A finitization of Littlewood's Tauberian theorem and an application in Tauberian remainder theory

Thomas Powell

February 7, 2022

Abstract

We analyse Littlewood's 1911 Tauberian theorem from a proof theoretic perspective. We first use Gödel's Dialectica interpretation to give a computational interpretation to the theorem, producing a finitary formulation of the result and, with the help of known quantitative results from approximation theory, extracting concrete bounds from the proof. Our finitary Tauberian theorem can be given an intuitive game semantics, with the bounds corresponding to a winning strategy. We then use our finitization to produce two general remainder theorems in terms of rates of convergence and metastability. We rederive the traditional remainder estimate for Littlewood's theorem as a special case of these.

1 Introduction

The extraction of programs from proofs is a central theme of logic and theoretical computer science. Modern research on this topic encompasses both foundational results, such as the correspondence between formal logic and programming languages, the computational semantics of proofs, and complexity theory, along with applications, including formal verification [20, 24] and the use of logical techniques to obtain numerical data from proofs in mathematics [10]. This paper forms a new contribution to the application of proof theory in mathematics, using Gödel's Dialectica interpretation to give a computational interpretation to Littlewood's classic Tauberian theorem.

Ever since the pioneering work of Kreisel [15, 16] on the "unwinding" of proofs, it has been clear that traditional logical methods, and in particular Gödel's Dialectica interpretation [6], have a deep mathematical significance. This significance lies not only in their power as techniques for obtaining new quantitative information from nonconstructive proofs (widely demonstrated in the last 30 years or so through the so-called *proof mining* program [10]), but their ability to yield qualitative generalisations of theorems and give us deep structural insights into mathematical phenomena. An example of the latter is the recent discovery that the Dialectica interpretation is connected to a fundamental correspondence principle between 'soft' and 'hard' statements in analysis as described by T. Tao [26, 5], and that the interpretation can be viewed as a method for 'finitizing' infinitary statements, particularly convergence properties. To summarise, using the Dialectica interpretation to explore computational aspects of mathematical proofs can yield both new quantitative results and novel connections between logic and pure mathematics.

In this work, we apply the Dialectica interpretation in this spirit to investigate the relationship between two forms of convergence, representing distinct methods for summing an infinite sequence of real numbers $\{a_n\}$:

$$(i)\ \sum_{i=0}^{\infty} a_i \quad \text{and} \quad (ii)\ \lim_{x\to 1^-} \sum_{i=0}^{\infty} a_i x^i$$

When the limit (ii) exists, we say that $\{a_n\}$ is *Abel summable*. While normal summability (i) implies Abel summability (ii), the converse is not true, although a partial converse can be proven by imposing a growth condition on $\{a_n\}$. In 1897 A. Tauber first showed that $a_n = o(1/n)$ suffices, but then in 1911 J. E. Littlewood established an optimal order condition such that convergence of (ii) implies (i), namely $a_n = O(1/n)$ [17]. This celebrated theorem marked the beginning of the lifelong collaboration between G. H. Hardy and Littlewood, which began with the development of further theorems of Tauberian type, and ultimately launched an area of analysis now knows as Tauberian theory [14].

There are three main reasons why we have chosen Littlewood's theorem as a candidate for applying the Dialectica interpretation:

1. The theorem is extremely simple to state, and from a logical perspective can be reformulated as a straightforward implication between Cauchy convergence properties (Section 3). Proofs of the theorem, on the other hand, are complex, and even the shortest method of proof discovered by Karamata [8] involves subtle ideas from approximation theory. As such, the Dialectica interpretation of Littlewood's theorem is elegant and intuitive, but extracting the corresponding witnessing terms is non-trivial.

2. There is an existing interest in quantitative versions of Tauberian theorems in the form of remainder estimates, which relate the convergence speed of (ii) to that of (i). We are able to not only provide a generalisation of the known estimate in the case of Littlewood's theorem, but also show that the latter falls out in a natural way from our proof-theoretic analysis.

3. Tauberian theory in general represents a new domain of application for proof theoretic methods, and is an area replete with simple convergence statements whose proofs make use of complex analytic techniques. As such, an analysis of Littlewood's theorem represents a step in a promising new direction with great potential for further study. Concrete suggestions for future research are given in the conclusion to this paper.

Relevance to computer science

Though the main results presented here concern infinite series, the central theme of this paper is the application of formal logical methods, and we consider this to be of broad relevance to the logic in computer science community in a number of different ways. For example, through our use of the Dialectica interpretation we are able impose a natural game semantics onto Tauberian theorems, along the lines of [3, 2], and the paper as a whole requires a careful analysis of the logical structure of convergence properties that could be of interest to researchers in constructive mathematics or formalization. Our remainder theorems in Section 6 involve concepts from computability theory such as Specker sequences, and are formulated in terms of higher-order functionals, so that Theorem 6.6 is essentially the specification of a type 3 functional program corresponding to a variant of Littlewood's theorem.

On a more general level, in recent years there has been a resurgence of interest in the Dialectica interpretation within theoretical computer science, encompassing formalization [24], category theory (starting with [4]), and most recently new connections with classical realizability[19] and the differential lambda calculus[9]. Therefore we believe that a self-contained case study presenting a novel example of how the Dialectica interpretation manifests in a natural way within a beautiful area of pure mathematics will be of interest in its own right, and we have deliberately written this paper without assuming any prior knowledge of Tauberian theory or proof interpretations.

Related work

Tauberian theory has grown into a large area of research, but applications of proof theory in this area are currently limited to [23], on which this paper builds significantly (though in the other direction, Tauberian theorems have been applied in proof theory by Weiermann to derive ordinal bounds cf. [28] and most recently [29]). More generally, the results of this paper represent one of very few applications of proof interpretations in number theory, though several of the earliest case-studies in applied proof theory touch on analytic methods in number theory, including [16] and [18].

Prerequisites

We have endeavoured to keep this paper as self-contained as possible, and we do not assume any prior knowledge of the two main topics that feature: Tauberian theory and Gödel's Dialectica interpretation. We provide a brief overview of Tauberian theory in the next section, and introduce the relevant aspects of the Dialectica interpretation as they are needed in later sections. Beyond this, we assume only a basic understanding of formal logic and a certain fluency in elementary analysis, specifically convergent series and integrals.

2 Tauberian theorems

Tauberian theory is an extensive area of research which, taken in a very general sense, is concerned with finding conditions under which summability methods converge. For a comprehensive overview of the field, including an account of its historical development and a survey of modern research in the area, the reader is encourage to consult the textbook [14], though everything that we require will be presented below. This paper involves just two simple summability methods, namely basic infinite series together with the power series they generate. To be more specific, let $\{a_n\}$ be a sequence of real numbers. For the remainder of this paper we will define $\{s_n\}$ and $F : [0,1) \to \mathbb{R}$ as follows:

$$s_n := \sum_{i=0}^{n} a_i \quad \text{and} \quad F(x) := \sum_{i=0}^{\infty} a_i x^i$$

Assuming that $\{|a_n|\}$ is bounded above by some $a > 0$, it is clear that F is well-defined on $x \in [0,1)$, since then

$$\sum_{i=0}^{\infty} |a_i x^i| \leq a \sum_{i=0}^{\infty} x^i = \frac{a}{1-x}$$

However, the question of whether or not $F(x)$ converges to some finite limit as $x \to 1^-$ is closely related to the convergence of $\{s_n\}$. In one direction we have a standard result:

Theorem 2.1 (Abel's theorem)**.** *If* $\lim_{n\to\infty} s_n = s$ *and thus* $F(x)$ *is well-defined on* $[0,1)$*, then* $\lim_{x\to 1^-} F(x) = s$.

The converse of this theorem is not true: Setting $a_n = (-1)^n$ we have

$$F(x) = \sum_{i=0}^{\infty} (-1)^i x^i = \frac{1}{1+x} \to \frac{1}{2}$$

as $x \to 1^-$, but $\sum_{i=0}^{\infty} (-1)^i$ does not converge. Tauber's theorem, from which Tauberian theory derives its name, establishes a simple growth condition under which a converse to Abel's theorem *does* hold:

Theorem 2.2 (A. Tauber 1897 [27])**.** *If* $\lim_{x\to 1^-} F(x) = s$ *and* $a_n = o(1/n)$ *then* $\lim_{n\to\infty} s_n = s$.

Both Abel's and Tauber's theorems can be easily proven with little more than elementary facts about convergent series (cf. [14, Chapter 1] and the corresponding proof theoretic analysis in [23]). However, in 1911, Littlewood established a 'big-O' strengthening of Tauber's theorem, a much deeper result which in some sense marked the beginning of Tauberian theory in earnest:

Theorem 2.3 (J. E. Littlewood 1911 [17])**.** *If* $\lim_{x\to 1^-} F(x) = s$ *and* $a_n = O(1/n)$ *then* $\lim_{n\to\infty} s_n = s$.

Littlewood showed that his growth condition $a_n = O(1/n)$ is optimal in the sense that for any sequence $\{b_n\}$ with $\lim_{n\to\infty} b_n = \infty$ there exists a sequence $\{a_n\}$ with $n|a_n| \leq b_n$ such that $\lim_{x\to 1^-} F(x)$ exists but $\sum_{i=0}^{\infty} a_i$ does not, although together with Hardy [7] his Tauberian theorem was further strengthened, in particular showing that a one-sided condition $na_n \geq -C$ is sufficient.

A key feature that distinguishes both Littlewood's and subsequent Tauberian theorems from Theorem 2.2 is the relative difficulty of proving them. Littlewood's original proof was complex and involved repeated differentiation, and while much simpler proofs were subsequently found (notably by Karamata [8] and then Wielandt [30]), all of these rely on analytic methods, specifically techniques for approximating continuous functions with polynomials.

Another good indication that Littlewood-style Tauberian theorems go fundamentally beyond elementary analysis is that optimal remainder estimates that relate the convergence speed of $F(x) \to s$ to that of $s_n \to s$ are established using numerical results from approximation theory, including bounds on the degree and coefficients of polynomials that approximate piecewise continuous functions. The canonical remainder estimate for Littlewood's theorem is as follows:

Theorem 2.4 (cf. Korevaar [14] page 346)**.** *Suppose that $a_n = O(1/n)$ and there is some $b > 0$ such that*

$$F(x) = s + O((1-x)^b)$$

as $x \to 1^-$. Then

$$s_n = s + O\left(\frac{1}{\log(n)}\right)$$

as $n \to \infty$.

Outline of the paper

Following this brief introduction to the relevant background in Tauberian theory, we are now in a position to give a more detailed outline of the main results of the present paper. In recent work [23], the comparatively elementary proofs of Abel's and Tauber's theorems were analysed, and finitary versions of these theorems established. In what follows, we extend this idea to the much more complex Littlewood Tauberian theorem, stating and proving a finitary version of Theorem 2.3 in Section 5 and showing how this can be clearly understood in terms of the Dialectica interpretation of implication. Before doing this, we give a Cauchy reformulation of Littlewood's theorem, and present a proof of this reformulation that is inspired by Wielandt's variation [30] of Karamata's method of proof [8].

We then combine our finitization with known bounds on polynomial approximations to present two new remainder theorems for Littlewood's theorem. The first deals with the case where $F(x) \to s$ with some *arbitrary* computable rate of convergence, while the second applies more generally still when $F(x) \to s$ as $x \to 1^-$ but without necessarily having a computable rate of convergence, instead converting a rate of *metastability* of the former to a rate of metastability for $s_n \to s$. Both of these form generalisations of Theorem 2.4, which we show can be derived as a special case. The use of proof theoretic methods to obtain Tauberian remainder estimates in this way is completely new.

As such, our results not only constitute a new application of the Dialectica to extract computational information from a highly non-trivial proof, but can then be used to produce concrete numerical results that are of relevance in the area of application.

3 A Cauchy variant of Littlewood's theorem

We start by giving a new presentation of Littlewood's theorem. In line with standard approaches to analysing convergence theorems using proof theoretic methods, we prefer to work with a reformulation of the theorem that minimises its quantifier complexity. In particular, we seek a version of Theorem 2.3 which, rather than referring directly to the limit s, expresses the relevant convergence properties in an equivalent Cauchy form.

Firstly, the assumption that $F(x) \to s$ as $x \to 1^-$ will be replaced with a natural Cauchy variant which says that for any $\delta > 0$ we have $|F(x) - F(y)| \leq \delta$ for x, y sufficiently close to 1. While it would be natural to then also replace the conclusion $s_n \to s$ as $n \to \infty$ with the standard Cauchy property for convergent sequences, the issue here is Littlewood's theorem doesn't just state that $\{s_n\}$ converges, but that it converges to the *same limit* as $F(x)$ as $x \to 1^-$. Therefore we instead formulate the conclusion as the following Cauchy property: for any $\varepsilon > 0$ we have $|s_n - F(x_m)| \leq \varepsilon$ for sufficiently large m, n, where x_m is some canonical sequence in $[0, 1)$ with $x_m \to 1$. From this we can then retrieve that $s_n \to \lim_{x \to 1^-} F(x)$. For reasons that will become clear when we give our proof of the theorem, we choose $x_m := e^{-1/m}$ as our canonical sequence. We present our Cauchy variant of Littlewood's theorem below, and then prove that the original result follows from it.

Definition 3.1. As usual we write $x \in (a, b)$ for $a < x < b$ and $x \in [a, b]$ for $a \leq x \leq b$, and so on. For integers l, m, n we also write $n \in [l, m]$ to denote $l \leq n \leq m$.

Theorem 3.2 (Littlewood's theorem, Cauchy variant)**.** *Suppose that there exists some $C > 0$ such that $n|a_n| \leq C$ for all $n \in \mathbb{N}$, and that*

$$\forall \delta > 0 \exists M \forall x, y \in [e^{-1/M}, 1)(|F(x) - F(y)| \leq \delta)$$

Then we have

$$\forall \varepsilon > 0 \exists N \forall m, n \geq N(|s_n - F(e^{-1/m})| \leq \varepsilon)$$

Before proving this, we show that it can be used to derive the original formulation of Littlewood's theorem:

Proof of Theorem 2.3 from Theorem 3.2. If $a_n = O(1/n)$ then by definition there exists some $C > 0$ such that $n|a_n| \leq C$ for all $n \in \mathbb{N}$. Furthermore, if $F(x) \to s$ as

$x \to 1^-$ then for any $\delta > 0$ there exists some $0 \le \mu < 1$ such that

$$x \in [\mu, 1) \implies |F(x) - s| \le \frac{\delta}{2}$$

Let M be sufficiently large that $e^{-1/M} \ge \mu$. Then $x, y \in [e^{-1/M}, 1)$ implies that

$$|F(x) - F(y)| \le |F(x) - s| + |s - F(y)| \le \delta$$

and so we have established the premise of Theorem 3.2. Therefore we can conclude that for any $\varepsilon > 0$ there exists some N such that

$$m, n \ge N \implies |s_n - F(e^{-1/m})| \le \frac{\varepsilon}{2}$$

But since $F(x) \to s$ as $x \to 1^-$ we can choose $m_0 \ge N$ large enough so that $|F(e^{-1/m_0}) - s| \le \varepsilon/2$, and then $n \ge N$ implies

$$|s_n - s| \le |s_n - F(e^{-1/m_0})| + |F(e^{-1/m_0}) - s| \le \varepsilon$$

and we've proven that $s_n \to s$ as $n \to \infty$. □

We now give a proof of Theorem 3.2, which will be analysed in the next section. This is an adaptation of Wielandt's proof [30] (see also [14, Chapter I.12]). Roughly speaking, this proof strategy, which dates back to Karamata [8], is based on representing the partial sums s_n as step functions, and using integral theory to show that these approach $F(e^{-1/m})$ as $m \to \infty$. Our use of integrals in this way requires us to replace the discontinuous step function with a polynomial approximation to it. That we are able to find such polynomials for arbitrary errors relies on a result from approximation theory, and we begin by stating this in the exact form in which it is needed:

Lemma 3.3. *Define $\chi : [0, \infty) \to \mathbb{R}$ to be*

$$\chi(t) := \begin{cases} 1 & \text{if } t \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

Then for any $\varepsilon > 0$ there exists a polynomial P with $P(0) = 0$ and $P(1) = 1$ satisfying

$$\int_0^\infty \frac{|\chi(t) - P(e^{-t})|}{t} \, dt < \varepsilon \tag{1}$$

Proof sketch. This is a standard result (cf. [13] or [14, Chapter I.11–12] for full details), and so we just sketch the idea. We first consider the function $f : [0, 1] \to \mathbb{R}$ defined by

$$f(t) := \frac{\chi(\log(1/t)) - t}{t(1 - t)}$$

which is Lipschitz continuous on both $[0, 1/e)$ and $[1/e, 1]$ and has a single jump discontinuity at $t = 1/e$. It is a well known fact that such functions can be approximated to arbitrary precision by polynomials, so we pick some polynomial p such that

$$\int_0^1 |f(t) - p(t)|\, dt < \varepsilon$$

and then set $P(t) := t + t(1-t)p(t)$, noting that $P(0) = 0$ and $P(1) = 1$. Substituting $t = e^{-u}$ in the above integral we have

$$\int_0^1 |f(t) - p(t)|\, dt = \int_0^\infty \frac{|\chi(u) - P(e^{-u})|}{1 - e^{-u}}\, du$$

and applying the standard inequality $1 + x \leq e^x$ in the form $1 - e^{-u} \leq u$ yields

$$\frac{|\chi(u) - P(e^{-u})|}{u} \leq \frac{|\chi(u) - P(e^{-u})|}{1 - e^{-u}}$$

for $0 < u$, and therefore

$$\int_0^\infty \frac{\chi(u) - P(e^{-u})}{u}\, du \leq \int_0^1 |f(t) - p(t)|\, dt < \varepsilon$$

which completes the proof. □

We are now ready to apply the above lemma to prove our Cauchy variant of Littlewood's theorem.

Proof. Fix $\varepsilon > 0$ for the remainder of the proof, and let $P(t)$ be as in Lemma 3.3 for error $\varepsilon/4C$. Define

$$I_P(n) := \int_0^\infty a(t)P(e^{-t/n})\, dt$$

Our main strategy is to show that

(i) $|s_n - I_P(n)| \leq \varepsilon/2$ for all $n \in \mathbb{N}$ and,

(ii) $|I_P(n) - F(e^{-1/m})| \leq \varepsilon/2$ for sufficiently large m, n.

Then putting these together we have

$$|s_n - F(e^{-1/m})| \leq |s_n - I_P(n)| + |I_P(n) - F(e^{-1/m})| \leq \varepsilon$$

for sufficiently large m, n, and since ε was arbitrary the theorem is proved. We tackle each of these in turn, using the growth condition $n|a_n| \leq C$ for (i) and convergence of $F(x)$ as $x \to 1^-$ for (ii).

For (i), let $a : [0, \infty) \to \mathbb{R}$ be the discontinuous step function corresponding to $\{a_n\}$, defined by $a(t) := a_n$ for $t \in [n, n+1)$ so that

$$s_n = \int_0^n a(t)\, dt = \int_0^\infty a(t)\chi(t/n)\, dt$$

and therefore

$$|s_n - I_P(n)| \leq \int_0^\infty |a(t)| \cdot |\chi(t/n) - P(e^{-t/n})| \, dt \tag{2}$$

Now for any $n \in \mathbb{N}$ we have $n|a_n| \leq C$ and therefore for $t \in [n, n+1)$:

$$|a(t)| = |a_n| \leq \frac{C}{n} \leq \frac{2C}{n+1} \leq \frac{2C}{t} \tag{3}$$

and thus it follows that $|a(t)| \leq 2C/t$ for all $t \in [0, \infty)$. Therefore from (2) we have

$$\begin{aligned} |s_n - I_P(n)| &\leq 2C \int_0^\infty \frac{|\chi(t/n) - P(e^{-t/n})|}{t} \, dt \\ &\leq 2C \int_0^\infty \frac{|\chi(u) - P(e^{-u})|}{u} \, du < \frac{\varepsilon}{2} \end{aligned} \tag{4}$$

using the substitution $u = t/n$ and the property of P.

To prove (ii), we first note that by Lemma 3.3 we know that $P(0) = 0$ and $P(1) = 1$, which implies that $P(t) = \sum_{i=1}^d c_i t^i$ for some $d \in \mathbb{N}$ and $c_1, \ldots, c_d$ with $\sum_{i=1}^d c_i = 1$. Writing out $I_P(n)$ as

$$I_P(n) = \sum_{i=1}^d c_i \int_0^\infty a(t) e^{-it/n} \, dt \tag{5}$$

and observing that for any $\alpha > 0$ we have

$$\begin{aligned} \int_0^\infty a(t) e^{-\alpha t} \, dt &= \sum_{k=0}^\infty a_k \int_k^{k+1} e^{-\alpha t} \, dt \\ &= \sum_{k=0}^\infty a_k e^{-\alpha k} \int_0^1 e^{-\alpha t} \, dt = F(e^{-\alpha}) \int_0^1 e^{-\alpha t} \, dt \end{aligned} \tag{6}$$

then combining (5) and (6) we have

$$I_P(n) = \sum_{i=1}^d c_i F(e^{-i/n}) \int_0^1 e^{-it/n} \, dt \tag{7}$$

Using now that $\sum_{i=1}^d c_i = 1$ and thus

$$F(e^{-1/m}) = \sum_{i=1}^d c_i F(e^{-1/m})$$

it follows from (7) that

$$
\begin{aligned}
&|I_P(n) - F(e^{-1/m})| \\
&= \left|\sum_{i=1}^{d} c_i \left(F(e^{-i/n}) \int_0^1 e^{-it/n}\, dt - F(e^{-1/m})\right)\right| \\
&\leq \sum_{i=1}^{d} |c_i| \cdot |F(e^{-i/n}) \int_0^1 e^{-it/n}\, dt - F(e^{-1/m})| \\
&\leq \sum_{i=1}^{d} |c_i| \cdot \left(L \cdot \left|1 - \int_0^1 e^{-it/n}\, dt\right| + |F(e^{-i/n}) - F(e^{-1/m})|\right)
\end{aligned}
\tag{8}
$$

where for the last step we use that $|F(x)|$ must be bounded above by some $L > 0$ as $x \to 1^-$. Now from

$$\lim_{n\to\infty} \int_0^1 e^{-it/n}\, dt = 1$$

and the convergence condition $|F(e^{-i/n}) - F(e^{-1/m})| \to 0$ as $m, n \to \infty$, it follows from (8) that we can choose m, n sufficiently large so that $|I_P(n) - F(e^{-1/m})| \leq \varepsilon/2$. This proves (ii), and we're done □

4 The Dialectica interpretation

Having motivated and presented Littlewood's Tauberian theorem, in a Cauchy form whose simplified logical structure makes it easier to interpret, we now briefly outline the logical technique on which our analysis is based: Gödel's Dialectica interpretation. It is important to note that this paper is not a rigorous study of the Dialectica itself – rather the interpretation acts as a guide in how to correctly formulate our finitary version of Littlewood's theorem, along with the subsequent remainder theorems. As such, exact details of the interpretation and the surrounding theory are not important here, and a prior familiarity with the interpretation not required to appreciate our quantitative theorems. With this in mind, we restrict our attention to key features of the interpretation that play a role in what follows. For detailed background on the Dialectica interpretation see e.g. [1] or [10].

4.1 The basic interpretation

Numerous different variants of the Dialectica interpretation have been explored over the years, ranging from Gödel's original interpretation to the specialised versions used in proof mining [10], and now including recent new varieties that incorporate concepts from programming languages [19, 21]. However, in its standard form the Dialectica assigns to any formula A in some formal intuitionistic theory P_I a logically equivalent formula of the form $A^D = \exists x \forall y A_D(x, y)$ in some higher-order variant P_ω of that theory, where

- $A_D(x, y)$ is 'computationally neutral' (typically quantifier free or decidable),

- the free variables of $\exists x \forall y A_D(x, y)$ are the same as those of A,

- x and y are (potentially empty) tuples of terms in all finite types, where these types depend on A.

To be more precise, $A_D(x, y)$ is defined by induction over the logical structure of A as follows:

$$
\begin{aligned}
A^D &:= A \quad \text{if } A \text{ is computationally neutral} \\
(A \wedge B)^D &:= \exists x, u \forall y, v (A_D(x, y) \wedge B_D(u, v)) \\
(A \vee B)^D &:= \exists b, x, u \forall y, v (A_D(x, y) \vee_b B_D(u, v)) \\
(A \implies B)^D &:= \exists f, g \forall x, v (A_D(x, gxv) \implies B_D(fx, v)) \\
(\exists z A[z])^D &:= \exists z, x \forall u\, A_D[z](x, u) \\
(\forall z A[z])^D &:= \exists f \forall z, u\, A_D[z](fz, y)
\end{aligned}
$$

where in the interpretation of disjunction b is a boolean or natural number and $P \vee_b Q$ is shorthand for

$$(b = 0 \implies P) \wedge (b \neq 0 \implies Q)$$

The most fundamental results concerning the Dialectica interpretation are *soundness theorems*, which guarantee that whenever A is provable in P_I, a witnessing term for A^D can be extracted from the proof:

Intuitionistic soundness theorem: If $P_I \vdash A$ then we can extract, from the proof of A, some term t in P_ω such that $P_\omega \vdash \forall y A_D(t, y)$.

The original soundness theorem due to Gödel instantiates P_I as Heyting arithmetic and P_ω as System T, though many more soundness theorems have been developed since. In the case of classical theories P_C, it is not always possible to extract computable terms that witness the Dialectica interpretation of formulas. However, in this case one can instead precompose the Dialectica with a negative translation A^N that embeds P_C into its intuitionistic variant P_I:

Classical soundness theorem: If $P_C \vdash A$ then we can extract, from the proof of A, some term t in P_ω such that $P_\omega \vdash \forall y (A^N)_D(t, y)$.

The value of the soundness theorems for applied proof theory lies primarily in the fact that they set out conditions under which it is theoretically possible to extract computational content from proofs, and also determine the precise shape that this computational information should take. However, outside of formal program synthesis, the soundness theorems are rarely applied in a brute-force manner to extract concrete realising terms from fully formalised proofs. Rather, the interpretation is used in an informal way, and in conjunction with ordinary mathematical intuition.

More concretely, in our case the Dialectica interpretation dictates the way in which we formulate our finitary Tauberian theorem below, and (in its classical variant) indicates how to obtain remainder theorems even when no computable rates of convergence exist in Section 6.2, but our quantitative results are all proven 'by-hand' and and without the rigorous application of logical methods, though the Dialectica interpretation implicitly guides our analysis of the proof of Theorem 3.2.

4.2 The interpretation of implication

The characterising feature of the Dialectica interpretation that sets it apart from similar proof interpretations such as modified realizability is its interpretation of implication. It is this that informs our computational interpretation of Littlewood's theorem, and so it is important to give some insight into its meaning.

In interpreting implication, we are required to choose a Skolemisation of the formula:

$$\exists x \forall y A_D(x,y) \implies \exists u \forall v B_D(u,v)$$

There are various options available, but the Dialectica selects the 'least nonconstructive' of these, which turns out to be:

$$\forall x \exists u \forall v \exists y (A_D(x,y) \implies B_D(u,v)) \tag{9}$$

(for the reasoning behind this choice see e.g. [10, pp. 128–129]). Bringing the existential quantifiers $\exists u$ and $\exists y$ to the front as functions gives us exactly $(A \implies B)^D$.

Happily, the formula (9) can be given an elegant reading in terms of game semantics, as a game between ∃loise and ∀belard where ∃loise seeks to prove $A^D \implies B^D$ while ∀belard attempts to contradict her by proving $A^D \wedge \neg B^D$:

1. ∀belard begins by proposing a witness x such that $\forall y A_D(x,y)$ holds, with the aim of disproving B^D.

2. ∃loise responds by proposing a witness u such that $\forall v B_D(u,v)$ is true.

3. ∀belard now tried to contradict ∃loise's witness for B^D by proposing a counterexample v such that $\neg B_D(u,v)$.

4. ∃loise responds by contradicting ∀belard's original claim that A^D is true by providing a counterexample y such that $\neg A^D(x,y)$.

Thus a pair of function f, g that witness $(A \implies B)^D$ represent nothing other than a winning strategy for ∃loise in the above game, and this intuitive reading of the Dialectica interpretation in this case is outlined in its abstract form here as it is important in understanding our game-theoretic reading of Littlewood's theorem in Section 5.3 below.

5 A finitary Littlewood Tauberian theorem

We now motivate and present our computational interpretation of Theorem 3.2. We characterise this as a 'finitary' Tauberian theorem because its statement only refers to finite parts of the input data: Rather than asking that $|F(x)-F(x)| \to 0$ as $x, y \to 1^-$ we only require that $|F(x)-F(y)|$ is sufficiently small within some fixed range $[l, r] \subset [0, 1)$. Similarly, the growth condition is replaced by an assumption that $n|a_n| \leq C$ for $n \leq p$ for some suitable p. Finally, our conclusion is also finitary in nature: instead of proving that $|s_n - F(e^{-1/m})| \to 0$ as $m, n \to \infty$ we establish that $|s_n - F(e^{-1/m})|$ is sufficiently small within some finite range $N \leq m, n \leq k$. In this sense, our notion of finitary coincides with what T. Tao regards as the 'hard' version of a soft analytical statement [26]. Nevertheless, despite only referring to finite parts of the data, our finitary theorem is equivalent to Theorem 3.2, and therefore also Littlewood's original formulation of the 'big-O' Tauberian theorem, in the same way that the Dialectica interpretation of implication (9) is equivalent to $A \implies B$.

5.1 The logical structure of Littlewood's theorem

In order to make the quantifier structure of Theorem 3.2 precise, we first introduce some predicate annotations.

Definition 5.1. Fixing some $\{a_n\}$ and $C > 0$, the formulas $A(p)$, $B(\delta, M, l)$ and $D(\varepsilon, N, k)$ are defined as follows:

$$\begin{aligned}
A(p) &:= \forall n \leq p (n|a_n| \leq C) \\
B(\delta, M, l) &:= \forall x, y \in [e^{-1/M}, e^{-1/(M+l)}](|F(x) - F(y)| \leq \delta) \\
D(\varepsilon, N, k) &:= \forall m, n \in [N, N+k](|s_m - F(e^{-1/m})| \leq \varepsilon)
\end{aligned}$$

The overall structure of Littlewood's theorem (and indeed Tauberian theorems in general) is an implication of the form

$$\text{(convergence)} \wedge \text{(growth condition)} \implies \text{(convergence)}$$

Technically, the quantifier structure of such an implication is rather complex, as each convergence property is a $\forall\exists\forall$ statement. However, just as in the prior analysis of simple Tauberian theorems in [23], an inspection of the proof of Theorem 3.2 reveals that we have in fact proven something stronger: Namely for any $\varepsilon > 0$ there is a concrete δ dependent only on ε, namely $\delta_\varepsilon := \varepsilon/4 \sum_{i=1}^{d} |c_i|$, such that $|F(x) - F(y)| \leq \delta_\varepsilon$ for all x, y sufficiently close to 1 implies that $|s_n - F(e^{-1/m})| \leq \varepsilon$ for all m, n sufficiently large (we do not justify this in detail at this point as this will be implicitly proven in Theorem 5.3 below). This is clearly a quantitative strengthening of the theorem, and renders it of the form

$$\forall \varepsilon [\forall p\, A(p) \wedge \exists M \forall l\, B(\delta_\varepsilon, M, l) \implies \exists N \forall k\, D(\varepsilon, N, k)] \tag{10}$$

Now, taking the Dialectica interpretation of the premise of (10) gives us

$$\forall \varepsilon [\exists M \forall p, l\, (A(p) \wedge B(\delta_\varepsilon, M, l)) \implies \exists N \forall k\, D(\varepsilon, N, k)]$$

and the Dialectica interpretation of the above implication as in (9) yields the following as a final Skolemisation of the Littlewood Tauberian theorem in its Cauchy formulation:

$$\forall \varepsilon, M \exists N \forall k \exists p, l [A(p) \wedge B(\delta_\varepsilon, M, l) \implies D(\varepsilon, N, k)] \tag{11}$$

5.2 The finitary theorem

Our finitary version of Littlewood's theorem corresponds to the Dialectica interpretation of Theorem 3.2 in the form (11). More importantly, we analyse the proof of Theorem 3.2 to provide concrete bounds on N in terms of ε and M and on p and l in terms of ε, M and k, parametrised by the big-O bound $C > 0$, together with a uniform bound $a > 0$ for the sequence $\{|a_n|\}$, the latter being necessary since we can no longer derive boundedness of $\{|a_n|\}$ from the growth condition, as we now only assume a finitary version of $a_n = O(1/n)$.

Theorem 3.2 relies crucially on the existence of approximating polynomials as set out in Lemma 3.3, and in order to give a computational interpretation to the former, we require numerical bounds on certain data from these polynomials, specifically a bound on both their degree d and the sum of the magnitude of their coefficients $\sum_{i=1}^{d} |c_i|$ in terms of an input error ε. Fortunately, such bounds can be derived using standard results from approximation theory, and were adapted to the precise form required in Lemma 3.3 by Korevaar [13].

Lemma 5.2. *There are constants $A, B > 0$ such that defining*

$$\Omega(\varepsilon) = (\Omega_0(\varepsilon), \Omega_1(\varepsilon)) = \left(\frac{A}{\varepsilon}, B^{1/\varepsilon}\right)$$

we have that for any $\varepsilon > 0$ there exists a polynomial $P(x) = \sum_{i=1}^{d} c_i x^i$ satisfying Lemma 3.3 with error ε, such that

$$d \leq \Omega_0(\varepsilon) \quad \textit{and} \quad \sum_{i=1}^{d} |c_i| \leq \Omega_1(\varepsilon)$$

Proof. We adapt Korevaar [13] in a straightforward way. Let $f : [0,1] \to \mathbb{R}$ be any function that is continuous except for a finite number of jump discontinuities, and such that there exists a constant a such that $|f(x) - f(y)| \leq a|x - y|$ for x, y in any subinterval on which f is continuous. Then it can be shown (cf. [13, Theorem 4.1]) that there exist constants H_1 and H_2 such that for any positive integer n there exists a polynomial $p_n(x) = \alpha_n x^n + \ldots + \alpha_1 x + \alpha_0$ of degree n and with $|\alpha_i| \leq H_2 3^n$ for all $i = 0, 1, \ldots, n$ such that

$$\int_0^1 |f(t) - p_n(t)|\, dt < \frac{H_1}{n+1}$$

Now let f be as in the proof of Lemma 5.2. Then, just as in that proof, we set $P_n(t) = t + t(1-t)p_n(t)$ and have

$$\int_0^\infty \frac{|\chi(t) - P_n(e^{-t})|}{t}\, dt < \frac{H_1}{n+1}$$

noting that $P_n(t) = c_{n+2}x^{n_2} + \ldots c_1 x$ has degree $n+2$, and that

$$\sum_{i=0}^{n+2} |c_i| \leq 2\sum_{i=0}^{n} |\alpha_i| + 1 \leq 2H_2 3^n + 1$$

Now, for any $\varepsilon > 0$ it is clear that setting $P(t) := P_n(t)$ for $n := \lceil H_1/\varepsilon \rceil$ we have

$$\int_0^\infty \frac{|\chi(t) - P(e^{-t})|}{t}\, dt < \varepsilon$$

and that

$$\deg(P) = \lceil H_1/\varepsilon \rceil + 2 \leq \frac{H_1 + 3}{\varepsilon}$$

and so we can set $A := H_1 + 3$. Finally,

$$\sum_{i=0}^{n+2} |c_i| \leq 2H_2 3^{\lceil H_1/\varepsilon \rceil} + 1 \leq B^{1/\varepsilon}$$

for a suitable choice of B. □

Theorem 5.3 (Finitary Tauberian theorem)**.** *Let $C > 0$ and suppose that $a > 0$ is a bound on $\{|a_n|\}$. Fix $\varepsilon > 0$ and let*

$$(b, v) := \Omega\left(\frac{\varepsilon}{8C}\right)$$

for Ω as defined in Lemma 5.2. Let

$$\delta := \frac{\varepsilon}{4v}$$

and given some $M \in \mathbb{N}$ define $N \in \mathbb{N}$ by

$$N := b \cdot \max\left\{\left\lceil \frac{L}{\delta} \right\rceil, M\right\} \quad \textit{for} \quad L := \frac{a}{1 - e^{-1/M}} + \delta$$

Finally, given $k \in \mathbb{N}$ define $p, l \in \mathbb{N}$ by

$$\begin{aligned} l &:= N + k - M \\ p &:= (N + k) \cdot \max\left\{\left\lceil \log\left(\frac{a(N+k)}{\delta}\right)\right\rceil, 1\right\} \end{aligned}$$

Then from

$$n|a_n| \leq C \quad \textit{for all } n \leq p \tag{12}$$

and

$$|F(x) - F(y)| \leq \delta \quad \textit{for all } x, y \in [e^{-1/M}, e^{-1/(M+l)}] \tag{13}$$

it follows that

$$|s_n - F(e^{-1/m})| \leq \varepsilon \quad \textit{for all } m, n \in [N, N+k]$$

Remark. We can actually say something more precise instead of (13): We in fact only need $|F(x)-F(y)| \leq \delta$ for all $x, y := e^{-i/n}, e^{-1/m}$ for $i \leq b$ and $m, n/i \in [M, M+l]$, and so technically the assumption only needs to hold for a finite number of points. However, we leave our slightly simpler formulation as it is, as this is sufficient for deriving remainder theorems in the next section.

Proof. We essentially follow the structure of the proof of Theorem 3.2, but backwards, and carrying out precise numerical calculations along the way. We suppose for contradiction that

$$\varepsilon < |s_n - F(e^{-1/m})|$$

for some $m, n \in [N, N+k]$. Now let $P(x) = c_d x^d + \ldots + c_1 x$ be the polynomial that satisfies Lemma 3.3 for error $\varepsilon/8C$, noting that $d \leq b$ and $\sum_{i=1}^{d} |c_i| \leq v$. Define I_P as in the proof of Theorem 3.2. We have

$$\varepsilon < |s_n - I_P(n)| + |I_P(n) - F(e^{-1/m})|$$

and therefore either

(i) $\varepsilon/2 < |s_n - I_P(n)|$ for some $n \leq N + k$ or,

(ii) $\varepsilon/2 < |I_P(n) - F(e^{-1/m})|$ for some $m, n \in [N, N+k]$.

We treat each possibility in turn, each leading to a contradiction in (12) and (13) respectively.

Case (i): $\varepsilon/2 < |s_n - I_P(n)|$ for $n \leq N + k$. We first observe that $p \geq N + k$ and therefore for any $t \geq p$ we have

$$\frac{t}{n} \geq \frac{t}{N+k} \geq 1$$

and thus $\chi(t/n) = 0$ for $t \in [p, \infty)$. Therefore using that $a_n \leq a$ for all $n \in \mathbb{N}$ we have

$$\int_p^\infty |a(t)| \cdot |\chi(t/n) - P(e^{-t/n})| \, dt \leq a \int_p^\infty |P(e^{-t/n})| \, dt$$

Now, for any $x \in [0, 1)$ we have

$$|P(x)| \leq \sum_{i=1}^{d} |c_i| x^d \leq x \cdot \sum_{i=1}^{d} |c_i| \leq xv$$

and therefore

$$a \int_p^\infty |P(e^{-t/n})| \, dt \leq av \int_p^\infty e^{-t/n} \, dt = avne^{-p/n} \tag{14}$$

Using again that $n \leq N + k$ and thus

$$p \geq n \cdot \log\left(\frac{an}{\delta}\right)$$

we obtain $\delta/an \geq e^{-p/n}$ and thus

$$avne^{-p/n} \leq \frac{\varepsilon}{4} \tag{15}$$

Using (2) together with (14) and (15) we have

$$|s_n - I_P(n)| \leq \int_0^p |a(t)| \cdot |\chi(t/n) - P(e^{-t/n})|\, dt + \frac{\varepsilon}{4}$$

and therefore

$$\frac{\varepsilon}{4} < \int_0^p |a(t)| \cdot |\chi(t/n) - P(e^{-t/n})|\, dt$$

But by (12) and (3) we have $|a(t)| \leq 2C/t$ for all $t \in [0, p]$ and thus using (4):

$$\begin{aligned}
\frac{\varepsilon}{4} &< 2C \int_0^p \frac{|\chi(t/n) - P(e^{-t/n})|}{t}\, dt \\
&\leq 2C \int_0^\infty \frac{|\chi(t/n) - P(e^{-t/n})|}{t}\, dt \\
&< 2C \cdot \frac{\varepsilon}{8C} = \frac{\varepsilon}{4}
\end{aligned}$$

a contradiction.

Case (ii): $\varepsilon/2 < |I_P(n) - F(e^{-1/m})|$ for $m, n \in [N, N+k]$. First of all, define

$$\epsilon_\alpha := 1 - \int_0^1 e^{-\alpha t}\, dt$$

For $\alpha > 0$ we have

$$e^{-\alpha} \leq \int_0^1 e^{-\alpha t}\, dt \leq 1$$

and therefore

$$0 \leq \epsilon_\alpha \leq 1 - e^{-\alpha} \leq \alpha \tag{16}$$

Now, similarly to (8) we see that

$$\begin{aligned}
\frac{\varepsilon}{2} &< |I_P(n) - F(e^{-1/m})| \\
&\leq \sum_{i=1}^d |c_i| \cdot |F(e^{-i/n})|(1 - \epsilon_{i/n}) - F(e^{-1/m})| \\
&\leq \sum_{i=1}^d |c_i| \cdot (\epsilon_{i/n}|F(e^{-i/n})| + |F(e^{-i/n}) - F(e^{-1/m})|)
\end{aligned} \tag{17}$$

We first aim to bound the term $\epsilon_{i/n}|F(e^{-i/n})|$ for $i = 1, \ldots, d$. Using $bM \leq N \leq n \leq N + k$ we have

$$\frac{1}{M+l} = \frac{1}{N+k} \leq \frac{i}{n} \leq \frac{b}{N} \leq \frac{b}{bM} = \frac{1}{M} \tag{18}$$

So to bound $|F(e^{-i/n})|$ we observe from (18) that $e^{-i/n} \in [e^{-1/M}, e^{-1/(M+l)}]$ and thus by condition (13)

$$|F(e^{-i/n})| \leq |F(e^{-1/M})| + \delta \leq \frac{a}{1 - e^{-1/M}} + \delta = L \tag{19}$$

where for the second inequality we have used that

$$F(x) \leq a \sum_{i=0}^{\infty} x^i = \frac{a}{1-x}$$

for $x \in [0, 1)$. Therefore from (16), (18) and (19) we have

$$\epsilon_{i/n}|F(e^{-i/n})| \leq \frac{i}{N} \cdot |F(e^{-i/n})| \leq \frac{b}{N} \cdot L \leq \delta \tag{20}$$

where the final inequality follows from the definition of N. Finally, analogously to (18) we have $1/(M+l) \leq 1/m \leq 1/M$ and thus $e^{-1/m} \in [e^{-1/M}, e^{-1/(M+l)}]$, and so from (13) we have

$$|F(e^{-i/n}) - F(e^{-1/m})| \leq \delta \tag{21}$$

Finally, putting together (17), (20) and (21) we have

$$\frac{\varepsilon}{2} < \sum_{i=1}^{d} |c_i| \cdot (\delta + \delta) \leq 2v\delta = \frac{\varepsilon}{2}$$

a contradiction. This completes the proof. □

5.3 A game semantics for Littlewood's theorem

Though the statement of our finitary theorem is somewhat complex, using the game-theoretic narrative from Section 4.2 we can give Theorem 5.3 a slightly more dynamic character, as setting out a winning strategy in a game corresponding to the Littlewood Tauberian theorem. Here, ∃loise sets out to foil ∀belard's attempt to disprove Littlewood's theorem by showing that $a_n = O(1/n)$, $F(x) \to s$ and $s_n \not\to s$ all hold together:

1. ∀belard starts by picking some $\varepsilon > 0$, assuming that $n|a_n| \leq C$ for all $n \in \mathbb{N}$, and proposing some $M \in \mathbb{N}$ such that $|F(x) - F(y)| \leq \delta$ for all $x, y \in [e^{-1/M}, 1)$. His aim is to show that it is now not the case that $|s_n - F(e^{-1/m})| \leq \varepsilon$ for sufficiently large m, n.

2. ∃loise responds by putting forward an $N \in \mathbb{N}$ such that $|s_n - F(e^{-1/m})| \leq \varepsilon$ for all $m, n \geq N$.

3. ∀belard rejects this by attempting to find a counterexample to the last move, playing $k \in \mathbb{N}$ and claiming that $\varepsilon < |s_n - F(e^{-1/m})|$ for some $n, m \in [N, N+k]$.

4. If ∀belard's attempt worked, then ∃loise responds by producing a pair $l, p \in \mathbb{N}$ which demonstrate that one of ∀belard's original assumptions was false:
   - either $C < n|a_n|$ for some $n \leq p$, or
   - $\delta < |F(x) - F(y)|$ for some $x, y \in [e^{-1/M}, e^{-1/(M+l)}]$.

A winning strategy for ∃loise constitutes a proof of the Littlewood Tauberian theorem, and such a winning strategy is provided by Theorem 5.3 in presenting bounds for winning moves for ∃loise in terms of any play from ∀belard .

6 Remainder theorems

Our final contribution in this paper is to use our finitisation of Littlewood's theorem to give a series of very general "remainder theorems". For us, a remainder theorem will be a quantitative form of the Tauberian theorem which differs from the finitary theorem of the previous section in the following ways:

(i) We assume that $a_n = O(1/n)$ rather than reformulating the growth condition in a finitary way,

(ii) For simplicity, we take as an additional parameter some $L > 0$ that bounds $|F(x)|$ on $x \in [0, 1)$ (which exists by the assumption that $F(x) \to s$),

(ii) We convert some quantitative measure of the convergence speed of $F(x) \to s$ as $x \to 1^-$ to a measure of the convergence speed of $s_n \to s$ as $n \to \infty$.

These results will form generalisations of Theorem 2.4, and are possible due to the careful analysis of the proof of Theorem 3.2 and the resulting witnesses of its Dialectica interpretation given in Theorem 5.3.

A key step towards these remainder theorems is the following corollary of Theorem 5.3, which we gain through our assumptions (i) and (ii) above:

Corollary 6.1. *Suppose that there exists some $C > 0$ such that $n|a_n| \leq C$ for all $n \in \mathbb{N}$, and let $L > 0$ be a bound on $|F(x)|$ for $x \in [0, 1)$. Define $\alpha : (0, \infty) \to (0, \infty)$, $\beta : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ and $\gamma : (0, \infty) \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ as follows:*

$$\begin{aligned}
\alpha(\varepsilon) &:= \frac{\varepsilon}{4\Omega_1(\varepsilon/8C)} \\
\beta(\varepsilon, M) &:= \Omega_0\left(\varepsilon/8C\right) \cdot \max\left\{\left\lceil \frac{L}{\alpha(\varepsilon)} \right\rceil, M\right\} \\
\gamma(\varepsilon, M, k) &:= \beta(\varepsilon, M) + k - M
\end{aligned}$$

for Ω as given in Lemma 5.2. Then we have

$$\forall \varepsilon, M, k[B(\alpha(\varepsilon), M, \gamma(\varepsilon, M, k)) \implies D(\varepsilon, \beta(\varepsilon, M), k)] \tag{22}$$

where the formulas $B(\delta, M, l)$ and $D(\varepsilon, N, k)$ are defined as in Section 5.1.

Proof. Directly from Theorem 5.3, noting that $\delta = \alpha(\varepsilon)$, $N = \beta(\varepsilon, M)$ and $\gamma(\varepsilon, M, k) = l$, with the only difference that in Theorem 5.3 we have $L = a/(1 - e^{-1/M}) + \delta$. However, its only role there is to act as a bound for $|F(x)|$ cf. (19), and so can be replaced by a general bound on $|F(x)|$. Note that the premise (12) trivially holds if we just assume that $n|a_n| \leq C$ for all $n \in \mathbb{N}$ as we do here, and so p plays no role. □

Letting P stand for the statement that $F(x)$ converges as $x \to 1^-$, and Q the statement that s_n converges as $n \to \infty$, (22) is essentially a Dialectica interpretation of $P \implies Q$. Our remainder theorems take the form of a computational interpretation of modus ponens:

$$\frac{P^* \qquad (P \implies Q)^D}{Q^*}$$

where P^* and Q^* are suitable computational interpretations of P and Q, specifically either direct rates of convergence or rates of metastability. Both of these will be carefully motivated and defined below.

6.1 A remainder theorem for rates of convergence

The known remainder estimate for Littlewood's theorem set out as Theorem 2.4 says that whenever $F(x)$ converges with exponential rate, then s_n converges with inverse logarithmic rate. This is a special example of a general phenomenon, which we make precise here, whereby we can convert *any* computable rate of convergence for $F(x)$ into a corresponding computable rate of convergence for s_n. We now state and prove this remainder theorem it using our finitary Tauberian theorem in the form of Corollary 6.1, and then rederive the known remainder estimate as a special case.

Definition 6.2. We define a rate of convergence for $F(x)$ as $x \to 1^-$ to be any function $\phi : (0, \infty) \to \mathbb{N}$ satisfying

$$\forall \delta > 0 \exists M \leq \phi(\delta) \forall x, y \in [e^{-1/M}, 1)(|F(x) - F(y)| \leq \delta)$$

Similarly, a rate of convergence for $s_n \to \lim_{x \to 1^-} F(x)$ as $n \to \infty$ is defined to be any function $\psi : (0, \infty) \to \mathbb{N}$ satisfying

$$\forall \varepsilon > 0 \exists N \leq \psi(\varepsilon) \forall m, n \geq N(|s_n - F(e^{-1/m})| \leq \varepsilon)$$

Theorem 6.3 (First remainder theorem)**.** *Suppose that there exists some $C > 0$ such that $n|a_n| \leq C$ for all $n \in \mathbb{N}$, and let $L > 0$ be a bound on $|F(x)|$ for $x \in [0, 1)$. Suppose that there exists a computable rate of convergence ϕ for $F(x)$ as $x \to 1^-$. Then a computable rate of convergence for $s_n \to \lim_{x \to 1^-} F(x)$ is given by*

$$\psi(\varepsilon) := \beta(\varepsilon, \phi(\alpha(\varepsilon)))$$

where α and β are defined as in Corollary 6.1.

Proof. Letting $B(\delta, M, l)$ and $D(\varepsilon, N, k)$ be defined as in Section 5.1, ϕ being a rate of convergence for $F(x)$ as $x \to 1^-$ is equivalent to the formula

$$\forall \delta > 0 \exists M \leq \phi(\delta) \forall l\, B(\delta, M, l)$$

In particular, for any $\varepsilon > 0$ and $k \in \mathbb{N}$ there exists $M \leq \phi(\alpha(\varepsilon))$ such that

$$\forall k\, B(\alpha(\varepsilon), M, \gamma(\varepsilon, M, k))$$

where γ is as in Corollary 6.1. Therefore from (22) we see that for $N := \beta(\varepsilon, M)$ we have $\forall k\, D(\varepsilon, N, k)$ or equivalently

$$\forall m, n \geq N(|s_m - F(e^{-1/m})| \leq \varepsilon).$$

Finally, we observe that β is monotone in its second argument and thus

$$N = \beta(\varepsilon, M) \leq \beta(\varepsilon, \phi(\alpha(\varepsilon))) = \psi(\varepsilon)$$

and since $\varepsilon > 0$ is arbitrary we have shown that ψ is a rate of convergence for $s_n \to \lim_{x \to 1^-} F(x)$. □

We can now give a concrete instance of this remainder theorem that corresponds directly to Theorem 2.4:

Corollary 6.4. *Suppose that there exists some $C > 0$ such that $n|a_n| \leq C$ for all $n \in \mathbb{N}$, and let $L > 0$ be a bound on $|F(x)|$ for $x \in [0, 1)$. Suppose that $F(x)$ converges with rate of convergence ϕ where*

$$\phi(\delta) \leq a\delta^{-b}$$

for some $a, b > 0$. Then there exists a rate of convergence ψ for $s_n \to \lim_{x \to 1^-} F(x)$ with

$$\psi(\varepsilon) \leq K^{1/\varepsilon}$$

where K is a suitable constant depending on C, a and b.

Proof. By our Theorem 6.3 and using the definition of Ω_0 from Lemma 5.2 we have

$$\begin{aligned} \psi(\varepsilon) = \beta(\varepsilon, \phi(\alpha(\varepsilon))) &\leq \beta\left(\varepsilon, \frac{a}{\alpha(\varepsilon)^b}\right) \\ &= \frac{8AC}{\varepsilon} \cdot \max\left\{\left\lceil \frac{L}{\alpha(\varepsilon)} \right\rceil, \frac{a}{\alpha(\varepsilon)^b}\right\} \end{aligned}$$

Substituting in the definition of Ω_1 together with the fact that $\varepsilon \geq e^{-1/\varepsilon}$ we have

$$\alpha(\varepsilon) = \frac{\varepsilon}{4B^{8C/\varepsilon}} \geq K_1^{-1/\varepsilon}$$

for a suitable constant K_1 dependent on B and C, and therefore

$$\psi(\varepsilon) \leq \frac{8AC}{\varepsilon} \cdot \max\{\lceil LK_1^{1/\varepsilon} \rceil, aK_1^{b/\varepsilon}\}$$

and assuming w.l.o.g. that $\varepsilon \leq 1$ we can find a sufficiently large constant K in terms of A, B, C, a and b such that

$$\psi(\varepsilon) \leq K^{1/\varepsilon}$$

and so we're done. □

We now show that this latter result is simply a reformulation, using 'proof-theoretic' rates of convergence, of the standard remainder estimate for Littlewood's theorem:

Proof of Theorem 2.4 from 6.4. Suppose that

$$F(x) = s + O((1-x)^b)$$

for some $b > 0$, or equivalently

$$|F(x) - s| \leq a(1-x)^b$$

for some $a > 0$. Let $M \in \mathbb{N}$ be arbitrary and suppose that $x, y \in [e^{-1/M}, 1)$. Then

$$\begin{aligned} |F(x) - F(y)| &\leq |F(x) - s| + |s - F(y)| \\ &\leq 2a(1 - e^{-1/M})^b \leq \frac{2a}{M^b} \end{aligned}$$

for the last line using the inequality $1 + x \leq e^x$ for $x = -1/M$. But then this implies that $|F(x) - F(y)| \leq \delta$ for any $x, y \in [e^{-1/M}, 1)$, provided that

$$M \geq (2a/\delta)^{-b}$$

and so $\phi(\delta) = (2a/\delta)^{-b}$ is a rate of convergence for $F(x)$ as $x \to 1^-$ in our sense. By Corollary 6.4 it follows that there is some constant $K > 0$ such that for any $\varepsilon > 0$ we have

$$\forall m, n \geq K^{2/\varepsilon}(|s_n - F(e^{-1/m})| \leq \varepsilon/2)$$

Choosing $m \in \mathbb{N}$ so that $|F(e^{-1/m}) - s| \leq \varepsilon/2$, we therefore have

$$\forall n \geq K^{2/\varepsilon}(|s_n - s| \leq \varepsilon)$$

for any $\varepsilon > 0$. We can invert this such that for any $n \in \mathbb{N}$, setting $\varepsilon := 2\log(K)/\log(n)$ given us $K^{2/\varepsilon} = n$ and thus

$$|s_n - s| \leq \frac{2\log(K)}{\log(n)}$$

and from this it follows that

$$s_n = s + O\left(\frac{1}{\log(n)}\right)$$

and so Theorem 2.4 is proved. □

6.2 A remainder theorem for rates of metastability

Our first remainder theorem applies to the situation where both $F(x)$ and s_n converge with computable rates. However, it is well known that it is not in general the case that sequences possess computable rates of convergence. The canonical counterexamples here are so-called Specker sequences [25], bounded and monotonically increasing sequences of rational numbers whose limit is not a computable real number.

In [23] (cf. Proposition 3.2) it was shown that for any Specker sequence $\{q_n\}$, defining

$$a_n := \frac{q_{m+1} - q_m}{2^{m-1}} \quad \text{for } m = \lceil \log_2(n) \rceil$$

we obtain a sequence satisfying $a_n = o(1/n)$ and $s_{2n} = q_{n+1}$, and so therefore $s_n \to q$ where q is the noncomputable limit of the Specker sequence, from which it can be shown that s_n has no computable rate of convergence. By Abel's theorem it follows that $F(x) \to q$ as $x \to 1^-$, but this also cannot possess a computable rate of convergence in the sense of Definition 6.2: If this were the case then because $a_n = o(1/n)$ and so in particular $a_n = O(1/n)$, by Theorem 6.3 a computable rate of convergence for $s_n \to q$ could be found.

As such, there are situations where Theorem 6.3 cannot apply, where $F(x) \to s$ but with no computable rate of convergence. However, we now show that we can use our finitary Tauberian theorem also in the case where we do not have computable rates, to instead convert computable rates of *metastability* for $F(x) \to s$ to corresponding rates of metastability for $s_n \to s$. First of all, we need to define what a rate of metastability means in this context.

Cauchy convergence in general is a $\forall\exists\forall$ statement i.e. of the form

$$\forall \varepsilon \exists N \forall k \, P(\varepsilon, N, k) \tag{23}$$

for some computationally neutral inner formula P. The existence of Specker sequences demonstrates that there are instances of formulas of this logical form where we cannot produce a computable bound on witnesses satisfying its Dialectica interpretation i.e. there is no computable ϕ satisfying

$$\forall \varepsilon \exists N \leq \phi(\varepsilon) \forall k \, P(\varepsilon, N, k)$$

Generally, this is because the proofs of such statements can use classical reasoning, and so the intuitionistic soundness theorem which would normally imply the existence of a computable ϕ does not apply. However, we can instead apply the Dialectica interpretation in its *classical* form, by precomposing formulas with a negative translation. In particular, it is typically the case that for $\forall\exists\forall$ formulas, the following is provable intuitionistically, even when (23) is not:

$$\forall \varepsilon \neg\neg \exists N \forall k \, P(\varepsilon, N, k) \tag{24}$$

Applying the Dialectica interpretation to (24), we are instead asked to find a witness for N in the following formula:

$$\forall \varepsilon, g \exists N \, P(\varepsilon, N, g(N))$$

In the case of Cauchy convergence, a functional $\Phi : (0,\infty)\times(\mathbb{N}\to\mathbb{N})\to\mathbb{N}$ that bounds N in terms of ε and g i.e.

$$\forall\varepsilon, g\exists N \leq \Phi(\varepsilon, g)\, P(\varepsilon, N, g(N)) \tag{25}$$

is known as a *rate of metastability*. Metastable convergence theorems and the key role that they play in analysis as finitizations of 'soft' convergence statements is discussed in an essay by T. Tao [26], and the connection with the Dialectica interpretation is explored in [10] and particularly [5]. It is usually possible to extract rates of metastability from convergence proofs even when direct rates of convergence are not possible, and the extraction of such metastable bounds is a standard result in applied proof theory [12, 11, 22]. We can define rates of metastability for the two relevant convergence properties here following the pattern described above, and these should both be viewed as bounds on witnessing terms for the combined negative translation plus Dialectica interpretation of the respective properties.

Definition 6.5. We define a rate of metastability for $F(x)$ as $x\to 1^-$ to be any functional $\Phi : (0,\infty)\times(\mathbb{N}\to\mathbb{N})\to\mathbb{N}$ satisfying

$$\begin{aligned}&\forall\delta>0, h:\mathbb{N}\to\mathbb{N}\exists M\leq\Phi(\delta,h)\\ &\forall x,y\in[e^{-1/M}, e^{-1/(M+hM)}](|F(x)-F(y)|\leq\delta)\end{aligned}$$

Similarly, a rate of metastability for $s_n\to\lim_{x\to 1^-}F(x)$ as $n\to\infty$ is defined to be any function $\Psi : (0,\infty)\times(\mathbb{N}\to\mathbb{N})\to\mathbb{N}$ satisfying

$$\begin{aligned}&\forall\varepsilon>0, g:\mathbb{N}\to\mathbb{N}\exists N\leq\Psi(\varepsilon,g)\\ &\forall m,n\in[N,N+gN](|s_n-F(e^{-1/m})|\leq\varepsilon)\end{aligned}$$

We now generalise our first remainder theorem so that it also applies in the case where $F(x)\to s$ as $x\to 1^-$, without a computable rate of convergence but with a computable rate of metastability.

Theorem 6.6 (Second remainder theorem)**.** *Suppose that there exists some $C>0$ such that $n|a_n|\leq C$ for all $n\in\mathbb{N}$, and let $L>0$ be a bound on $|F(x)|$ for $x\in[0,1)$. Suppose that Φ is a computable rate of metastability for $F(x)$ as $x\to 1^-$. Then a computable rate of metastability for $s_n\to\lim_{x\to 1^-}F(x)$ is given by*

$$\Psi(\varepsilon,g) := \beta(\varepsilon,\Phi(\alpha(\varepsilon),h_{\varepsilon,g}))$$

for $h_{\varepsilon,g}:\mathbb{N}\to\mathbb{N}$ defined by

$$h_{\varepsilon,g}(k) := \gamma(\alpha(\varepsilon), k, g(\beta(\varepsilon,k)))$$

where α,β and γ are defined as in Corollary 6.1.

Proof. Again, letting $B(\delta,M,l)$ and $D(\varepsilon,N,l)$ be defined as in Section 5.1, if Φ is a rate of metastability for $F(x)$ as $x\to 1^-$ then

$$\forall\delta>0, h\exists M\leq\Phi(\delta,h)\, B(\delta,M,h(M))$$

In particular, for any $\varepsilon > 0$ and $g : \mathbb{N} \to \mathbb{N}$, setting $\delta = \alpha(\varepsilon)$ and $h = h_{\varepsilon,g}$ there exists $M \leq \Phi(\alpha(\varepsilon), h_{\varepsilon,g})$ such that

$$B(\alpha(\varepsilon), M, \gamma(\alpha(\varepsilon), M, g(\beta(\varepsilon, M))))$$

and therefore setting $N := \beta(\varepsilon, M)$, it follows from (22) that

$$D(\varepsilon, N, g(N))$$

By monotonicity of β we have

$$N = \beta(\varepsilon, M) \leq \beta(\varepsilon, \Phi(\alpha(\varepsilon), h_{\varepsilon,g})) = \Psi(\varepsilon, g)$$

and since ε and g were arbitrary we have shown that Ψ is a rate of metastability for $s_n \to \lim_{x\to 1^-} F(x)$. □

7 Concluding remarks

In this paper, we have given a computational interpretation via Gödel's Dialectica interpretation to Littlewood's celebrated Tauberian theorem. The immediate relevance of this computational Tauberian theorem is demonstrated by deriving two remainder theorems which generalise existing remainder estimates to arbitrary rates of convergence, and even the case where no computable rate of convergence exists. But as a side product, we gain a intuitive constructive reading of Littlewood's theorem in terms of a winning strategy in a two player game, and we consider this to be of independent interest. It is also hoped that our case study is self-contained enough that it will form an useful illustration of how the Dialectica interpretation can be applied in mathematics to obtain quantitative information from proofs.

However, above all we see this paper as a forming a stepping stone to deeper results in quantitative Tauberian theory, bringing initial ideas sketched in [23] to bear on a much more complex Tauberian theorem, and demonstrating in turn that the Dialectica can be used to both rederive and generalise known numerical results in this area. We propose Tauberian theory as a area where there is a great deal of potential for applying proof-theoretic methods, and conclude with the following questions:

1. Can we extend the ideas presented here to the more complex Tauberian theorems later proved by Hardy and Littlewood in e.g. [7], to integral analogues of Tauberian theorems using Karamata's method as discussed in [14, Chapter I.13–14], or to even deeper results in Tauberian theory involving Fourier transformations and Wiener kernels cf. [14, Chapter II]?

2. Can we use techniques from proof theory to make further contributions to Tauberian remainder theory? In particular, are there Tauberian theorems with no known remainder estimates, for which the application of proof-theoretic methods could produce not just generalisations of existing remainder estimates as in this case, but the first ever remainder theorems?

References

[1] Avigad, J., and Feferman, S. Gödel's functional ("Dialectica") interpretation. In *Handbook of Proof Theory*, S. R. Buss, Ed., vol. 137 of *Studies in Logic and the Foundations of Mathematics*. North Holland, Amsterdam, 1998, pp. 337–405.

[2] Berardi, S., Bezem, M., and Coquand, T. On the computational content of the axiom of choice. *Journal of Symbolic Logic 63*, 2 (1998), 600–622.

[3] Coquand, T. A semantics of evidence for classical arithmetic. *Journal of Symbolic Logic 60* (1995), 325–337.

[4] de Paiva, V. *The Dialectica categories*. PhD thesis, University of Cambridge, 1991. Published as Technical Report 213, Computer Laboratory, University of Cambridge.

[5] Gaspar, J., and Kohlenbach, U. On Tao's "finitary" infinite pigeonhole principle. *Journal of Symbolic Logic 75* (2010), 355–371.

[6] Gödel, K. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *dialectica 12*, 3–4 (1958), 280–287.

[7] Hardy, G. H., and Littlewood, J. E. Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive. *Proceedings of the London Mathematical Society 2*, 13 (1914), 174–191.

[8] Karamata, J. Über die Hardy-Littlewoodschen Umkehrungen des Abelschen Stetigkeitssatzes. *Mathematische Zeitschrift 32* (1930), 319–320.

[9] Kerjean, M., and Pédrot, P.-M. δ is for Dialectica: Typing differentiable programming. Preprint, available at `https://hal.archives-ouvertes.fr/hal-03123968/`, 2021.

[10] Kohlenbach, U. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer Monographs in Mathematics. Springer, 2008.

[11] Kohlenbach, U., and Koutsoukou-Argyraki, A. Rates of convergence and metastability for abstract cauchy problems generated by accretive operators. *Journal of Mathematical Analysis and Applications 423* (2015), 1089–1112.

[12] Kohlenbach, U., and Leuştean, L. Effective metastability of halpern iterates in CAT(0) spaces. *Advances in Mathematics 321* (2012), 2526–2556.

[13] Korevaar, J. Best L_1 approximation and the remainder in Littlewood's Tauberian theorem. *Indagationes Mathematicae 15* (1953), 281–293.

[14] Korevaar, J. *Tauberian Theory*. Grundlehren der mathematischen Wissenschaften. Springer-Verlag Berlin Heidelberg, 2004.

[15] Kreisel, G. On the interpretation of non-finitist proofs, Part I. *Journal of Symbolic Logic 16* (1951), 241–267.

[16] Kreisel, G. On the interpretation of non-finitist proofs, Part II: Interpretation of number theory. *Journal of Symbolic Logic 17* (1952), 43–58.

[17] Littlewood, J. E. The converse of Abel's theorem on power series. *Proceedings of the London Mathematical Society 2*, 9 (1911), 434–448.

[18] Luckhardt, H. Herbrand-Analysen zweier Beweise des Satzes von Roth: Polynomiale Anzahlschranken. *Journal of Symbolic Logic 54* (1989), 234–263.

[19] Pédrot, P.-M. *A Materialist Dialectica*. PhD thesis, Paris Diderot University, 2015.

[20] Poernomo, I., Crossley, J. N., and Wirsing, M. *Adapting Proofs-as-Programs*. Monographs in Computer Science. Springer, 2005.

[21] Powell, T. A functional interpretation with state. In *Proceedings of Logic in Computer Science (LICS '18)* (2018), ACM, pp. 839–848.

[22] Powell, T. A new metastable convergence criterion and an application in the theory of uniformly convex Banach spaces. *Journal of Mathematical Analysis and Applications 478*, 2 (2019), 790–805.

[23] Powell, T. A note on the finitization of Abelian and Tauberian theorems. *Mathematical Logic Quarterly 66*, 3 (2020), 300–310.

[24] Schwichtenberg, H., and Wainer, S. *Proofs and Computations*. Perspectives in Logic. Cambridge University Press, 2011.

[25] Specker, E. Nicht konstruktiv beweisbare Sätze der Analysis. *Journal of Symbolic Logic 14* (1949), 145–158.

[26] Tao, T. Soft analysis, hard analysis, and the finite convergence principle. Essay posted 23 May 2007, 2007. Appeared in: 'T. Tao, Structure and Randomness: Pages from Year One of a Mathematical Blog. AMS, 298pp., 2008'.

[27] Tauber, A. Ein Satz aus der Theorie der unendlichen Reihen. *Monatshefte für Mathematik und Physik 8* (1897), 273–277.

[28] Weiermann, A. Analytic combinatorics for a certain well-ordered class of iterated exponential terms. In *Proceedings of International Conference on Analysis of Algorithms* (2005), pp. 409–416.

[29] Weiermann, A. Monadic second order limit laws for natural well orderings. Preprint, available at `https://arxiv.org/abs/2007.14111`, 2020.

[30] Wielandt, H. Zur Umkehrung des Abelschen Stetigkeitssatzes. *Mathematische Zeitschrift 56* (1952), 206–207.