THE SIZE OF THE SET OF THE BADLY APPROXIMABLE NUMBERS

ANDREW FOOTE

Abstract. This project is a survey of some established results in the field of Diophantine approximation, starting with the most basic results, such as Dirichlet's theorem, but focusing towards the end on two results concerning the size of the set of the badly approximable numbers, which together state that this set is large in one sense and small in another. The first result is that this set is uncountable, and it is proven using a generalized Cantor set construction. The second result is that the Lebesgue measure of this set is 0, and it is proven using the Lebesgue density theorem. The other results mentioned provide background for these two main results.

Contents

Date: 3 May 2016.

Notation

Varieties of numbers. The expressions $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$ denote the sets of the positive integers, integers, rational numbers, real numbers and complex numbers, respectively. The phrase "natural number" is avoided in favour of "positive integer". The word "number", when it is not modified by any of the adjectives "rational", "real" or "complex", is synonymous with "real number".

Denominators. For every $x \in \mathbf{Q}$, the *denominator* of x is the smallest $q \in \mathbf{N}$ with $qx \in \mathbf{Z}$. It can also be characterised as the unique $q \in \mathbf{N}$ with $x = p/q$ for some $p \in \mathbf{Z}$ coprime to q.

Intervals. For every pair $\langle a, b \rangle$ of real numbers such that $a < b$,

$$\begin{aligned}(a, b) &= \{x \in \mathbf{R} : a < x < b\},\\ (a, b] &= \{x \in \mathbf{R} : a < x \leq b\},\\ [a, b) &= \{x \in \mathbf{R} : a \leq x < b\},\\ [a, b] &= \{x \in \mathbf{R} : a \leq x \leq b\}.\end{aligned}$$

Pointwise operations on functions. For every set S, every real-valued function f on S and every $c \in \mathbf{R}$, the expressions id_S, c_S and cf denote functions on S called the *identity function on S*, the *constant function on S of value c*, and the *pointwise product of c and f*, respectively, and for every $x \in S$,

$$\begin{aligned}\mathrm{id}_S(x) &= x,\\ c_S(x) &= c,\\ (cf)(x) &= c(f(x)).\end{aligned}$$

Also, if $f(x) > 0$ for every $x \in S$, then f^c denotes a function on S and for every $x \in S$,

$$(f^c)(x) = (f(x))^c.$$

Pointwise operations on sets. For every set $S \subseteq \mathbf{R}$ and every $c \in \mathbf{R}$,

$$cS = \{cx : x \in S\}.$$

Limits inferior and limits superior. For every sequence $\langle S_n \rangle_{n \in \mathbf{N}}$ of sets,

$$\begin{aligned}\liminf_{n \to \infty} &= \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} S_n,\\ \limsup_{n \to \infty} &= \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} S_n.\end{aligned}$$

The common denotation of the expressions on both sides of the former equation is called the *limit inferior* of $\langle S_n \rangle_{n \in \mathbf{N}}$. The common denotation of the expressions on both sides of the latter equation is called the *limit superior* of $\langle S_n \rangle_{n \in \mathbf{N}}$.

The members of the limit inferior of $\langle S_n \rangle_{n \in \mathbf{N}}$ are the entities x such that $x \in S_n$ for every sufficiently large $n \in \mathbf{N}$. The members of the limit superior of $\langle S_n \rangle_{n \in \mathbf{N}}$ are the entities x such that there are arbitrarily large $n \in \mathbf{N}$ with $x \in S_n$.

Integer parts and fractional parts. For every $x \in \mathbf{R}$,

$$[x] = \min_{\substack{n \in \mathbf{Z} \\ n \leq x}} n,$$

$$\langle x \rangle = x - [x],$$

$$\|x\| = \begin{cases} \langle x \rangle & \text{if } \langle x \rangle \leq 1/2, \\ 1 - \langle x \rangle & \text{if } \langle x \rangle > 1/2. \end{cases}$$

The expressions $[x]$ and $\langle x \rangle$ are called the *integer part* and *fractional part* of x, respectively. Note that $x = [x] + \langle x \rangle$, which is why the word "part" is used. The expression $\|x\|$ has no name of its own, but it may referred to as the distance between x and the closest integer to x, because $[x]$ and $[x] + 1$ are clearly the closest two integers to x, and the distance between x and $[x]$ is $\langle x \rangle$ and the distance between x and $[x] + 1$ is $1 - \langle x \rangle$.

Note that:

(1) $[x]$ is the unique $n \in \mathbf{Z}$ such that $n \leq x < n + 1$.
(2) $\langle x \rangle$ is the unique $d \in \mathbf{R}$ such that $d < 1$ and $x - d \in \mathbf{Z}$.
(3) $\|x\|$ is the unique $d \in \mathbf{R}$ such that $d \leq 1/2$ and at least one of $x - d$ and $d - x$ is integral.

1. Introduction

A *Diophantine equation* is an equation of the form

$$pu = qv, \tag{1}$$

where u and v are real numbers and p and q are unknown integers. The term "Diophantine equation" can also be used to refer to members of other classes of equations, but it will be used to refer to members of this specific class of equations in this document.

Let's briefly remark on the solutions of (1). If $a = b = 0$ then clearly every pair $\langle p, q \rangle$ of integers is a solution. Otherwise, we can assume without loss of generality that $a \neq 0$, in which case (1) is equivalent to the equation

$$p = qx, \tag{2}$$

where $x = v/u$. Clearly $\langle 0, 0 \rangle$ is a solution of this equation, and it is in fact the only solution of the form $\langle p, 0 \rangle$ where $p \in \mathbf{R}$. Let's call $\langle 0, 0 \rangle$ the *trivial solution* of (2). For every non-trivial solution $\langle p, q \rangle$ of (2), we have $q \neq 0$, and therefore the non-trivial solutions of (2) are exactly the solutions of

$$\frac{p}{q} = x. \tag{3}$$

The set of values which the left-hand side of (3) may evaluate to, given that p and q are integers, is exactly $\mathbf{Q}$. Therefore, (3) has a solution if and only if $x \in \mathbf{Q}$. In this case, there is a unique solution $\langle p, q \rangle$ of (3) with $\gcd(p, q) = 1$ and $q > 0$, and every other solution of (3) is of the form $\langle mp, mq \rangle$ with $m \in \mathbf{Z}$ and $m \neq 0$.

If $x \notin \mathbf{Q}$, then (3) has no solution. But it does have the next best thing—an *approximate* solution. For every positive $\varepsilon \in \mathbf{R}$, let's call a pair $\langle p, q \rangle$ of integers an *approximate solution of* (3) *with error* ε if and only if

$$\left| x - \frac{p}{q} \right| = \varepsilon. \tag{4}$$

For every such pair $\langle p, q \rangle$, the rational number p/q is at a distance of ε from x. In other words, it is a *rational approximation of x with error ε*. Moreover, for every rational number p/q at a distance of ε from x, where p and q are integers and $p \neq 0$, the pair $\langle p, q \rangle$ is a solution of (4). The problem of finding solutions of (4) for a given set of values of ε, which is an example of a problem of *Diophantine approximation*, is therefore essentially the same as the problem of finding rational approximations of x with errors in a given set, which is the problem of *rational approximation*. Thus the terms "Diophantine approximation" and "rational approximation" are used as synonyms.

There is in fact a whole field of mathematics which is known as Diophantine approximation or rational approximation. The problem just described is really only a very specific example of a problem that this field deals with, and it has a very simple solution which we shall describe in the next section. However, we shall move on to more general forms of the problem afterwards, which are more mathematically interesting. Our examination of these more general forms will naturally lead to the definition of badly approximable numbers, and the final two sections of the project will give proofs of two results on the size of the set of the badly approximable numbers, namely that it is uncountable and that its Lebesgue measure is 0.

2. The density of $\mathbf{Q}$ in $\mathbf{R}$

One of the well-known properties of the real line is that rational numbers are densely distributed along it. That is, for every pair $\langle a, b\rangle$ of real numbers such that $a < b$, there is an $x \in \mathbf{Q}$ with $a < x < b$. Now, for every $x \in \mathbf{R}$ and every positive $\varepsilon \in \mathbf{R}$, we have $x - \varepsilon < x + \varepsilon$, and therefore there is an $r \in \mathbf{Q}$ with $x - \varepsilon < r < x + \varepsilon$. This chain of inequalities is equivalent to the single inequality

$$|x - r| < \varepsilon.$$

Given that $r \in \mathbf{Q}$, there are integers p and q with $r = p/q$, and therefore $\langle p, q\rangle$ is an approximate solution of (3) with error less than ε. It follows that (3) has approximate solutions with arbitrarily small errors. In other words, every real number is arbitrarily closely approximable by rational numbers.

For the integers, the situation is different. The integers are not densely distributed along the real line. In fact, they have the opposite property: for every $x \in \mathbf{R} \setminus \mathbf{Z}$, there is a punctured neighbourhood of x which is disjoint from $\mathbf{Z}$, namely $(x - \|x\|, x + \|x\|) \setminus \{x\}$, and therefore x is not arbitrarily closely approximable by integers. The set $\mathbf{Z}$ is therefore said to be *nowhere dense* in $\mathbf{R}$.

More generally, for every $q \in \mathbf{N}$ and every $x \in \mathbf{R}$, there is a $p \in \mathbf{Z}$ with $\|qx\| = |qx - p|$, and we have

$$\frac{\|qx\|}{q} = \frac{|qx - p|}{q} = \left|\frac{qx - p}{q}\right| = \left|x - \frac{p}{q}\right|.$$

For every other $p' \in \mathbf{Z}$, we have $|qx - p'| \geq \|qx\|$ which implies $|x - p'/q'| \geq |x - p/q|$. It follows that the closest rational numbers to x of the form p/q with $p \in \mathbf{Z}$ are at a distance of $\|qx\|/q$ from x.

Now, for every $Q \in \mathbf{N}$, let

$$S_Q = \bigcup_{q=1}^{Q} \mathbf{Z}/q.$$

Then the members of S_Q are the rational numbers of the form p/q, where $p \in \mathbf{Z}$ and $0 < q \leq Q$, and for every $x \in \mathbf{R} \setminus S_Q$, the punctured neighbourhood

$$\left(x - \min_{q=1}^{Q} \frac{\|qx\|}{q}, x + \min_{q=1}^{Q} \frac{\|qx\|}{q}\right) \setminus \{x\},$$

is disjoint from S_Q. Therefore, S_Q is nowhere dense, just like $\mathbf{Z}$.

The upshot of this is that although every irrational number x is arbitrarily closely approximable by rational numbers, the denominator of every sufficiently close rational approximation of x is arbitrarily large. That is, for every sequence $\langle p_n/q_n \rangle_{n \in \mathbf{N}}$ of rational approximations of x with $|x - r_n| \to 0$ as $n \to \infty$, where p_n/q_n is an irreducible fraction for every $n \in \mathbf{N}$, we also have $q_n \to \infty$ as $n \to \infty$. Otherwise, there is a $Q \in \mathbf{N}$ such that x is arbitrarily closely approximable by rational numbers whose denominators are less than or equal to Q, and this is impossible because S_Q is nowhere dense. There is thus a trade-off between error and denominator magnitude in the rational approximation of irrational numbers. The closer the approximation has to be, the greater its denominator has to be. Of course, the same is not true for rational numbers because every rational number approximates itself perfectly with error 0.

3. Approximating functions

Suppose $x \in \mathbf{R}$, Ψ is a positive real-valued function on $\mathbf{N}$, $p \in \mathbf{Z}$ and $q \in \mathbf{N}$. In light of the trade-off between error and denominator magnitude described in the previous section, the ratio

$$\frac{|x - p/q|}{\Psi(q)}$$

is interesting. Rational numbers with smaller denominators can be considered simpler than those with larger denominators, so, if two rational approximations have the same errors, the one with the smaller denominator can be considered better. But even if the errors are different, with the more complex approximation having a smaller error, we might still prefer the simpler approximation if the difference is sufficiently small. One way of making a definite choice between the two approximations is to compare the ratios $|x - p_1/q_1|/\Psi(q_1)$ and $|x - p_2/q_2|/\Psi(q_2)$, where p_1/q_1 and p_2/q_2 are the irreducible fraction expressions of the two approximations, and to prefer the approximation on which the smaller ratio depends. These ratios can be thought of as the "relative errors" of the two approximations, as opposed to the absolute errors $|x - p_1/q_1|$ and $|x - p_2/q_2|$ which do not take into account the sizes of the denominators. Exactly *how* the sizes of the denominators are taken into account by the relative errors is determined by the nature of the function Ψ (about which we have supposed nothing besides that its range is a set of positive real numbers). For example, if $\Psi(q) = 1/q^2$ for every $q \in \mathbf{N}$, then it takes a greater difference to cause us to prefer the more complex approximation than if $\Psi(q) = 1/q$ for every $q \in \mathbf{N}$.

These considerations lead us to a more general formulation of the problem of Diophantine approximation. Diophantine approximation is the problem of finding solutions of the inequality

$$\left|x - \frac{p}{q}\right| < \Psi(q) \tag{5}$$

where $x \in \mathbf{R}$, Ψ is a positive real-valued function on $\mathbf{N}$, p and q are unknown integers and $q > 0$. The solutions of (5) are the pairs $\langle p, q\rangle$ of integers with $q > 0$ such that the relative error of p/q with respect to Ψ, i.e. $|x - p/q|/\Psi(q)$, is less than 1. Of course, there is nothing special about relative errors less than 1 in particular. But for every positive $\varepsilon \in \mathbf{R}$, the solutions of the inequality

$$\left|x - \frac{p}{q}\right| < \varepsilon\Psi(q), \tag{6}$$

i.e. the pairs (p, q) of integers with $q > 0$ and $|x - p/q|/\Psi(q) < \varepsilon$, are exactly the solutions of the inequality (5) under the subsitution of the function $\varepsilon\Psi$ in place of Ψ.

In this document we shall concentrate on the question of *how many* solutions (5) has for given choices of x and Ψ. Let's first note that for every $n \in \mathbf{Z}$ and every pair (p, q) of integers such that $q > 0$,

$$\left|(x + n) - \frac{p}{q}\right| = \left|x - \frac{p - qn}{q}\right|,$$

so if we let S be the set of the pairs $\langle p, q\rangle$ of integers such that $q > 0$ and

$$\left|(x + n) - \frac{p}{q}\right| < \Psi(q),$$

and we let T be the set of the solutions of (5), then for every $p \in \mathbf{Z}$ and every $q \in \mathbf{Z}$, we have $\langle p, q\rangle \in S$ if and only if $\langle p - qn, q\rangle \in T$. In other words, S is the image of T under the function f on $\mathbf{Z} \times \mathbf{N}$ such that $f(p, q) = \langle p - qn, q\rangle$ for every $p \in \mathbf{Z}$ and every $q \in \mathbf{N}$. Now, this function f is injective, because for every pair $\langle\langle p, q\rangle, \langle p', q'\rangle\rangle$ of members of $\mathbf{Z} \times \mathbf{N}$, the equation $f(p, q) = f(p', q')$ is equivalent to $\langle p - qn, q\rangle = \langle p' - q'n, q'\rangle$ and this equation is in turn equivalent to the conjunction of the equations $p - qn = p' - q'n$ and $q = q'$. The latter equation in the conjunction implies that $qn = q'n$ and therefore the former equation is equivalent to $p = p'$. It follows that the restriction of f to the domain T is a bijection onto S, and therefore $\#S = \#T$. This means that the number of solutions of (5) is invariant under substitution of expressions of the form $x - n$, with $n \in \mathbf{Z}$, in place of x. In particular, for every

$x \in \mathbf{R}$ with $x \notin \mathbf{I}$, where

$$\mathbf{I} = [0, 1),$$

we have $\langle x \rangle \in \mathbf{I}$, and the number of solutions of (5) is invariant under substitution of $\langle x \rangle$ in place of x. In our investigation of how many solutions (5) has for given choices of x and Ψ, we may therefore restrict our attention to choices of x with $x \in \mathbf{I}$. This conveniently simplifies the investigation in some respects, due to the property of boundedness possessed by $\mathbf{I}$ but not $\mathbf{R}$.

Let $W_1(\Psi)$ denote the set of the $x \in \mathbf{I}$ such that there is at least one solution of (5). Clearly, $W_1(\Psi)$ is simply the intersection of $\mathbf{I}$ and the union of the open intervals of the form

$$\left]\frac{p}{q} - \Psi(q), \frac{p}{q} + \Psi(q)\right[,$$

where $q \in \mathbf{N}$ and $0 \leq p \leq q$. Therefore, $W_1(\Psi)$ is an open set.

More generally, for every $n \in \mathbf{N}$, let $W_n(\Psi)$ denote the set of the $x \in \mathbf{I}$ such that there are at least n solutions of (5). Then $W_n(\Psi)$ is always the the intersection of $\mathbf{I}$ and the union of all of the n-ary intersections of pairwise distinct open intervals of the form above. Because finite intersections of open sets are open themselves it follows that $W_n(\Psi)$ is open, and therefore $W_n(\Psi)$ is a union of open intervals in $\mathbf{R}$.

In this document we shall concentrate on the question of for which choices of x there are *infinitely many* solutions of (5). The $x \in \mathbf{R}$ for which this is the case are said to be Ψ*-approximable* or *approximable to the order of* Ψ. In addition, the function Ψ is said to be an *approximating function* of every such number, and the set of the Ψ-approximable members of $\mathbf{I}$ is denoted $W(\Psi)$.

There is a slightly different characterisation of the Ψ-approximable numbers which it is useful to be familiar with. First, note that for every $x \in \mathbf{R}$, every $p \in \mathbf{Z}$ and every $q \in \mathbf{N}$, (5) is equivalent to

$$|qx - p| < q\Psi(q), \tag{7}$$

and therefore the the set S of the $p \in \mathbf{Z}$ such that (5) holds is $(qx - q\Psi(q), qx + q\Psi(q)) \cap \mathbf{Z}$, which is finite (in fact, it has at most $[2q\Psi(q)] + 1$ members). It follows that (5) has infinitely many solutions if and only if there are infinitely many $q \in \mathbf{N}$ such that (5) has a solution of the form $\langle p, q \rangle$ with $p \in \mathbf{Z}$.

Proposition 1. *For every $x \in \mathbf{I}$ and every positive real-valued function Ψ on $\mathbf{N}$, x is Ψ-approximable if and only if there are infinitely many $q \in \mathbf{N}$ such that*

$$\|qx\| < q\Psi(q). \tag{8}$$

Proof. Suppose $x \in \mathbf{I}$ and Ψ is a positive real-valued function on $\mathbf{N}$. If x is Ψ-approximable, then there are infinitely many $q \in \mathbf{N}$ such that (5) has a solution of the form $\langle p, q\rangle$, where $p \in \mathbf{Z}$. Given that $q > 0$, (5) is equivalent to (7), and we have $\|qx\| \leq |qx - p|$ by the definition of $\|qx\|$, which together with (7) implies that (8) holds. And this is the case for all of the infinitely many such q.

If there are infinitely many $q \in \mathbf{N}$ such that (8) holds, then for every such q there is a $p \in \mathbf{Z}$ with that $\|qx\| = |qx - p|$ and therefore (7) holds, which is equivalent to (5). □

Note that $W(\Psi)$ can be expressed as a limit superior of a sequence of finite unions of open intervals:

$$W(\Psi) = \limsup_{q\to\infty} \mathbf{I} \cap \bigcup_{p=0}^{q} \left(\frac{p}{q} - \Psi(q), \frac{p}{q} + \Psi(q)\right).$$

However, the limit superior of a sequence of finite unions of open intervals is not necessarily open itself, so $W(\Psi)$ has a more interesting structure than the sets of the form $W_n(\Psi)$ with $n \in \mathbf{N}$, which are merely unions of open intervals.

3.1. **Strongly approximating functions.** Another property of real numbers x relating to solutions of (5) that we might naturally be interested in is *strong Ψ-approximability.* We shall say that x is strongly Ψ-approximable, or that x is strongly approximable to the order of Ψ, or that Ψ is a strongly approximating function of x, if and only if for every positive real number ε, the equation (6)

$$\left|x - \frac{p}{q}\right| < \varepsilon\Psi(q)$$

has at least one solution. The set of the strongly Ψ-approximable members of $\mathbf{I}$ is denoted $W^*(\Psi)$. The members of $W^*(\Psi)$ are exactly the members of $\mathbf{I}$ approximable by rational numbers with arbitrarily small relative errors with respect to Ψ.

It is not too difficult to see that strong Ψ-approximability implies Ψ-approximability.

Proposition 2. *For every positive real-valued function Ψ on $\mathbf{N}$,*

$$\mathbf{I} \cap \mathbf{Q} \subseteq W(\Psi).$$

Proof. Suppose $x \in \mathbf{I} \cap \mathbf{Q}$ and Ψ is a positive real-valued function on $\mathbf{N}$. Then there is a $q \in \mathbf{N}$ such that $q > 0$ and $qx \in \mathbf{Z}$. Because $q > 0$, we have

$$q < 2q < 3q < \dots$$

and for every $n \in \mathbf{Z}$, we have $nqx \in \mathbf{Z}$, because $qx \in \mathbf{Z}$, and therefore $\|nqx\| = 0$. It follows that x is Ψ-approximable. $\square$

Proposition 3. *For every positive real-valued function Ψ on $\mathbf{N}$,*

$$W^*(\Psi) \subseteq W(\Psi).$$

Proof. Suppose Ψ is a positive real-valued function on $\mathbf{N}$ and x is a strongly Ψ-approximable number. If $x \in \mathbf{Q}$ then it is Ψ-approximable by Proposition 2; we do not even need to use the assumption of strong Ψ-approximability of x in this case. So let us now suppose that $x \notin \mathbf{Q}$.

By the definition of strong Ψ-approximability there is certainly at least one solution $\langle P, Q\rangle$ of (5), because 1 is a positive real number like any other. Now, for every integer $n \geq Q$, if we let

$$c = \min_{\substack{\langle p,q\rangle \in \mathbf{Z}^2 \\ 0<q\leq n}} \frac{|x - p/q|}{\Psi(q)},$$

then $c \leq |x - P/Q|/\Psi(Q) < 1$. Given that $x \notin \mathbf{Q}$, we have $c > 0$, and therefore, by the strong Ψ-approximability of x, the inequality

$$\left| x - \frac{p}{q} \right| < c\Psi(q).$$

has a solution $\langle p, q\rangle$. This inequality implies that $|x - p/q| < \Psi(q)$ (because $c < 1$) and is equivalent to $|x - p/q|/\Psi(q) < c$, from which it follows that $q > n$. Therefore, there is a solution $\langle p, q\rangle$ of (5) with $q > n$. This is the case for every $n \geq Q$, so (5) has solutions $\langle p, q\rangle$ with q arbitrarily large. Therefore, x is Ψ-approximable. $\square$

It is also easy to see that strong Ψ-approximability implies strong $c\Psi$-approximability for every $c \in \mathbf{R}$ with $c > 0$.

Proposition 4. *For every positive real-valued function Ψ on $\mathbf{N}$ and every $c \in \mathbf{R}$ with $c > 0$,*

$$W^*(\Psi) \subseteq W^*(c\Psi).$$

Proof. Suppose Ψ is a positive real-valued function on $\mathbf{N}$, $c \in \mathbf{R}$, $c > 0$ and x is a strongly Ψ-approximable number. Then for every positive $\varepsilon \in \mathbf{R}$, $\varepsilon c > 0$ and therefore the inequality

$$\left| x - \frac{p}{q} \right| < \varepsilon c\Psi(q)$$

has a solution, so x is strongly $c\Psi$-approximable. $\square$

From Proposition 4 together with Proposition 3 it follows that strong Ψ-approximability implies $\varepsilon\Psi$-approximability for every positive $\varepsilon > 0$. Moreover, the converse is obviously true because the existence of infinitely many of something implies the existence of one of it. Therefore, we have a characterisation of strong Ψ-approximability in terms of Ψ-approximability.

Proposition 5. *For every positive real-valued function Ψ on $\mathbf{N}$,*

$$W^*(\Psi) = \bigcap_{\substack{\varepsilon \in \mathbf{R} \\ \varepsilon > 0}} W(\varepsilon\Psi).$$

3.2. **The approximability hierarchy.** Let's now consider how the Ψ-approximability conditions for different positive real-valued functions Ψ on $\mathbf{N}$ are logically related to another. In particular, we shall describe how the conditions can be ordered by their "strength" (that is, which conditions they imply and are implied by).

Proposition 6. *For every pair of positive real-valued functions Φ and Ψ on $\mathbf{N}$ such that $\Phi(q) \leq \Psi(q)$ for every sufficiently large $q \in \mathbf{N}$,*

$$W(\Phi) \subseteq W(\Psi).$$

Proof. Suppose Φ and Ψ are positive real-valued functions on $\mathbf{N}$, $Q \in \mathbf{N}$, $\Phi(q) \leq \Psi(q)$ for every $q \in \mathbf{N}$ such that $q > Q$, and x is a Φ-approximable number. Then there are infinitely many $q \in \mathbf{N}$ such that

$$\|qx\| < q\Phi(q),$$

and at most Q of these are less than or equal to Q, so for the remaining $q \in \mathbf{N}$, of which there are infinitely many, (5) holds in addition to the inequality above because $q\Phi(q) \leq q\Psi(q)$. It follows that x is Ψ-approximable. $\square$

If two positive real-valued functions Φ and Ψ on $\mathbf{N}$ satisfy the condition that $\Phi(q) \leq \Psi(q)$ for every sufficiently large $q \in \mathbf{N}$, then, and only then, we shall say that Φ is less than or equal to Ψ and write $\Phi \leq \Psi$. Thus we partially order the positive real-valued functions on $\mathbf{N}$. Proposition 6 says that if $\Phi \leq \Psi$, then $W(\Phi) \subseteq W(\Psi)$. The positive real-valued functions on $\mathbf{N}$ are thus ordered by the strength of the approximability conditions defined with reference to them.

Note also that Φ and Ψ are related by the asymptotic equation

$$\Phi(q) = o(\Psi(q)) \quad (q \to \infty) \tag{9}$$

if and only if $\Phi \leq \varepsilon\Psi$ for every positive $\varepsilon > 0$, and therefore (9) implies that $W(\Phi) \subseteq W^*(\Psi)$.

The order relation that we have just defined is only partial. But it does totally order some sets of positive real-valued functions on $\mathbf{N}$. One example is the set of all functions on $\mathbf{N}$ of the form $\mathrm{id}_{\mathbf{N}}^{\tau}$, where $\tau \in \mathbf{R}$. Let's introduce some convenient terminology relating to functions of this form specifically. For every $x \in \mathbf{R}$ and every $\tau \in \mathbf{R}$, x is said to be *approximable to the order* τ if and only if it is $\mathrm{id}_{\mathbf{N}}^{-\tau}$-approximable. Likewise, x is said to be*strongly approximable to the order* τ if and only if it is strongly $\mathrm{id}_{\mathbf{N}}^{\tau}$-approximable.[1]

The approximating functions that we shall refer to in the rest of this document are mainly of the form $c\,\mathrm{id}_{\mathbf{N}}^{-\tau}$, where c is a positive real number and n is a non-negative real number. It may be helpful to write down a list of representative approximability conditions defined with reference to functions of this form, listed in order from strongest to weakest, so that the list can be thought of as a hierarchy of approximability conditions. Note that this list is by no means complete; other conditions could be added at the end, or in between items.

(1) $c_{\mathbf{N}}$-approximability for some $c \in \mathbf{R}$ such that $c > 0$
   (a) $c_{\mathbf{N}}$-approximability for some $c \in \mathbf{R}$ such that $c > 1$
   (b) approximability to the order 0
   (c) $c_{\mathbf{N}}$-approximability for some $c \in \mathbf{R}$ such that $0 < c < 1$
(2) strong approximability to the order 0
(3) $c\,\mathrm{id}_{\mathbf{N}}^{-1}$-approximability for some $c \in \mathbf{R}$ such that $c > 0$
   (a) $c\,\mathrm{id}_{\mathbf{N}}^{-1}$-approximability for some $c \in \mathbf{R}$ such that $c > 1$
   (b) approximability to the order 1
   (c) $c\,\mathrm{id}_{\mathbf{N}}^{-1}$-approximability for some $c \in \mathbf{R}$ such that $0 < c < 1$
(4) strong approximability to the order 1
(5) $c\,\mathrm{id}_{\mathbf{N}}^{-2}$-approximability for some $c \in \mathbf{R}$ such that $c > 0$
   (a) $c\,\mathrm{id}_{\mathbf{N}}^{-2}$-approximability for some $c \in \mathbf{R}$ such that $c > 1$
   (b) approximability to the order 2
   (c) $c\,\mathrm{id}_{\mathbf{N}}^{-2}$-approximability for some $c \in \mathbf{R}$ such that $0 < c < 1$
(6) strong approximability to the order 2

A member of $\mathbf{R}$ can be considered to be "positioned at" some place in the hierarchy of approximability conditions if it satisfies the approximability condition at that place (which implies that it satisfies all the approximability conditions given in the previous items), but it does not satisfy the approximability condition at the next place. In the next section, we describe the members of $\mathbf{I}$ that are positioned at some of the places in the hierarchy.

[1]Note that this terminology is different from that of Hardy & Wright in [3] (p. 158); by their definition an $x \in \mathbf{R}$ is approximable to the order τ if and only if it is $c\,\mathrm{id}_{\mathbf{N}}^{\tau}$-approximable for some positive $c \in \mathbf{R}$.

4. Dirichlet's theorem

We have already seen that for every $x \in \mathbf{R}$ and every positive $\varepsilon \in \mathbf{R}$, the inequality

$$\left|x - \frac{p}{q}\right| < \varepsilon$$

always has at least one solution by the density of $\mathbf{Q}$ in $\mathbf{R}$. This means that $W^*(1_{\mathbf{N}}) = \mathbf{I}$. It is also very easy to see that for every $c \in \mathbf{R}$ such that $c > 1/2$, we have $W(c\,\mathrm{id}_{\mathbf{N}}^{-1}) = \mathbf{I}$. In fact, for every $x \in \mathbf{R}$ and *every* $q \in \mathbf{N}$ (not just infinitely many $q \in \mathbf{N}$), we have

$$\|qx\| \leq \frac{1}{2} < c \tag{10}$$

by the definition of $\|qx\|$.

This is a statement about *every* $q \in \mathbf{N}$, but a much stronger statements can be made about *infinitely many* $q \in \mathbf{N}$. A theorem proven by Dirichlet in 1842 implies that

$$W^*(\mathrm{id}_{\mathbf{N}}^{-1}) = \mathbf{I}. \tag{11}$$

Dirichlet's theorem makes uses of a principle called the *pigeonhole principle*, which we formally state below as Proposition 7. Informally, the principle states that for every pair of positive integers m and n such that $m < n$, if the members of a set S of n objects are distributed among the sets in a family of m sets, then at least one of the sets in the family will contain two or more members of S. It should be intuitively obvious to the reader, but it can be easily proved if it is not.

Proposition 7 (Pigeonhole principle)**.** *For every set S, every family $\mathcal{F}$ of pairwise disjoint sets such that $\#S > \#\mathcal{F}$ and every function f from S to the union of $\mathcal{F}$, there is at least one member of $\mathcal{F}$ which contains the images under f of two distinct members of S.*

Theorem 1 (Dirichlet, 1842)**.** *For every $x \in \mathbf{R}$ and every $n \in \mathbf{N}$, there is a $q \in \mathbf{N}$ such that*

$$\|qx\| < \frac{1}{n}. \tag{12}$$

Proof. Suppose $x \in \mathbf{R}$ and $n \in \mathbf{N}$. Then the function f on $\{0, 1, \ldots, n\}$ such that $f(k) = \langle kx \rangle$ for every $k \in \mathbf{Z}$ such that $0 \leq k \leq n$ has the codomain $[0, 1)$, which can be partitioned into the n subsets $[0, 1/n)$, $[1/n, 2/n)$, ... and $[(n-1)/n, 1)$. The cardinality of the domain of f is $n+1$, which is greater than n, so at least one of these n subsets contains the images under f of two distinct members r and s of the domain of

f by Proposition 7. Because each of the subsets in the partition is an interval of length $1/n$, and not closed, it follows that

$$\begin{aligned}\frac{1}{n} &> |\langle sx\rangle - \langle rx\rangle| \\ &= |(sx - [sx]) - (rx - [rx])| \\ &= |(s-r)x - ([sx] - [rx])|\end{aligned}$$

and therefore (12) holds if we let $q = s - r$. If we assume without loss of generality that $r < s$, then $q > 0$, and moreover, because $0 \le r \le n$ and $0 \le s \le n$, we have $q \le n$. □

For every positive $\varepsilon \in \mathbf{R}$, if we let $n = [1/\varepsilon] + 1$, then $n \in \mathbf{N}$ and

$$\frac{1}{n} = \frac{1}{[1/\varepsilon]+1} < \frac{1}{1/\varepsilon} = \varepsilon.$$

Therefore, Dirichlet's theorem implies (11). However, Dirichlet's theorem actually implies something stronger than this. It doesn't just say that for every $n \in \mathbf{N}$, there is a $q \in \mathbf{N}$ such that (12) holds; it also says that $q \le n$. Thus it gives us an upper bound on the size of the smallest $q \in \mathbf{N}$ such that $\|qx\| < \varepsilon$, namely $[1/\varepsilon] + 1$. Using this upper bound we can conclude that every real number is in fact approximable to the order 2.

Theorem 2. *Every real number is approximable to the order 2.*

Proof. By Proposition 2, every rational number is approximable to the order 2. Suppose, then, that x is an irrational number and $\langle q_k\rangle_{k=1}^n$ is a finite sequence of positive integers such that for every $k \in \mathbf{Z}$ with $0 < k \le n$,

$$\|q_k x\| < \frac{1}{q_k}.$$

We shall show that there is another $q \in \mathbf{N}$ such that $\|qx\| < 1/q$ which is distinct from all of the terms of $\langle q_k\rangle_{k=1}^n$.

Let

$$c = \min_{k=1}^{n} \|q_k x\|.$$

By the irrationality of x, we have $c > 0$. Let $n = [1/c] + 1$. Then $n \in \mathbf{N}$, so by Dirichlet's theorem there is a $q \in \mathbf{N}$ such that

$$\|qx\| < \frac{1}{n}$$

Because $1/n = 1/([1/c] + 1) < 1/(1/c) = c$ it follows that $q \ne q_k$ for every $k \in \mathbf{Z}$ with $0 < k \le n$. And because $q \le n$, i.e. $1/n \le 1/q$, it follows that $\|qx\| < 1/q$. □

Even Theorem 2 does not give the full picture. It was proven by Hurwitz in 1891 that the smallest positive real number ε such that every real number is $\varepsilon\,\mathrm{id}_{\mathbf{N}}^2$-approximable is in fact $1/\sqrt{5}$. However, we shall omit this proof in this document, and move straight on to the investigation of which real numbers are *strongly* approximable to the order 2.

5. Badly approximable numbers

It is clear that there are real numbers which are strongly approximable to the order 2, and in fact to any order. Indeed, suppose $S \subseteq \mathbf{N}$ and let

$$x = \sum_{k \in S} 2^{-(k!)},$$

so that for every $k \in \mathbf{N}$, the digit at the $(k!)$th place after the radix point of the binary expansion of x is 1 if $k \in S$ and 0 if $k \notin S$, and every digit at a place after the radix point of non-factorial index is 0. For every $n \in \mathbf{N}$, let

$$q_n = 2^{n!},$$
$$p_n = q_n \sum_{k=1}^{n} 2^{-(n!)},$$

so that

$$q_n^n = (2^{n!})^n = 2^{n \cdot n!} = 2^{(n+1) \cdot n! - n!} = 2^{(n+1)! - n!}$$

and

$$\begin{aligned}
\left| x - \frac{p}{q} \right| &= \left| \sum_{k \in S} 2^{-(k!)} - \sum_{k=1}^{n} 2^{-(k!)} \right| = \sum_{\substack{k > n \\ k \in S}} 2^{-(k!)} \\
&\leq \sum_{k=n+1}^{\infty} 2^{-(k!)} \\
&< \sum_{k=(n+1)!}^{\infty} 2^{-k} = 2^{-((n+1)!-1)} \\
&\leq 2^{-((n+1)!-n!)} = \frac{1}{q_n^n}.
\end{aligned}$$

Then for every $\tau \in \mathbf{R}$ with $\tau > 0$, if we let $m = [\tau]$, then for every integer $n > m$ we have $\tau < n$ and therefore $|x - p_n/q_n| < 1/q_n^n < 1/q_n^\tau$. It follows that x is approximable to the order τ. Real numbers like x, which are approximable to every order, are called *Liouvillle numbers.* Note that the choice of S is arbitrary and determines x

uniquely (because the binary expansion of x is certainly not recurring), so there are in fact uncountably many Liouville numbers.

However, there are also real numbers that are not strongly approximable to the order 2. These real numbers are called the *badly approximable numbers*. We shall denote the set of the badly approximable members of $\mathbf{I}$ by $\mathbf{B}$. In order to prove the existence of badly approximable numbers, we shall prove that every quadratic irrational number is badly approximable. There are a number of different ways to do this, but in this document we shall do it by proving a more general theorem, first proven by Liouville in 1844, which says that for every integer $n > 1$, no algebraic number of degree n is strongly approximable to the order n.

Theorem 3 (Liouville, 1844)**.** *For every integer $n > 1$, no algebraic member number of degree n is strongly approximable to the order n.*

Proof. Suppose $n \in \mathbf{Z}$, $n > 1$ and x is an algebraic number of degree n. Then there is a finite sequence $\langle a_k \rangle_{k=1}^{n}$ of integers such that $a_n \neq 0$ and

$$a_0 + a_1 x + \cdots + a_n x^n = 0.$$

Let f be the function on $\mathbf{R}$ such that for every $y \in \mathbf{R}$,

$$f(y) = a_0 + a_1 y + \cdots + a_n y^n,$$

If x is the only root of f, let $c = 1$. Otherwise, let

$$c = \min_{\substack{y \in \mathbf{R} \\ x \neq y \\ f(y)=0}} |x - y|.$$

This minimum exists because f has at most n roots. It is also positive by its definition. Note that f' is continuous and therefore bounded on every bounded interval in $\mathbf{R}$. In particular it is bounded on $(x-c, x+c)$ by some positive $M \in \mathbf{R}$.

For every pair (p, q) of integers such that $q > 0$ and

$$\left| x - \frac{p}{q} \right| < c,$$

the rational number p/q is closer to x than every root of f distinct from x by the definition of c, but it is not equal to x because $n > 1$ and therefore $x \notin \mathbf{Q}$ (the algebraic numbers of degree less than or equal to

1 are the rational numbers). Therefore,

$$\begin{aligned} 0 &\neq f\left(\frac{p}{q}\right) \\ &= a_0 + \frac{a_1 p}{q} + \cdots + \frac{a_n p^n}{q^n} \\ &= \frac{a_0 q^n + a_1 p q^{n-1} + \cdots + a_n p^n}{q^n}. \end{aligned}$$

Given that p and q are integers, the numerator of the fraction at the end of this chain of equations and inequalities is an integer, and given that $f(p/q) \neq 0$ it is nonzero; therefore,

$$\frac{1}{q^n} \leq \left| f\left(\frac{p}{q}\right) \right| = \left| -f\left(\frac{p}{q}\right) \right| = \left| 0 - f\left(\frac{p}{q}\right) \right| = \left| f(x) - f\left(\frac{p}{q}\right) \right|.$$

Let

$$I = \begin{cases} [p/q, x] & \text{if } p/q < x \\ [x, p/q] & \text{if } p/q > x, \end{cases}$$

so that I is a closed interval on which f is differentiable, and which is a subset of $(x - c, x + c)$ (so that $|f'(y)| < M$ for every $y \in I$). Then by the mean value theorem there is a $y \in \mathbf{R}$ strictly between p/q and x such that

$$\frac{|f(x) - f(p/q)|}{|x - p/q|} = |f'(y)|,$$

i.e.

$$|f'(y)| \left| x - \frac{p}{q} \right| = \left| f(x) - f\left(\frac{p}{q}\right) \right| \geq \frac{1}{q^n}.$$

It follows that

$$\left| x - \frac{p}{q} \right| \geq \frac{1}{|f'(y)| q^n} > \frac{1}{M q^n}.$$

This is the case for every pair $\langle p, q \rangle$ of integers with $q > 0$, so x is not strongly approximable to the order n. □

It is worth noting that Liouville's theorem does not say that for every $n \in \mathbf{Z}$ such that $n > 1$, n is the *least* $\tau \in \mathbf{R}$ such that no algebraic member of $\mathbf{I}$ of degree n is strongly approximable to the order τ. In fact, that statement is true only in the case where $n = 2$. In 1955, K. F. Roth proved [4] that no algebraic number of degree greater than 2 is strongly approximable to any order greater than 2. The same is of course the case for algebraic numbers of degree exactly 2, but this was already implied by Liouville's theorem. Now, it follows from Roth's theorem that for every $n \in \mathbf{Z}$ such that $n > 2$, the least $\tau \in \mathbf{R}$ such that no algebraic member of $\mathbf{I}$ of degree n is strongly approximable to the order τ either does not exist, or is equal to 2. (It cannot be less than 2 by Theorem 2 and the fact that for every $\tau \in \mathbf{R}$ such that $\alpha < 1$, $q^{-2} = o(q^{-\tau})$ as q approaches ∞.) Remarkably, it remains an open question which of these two possibilities is the case. A proof of the existence of just one badly approximable algebraic number of degree greater than 2 would imply the former was the case, and a proof of the nonexistence of any such number would imply the latter was the case. Nobody has even been able to prove that any algebraic number of degree greater than 2 is *not* badly approximable.

It is, however, easy enough to prove that there are badly approximable numbers that are not quadratic, and to exhibit examples of such numbers. In fact, *most* badly approximable numbers are not quadratic, because $\mathbf{B}$ is uncountable but the set of the quadratic numbers is countable. We shall prove the uncountability of $\mathbf{B}$ in the next section by making use of a generalized Cantor set construction.

6. The cardinality of $\mathbf{B}$

6.1. **The Cantor set.** The *Cantor set* $\mathbf{K}$ is the set of the members of $[0, 1]$ whose ternary expansions do not contain the digit 1. It can also be expressed as

$$\bigcap_{n=0}^{\infty} \mathbf{K}_n,$$

where for every integer $n \geq 0$, $\mathbf{K}_n$ is the set of the members of $\mathbf{I}$ whose first n digits in base-3 positional notation after (and not including) the unit digit are all distinct from 1. The sequence $\langle \mathbf{K}_n \rangle_{n \in \mathbf{Z}, n \geq 0}$ can also be defined by induction. In order to state this definition, let us introduce some convenient notation.

Definition 1. For every $b \in \mathbf{N}$, every integer $n \geq 0$ and every $k \in \mathbf{Z}$ such that $0 \leq k < b^n$,

$$I_{b,n,k} = \left[\frac{k}{b^n}, \frac{k+1}{b^n} \right].$$

The common denotation of the expressions on both sides of the above equation is said to be of *base* b, *order* n and *index* k.

Note that for every $b \in \mathbf{N}$, we have:

(1) $I_{b,0,0} = [0, 1]$.
(2) For every integer $n \geq 0$ and every $k \in \mathbf{Z}$ such that $0 \leq k < b^n$, $I_{b,n,k}$ is the finite union of the closed intervals $I_{b,n+1,kb}$, $I_{b,n+1,kb+1}$, ... and $I_{b,n+1,kb+(b-1)}$, and no intersection of two of

these intervals has more than one member. Also, the length of $I_{b,n,k}$ is $1/b^n$.

The definition by induction of $\langle \mathbf{K}_n \rangle_{n\in\mathbf{N}}$ is as follows.

(1) $\mathbf{K}_0 = [0, 1]$.
(2) For every integer $n \geq 0$,

$$\mathbf{K}_{n+1} = \bigcup_{\substack{k\in\mathbf{Z} \\ 0\leq k<3^n \\ I_{n,k}\subseteq\mathbf{K}_n}} I_{n+1,3k} \cup I_{n+1,3k+2}.$$

In other words, $\mathbf{K}_{n+1}$ is the result of partitioning every interval in $\mathbf{K}_n$ of the form $I_{3,n,k}$ where $0 \leq k < 3^n$ into its three subintervals of order $n+1$, namely $I_{3,n+1,3k}$, $I_{3,n+1,3k+1}$ and $I_{3,n+1,3k+2}$, and removing the middle interval $I_{3,n+1,3k+1}$.

It is easily seen that $\mathbf{K}$ is uncountable. Let f be the function on the set of the non-cofinite subsets of $\mathbf{N}$ such that for every $S \subseteq \mathbf{N}$,

$$f(S) = \sum_{n\in S} 2 \cdot 3^{-n}.$$

Then the expansion of $f(S)$ in base-3 positional notation does not contain the digit 1, so $f(S) \in \mathbf{K}$. And because S is not cofinite, the expansion contains the digit 0 at infinitely many places—it does not end in a recurring sequence of unit digits. It follows that f is injective and hence there are at least as many members of $\mathbf{K}$ as there are non-cofinite subsets of $\mathbf{N}$. There are uncountably many such subsets of $\mathbf{N}$ because there are no more cofinite subsets of $\mathbf{N}$ than there are finite subsets of $\mathbf{N}$, and there are only countably many finite subsets of $\mathbf{N}$.

In the next section, we generalize this construction in order to construct a subset of $\mathbf{B}$ which can be easily shown to be uncountable, and thereby prove the uncountability of $\mathbf{B}$. There are lots of different ways the construction of the Cantor set can be generalized, but what we want to do, in particular, is to generalize the base of the intervals referred to in the definition, and the choice of which intervals are removed.

6.2. **Generalized Cantor sets.** The following definition is adapted from [1], although it is presented a little differently. The proof that $\#\mathbf{B} > \aleph_0$ on the next page is also adapted from the same source.

Definition 2. For every $b \in \mathbf{N}$ and every family $\mathcal{F}$ of intervals of base b and positive order,

$$\mathbf{K}(b, \mathcal{F}) = \bigcap_{n=0}^{\infty} \mathbf{K}_n(b, \mathcal{F})$$

where:

(1) $\mathbf{K}_0(b,\mathcal{F}) = [0,1]$.
(2) For every integer $n \geq 0$,

$$\mathbf{K}_{n+1}(b,\mathcal{F}) = \bigcup_{\substack{k\in\mathbf{Z}\\ 0\leq k<b^n\\ I_{b,n,k}\subseteq \mathbf{K}_n(b,\mathcal{F})}} \bigcup_{\substack{l\in\mathbf{Z}\\ 0\leq l<b\\ I_{b,n+1,kb+l}\in\mathcal{F}}} I_{b,n+1,kb+l}$$

In other words, $\mathbf{K}_{n+1}(b,\mathcal{F})$ is the result of partitioning every interval in $\mathbf{K}_n(b,\mathcal{F})$ of the form $I_{b,n,k}$, where $0 \leq k < b^n$, into its b subintervals of order $n+1$, namely $I_{b,n+1,kb}$, $I_{b,n+1,kb+1}$, $\ldots$ and $I_{b,n+1,kb+(b-1)}$, and removing the subintervals that are not members of $\mathcal{F}$.

The common denotation of the expressions on both sides of the equation above is called the *generalized Cantor set* of base b and *filter* $\mathcal{F}$.

Not every generalized Cantor set as defined by Definition 2 is uncountable. For example, $\#\mathbf{K}(b,\emptyset) = 0$ for every $b \in \mathbf{N}$. But we can easily establish a criterion for uncountability. In informal terms, if, for every integer $n \geq 0$, at least two subintervals always remain unremoved from every interval in $\mathbf{K}_n(b,\mathcal{F})$ during the construction of $\mathbf{K}_{n+1}(b,\mathcal{F})$, and these two intervals are not adjacent to one another (so that they are disjoint), then $\mathbf{K}(b,\mathcal{F})$ is uncountable.

Definition 3. For every $b \in \mathbf{N}$ and every family $\mathcal{F}$ of intervals of base b and positive order, $\mathbf{K}(b,\mathcal{F})$ is said to be *strictly branching* if and only if for every integer $n \geq 0$ and every $k \in \mathbf{Z}$ such that $0 \leq k < b^n$ and $I_{b,n,k} \subseteq \mathbf{K}_n(b,\mathcal{F})$, there is a pair $\langle i,j\rangle$ of integers such that $0 \leq i+1 < j < b$ and

$$I_{b,n+1,kb+i} \in \mathcal{F},$$
$$I_{b,n+1,kb+j} \in \mathcal{F}.$$

Theorem 4. *For every $b \in \mathbf{N}$ and every family $\mathcal{F}$ of intervals of base b and positive order such that $\mathbf{K}(b,\mathcal{F})$ is strictly branching,*

$$\#\mathbf{K}(b,\mathcal{F}) > \aleph_0. \tag{13}$$

Proof. Suppose $b \in \mathbf{N}$, $\mathcal{F}$ is a family of intervals of base b and positive order, $\mathbf{K}(b,\mathcal{F})$ is strictly branching and $\langle x_n\rangle_{n\in\mathbf{N}}$ is a sequence of real numbers such that $x_m \neq x_n$ for every pair (m,n) of positive integers. We shall prove the existence of an $x \in \mathbf{K}(b,\mathcal{F})$ such that $x \neq x_n$ for every $n \in \mathbf{N}$, and thus prove that every countable subset of $\mathbf{K}(b,\mathcal{F})$ is a proper subset of $\mathbf{K}(b,\mathcal{F})$.

Let

$$I = \bigcap_{n=1}^{\infty} I_{b,n,k_n}$$

where $\langle k_n \rangle_{n \in \mathbf{Z}, n \geq 0}$ is a sequence defined inductively as follows.

(1) $k_0 = 0$.

(2) For every integer $n \geq 0$, if we assume as an inductive hypothesis that $0 \leq k_n < b^n$ and $I_{b,n,k_n} \subseteq \mathbf{K}_n(b, \mathcal{F})$, then by the strict branchingness of $\mathbf{K}(b, \mathcal{F})$ there is a pair $\langle i, j \rangle$ of integers such that $0 \leq i + 1 < j < b$ and $I_{b,n+1,k_n b+i}$ and $I_{b,n+1,k_n b+j}$ are both members of $\mathcal{F}$. Given that $I_{b,n,k_n} \subseteq \mathbf{K}_n(b, \mathcal{F})$, both $I_{b,n+1,k_n b+j}$ and $I_{b,n+1,k_n b+j}$ are subsets of $\mathbf{K}_{n+1}(b, \mathcal{F})$. Given that $i+1 < j$, the intersection $I_{b,n+1,k_n b+i} \cap I_{b,n+1,k_n b+j}$ is empty, and therefore x_{n+1} is not a member of both $I_{b,n+1,k_n b+i}$ and $I_{b,n+1,k_n b+j}$. Therefore, by letting k_{n+1} be $k_n b + i$ or $k_n b + j$ as appropriate, we have $x_{n+1} \notin I_{b,n,k_n}$. We also have $0 \leq k_{n+1} < b^{n+1}$ and $I_{b,n+1,k_{n+1}} \subseteq \mathbf{K}_{n+1}(b, \mathcal{F})$, so the assumption of the inductive hypothesis is valid. Moreover, $I_{b,n+1,k_{n+1}} \subseteq I_{b,n,k_n}$ and the length of I_{b,n,k_n} is b times the length of $I_{b,n+1,k_{n+1}}$.

It is easily seen that the length of I_{b,n,k_n} approaches 0 as n approaches ∞, and because we also have that $I_{b,n+1,k_{n+1}} \subseteq I_{b,n,k_n}$ for every integer $n \geq 0$, it follows by the Cantor intersection theorem that I is a singleton. Let x be its unique member. Then $x \neq x_n$ for every $n \in \mathbf{N}$, because $x \in I_{b,n,k_n}$ and $x_n \notin I_{b,n,k_n}$, and $x \in \mathbf{K}(b, \mathcal{F})$, because $I_{b,n,k_n} \subseteq \mathbf{K}_n(b, \mathcal{F})$ for every integer $n \geq 0$. □

6.3. **Proof that B is uncountable.** In order to prove that $\#\mathbf{B} > \aleph_0$ we need to find a generalized Cantor set $\mathbf{K}' = \mathbf{K}(b, \mathcal{F})$ such that:

(1) $\mathbf{K}' \subseteq \mathbf{B}$.

(2) $\mathbf{K}'$ is strictly branching.

In order for (1) to hold the filter $\mathcal{F}$ must somehow remove every member of $\mathbf{I}$ which is not badly approximable from the set at some stage in the construction. A simple option would be to choose $\mathcal{F}$ to be the set of the intervals of base b and positive order whose members are all badly approximable. But it is easy to see from the construction of the Liouville numbers described in the previous section that there are Liouville numbers, which are of course not badly approximable, in every open interval in $\mathbf{R}$ (and therefore in every closed interval in $\mathbf{R}$ of positive length, including every interval of base b). So this choice of $\mathcal{F}$ gives us that $\mathbf{K}(b, \mathcal{F}) = \emptyset$, and therefore (2) does not hold. A more complicated choice of $\mathcal{F}$ is therefore necessary.

In order to see how we should choose $\mathcal{F}$, first, note that

$$\mathcal{F} = \bigcup_{n=1}^{\infty} \mathcal{F}_n,$$

where for every $n \in \mathbf{N}$, $\mathcal{F}_n$ is the set of the intervals of order n that are members of $\mathcal{F}$. The task at hand is best thought of as the task of choosing $\mathcal{F}$ "piece by piece", i.e. choosing $\mathcal{F}_n$ for every $n \in \mathbf{N}$, so that the eventual infinite union $\mathcal{F}$ filters out every non-badly approximable member of $\mathbf{I}$.

Recall that an $x \in \mathbf{R}$ is badly approximable if and only if there is a $c \in \mathbf{R}$ such that $c > 0$ and

$$\left| x - \frac{p}{q} \right| \geq \frac{c}{q^2} \tag{14}$$

for every pair $\langle p, q \rangle$ of integers with $q > 0$. Suppose that for every $n \in \mathbf{N}$, the members x of every interval in $\mathcal{F}_n$ satisfy (14) for *some* pairs $\langle p, q \rangle$ of integers such that $q > 0$, but this is guaranteed only for the pairs in a subset Q_n of $\mathbf{Z} \times \mathbf{N}$, rather than the pairs in the whole set $\mathbf{Z} \times \mathbf{N}$. Suppose further that

$$\bigcup_{n=1}^{\infty} Q_n = \mathbf{Z} \times \mathbf{N}. \tag{15}$$

For every $x \in \mathbf{K}(b, \mathcal{F})$, there is a sequence $(I_n)_{n \in \mathbf{N}}$ such that for every $n \in \mathbf{N}$, I_n is an interval of base b and order $n + 1$, $I_n \in \mathcal{F}_n$ and I_n contains x. (This can be easily seen from Definition 2.) Therefore, if $\mathcal{F}$ is chosen as just described, (14) holds for every $n \in \mathbf{N}$ and every $\langle p, q \rangle \in Q_n$. Given (15), it follows that x is badly approximable (in fact, we know a specific real number $c > 0$ for which it is not $c\,\mathrm{id}_{\mathbf{N}}^{-2}$-approximable) and hence $\mathbf{K}(b, \mathcal{F}) \subseteq \mathbf{B}$.

Now, we have not said anything about how c and $\langle Q_n \rangle_{n \in \mathbf{N}}$ are chosen. Given (15), every pair (p, q) of integers with $q > 0$ is a member of Q_n for some $n \in \mathbf{N}$. A natural way to choose $\langle Q_n \rangle_{n \in \mathbf{N}}$ so that this is the case is to let f be a monotonic positive real-valued function on $\mathbf{N}$ such that the range of f contains 1 and $f(n+1) - f(n) > 1$ for every $n \in \mathbf{N}$ with $f(n) \geq 1$, and to let

$$Q_n = \{\langle p, q \rangle \in \mathbf{Z} \times \mathbf{N} : \gcd(p, q) = 1 \wedge f(n) \leq q < f(n + 1)\}$$

for every $n \in \mathbf{N}$. This choice of $\langle Q_n \rangle_{n \in \mathbf{N}}$ has the additional convenient properties that the sets Q_1, Q_2, Q_3, … are pairwise disjoint, and Q_n is non-empty for every sufficiently large $n \in \mathbf{N}$.

If $\langle Q_n \rangle_{n\in\mathbf{N}}$ is chosen thus, then for every $n \in \mathbf{N}$ and every $\langle p, q\rangle \in Q_n$ as defined above, we have $f(n)^2 \leq q^2$ and therefore for every $x \in \mathbf{R}$,

$$\left| x - \frac{p}{q} \right| \geq \frac{c}{f(n)^2}$$

implies (14). Therefore, for every $k \in \mathbf{Z}$ such that $0 \leq k < b^n$, $I_{b,n,k}$ satisfies the necessary condition for inclusion in $\mathcal{F}_n$ described above (that (14) holds for every $x \in I_{b,n,k}$ and every $\langle p, q\rangle \in Q_n$) if every member of $I_{b,n,k}$ is at a distance of at least $c/f(n)^2$ from every member of Q_n. Given that the necessary condition for membership in $\mathcal{F}_n$ described above exists, the choice of $\mathcal{F}_n$ is arbitrary, so we may take $\mathcal{F}_n$ to be the set of the intervals of order n whose members are all at distances of at least $c/f(n)^2$ from every member of Q_n. So that we can state this sufficient condition for inclusion is a slightly more concise way, let

$$\Delta(n, p/q) = \left(\frac{p}{q} - \frac{c}{f(n)^2}, \frac{p}{q} + \frac{c}{f(n)^2} \right)$$

for every $\langle p, q\rangle \in Q_n$. We shall call the intervals of the form $\Delta(n, p/q)$, where $\langle p, q\rangle \in Q_n$, the *dangerous intervals* of order n. Then we have $I_{b,n,k} \in \mathcal{F}_n$ if and only if $I_{b,n,k}$ is disjoint from every dangerous interval of order n.

Now, for every $\langle p, q\rangle \in Q_n$, the dangerous interval $\Delta(n, p/q)$ has length $(2c)/(f(n)^2)$, and for every $k \in \mathbf{Z}$ such that $0 \leq k < b^n$, the interval $I_{b,n,k}$ has length $1/b^n$. Therefore, if

$$\frac{2c}{f(n)^2} \leq \frac{1}{b^n},$$

i.e.

$$c \leq \frac{f(n)^2}{2b^n}, \tag{16}$$

then we can be assured that there are no more than two $k \in \mathbf{Z}$ such that $0 \leq k < b^n$ and $I_{b,n,k} \cap \Delta(n, p/q) \neq \emptyset$. It follows that if (16) holds, then among any 5 or more intervals of the form $I_{b,n,k}$, where $k \in \mathbf{Z}$ and $0 \leq k < b^n$, there are two which are disjoint from each other and from $\Delta(p, q)$. Note that in order for it to be possible to choose c and f so that (16) holds for *every* $n \in \mathbf{N}$, the asymptotic equation

$$\frac{1}{f(n)^2} = O\left(\frac{1}{b^n}\right) \quad (n \to \infty)$$

must hold. Clearly we can choose f so that it satisfies this asymptotic equation; for example, if $f(n) = b^{n/2}$ for every $n \in \mathbf{N}$, then f satisfies the equation, f is monotonic and real-valued, $f(0) = 1$, and $f(n+1) -$

$f(n) = b^{n/2}(b-1) \geq 5^{1/2}(5-1) = 4\sqrt{5} > 1$ for every $n \in \mathbf{N}$, given that $b \geq 5$ as we assumed above. For this particular choice of f, any choice of c such that $c \leq 1/2$ will give us (16) for every $n \in \mathbf{N}$.

Choosing c and f so that (16) holds for every $n \in \mathbf{N}$ *almost* makes property (2) to hold. Suppose $n \in \mathbf{Z}$, $n \geq 0$, $k \in \mathbf{Z}$, $0 \leq k < b^n$ and $I_{b,n,k} \subseteq \mathbf{K}_n(b, \mathcal{F})$. Property (2) requires that there be a pair $\langle i, j \rangle$ of integers such that such that $0 \leq i + 1 < j < b$ and

$$I_{b,n+1,kb+i} \in \mathcal{F}_{n+1},$$
$$I_{b,n+1,kb+j} \in \mathcal{F}_{n+1}.$$

Every dangerous interval $\Delta(n, p/q)$, where $\langle p, q \rangle \in {Q_n}^2$, which is disjoint from $I_{b,n,k}$ is also disjoint from $I_{b,n+1,kb+l}$ for every $l \in \mathbf{Z}$ such that $0 \leq l < b$, because $I_{b,n+1,kb+l} \subseteq I_{b,n,k}$. So if it happens that there is at most one dangerous interval $\Delta(n, p/q)$, where $(p, q) \in Q_n$, such that $I_{b,n,k} \cap \Delta(n, p/q) \neq \emptyset$, and we also have $b \geq 5$, then by the reasoning in the previous paragraph there is a pair $\langle i, j \rangle$ of integers such that $0 \leq i + 1 < j < b$ and both $I_{b,n+1,kb+i}$ and $I_{b,n+1,kb+j}$ are disjoint from $\Delta(n, p/q)$ as well as all the other dangerous intervals of order n. Therefore, $I_{b,n+1,kb+i} \in \mathcal{F}$ and $I_{b,n+1,kb+j} \in \mathcal{F}$. But without these additional assumptions, property (2) does not follow. To complete the proof, then, we must prove that c and f can be chosen so that for every $n \in \mathbf{N}$ and every $k \in \mathbf{Z}$ such that $0 \leq k < b^n$, there is indeed at most one dangerous interval $\Delta(n, p/q)$ such that $I_{b,n,k} \cap \Delta(n, p/q) \neq \emptyset$.

Let's first note that for every $\langle p, q \rangle \in Q_n$, we have $q < f(n+1)$. From this we see that for every pair of distinct rational numbers of the forms p/q and p'/q, where $\langle p, q \rangle \in Q_n$ and $\langle p', q' \rangle \in Q_n$,

$$\left| \frac{p}{q} - \frac{p'}{q'} \right| = \left| \frac{pq' - p'q}{qq'} \right| = \frac{|pq' - p'q|}{qq'} \geq \frac{1}{qq'} > \frac{1}{f(n+1)^2},$$

because the equations $|pq' - p'q| = 0$, $pq' = p'q$ and $p/q = p'/q'$ are equivalent, and pq' and $p'q$ are both integers. The distance between the midpoints of every pair $\langle \Delta(p/q), \Delta(p'/q') \rangle$ of distinct dangerous intervals of order n is therefore greater than $1/f(n+1)^2$.

Because the length of every dangerous interval of order n is $(2c)/(f(n)^2)$, it follows that for every $x \in \Delta(p/q)$ and every $y \in \Delta(p'/q')$,

$$|x - y| \geq \frac{1}{f(n+1)^2} - \frac{2c}{f(n)^2} = \frac{f(n)^2 - 2cf(n+1)^2}{(f(n)f(n+1))^2}.$$

Therefore, every interval in $\mathbf{R}$ of length less than $(f(n)^2 - 2cf(n+1)^2)/(f(n)f(n+1))^2$ has a non-empty intersection with at most one

[2]Technically we have not defined Q_n in the case where $n = 0$, but there is no problem with taking Q_n to be empty in that case.

dangerous interval of length n. The interval $I_{b,n,k}$ has length $1/b^n$, so the desired property exists if

$$\frac{1}{b^n} < \frac{f(n)^2 - 2cf(n+1)^2}{(f(n)f(n+1))^2},$$

i.e.

$$c < \frac{f(n)^2(b^n - f(n+1)^2)}{2b^n f(n+1)^2} = \frac{f(n)^2}{2b^n} \cdot \frac{b^n - f(n+1)^2}{f(n+1)^2} \tag{17}$$

Given (16), this inequality is implied by $(b^n - f(n+1)^2)/(f(n+1)^2) > 1$, which is equivalent to

$$f(n+1)^2 < \frac{b^n}{2}. \tag{18}$$

And f can indeed be chosen so that (18) holds for every $n \in \mathbf{N}$. For example, consider the function f on $\mathbf{N}$ such that for every $n \in \mathbf{N}$,

$$f(n) = b^{(n-3)/2}.$$

This choice of f is monotonic, positive real-valued, and we have $f(3) = 1$. For every integer $n \geq 3$, $f(n+1) - f(n) = b^{(n-3)/2}(b-1) \geq 5^{(3-3)/2}(5-1) = 4 > 1$, given that $b \geq 4$. And, for every $n \in \mathbf{N}$, we have

$$f(n+1)^2 = (b^{(n-2)/2})^2 = b^{n-2} = \frac{b^n}{b^2} \leq \frac{b^n}{25} < \frac{b^n}{2},$$

again given that $b \geq 5$, so (18) holds. Moreover, this function satisfies (16) for every $n \in \mathbf{N}$, because

$$\frac{f(n)^2}{2b^n} = \frac{(b^{(n-3)/2})^2}{2b^n} = \frac{b^{n-3}}{2b^n} = \frac{1}{2b^3} \geq c,$$

if we choose c so that $c \leq 1/(2b^3)$. This completes the proof that $\mathbf{B}$ is uncountable.

The result is stated again below in a more concise form.

Theorem 5. *Suppose $b \in \mathbf{Z}$, $b \geq 5$, $c \in \mathbf{R}$ and $0 < c \leq 1/(2b^3)$. For every $n \in \mathbf{N}$, let*

$$Q_n = \{\langle p, q\rangle \in \mathbf{Z} \times \mathbf{N} : \gcd(p,q) = 1 \wedge b^{(n-3)/2} \leq q < b^{(n-2)/2}\}.$$

For every $\langle p, q\rangle \in Q_n$, let

$$\Delta(n, p/q) = \left(\frac{p}{q} - \frac{1}{2b^n}, \frac{p}{q} + \frac{1}{2b^n}\right).$$

And, let $\mathcal{F}_n$ be the set of the intervals of the form $I_{b,n,k}$, where $0 \le k < b^n$, that are disjoint from every interval of the form $\Delta(n, p/q)$, where $\langle p, q\rangle \in Q_n$. Then

$$\mathbf{K}(b, \mathcal{F}) \subseteq \mathbf{B},$$

where

$$\mathcal{F} = \bigcup_{n=1}^{\infty} \mathcal{F}_n.$$

In fact, no member of $\mathbf{K}(b, \mathcal{F})$ is $c\,\mathrm{id}_{\mathbf{N}}^{-2}$-approximable.

7. The Lebesgue measure of B

In the previous section we proved something about the size of **B** in terms of cardinality. But there are other meanings which the word "size" can have when applied to a set of real numbers. One of the most important of these notions of "size" is Lebesgue measure. We shall show that even though **B** is uncountable and thus large, it is small in another sense, because its Lebesgue measure is 0, which is the smallest Lebesgue measure possible.

7.1. Lebesgue measure.

A thorough definition and description of the properties of Lebesgue measure can be found in textbooks such as [5]. In this section, we note only the properties that we use to prove that **B** is uncountable, and we do not prove these propositions.

For every Lebesgue measurable $S \subseteq \mathbf{R}$, the Lebesgue measure of S is denoted $\mu(S)$. The modifier "Lebesgue measurable" is necessary in this statement because there are subsets of **R** which are not Lebesgue measurable. The following basic properties exist.

(1) For every Lebesgue measurable $S \subseteq \mathbf{R}$, $\mu(S)$ is either a non-negative real number or ∞.
(2) For every pair of real numbers a and b such that $a < b$, the measure of every interval in **R** whose endpoints are a and b is $b - a$.
(3) For every $x \in \mathbf{R}$, $\mu(\{x\}) = 0$ (because $\{x\} = [x, x]$ and $x - x = 0$).
(4) For every countable family $\mathcal{F}$ of pairwise disjoint Lebesgue measurable subsets of **R**, the union of $\mathcal{F}$ is Lebesgue measurable and

$$\mu\left(\bigcup_{S\in\mathcal{F}} S\right) = \sum_{S\in\mathcal{F}} \mu(S).$$

(5) For every countable $S \subseteq \mathbf{R}$, $\mu(S) = 0$ (because S is the union of the sets of the form $\{x\}$, where $x \in S$, and this union is countable).
(6) For every Lebesgue measurable $S \subseteq \mathbf{R}$, $\mathbf{R} \setminus S$ is Lebesgue measurable.
(7) For every pair of Lebesgue measurable sets S and T of real numbers such that $S \subseteq \mathbf{R}$, $\mu(S) \leq \mu(T)$ and $\mu(T \setminus S) = \mu(T) - \mu(S)$ (because $\mu(T) = \mu(S \cup (T \setminus S)) = \mu(S) + \mu(T \setminus S)$ and $\mu(T \setminus S) \geq 0$).
(8) For every countable family $\mathcal{F}$ of Lebesgue measurable subsets of $\mathbf{R}$ (which are not necessarily pairwise disjoint), the union and intersection of $\mathcal{F}$ are Lebesgue measurable.
(9) For every Lebesgue measurable $S \subseteq \mathbf{R}$ and every $a \in \mathbf{R}$, $\mu(\{x + a : x \in S\} = \mu(S)$.
(10) For every countable family $\mathcal{F}$ of Lebesgue measurable subsets of $\mathbf{R}$ (which are not necessarily pairwise disjoint),

$$\mu\left(\bigcup_{S \in \mathcal{F}} S\right) \leq \sum_{S \in \mathcal{F}} \mu(S).$$

There is also a more complex property that we need to describe. For every Lebesgue measurable $S \subseteq \mathbf{R}$, every $x \in S$ and every $r \in \mathbf{R}$ such that $r > 0$,

$$\frac{\mu((x - r, x + r) \cap S}{2r}$$

is the ratio of the measure of the portion of S that overlaps $(x - r, x - r)$ to the measure of $(x - r, x - r)$ (which is $2r$). It can be thought of as the proportionate measure of S in $(x - r, x + r)$. The limit

$$\lim_{r \to 0} \frac{\mu((x - r, x + r) \cap S)}{2r}$$

is called the *density* of S at x. There is a theorem on the densities of members of Lebesgue measurable sets known as the Lebesgue density theorem, which we shall use in the next section. It is stated below without proof, but its proof can be found in any measure theory textbook such as [5].

Theorem 6. *For every Lebesgue measurable $S \subseteq \mathbf{R}$, the density of S at x exists and is equal to 1 for almost every member x of S.*

Note that the phrase “almost every member x of S” here refers to every member of S except for the members of a particular subset of S which has Lebesgue measure 0. This is standard terminology. For example, $\mu(\mathbf{Q}) = 0$, because $\mathbf{Q}$ is countable, and therefore we can say that almost every real number is irrational.

7.2. **Proof that the Lebesgue measure of B is 0.** Recall that for every positive real-valued function Ψ on $\mathbf{N}$,

$$\begin{aligned} W(\Psi) &= \limsup_{q\to\infty} \mathbf{I} \cap \left(\frac{p}{q} - \Psi(q), \frac{p}{q} + \Psi(q)\right) \\ &= \bigcap_{t=1}^{\infty} \bigcup_{q=t}^{\infty} \mathbf{I} \cap \left(\frac{p}{q} - \Psi(q), \frac{p}{q} + \Psi(q)\right). \end{aligned}$$

It is helpful to introduce some notation that allows us to express this equation more concisely.

(1) For every $q \in \mathbf{N}$, let

$$A_q(\Psi) = \mathbf{I} \cap \bigcup_{p=0}^{q} \left(\frac{p}{q} - \Psi(q), \frac{p}{q} + \Psi(q)\right).$$

$A_q(\Psi)$ is the set of the $x \in \mathbf{I}$ such that $\|qx\| < q\Psi(q)$.

(2) For every $t \in \mathbf{N}$, let

$$A_t^{\infty}(\Psi) = \bigcup_{q=t}^{\infty} A_q(\Psi),$$

$A_t^{\infty}(\Psi)$ is the set of the $x \in \mathbf{I}$ such that $\|qx\| < q\Psi(q)$ for some $q \in \mathbf{Z}$ with $q \geq t$.

Using this notation, we have

$$W(\Psi) = \bigcap_{t=1}^{\infty} A_t^{\infty}.$$

The following theorem, which appears in [2], relates the Lebesgue measure of $W(\Psi)$ to the Lebesgue measures of the sets of the form $W(c\Psi)$, where $c \in \mathbf{R}$ and $c > 0$; more specifically, it says that all of these Lebesgue measures are the same.

Theorem 7. *For every positive real-valued function Ψ on $\mathbf{N}$ such that $\Psi(q) \to 0$ as $q \to \infty$ and every $c \in \mathbf{R}$ such that $c > 0$,*

$$\mu(W(c\Psi)) = \mu(W(\Psi)). \tag{19}$$

Proof. Suppose Ψ is a positive real-valued function on $\mathbf{N}$, $\Psi(q) \to 0$ as $q \to \infty$, $c \in \mathbf{R}$ and $c > 0$. We may assume without loss of generality that $c \leq 1$, because if $c \geq 1$, then $1/c \leq 1$ and we can take Ψ and c to denote the entities originally denoted by $c\Psi$ and $1/c$, respectively.

Given that $c \leq 1$, we have $W(c\Psi) \subseteq W(\Psi)$ by Proposition 3. Therefore, (19) is equivalent to

$$\mu(W(\Psi) \setminus W(c\Psi)) = 0. \tag{20}$$

Note that

$$W(\Psi) \setminus W(c\Psi) = W(\Psi) \setminus \bigcap_{t=1}^{\infty} A_t^{\infty}(c\Psi) = \bigcup_{t=1}^{\infty} W(\Psi) \setminus A_t^{\infty}(c\Psi)$$

and therefore (20) holds if and only if

$$\mu(W(\Psi) \setminus A_t^{\infty}(c\Psi) = 0$$

for every $t \in \mathbf{N}$.

Suppose to the contrary that there is a $t \in \mathbf{N}$ such that

$$\mu(W(\Psi) \setminus A_t^{\infty}(c\Psi)) > 0,$$

and let $S = W(\Psi) \setminus A_t^{\infty}(c\Psi)$, so that S is the set of the $x \in \mathbf{R}$ such that $|x - p/q| \geq c\Psi(q)$ for every pair $\langle p, q\rangle$ of integers with $q \geq t$. Then because $\mu(S) > 0$, there is an $x \in S$ at which the density of S is 1 by Theorem 6. Therefore, for every $\varepsilon \in \mathbf{R}$ such that $\varepsilon > 0$, we have

$$\begin{aligned}\varepsilon &> \left|\frac{\mu(]x-r, x+r[\cap S)}{2r} - 1\right| \\ &= \left|\frac{\mu((x-r, x+r) \cap S) - 2r|}{2r}\right| \\ &= \frac{2r - \mu((x-r, x+r) \cap S)}{2r},\end{aligned}$$

i.e.

$$\mu(]x-r, x+r[\cap S) > 2r(1-\varepsilon) \tag{21}$$

for every sufficiently small $r \in \mathbf{R}$ such that $r > 0$. Given that $S = W(\Psi) \setminus A_t^{\infty}(c\Psi)$, we have $S \subseteq \mathbf{I} \setminus A_t^{\infty}(c\Psi)$ and therefore (21) implies

$$2r(1-\varepsilon) < \mu((x-r, x+r) \setminus A_t^{\infty}(c\Psi)).$$

By rearranging this equation, we see that it is equivalent to

$$\begin{aligned}2r\varepsilon &> 2r - \mu((x-r, x+r) \setminus A_t^{\infty}(c\Psi)) \\ &= \mu((x-r, x+r) \setminus ((x-r, x+r) \setminus A_t^{\infty}(c\Psi))) \\ &= \mu((x-r, x+r) \cap A_t^{\infty}(\Psi))\end{aligned}$$

and therefore we have

$$\mu((x-r, x+r) \cap A_t^{\infty}(c\Psi) < 2r\varepsilon \tag{22}$$

for every sufficiently small $r \in \mathbf{R}$ with $r > 0$. Note that this is the case for every $\varepsilon \in \mathbf{R}$ with $\varepsilon > 0$.

Now, for every $r \in \mathbf{R}$ with $r > 0$, we have $\Psi(q) < r/2$ for every sufficiently large $q \in \mathbf{N}$ because $\Psi(q) \to 0$ as $n \to \infty$. And, because $x \in S \subseteq W(\Psi)$, we have $|x - p/q| < \Psi(q)$ for pairs $\langle p, q \rangle$ of integers with q arbitrarily large. It follows that there are pairs $\langle p, q \rangle$ of integers with q arbitrarily large such that

$$\left| x - \frac{p}{q} \right| < \Psi(q) < \frac{r}{2}$$

and in particular there is such a pair (p, q) with $q \geq t$.

If r is small enough that (22) holds for every $s \in \mathbf{R}$ with $0 < s < r$, then if we let $s = 2\Psi(q)$, we have

$$\mu((x - s, x + s) \cap A_t^\infty(c\Psi) < 2s\varepsilon = 4\varepsilon\Psi(q).$$

In the particular case where $\varepsilon = c/4$, this equation is equivalent to

$$\mu((x - s, x + s) \cap A_t^\infty(c\Psi)) < c\Psi(q).$$

But this results in a contradiction, because we also have that $\mu((x - s, x + s) \cap A_t^\infty(c\Psi)) \geq c\Psi(q)$ (in fact, it is greater than or equal to $2c\Psi(q)$), by the following reasoning. Let

$$I = \left(\frac{p}{q} - c\Psi(q), \frac{p}{q} + c\Psi(q) \right).$$

The Lebesgue measure of I is $2c\Psi(q)$. Given that $q \geq t$, we have $I \subseteq A_t^\infty(c\Psi)$. And for every $y \in I$,

$$|y - x| \leq \left| x - \frac{p}{q} \right| + \left| y - \frac{p}{q} \right| < \Psi(q) + c\Psi(q) \leq 2\Psi(q) = s,$$

from which it follows that $y \in (x - s, x + s)$. Therefore, I is a subset of $(x - s, x + s)$ as well as $A_t^\infty(c\Psi)$. It follows that I is a subset of the intersection $(x - s, x + s) \cap A_t^\infty(c\Psi)$, and therefore $\mu(I) \leq \mu((x - s, x + s) \cap A_t^\infty(c\Psi))$. □

Given Theorem 7, the proof that $\mu(\mathbf{B}) = 0$ is very straightforward. First, note that for every positive real-valued function Ψ on $\mathbf{N}$,

$$W^*(\Psi) = \bigcap_{\substack{\varepsilon \in \mathbf{R} \\ \varepsilon > 0}} W(\varepsilon\Psi) = \bigcap_{n=1}^{\infty} W(\Psi/n)$$

because $1/n \to 0$ as $n \to \infty$. Therefore,

$$\begin{aligned}\mu(\mathbf{I} \setminus W^*(\Psi)) &= \mu\left(\mathbf{I} \setminus \bigcap_{n=1}^{\infty} W(\Psi/n)\right) \\ &\leq \sum_{n=1}^{\infty} \mu(\mathbf{I} \setminus W(\Psi/n)) \\ &= \sum_{n=1}^{\infty} (\mu(\mathbf{I}) - \mu(W(\Psi/n))) \\ &= \sum_{n=1}^{\infty} (1 - \mu(W(\Psi))).\end{aligned}$$

It follows that if $1 - \mu(W(\Psi)) = 0$, i.e. $\mu(W(\Psi)) = 1$, then we have $\mu(\mathbf{I} \setminus W^*(\Psi)) = 0$, i.e. $\mu(W^*(\Psi)) = 1$. In the particular case where $\Psi = \mathrm{id}_{\mathbf{N}}^{-2}$, we have

$$W(\mathrm{id}_{\mathbf{N}}^{-2}) = \mathbf{I}$$

by Theorem 2, from which it follows that $\mu(W(\mathrm{id}_{\mathbf{N}}^{-2})) = 1$. Therefore,

$$\mu(\mathbf{B}) = \mu(\mathbf{I} \setminus W^*(\mathrm{id}_{\mathbf{N}}^{-2})) = 0.$$

If we use the terminology established in the previous section, what this result means is that *almost every real number is not badly approximable*. Or, to put it another way, almost every real number is strongly approximable to the order 2. Thus the result is "almost" a strengthening of Dirichlet's theorem.

8. Acknowledgements

I would like to thank my supervisor Sanju Velani for his guidance over the course of the year.

References

[1] Beresnevich, V., Ramírez, F., & Velani, S. 2016. *Metric Diophantine Approximation: aspects of recent work*. eprint arXiv:1601.01948.

[2] Cassels, J. W. S., 1950. Some metrical theorems in Diophantine approximation. *Mathematical Proceedings of the Cambridge Philosophical Society* 46, pp. 209-218. Cambridge University Press.

[3] Hardy, G. H. & Wright, E. M., 1960. *An introduction to the theory of numbers*. Oxford University Press.

[4] Roth, K.F., 1955. Rational approximations to algebraic numbers. *Mathematika*, 2(01), pp. 1-20.

[5] Tao, T., 2011. *An introduction to measure theory*. American Mathematical Society.