n_s)\,\theta = \lim_r \left(\frac{1}{2}\alpha_1 + n_r\,\theta\right) = \alpha_1$$

$$\lim_{r,s}(n_r + n_s)^2\,\theta = \lim_r (n_r\,\alpha_1 + n_r^2\,\theta) = \alpha_2,$$

$$\lim_{r,s}(n_r + n_s)^3\,\theta = \lim_r \left(\frac{3}{2} n_r^2\alpha_1 + n_r^3\theta\right) = \alpha_3;$$

with similar results for φ. It follows by the principle **1.20** that there is a sequence giving the desired scheme, and the proof of the induction, and therefore that of the theorem, is completed.

1.3. — The order of the approximation.

1.30. We have proved that under certain conditions we can find a sequence (n_r) such that

$$(n_r^\varkappa\,\theta_p) \to \alpha_{\varkappa p} \quad (\varkappa = 1, 2, \cdots k;\ p = 1, 2, \cdots m). \tag{1.301}$$

There are a number of interesting questions which may be asked with regard to the *rapidity* with which the scheme of limits is approached.

The relations (1.301) assert that, if we are given λ, there is a function $\Phi(k, m;\ \theta_1, \theta_2, \cdots \theta_m;\ \alpha_{11}, \alpha_{12}, \cdots \alpha_{km};\ \lambda)$[1] such that

$$|(n^\varkappa\,\theta_p) - \alpha_{\varkappa p}| < 1/\lambda$$

for some $n < \Phi$. It is hardly necessary to observe, after the explanations of 1.00, that this inequality requires a modification when $\alpha_{\varkappa p} = 0$, which may be expressed roughly by saying that $\alpha_{\varkappa p}$ is then to be regarded as a two-valued symbol capable of assuming indifferently the values 0 and 1.

(i) Does Φ necessarily depend on the θ's and α's: can we for example, find a Φ independent of the α's? It will be seen that this last question is answered in the affirmative.

(ii) Can we assert anything concerning the order of Φ *qua* function of λ, the variables θ and α being supposed fixed? The same question may be asked concerning any Φ which is independent of the α's; it should be observed, moreover, that the best answer to the latter question does not necessarily give the best answer to the former.

[1] For shortness we shall write this $\Phi(k, m, \theta, \alpha, \lambda)$.

Our attempts to answer these questions have not been successful, and such results as we have been able to obtain are of a negative character. The question then arises as to whether we can obtain more definite results by imposing restrictions on the θ's or the α's, by supposing for example that all the α's are zero, or that the θ's belong to some special class of irrationals.

(iii) The relations (1.301) imply the truth of the following assertion: there is a function $\varphi(k, m, \theta, \alpha, n)$ which tends to infinity with n, and is such that

$$|(n^{\varkappa}\theta_p) - \alpha_{\varkappa p}| < 1/\varphi$$

for an infinity of values of n. A series of questions may then be asked concerning φ similar to those which we have stated with reference to Φ.

1.31. We shall begin by proving two theorems which are connected with the questions (i). The first of them deals with the case in which all the α's are zero, and it will be convenient to use in its statement, as in 1.10, not the function (x), but the allied function $\bar{x}$.

Theorem 1.31. *There is a function $\Phi(k, m, \lambda)$, depending only on k, m, and λ, such that*

$$|\overline{n^{\varkappa}\theta_p}| < 1/\lambda, \ (\varkappa = 1, 2, \cdots k; \ p = 1, 2, \cdots m),$$

for some $n < \Phi$.

For suppose that this theorem is false. Then to every r corresponds a set of θ's, say ${}_r\theta_1, {}_r\theta_2, \cdots {}_r\theta_m$, such that the inequalities

$$(1.311) \qquad |\overline{n^{\varkappa}\,{}_r\theta_p}| < 1/\lambda$$

are not all true unless $n > r$. The set of points $({}_r\theta_1, {}_r\theta_2, \cdots {}_r\theta_m)$ has at least one limiting point $(\Theta_1, \Theta_2, \cdots \Theta_m)$, and by restricting ourselves to a subsequence of r's we can make

$${}_r\theta_p \to \Theta_p, \ (p = 1, 2, \cdots m).$$

From this it follows that we can choose a number n_r which tends to infinity with r but so slowly that

$$(1.312) \qquad n_r^k |{}_r\theta_p - \Theta_p| < 1/2\lambda, \ (p = 1, 2, \cdots m).$$

Clearly we may suppose that $n_r \leq r$, and so we have, for an infinity of values of r, $n_r \leq r$ and

$$(1.313) \qquad n^{\varkappa} |{}_r\theta_p - \Theta_p| < 1/2\lambda, \ (n \leq n_r; \ \varkappa = 1, 2, \cdots k; \ p = 1, 2, \cdots m).$$

From (1.311) and (1.313) it follows that the inequalities

$$|\overline{n^{\varkappa}\theta_p}| < 1/2\lambda$$

cannot all be true unless $n > n_r$, and so, since $n_r \to \infty$, cannot be true for any value of n. This contradicts Theorem **1.011.**

In the case $k = 1$ it is possible to assert much more than this. It is known, and is proved in 1.10, that in this case we may take

$$\Phi = ([\lambda] + 1)^m \tag{1.314}$$

This problem, in fact, may be regarded as completely solved. When $k > 1$, however, the case is very different. We have not even succeeded in finding a definite function $\Phi(\lambda)$, the same for all θ's, such that

$$|\overline{n^2\theta}| \leq 1/\lambda$$

for $n \leq \Phi$. It would be not unnatural to suppose that the »best possible» function[1] Φ is less than $K\lambda$, where K is an absolute constant. But we have been unable to prove this or indeed any definite result as to its order in λ.

1.32. **Theorem 1.32.** *If the θ's are linearly independent irrationals, it is possible to find a function $\Phi(k, m, \theta, \lambda)$, independent of the α's, such that*

$$|(n^{\varkappa}\theta_p) - \alpha_{\varkappa p}| < 1/\lambda, \quad (\varkappa = 1, 2, \cdots k; \; p = 1, 2, \cdots m)$$

for some $n < \Phi$.

That this theorem is true for the special case $k = 1$, $m = 1$, follows from the argument (a) in 1.11. It is easily proved in the most general case by an argument resembling, but simpler than, that of 1.31.

If the theorem is untrue, it is possible to find a sequence of sets $({}_r\alpha_{\varkappa p})$ $(r = 1, 2, \cdots)$ for which the inequalities of the theorem do not all hold unless $n > r$. The sequence of sets has at least one limiting set $(\bar{\alpha}_{\varkappa p})$: let us choose r so that

$$|{}_r\alpha_{\varkappa p} - \bar{\alpha}_{\varkappa p}| < 1/2\lambda, \quad (\varkappa = 1, 2, \cdots k; \; p = 1, 2, \cdots m).$$

Then clearly the inequalities

$$|(n^{\varkappa}\theta_p) - \bar{\alpha}_{\varkappa p}| < 1/2\lambda$$

cannot all be true unless $n > r$, and so, since r is arbitrarily large, cannot all be true for any n. This contradicts Theorem **1.011.**

[1] That is, the function which has, for each value of λ, the least possible value. For the existence of this function it is necessary that the sign $\leq$ above should not be replaced by $<$.

1.33. Let us consider more particularly the case in which $k = 1$.

The equation (1.314) suggests that it may in this case be possible to choose for Φ a function of the form

$$\Omega(m, \theta, \alpha)\lambda^m.$$

This we believe to be improbable, but we have not succeeded, even when $m = 1$, in obtaining a definite proof. What is certain is that no corresponding result is true of the Φ of Theorem **1.32**. It is impossible to choose a function $\Omega(m, \theta)$ independent of λ, and a function $\psi(m, \lambda)$ independent of the θ's, in such a way that the Φ of this theorem may be taken to be of the form

$$\Phi = \Omega(m, \theta)\psi(m, \lambda).$$

This is shown by the following theorem.[1]

Theorem 1.33. *Let $\psi(\lambda)$ be an arbitrary function of λ which tends steadily to infinity with λ. Then it is possible to find irrational numbers θ for which the assertion 'there is a function*

$$\Phi(\theta, \lambda) = \Omega(\theta)\psi(\lambda)$$

such that, when λ is chosen, the inequality

$$|(n\theta) - \alpha| < 1/\lambda$$

is satisfied, for every α, by some n less than Φ' is false.

Suppose that the assertion in question is true. Taking $\alpha = 1/\lambda$, we see that

$$0 < (n\theta) < 2/\lambda \tag{1.331}$$

for some n less then Φ.

Let p_ν/q_ν be the ν-th convergent to the simple continued fraction

$$\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} + \cdots$$

which represents θ, so that $p_1 = 1$, $q_1 = a_1$; and let us consider the system of 'intermediate convergents'

$$\frac{p_{2n,r}}{q_{2n,r}} = \frac{p_{2n} + r\,p_{2n+1}}{q_{2n} + r\,q_{2n+1}}, \quad (0 \leq r \leq a_{2n+2}),$$

[1] In proving a result of this negative character we may evidently confine ourselves to the special case in which $m = 1$.

intercalated between p_{2n}/q_{2n} and p_{2n+2}/q_{2n+2}. These fractions are all less than θ and increase with r. Also

$$(1.332) \qquad \theta - \frac{p_{2n,r}}{q_{2n,r}} = \frac{a'_{2n+2} - r}{q_{2n,r}\, q'_{2n+2}},$$

where a'_{2n+2} is the complete quotient corresponding to a_{2n+2}, and

$$q'_{2n+2} = a'_{2n+2}\, q_{2n+1} + q_{2n}.$$

Let

$$(1.333) \qquad \lambda_n = \frac{2\, q'_{2n+2}}{a'_{2n+2} - s},$$

where s is a particular value of r which we shall fix in a moment. We shall suppose a_{2n+2} large, and s also large, but small in comparison with a_{2n+2}. In these circumstances λ_n will be approximately equal to $2\, q_{2n+1}$.

We shall now prove that if

$$(1.334) \qquad 0 < (Q\theta) < 2/\lambda_n$$

then

$$(1.335) \qquad Q > q_{2n,s}.$$

From (1.334) it follows that there is a fraction P/Q such that

$$(1.336) \qquad 0 < \theta - \frac{P}{Q} < \frac{2}{\lambda_n Q}.$$

On the other hand

$$(1.337) \qquad \theta - \frac{p_{2n,s}}{q_{2n,s}} = \frac{2}{\lambda_n\, q_{2n,s}}$$

If P/Q actually gave a better approximation by defect to θ than $p_{2n,s}/q_{2n,s}$, it would follow at once that $Q > q_{2n,s}$. We may therefore suppose the contrary; and then it follows from (1.336) and (1.337) that

$$0 < \frac{p_{2n,s}}{q_{2n,s}} - \frac{P}{Q} < \frac{2}{\lambda_n Q},$$

Hence

$$0 < p_{2n,s}\, Q - q_{2n,s}\, P < 2\, q_{2n,s}/\lambda_n.$$

But

$$\frac{1}{2}\lambda_n = \frac{a'_{2n+2}\, q_{2n+1} + q_{2n}}{a'_{2n+2} - s} > q_{2n+1},$$

and

$$q_{2n,s} = q_{2n} + s q_{2n+1} < (s+1)\, q_{2n+1}.$$

Hence $p_{2n,s} Q - q_{2n,s} P$ is less than $s + 1$, and so

$$(1.338) \qquad p_{2n,s} Q - q_{2n,s} P = \varrho \ (\theta \leq \varrho \leq s).$$

On the other hand

$$p_{2n,s}\, q_{2n,s-\varrho} - q_{2n,s}\, p_{2n,s-\varrho} = \varrho;$$

and so

$$p_{2n,s}(Q - q_{2n,s-\varrho}) = q_{2n,s}(P - p_{2n,s-\varrho}).$$

Hence either $Q = q_{2n,s-\varrho}$, or $Q - q_{2n,s-\varrho}$ is divisible by $q_{2n,s}$; and the latter hypothesis plainly involves that $Q > q_{2n,s}$.

On the other hand, if $Q = q_{2n,s-\varrho}$, then $P = p_{2n,s-\varrho}$, and

$$\theta - \frac{P}{Q} = \frac{a'_{2n+2} - s + \varrho}{q_{2n,s-\varrho}\, q'_{2n+2}} \geq \frac{2}{\lambda_n\, q_{2n,s-\varrho}},$$

$$(Q\theta) \geq 2/\lambda_n,$$

which contradicts (1.337). Hence in any case $Q > q_{2n,s}$.

It is now easy to complete the proof of the theorem. We have *a fortiori* $Q > s$. Also, if a_{2n+2} is large, and s large, but small in comparison with a_{2n+2}, λ_n will clearly be less than $4 q_{2n+1}$. We may suppose for definiteness that $s = [\sqrt{a_{2n+2}}]$.

We choose a value of θ such that the inequality

$$a_{2n+2} > \{\psi(4\, q_{2n+1})\}^4$$

is satisfied for an infinity of values of n. Then

$$s > \frac{1}{2}\{\psi(4\, q_{2n+1})\}^2.$$

But if Q, and *a fortiori* s, is less than Φ, we must have

$$\frac{1}{2}\{\psi(4\, q_{2n+1})\}^2 < \Omega(\theta)\, \psi(\lambda_n) < \Omega(\theta)\, \psi(4\, q_{2n+1});$$

and this is obviously impossible when n is sufficiently large. This completes the proof of the theorem.

It should be observed that the success of our argument depends entirely on our initial choice of α in such a way that $(n\theta)$ is small. It would not be enough that $\overline{n\theta}$ should be small, that is to say that $(n\theta)$ should be nearly equal to either 0 or 1: this can of course be secured by choice of an n less than Φ, Φ being indeed independent of θ.

1.34. We turn now for a moment to the questions concerning φ. If we have found a function $\Phi(\lambda)$ which is continuous and monotonic, the inverse function is plainly a φ. The converse, however, is not true, and we cannot, from the existence of a φ of given form, draw any conclusion as to the order of Φ for *all* values of λ. This is clear from the fact that, to put it roughly, the existence of φ asserts an inequality which need only hold very occasionally, and which therefore gives us information as to the behaviour of Φ only for occasional values of λ. Thus the existence of a Φ asserts much more than that of the corresponding φ. Since moreover it will appear (in the third paper of the series) that in applications of the present theory it is always the properties of Φ, and not those of φ, which are relevant, we are justified in regarding theorems concerning φ as of rather minor importance. There are, however, one or two results which are worth noticing, and which are not deductions from the corresponding results concerning Φ. It should be observed that whereas we wish Φ to increase as slowly as possible, we wish φ to increase as rapidly as possible.

Theorem 1.340. *It is possible to choose the α's so that $\varphi(m,\theta,\alpha,n)$ increases with arbitrary rapidity. Moreover the α's may be chosen in an arbitrarily small neighbourhood of any set $(\alpha_1,\alpha_2,\cdots\alpha_m)$.*

We omit the proof of this theorem, which is easy.

Theorem 1.341. *If $k=1$ and $m=1$, then, provided only that θ is irrational, we may take*

$$\varphi(n)=\frac{1}{3}n$$

(a function independent of both θ and α).

This follows at once from the argument (*a*) of 1.11. It is natural to suppose that, when $m>1$, we may take

$$\varphi(n)=\omega(m)\sqrt[m]{n},$$

where $\omega(m)$ depends only on m. But this we have not been able to prove.

A comparison of Theorems **1.33** and **1.341** shows very clearly the difference between theorems involving Φ and those involving φ, and the greater depth and difficulty of the former.

1.35. Theorem **1.33** shows that it is hopeless to expect any such simple result concerning Φ as is asserted concerning φ in Theorem **1.341**. It is however possible to obtain theorems which involve Φ and correspond to Theorem **1.341**, if we suppose that certain classes of irrationals (as well as the rationals) are excluded from the range of variation of θ. In the two theorems which follow it is supposed that $m = 1$ and $k = 1$.

Theorem 1.350. *Let θ be confined to the class of irrationals whose partial quotients are limited, a set which is everywhere dense. Then we may take*

$$\Phi = \lambda \Omega(\theta).$$

Theorem 1.351. *Let θ be confined to the class of irrationals whose partial quotients a_n satisfy, from a certain value of n onwards, the inequality*

$$a_n < n^{1+\delta} \ (\delta > 0).$$

Then we may take,

$$\Phi = \lambda (\log \lambda)^{1+\delta'} \Omega(\theta)$$

where δ' is any number greater than δ.

The interest of the last theorem lies in the fact that the set in question is of measure 1,[1] so that we may take Φ to be of the form $\lambda (\log \lambda)^{1+\varepsilon} \Omega(\theta)$,[2] where ε is an arbitrarily small positive number, for *almost all* values of θ.

The proofs of these theorems are simple and depend merely on an adaptation of KRONECKER's argument reproduced in 1.11. Suppose first that the partial quotients of θ are limited. We can choose H so that, when λ is assigned, there is always a denominator q_m of a convergent to θ such that

$$2\lambda \leq q_m < H\lambda \tag{1.350}$$

We take $q = q_m$. It follows from KRONECKER's argument that there is for any α a number ν such that

$$|(\nu\theta) - \alpha| < 2/q, \quad q/2 < \nu < 3q/2,$$

and so

$$|(\nu\theta) - \alpha| < 1/\lambda$$

for some ν less than a constant multiple of λ.

---

[1] By a theorem of BOREL and BERNSTEIN. See BOREL, *Rendiconti di Palermo*, vol. 27, p. 247, and *Math. Ann.*, vol. 72, p. 578; BERNSTEIN, *Math. Ann.*, vol. 71, p. 417.

[2] It is not difficult to replace $\lambda(\log \lambda)^{1+\varepsilon}$ by $\lambda \log \lambda (\log\log \lambda)^{1+\varepsilon}$, or by the corresponding but more complicated functions of the logarithmic scale.

The proof of Theorem **1.351** is very similar. We suppose that

$$q_{m-1} \leq 2\lambda < q_m,$$

and so

$$q_m/m^{1+\delta} < 2\lambda < q_m.$$

There is a constant ϱ such that $q_m > e^{\varrho m}$; and from these facts it follows easily that

$$q_m < \lambda (\log \lambda)^{1+\delta'}$$

for sufficiently large values of λ. The proof may now be completed in the same manner as that of Theorem **1.350**.

It is natural to suppose that these theorems have analogues when $m > 1$. But our arguments, depending as they do on the theory of continued fractions, do not appear to be capable of extension.

1.4. — The general sequence $(f(n)\theta)$ and the particular sequence $(a^n\theta)$

1.40. We return now to the general sequence $(f(n)\theta)$: it will be convenient to write λ_n for $f(n)$. We suppose then that (λ_n) is an arbitrary increasing sequence of numbers whose limit is infinity.[1]

It would be natural to attempt to prove that, if θ is irrational and α is any number such that $0 \leq \alpha < 1$, a sequence (n_r) can be found such that

$$(\lambda_{n_r}\theta) \rightarrow \alpha;$$

but we saw in 1.00 that this statement is certainly false, for example when $\lambda_n = 2^n$ or $\lambda_n = n!$

The result which is in fact true was suggested to us by a theorem of Bernstein,[2] which runs as follows:

If λ_n is always an integer, then the set of values of θ for which

$$(\lambda_n\theta) \rightarrow 0$$

is of measure zero.

This result, when considered in conjunction with what we have already proved, at once suggests the following theorem.

[1] In the introductory remarks of 1.00 we stated our main problem subject to the restriction that λ_n is an integer. No such restriction, however, is required in what follows.

[2] F. Bernstein, *loc. cit.*

Theorem 1.40. *The set of values of θ, for which the set of points $(\lambda_n \theta)$ is not everywhere dense in the interval* (0, 1), *is of measure zero.*

In other words, the main question asked in 1.00 may be answered affirmatively if we make exception of a set of measure zero.

1.41. The proof will be based upon the following lemma.

Lemma 1.41. *Suppose that a finite number of intervals are excluded from the continuum* (0, 1), *and that the length of the remainder S is l. Let a be any number betwen* 0 *and* 1, *and consider the set T of $[\lambda]$ intervals of length δ/λ $(\delta < 1)$ whose centres are at the points*

$$\frac{a}{\lambda}, \frac{a+1}{\lambda}, \ldots\ldots, \frac{a+[\lambda]-1}{\lambda}.^{1}$$

Then the length of the common part of S and T is

$$\delta l + \varepsilon_\lambda,$$

where $\varepsilon_\lambda \to 0$ as $\lambda \to \infty$.

The truth of the lemma is almost obvious. A formal proof may be given as follows. Let the lengths of the intervals excluded from S be $l_1, l_2, \ldots, l_p$. If now we extend each of these intervals a distance $1/2\lambda$ at each end,[2] we obtain a system of p intervals of length

$$l'_s = l_s + \frac{1}{\lambda}, \ (s = 1, 2, \cdots p).$$

We denote what is left of (0, 1) by S'.

If $(a+r)/\lambda$ falls in S', the whole of the corresponding interval of T falls in S. Hence the part of S inside T has a length not less than $\mu\delta/\lambda$, where μ is the number of points $(a+r)/\lambda$ in S'. If $\nu_1, \nu_2, \ldots, \nu_p$ are the numbers of these points which fall in the intervals excluded from S', we have

$$\mu + \Sigma \nu_s = [\lambda], \ (\nu_s - 1)/\lambda \leq l'_s;$$

and so

$$\begin{aligned} \mu = [\lambda] - \Sigma \nu_s &> \lambda - 1 - p - \lambda \Sigma l'_s \\ &= \lambda - p - 1 - \lambda \Sigma \left(l_s + \frac{1}{\lambda}\right) \\ &= l\lambda - 2p - 1, \end{aligned}$$

[1] It is of course to be understood that an interval, or a part of an interval, which falls outside (0, 1), is to be replaced by the congruent interval inside.

[2] We suppose λ large enough to ensure that this extension does not cause any overlapping. If any part of an extended interval should fall outside (0, 1), as will happen if an interval contains 0 or 1, we of course replace this part by the congruent part of (0, 1).

since $\Sigma l_s = 1 - l$. Hence the length in question is greater than

$$\delta l - \frac{(2p+1)\delta}{\lambda}.$$

A similar argument, which we may leave to the reader, furnishes a corresponding upper limit for the length; and the lemma follows. It is plain that $\varepsilon_\lambda = O(1/\lambda)$.

1.42. We can now prove the following theorem, which is a generalisation of BERNSTEIN's, but is itself contained in Theorem **1.40**.

Theorem 1.42. *If* I *is any interval contained in* (0, 1), *the set* Θ *of points* θ *such that no one of the points* $(\lambda_n \theta)$ *falls inside* I, *is of measure zero.*

Let a be the centre of I and δ its length; and let T_m be the set T of the lemma, with $\lambda = \lambda_m$. If, for any value of m, θ falls in T_m, then $(\lambda_m \theta)$ falls in I, and so θ belongs to the set complementary to Θ.

Let

$$S_n = T_1 + T_2 + \cdots + T_n,$$

and let l_n be the length of S_n. Finally let $l_n \to l$ as $n \to \infty$. We have to show that $l = 1$.

We now apply the lemma, taking S to be the set $\bar{S}_n$ complementary to S_n, and T to be T_m. If m is large enough, the length of the common part $(\bar{S}_n, T_m)$ of $\bar{S}_n$ and T_m is greater than

$$\delta(1 - l_n) - \varepsilon.$$

Any point which belongs either to this set or to S_n itself belongs to *some* S_ν. Hence

$$l \geq l_n + \delta(1 - l_n) - \varepsilon;$$

and so

$$l \geq l + \delta(1 - l) - \varepsilon:$$

which is impossible unless $l = 1$.

1.43. We can now complete the proof of Theorem **1.40**. Let E_n be the set of values of θ such that some one of the intervals

$$\left(0, \frac{1}{n}\right), \left(\frac{1}{n}, \frac{2}{n}\right), \ldots\ldots, \left(\frac{n-1}{n}, 1\right)$$

contains no point $(\lambda_n \theta)$. Then E_n is of measure zero, and so

$$E = E_1 + E_2 + E_3 + \ldots\ldots$$

is of measure zero.

If now the set $(\lambda_m \theta)$ is not everywhere dense in $(0, 1)$, there is an interval i which contains no $(\lambda_m \theta)$. We can choose n so that some interval $\left(\frac{r}{n}, \frac{r+1}{n}\right)$ falls inside i. Then θ belongs to E_n and so to E. Thus the theorem is established.

1.44. Perhaps the most interesting special sequence falling under the general type $(f(n)\theta)$ is that in which $f(n) = a^n$, where a is a positive integer. When θ is expressed as a decimal in the scale of a, the effect of multiplication by a is merely to displace the digits. To study the properties of the sequence $(a^n \theta)$ is therefore equivalent to studying the distribution of the digits in the expression of θ in the scale of a: it is to this fact that this form of $f(n)$ owes its peculiar interest.

Let b be one of the possible digits $0, 1, 2, \ldots, a-1$, and let $p(n, m)$ denote the number of decimals of n figures whose digits include exactly m b's. Then

$$p(n, m) = \frac{n!}{m!\,(n-m)!}(a-1)^{n-m}. \tag{1.441}$$

We write

$$\mu = m - \frac{n}{a}; \tag{1.442}$$

so that μ is the excess of the number of b's above the average.

We shall base our investigation on a series of lemmas.

Lemma 1.441. *Given any positive number δ, we can find a positive number ε such that*

$$p(n, m) < \frac{a+\delta}{\sqrt{2\pi(a-1)n}} e^{-(\alpha-\delta)\mu^2/n} a^n \tag{1.443}$$

where

$$\alpha = \frac{a^2}{2(a-1)},$$

for $|\mu| < \varepsilon n$ and all sufficiently large values of n.

We omit the proof of this lemma, which depends merely on a straightforward application of STIRLING's Theorem.

Lemma 1.442. *Given any positive number ε, we can find a positive number ζ such that*

$$p(n, m) < a^n e^{-\zeta n}$$

for $|\mu| \geq \varepsilon n$ *and all sufficiently large values of* n.

Suppose, *e. g.*, $\mu > \frac{1}{2}\varepsilon n$. Then

$$\frac{p(n, m+1)}{p(n, m)} = \frac{n-m}{(a-1)(m+1)} < \frac{a - 1 - \frac{1}{2}a\varepsilon}{(a-1)(1+\frac{1}{2}a\varepsilon)} < 1;$$

and from this it is easy to deduce the truth of the lemma when $\mu > \varepsilon n$. A similar proof applies when $\mu < -\varepsilon n$.

Lemma 1.443. *Let c be a positive constant. Then*

$$\overline{\lim_{n\to\infty}}\, a^{-n} \sum_{|\mu| < c\sqrt{n}} p(n, m) < 1, \tag{1.4431}$$

$$\overline{\lim_{n\to\infty}}\, a^{-n} \sum_{\mu > -c\sqrt{n}} p(n, m) < 1, \tag{1.4432}$$

$$\overline{\lim_{n\to\infty}}\, a^{-n} \sum_{\mu < c\sqrt{n}} p(n, m) < 1. \tag{1.4433}$$

Of these three inequalities the first is plainly a consequence of either the second or third. It will be enough to prove the second.

We have

$$a^{-n} \sum_{\mu > -c\sqrt{n}} p(n, m) = a^{-n} \sum_{-c\sqrt{n}}^{\varepsilon n} + a^{-n} \sum_{\varepsilon n}^{(a-1)n/a} = S_1 + S_2,$$

say. By Lemma **1.442**,

$$S_2 < \frac{(a-1)n}{a} e^{-\zeta n} \to 0.$$

And by Lemma **1.441**,

$$S_1 < \frac{a+\delta}{\sqrt{2\pi(a-1)n}} \sum_{-c\sqrt{n}}^{\varepsilon n} e^{-(a-\delta)\mu^2/n}$$

$$< \frac{a+\delta}{\sqrt{2\pi(a-1)n}}\{2+\int_{-c\sqrt{n}}^{\infty} e^{-(a-\delta)\mu^2/n}\,d\mu\}.$$

The term of order $1/\sqrt{n}$ may be ignored. The remainder is less than

$$\frac{a+\delta}{\sqrt{2\pi(a-1)}}\int_{-c}^{\infty} e^{-(a-\delta)\xi^2}\,d\xi,$$

which is less than 1. Thus the lemma is proved. In a similar manner we can prove

Lemma 1.444. *If ν is a function of n such that $\nu/\sqrt{n}\to\infty$, then*

$$a^{-n}\sum_{|\mu|<\nu} p(n,m) < K\left\{\frac{\sqrt{n}}{\nu}e^{-(a-\delta)\nu^2/n}+a^{-n}\right\},$$

where K depends only on a.

1.45. We are now in a position to prove our main theorems. We observe first that all irrational[1] numbers θ between 0 and 1, whose decimals have just m b's in their first n figures, may be included in a set of intervals whose total length is

$$a^{-n}p(n,m).$$

For let $\theta_1, \theta_2, \ldots, \theta_q$, where $q=a^n$, denote the terminating decimals of n figures. The set of intervals $(\theta_r, \theta_r+a^{-n})$ just fills up the whole interval $(0, 1)$. Among the numbers θ_r there are $p(n,m)$ which have just m b's, which we may call $\xi_1, \xi_2, \ldots, \xi_p$; and the set of intervals (ξ_s, ξ_s+a^{-n}) fulfils our requirements.

Theorem 1.45. *Let δ be any positive number. Then the set of numbers θ for which*

$$\overline{\lim_{n\to\infty}}\frac{|\mu|}{\sqrt{n\log n}} < \sqrt{\frac{1}{\alpha}}+\delta$$

is of measure 1.

Let $\bar{S}$ denote the complementary set. Any number belonging to $\bar{S}$ satisfies

$$\mu > \left(\sqrt{\frac{1}{\alpha}}+\delta'\right)\sqrt{n\log n}=\nu_n$$

for an infinity of values of n, δ' being any positive number less than δ.

[1] The end points of the intervals will be *rational* numbers satisfying the condition. In what follows we may confine ourselves to irrational values of θ, since the rational values form in any case a set of measure zero.

All θ's for which this inequality is true for a particular n may be enclosed in a set of intervals whose total length is

$$a^{-n} \sum_{|\mu| < \nu_n} p(n, m). \tag{1.451}$$

We can choose a positive number δ'' such that

$$2\delta' \sqrt{\alpha} - \frac{\delta''}{\alpha} - \frac{2\delta'\delta''}{\sqrt{\alpha}} > 0,$$

and then choose n_1 so that the expression (1.451) is less than

$$K\left\{\frac{\sqrt{n}}{\nu_n} e^{-(\alpha-\delta'')\nu_n^2/n} + a^{-n}\right\}$$

for $n \geq n_1$. To prove the theorem it is enough to show that the result of summing this expression for $n = n_1, n_1 + 1, \ldots\ldots$ can be made as small as we please by choice of n_1; and it is obvious that this conclusion cannot be affected by the presence of the term a^{-n}. But

$$\frac{\sqrt{n}}{\nu_n} e^{-(\alpha-\delta'')\nu_n^2/n} < e^{-\frac{1}{2}(\alpha-\delta'')\left(\frac{1}{\sqrt{\alpha}}+\delta'\right)^2 \log n}$$

$$= n^{-1-\delta'''},$$

where

$$\delta''' > (\alpha - \delta'')\left(\frac{1}{\alpha} + \frac{2\delta'}{\sqrt{\alpha}}\right) - 1$$

$$= 2\delta' \sqrt{\alpha} - \frac{\delta''}{\alpha} - \frac{2\delta'\delta''}{\sqrt{\alpha}} > 0;$$

and plainly

$$\sum_{n_1}^{\infty} n^{-1-\delta'''}$$

can be made as small as we please by choice of n_1. Thus the theorem is proved.

Theorem **1.45** includes as a particular case

Theorem 1.451. *If n_b is the number of b's in the first n figures of the expression of θ as a decimal in the scale of a, then*

$$n_b \sim n/a$$

for almost all values of θ.

1.46. Theorem **1.45** shows that the deviation, from the average n/a, of the number of occurrences of a particular figure b in the first n places, is not in general of an order materially greater than $\sqrt{n}$.[1] If we were to suppose that there was a *steady* deviation from the average (instead of a merely occasional deviation), we would naturally obtain a more precise result. Thus reasoning analogous to, but simpler than, that which led to theorem 1.45, leads also to

Theorem 1.46. *If $\varphi(n) \to \infty$ with n, then the set of θ's for which*

$$|\mu(n)|/\sqrt{n}\,\varphi(n) \to \infty$$

is of measure zero.

This theorem, however, is included in a much more interesting and general theorem which we shall now proceed to prove, which, to put it roughly, assigns a *lower* limit for the deviation in either direction.

1.47. **Theorem 1.47.** *If c is any positive constant, the set of θ's for which*

$$\mu(n) > -c\sqrt{n},$$

and the set for which $\mu(n) < c\sqrt{n}$, are of measure zero.

Let

$$c_0 = c \prod_{m=1}^{\infty} \left(1 + \frac{1}{2^m}\right).$$

By Lemma **1.443**, there is a positive number δ_{c_0} such that

$$\overline{\lim_{n \to \infty}}\, a^{-n} \sum_{\mu > -c_0\sqrt{n}} p(n, m) = 1 - \delta_{c_0}.$$

And if $c \leq c_1 < c_0$, it is clear that

$$\overline{\lim_{n \to \infty}}\, a^{-n} \sum_{\mu > -c_1\sqrt{n}} p(n, m) = 1 - \delta_{c_1},$$

where

$$\delta_c \geq \delta_{c_1} \geq \delta_{c_0}.$$

Let E_c be the set of the theorem. We can enclose E_c in a set of intervals of total length

[1] It follows from the elements of the theory of errors that the 'most probable error' is of order $\sqrt{n}$.

$$a^{-n_1} \sum_{\mu > -c\sqrt{n_1}} p(n_1, m) < 1 - \frac{1}{2}\delta_c.$$

Consider now any one of the

$$N = \Sigma p(n, m)$$

intervals of this set, each of which is of length a^{-n_1}; and let $\xi = (a^{n_1}\theta)$. As θ ranges in the interval in question, ξ ranges in the whole interval $(0, 1)$.

If θ belongs to E_c, the corresponding ξ has the property that

$$\mu(n') > -c\sqrt{n_1 + n'}$$

for all values of n'; and so, if n' is large enough compared with n_1,

$$\mu(n') > -c'\sqrt{n'},$$

where

$$c' = c(1 + 2^{-n_1-1}).$$

We may now enclose the ξ's in a set of intervals whose total length is less than

$$1 - \frac{1}{2}\delta_{c'};$$

and therefore we may enclose the θ's which lie in the particular interval under consideration in a set of intervals whose total length is less than $a^{-n_1}(1 - \frac{1}{2}\delta_{c'})$. If we do this for each of the N intervals, we have enclosed the θ's in a set of intervals of length less than

$$\left(1 - \frac{1}{2}\delta_c\right)\left(1 - \frac{1}{2}\delta_{c'}\right).$$

Repeating this argument, it is clear that we can enclose the θ's in a set of intervals of total length less than

$$\left(1 - \frac{1}{2}\delta_c\right)\left(1 - \frac{1}{2}\delta_{c'}\right)\left(1 - \frac{1}{2}\delta_{c''}\right)\cdots\left(1 - \frac{1}{2}\delta_{c^{(\nu)}}\right),$$

where

$$c^{(\nu)} = c(1 + 2^{-n_1-1})(1 + 2^{-n_2-1})\ldots(1 + 2^{-n_\nu-1}),$$

the indices n_ν being integers which tend to infinity with ν, as rapidly as we please. Plainly $c^{(\nu)} < c_0$ and so

$$\delta_{c^{(\nu)}} \geq \delta_{c_0}, \ 1 - \frac{1}{2}\delta_{c^{(\nu)}} \leq 1 - \frac{1}{2}\delta_{c_0},$$

$$\left(1 - \frac{1}{2}\delta_c\right)\left(1 - \frac{1}{2}\delta_{c'}\right)\cdots\left(1 - \frac{1}{2}\delta_{c^{(\nu)}}\right) \leq \left(1 - \frac{1}{2}\delta_{c_0}\right)^{\nu+1}.$$

As this tends to zero as $\nu \to \infty$, our theorem is proved.

From Theorem **1.47** we can at once deduce

Theorem 1.471. *The set of θ's such that to each θ corresponds a c for which $\mu(n) > -c\sqrt{n}$ is of measure zero.*

Let E_c denote the set of Theorem **1.47**. The set of this theorem is plainly the sum of the sets $E_1, E_2, E_3, \ldots$; and so is of measure zero.

1.48. So far we have considered merely the occurrence of a particular digit b in the decimal which represents θ. But our results are easily extended so as to give analogous information concerning the occurrence of any combination of digits. The method by which this extension is effected is quite simple in principle, and it will be sufficient to show its working in a special case.

Consider the succession 317 of digits, in the scale of 10. In the scale of 1000, the number 317 corresponds to a single digit τ; and, if θ is expressed in the scale of 1000, it will, by theorem **1.451**, be almost always true that the number n_τ of occurrences of τ, among the first n figures, satisfies the relation

$$n_\tau \sim \frac{n}{1000}.$$

Now the combination 317, in the expression of θ in the scale of 10, will occur when, and only when, the digit τ occurs in the expression of one or other of the three numbers

$$\theta, 10\theta, 100\theta$$

in the scale of 1000. Hence it is almost always true that the number of occurrences of the combination 317, in the first n digits of the expression of θ in the scale of 10, is asymptotically equivalent to

$$\frac{n}{1000}.$$

We may now, without further preface, enunciate the following theorems.

Theorem 1.48. *It is almost always true that, when a number θ is expressed in any scale of notation, the number of occurrences of any digit, or any combination of digits, is asymptotically equivalent to the average number which might be expected.*

Theorem 1.481. *It is almost always true that the deviation from the average, in the first n places, is not of order exceeding $\sqrt{n \log n}$.*

Theorem 1.482. *It is almost always true that the deviation, in both directions, is sometimes of order exceeding $\sqrt{n}$.*

Theorem 1.483. *The number of the first n numbers $(a^\nu \theta)$ which fall inside an interval of length δ included in the interval* (0, 1) *is almost always asymptotically equivalent to δn.*

The last theorem is merely a translation of theorem **1.48** into different language, and a corresponding form may of course be given to theorems **1.481** and **1.482**.

1.49. Throughout this section (1.4) we have confined ourselves to results concerning a single irrational θ. Some of our theorems, however, have obvious many-dimensional analogues. It will be sufficient, for the present, to mention the following, which are generalisations of Theorems **1.40** and **1.483** respectively.

The interval (0, 1) is now replaced by an m-dimensional 'square'.

Theorem 1.49. *The set of values $(\theta_1, \theta_2 \cdots \theta_m)$, for which the points $(\lambda_n \theta_1, \lambda_n \theta_2, \cdots \lambda_n \theta_m)$ are not everywhere dense in the square, is of measure zero.*

Theorem 1.491. *The number of the first n points $(a^\nu \theta_1, a^\nu \theta_2, \cdots a^\nu \theta_m)$, which fall inside a portion of the square, of area δ, is almost always asymptotically equivalent to δn.*

We leave the proofs to the reader. The first theorem may be proved by an obvious adaptation of the proof of Theorem **1.40**, and the second deduced from Theorem **1.483** by a process of correlation very similar to that employed n 1.48.

Contents.