THE EARLY CAREER OF G.H. HARDY

by

Brenda Davison
B.A.Sc.(hons), University of British Columbia, 1989

A Thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
in the Department
of
Mathematics

SIMON FRASER UNIVERSITY
Summer 2010

APPROVAL

Name: Brenda Davison

Degree: Master of Science

Title of Thesis: The Early Career of G.H. Hardy

Examining Committee: Dr. Veselin Jungić
Chair

Dr. W. Tom Archibald, Senior Supervisor

Dr. Len Berggren, Supervisor

Dr. Nilima Nigam, External Examiner

Date Approved: July 29, 2010

Abstract

G.H. Hardy's early career can be demarcated by his election to Trinity College in 1901 and by the beginning of his collaboration, in 1911, with J.E. Littlewood. During this time he wrote a textbook of enduring importance, established a reputation as an analyst, wrote five papers on set theory, contributed to the *Educational Times* and wrote several book reviews. He also began to play a role in political and social issues via his membership in the Apostles and the London Mathematical Society, as well as through his work to abolish the Tripos examinations.

I will discuss Hardy's mathematical work during this early period including his work on integration, his textbook, *A Course of Pure Mathematics*, and his five set theory papers.

For Nadia

"The function of a mathematician, then, is simply to observe the facts about his own intricate system of reality, that astonishingly beautiful complex of logical relations which forms the subject-matter of his science, as if he were an explorer looking at a distant range of mountains, and to record the results of his observations in a series of maps, each of which is a branch of pure mathematics. Many of these maps have been completed, while in others, and these, naturally, are the most interesting, there are vast uncharted regions."

— *G.H. Hardy,* NATURE, *1922*

Acknowledgments

Foremost, I thank my supervisor Dr. Tom Archibald for accepting me as his student and providing the excellent instruction that went with that responsibility.

Additionally, I thank Dr. Len Berggren from whom I took my first course in the History of Mathematics and who encouraged me to continue in this field.

Finally, I thank my family for giving me the time I needed to complete this project.

Contents

Chapter 1

Introduction

Godfrey Harold Hardy (1877-1947) entered Trinity College, Cambridge in 1896 with an open scholarship after attending Winchester College[1] public school. He was 4^{th} wrangler in the Tripos examination of 1898 and was elected as a fellow of Trinity College in 1900. He was subsequently awarded a Smith's prize in 1901.

The first period of Hardy's career can be delineated by his election as a fellow of Trinity College and by the beginning of his collaboration, in 1911, with John Edensor Littlewood (1885-1977).[2]

Hardy started to publish in 1899, prior to his election as a fellow. In the first two years, all but one of his 18 papers appeared in the *Educational Times*. This was a venue for posing questions and solutions to stated mathematical problems. The remaining paper, on integration, was his "first substantial paper" [92, p. 6] and the first of many papers on integration – most of which were published in the early period of Hardy's career.

Integration was a mathematical interest throughout Hardy's career and it was the main interest of Hardy's early career. His reputation as an analyst was established with these early integration papers. In 1905, Hardy's first book, titled *The Integration of Functions of a Single Variable* was published as part of the *Cambridge Tracts in Mathematics and Mathematical Physics*. This book, which was not an extension of or a summary of the work present in the concurrent papers, presented integration in finite terms as a methodical, algorithmic topic. This first mathematical interest of Hardy's is discussed in chapter 2.

[1]widely regarded as the best school in England for mathematical training

[2]This collaboration, which will not be discussed in this thesis, was one of the most productive collaborations in the mathematical history. Hardy and Littlewood wrote 94 joint papers.

A second early mathematical interest of Hardy's was the topic of set theory. He was one of a few British mathematicians to read and incorporate the work of Georg Cantor (1845-1918) and Richard Dedekind (1831-1916) into his own. This interest led to five papers on set theory published in the years between 1904 and 1910. Hardy's role in bringing this continental mathematics from Germany to Britain in the first decade of the 20^{th} century and the content of his five set theory papers is discussed in chapter 3.

The third theme in Hardy's early career was his dominant role in the shift in Cambridge mathematics that occurred during the early part of the 20^{th} century. In 1908, Hardy wrote a textbook titled *A Course of Pure Mathematics*. This book, in conjunction with Hardy himself, is often credited[3] with bringing rigour to and transforming British analysis. Some of the new set theory material was introduced to British mathematicians through this textbook.

Another facet of Hardy's impact on Cambridge mathematics was his role in the reform of the mathematical examinations, called the Tripos examinations, administered at Cambridge. Teaching was dominated by the preparation for these examinations. Hardy spent considerable time and energy to bring about reform to lessen the impact of these examinations on the development of mathematical talent at Cambridge.

Both the role of Hardy's text, *A Course of Pure Mathematics*, and the role of Hardy himself in changing analysis at Cambridge is discussed in chapter 4.

Hardy said of the early period of his career that he "wrote a great deal... but very little of any importance; there are not more than four or five papers which I can still remember with some satisfaction" [54, p.147]. However, John Burkhill (1900-1993), who wrote the entry on Hardy in the Dictionary of Scientific Biography, said of Hardy's early career that he

> wrote many papers on the convergence of series and integrals and allied topics. Although this work established his reputation as an analyst, his greatest service to mathematics in this early period was *A Course of Pure Mathematics (1908).* This work was the first rigorous exposition of number, function, limit, and so on, adapted to the undergraduate, and thus it transformed university teaching [10].

I will establish that these first 10 years of Hardy's career were mathematically interesting, productive, and that they justify Hardy's reputation as a vector for continental mathematics into Britain. During the 19th century,

[3]This will be established via a variety of quotations from numerous sources.

> in pure mathematics, England produced only a handful of algebraists, among them Cayley and Sylvester, and failed to produce any notable analysts. This sorry state of affairs was changed by Littlewood and Hardy: by 1930 the school of analysis established by them was second to none." [5, p. 2]

In this thesis I cover part of the story of how this happened – detailing the influence of Hardy.

Chapter 2

Integration

In 1905, at the age of 28, Hardy published his first book – *The Integration of Functions of a Single Variable* (subsequently called *Integration*). This publication represents one of the major themes in Hardy's early career as his reputation as an analyst was established with a series of papers on series and integral convergence. This early interest in integration is discussed here using Hardy's *Integration* as a touchstone.

Integration seeks to present integration in finite terms as a coherent, methodical topic, in contrast to other integration texts which Hardy felt presented integration as a series of disconnected, clever tricks that must be learned. Without referring to specific textbooks[1] on the integral calculus, Hardy said

> the student who is only familiar with the latter [i.e. textbooks on integral calculus] is apt to be under the impression that the process of integration is essentially 'tentative' in character, and that its performance depends on a large number of disconnected though ingenious devices. [41, p. v]

The view of indefinite integration as a collection of tricks is one that has prevailed. A 1992 text devoted to algorithmic solutions of algebraic problems stated:

> Integration is typically not viewed as an algorithmic process, but rather as a collection of tricks which can only solve a limited number of integration problems. [25, p.18]

[1] However, in the footnotes of *Integration*, Hardy refers to both Chrystal's *Algebra* and A.G. Greenhills' 1888 *A Chapter in the Integral Calculus*.

Hardy's insistence on rigorous proof, his clear expository style, and his confidence and clarity in expressing what he felt to be the important issues about a topic are all easily seen in this early work.

2.1 Background Information

2.1.1 The Cambridge Tracts

Hardy's *Integration* is the second in a series of books collectively called the *Cambridge Tracts in Mathematics and Mathematical Physics.* The general editors of the series were J.G. Leathem (1871-1923) and E.T. Whittaker (1873-1956).[2] Leathem, a fellow of St. John's College, was also the author of the first book in the series, a book of mathematical physics titled *Volume and Surface Integrals Used in Physics.* This first tract was also published in 1905.

Whittaker, a fellow of Trinity College, Cambridge, was a contemporary of Hardy's during this early period of Hardy's career. In fact, Hardy was originally one of Whittaker's first students. As such, Whittaker was undoubtedly aware of Hardy's interests and abilities and his appropriateness as the author of a tract on integration.

Starting in 1905 with tract one and ending in 1972 with tract 63, the Cambridge tracts were intended to be concise summaries of a single topic. These were later supplanted by the *Cambridge Tracts in Mathematics*[3], which are still being published by Cambridge University Press. As of 2010 there are 166 titles in the new series.

Hardy went on to author three additional Cambridge tracts. Tract number 12, titled *Orders of Infinity: The Infinitärcalcül of Paul du Bois-Reymond*, was published in 1910. Its purpose was to bring the work and symbolism of Paul du Bois-Reymond (1831-1889) to a broader audience – an attempt that appears to have failed.[4] The tract mostly concerns

---

[2]It is claimed in Leathem's obituary that, under the care of Leathem and Whittaker, the Cambridge tracts had become "an important survey, almost an encyclopedia, of domains of recent higher mathematics." [67, p. 437]

[3]This publication has the stated purpose of being "devoted to thorough, yet reasonably concise treatments of topics in any branch of mathematics. Typically, a Tract takes up a single thread in a wide subject, and follows its ramifications, thus throwing light on various of its aspects. Tracts are expected to be rigorous, definitive and of lasting value to mathematicians working in the relevant disciplines. Exercises can be included to illustrate techniques, summarize past work and enhance the book's value as a seminar text." [4]

[4]For example, this is a quote from a review done by Hurwitz [61, p. 202] "One is impelled to wonder how much of the fairly extensive notation introduced will be found really desirable in actual use of the results.

what today would be referred to as big O notation – the analysis of how fast a function of a variable grows with increasing values of the variable.

A review of this book in 1915 noted Hardy's high standards of proof, clarity and accuracy:

> A great many theorems dealing with the limit-behavior of functions were obtained by Du Bois-Reymond in a series of papers dating from 1871 to 1880. These theorems have been collected, recast according to modern requirements of rigor, and amplified by Hardy in No. 12 of the Cambridge Tracts in Mathematics and Mathematical Physics. The work has been done in a manner admitting of no criticism; the treatment is clear and readable; the proofs are accurate and carefully worded. [61, p. 202]

Hardy's third book in the series is Cambridge Tract number 18, titled *The General Theory of Dirichlet's Series*, was published in 1915 and coauthored by Marcel Riesz (1886-1969). Riesz was a Hungarian born mathematician who spent nearly all of his working life in Sweden.

Hardy's fourth book in the series, coauthored by W. W. Rogosinski (1894-1964) and published in 1944, is titled *Fourier Series*. It is of less interest here since it comes very late in Hardy's career.

2.1.2 Hardy's Work on Integration prior to his Cambridge tract

A significant portion of the work that Hardy had done prior to publishing *Integration* was, unsurprisingly, on integration. He published approximately thirty papers on integration between 1900 and 1905 (see table 2.1 below)

Table 2.1: Number of Hardy's Integration Papers by Year

Year	1900	1901	1902	1903	1904	1905
Papers	1	4	9	7	2	3

mostly in the *Messenger of Mathematics* as well as in the *Quarterly Journal* and in the *Proceedings of the London Mathematical Society*. Integration was

The notions of inferior and superior limit have won a permanent place; Landau's symbols $O(f)$ and $o(f)$ have in a short time come into such wide use as apparently to insure their retention. It is doubtful whether any further notation will be found necessary."

> a subject which turned out to be one of his permanent minor interests, and on which he was still writing in the last year of his life. [93, p. 448]

Hardy's collected works span seven volumes and all of the fifth volume is devoted to the integral calculus. The fifth volume is close to seven hundred pages, and half of the papers in it were published prior to 1908 and all but two of them were published prior to 1920. For example, Hardy's first paper on integration, *On a class of definite integrals containing hyperbolic functions* of 1900, concerns the evaluation of definite integrals of the form

$$\int_{-\infty}^{\infty} \frac{e^{\alpha+i\cdot\beta}}{ae^x + b + ce^{-x}} R(x)dx$$

where $a = \pm c$ and $R(x)$ is a rational function. Contour integration around a large semicircle is used.

This type of paper seems to be typical of Hardy's early work on integration and the topic of indefinite integration in finite terms appears to be confined to *Integration* and does not appear to be an extension of the work that Hardy was concurrently publishing in his papers. An editorial comment appended to this 1900 paper stated

> it already shows his ingenuity in devising methods for the evaluation of definite integrals, a topic to which he returned again and again with obvious enjoyment[5]

2.1.3 Works to which Hardy refers

In the preface of *Integration*, Hardy acknowledged three authors to whom he owed a debt. They are Charles Hermite (1822-1901), Edouard Jean-Baptiste Goursat (1858-1936) and Joseph Liouville (1809-1882). Regarding both Hermite and Goursat, he referred to their identically titled books *Cours d'Analyse*. About Liouville, he said:

> my greatest debt is to Liouville, who published in the years 1830-40 a series of remarkable memoirs on the general problem of integration which appear to have fallen into an oblivion which they certainly do not deserve. [41, p.vi]

Hardy was correct in his analysis of Liouville's papers. Liouville discontinued his research into integration in finite terms early in the 1840's, possibly discouraged by the limited impact his theory had in his own time. Lützen, in his comprehensive biography of Liouville [69],

[5] See Hardy's collected papers, volume 7, page 27.

could only find two traces of contemporary interest in Liouville's papers on integration in finite terms.

The Lützen biography of Liouville claims that Hardy's *Integration* was

> the first work after Liouville to deal in full generality with the question of integration in finite terms. With its praise of Liouville, it probably recreated an interest in this almost forgotten field. [69, p. 419]

This is a topic that gained increasing importance over the 20th century with the development of computer algebra systems.

2.2 The Contents of *The Integration of Functions of a Single Variable*

2.2.1 Overview of the text

Integration is a short book of 53 pages – Hardy refers to it as a pamphlet. His stated aim is to "find a function whose differential coefficient is a given function"[6] [41, p.1] in a manner that shows that the

> solution of any elementary problem of integration may be sought in a perfectly definite and systematic way. [41, p. v]

The book has a table of contents but no index. The appendix lists the works of Abel, Liouville and Chebyshev that Hardy refers to in the text, and to which the reader may refer for further study. There is also a longer, but less detailed, list of works that relate to the integration of algebraical functions.

Hardy then delineates the topic. In his view, the theory of the integral calculus is not a subset of the theory of differential equations and it is not a subset of the theory of the definite integral. He is not concerned with limits, continuity, or convergence. What is important is

> the *form* of the solution, and the only proof of its existence which is of any value to us is that which consists in actually expressing it in terms of x. [41, p.2]

[6] E.g. to solve the differential equation $\frac{dy}{dx} = f(x)$

2.2.2 Elementary Functions

Hardy first defines an elementary function as a "member of a class of functions" [Har05, p. 3] comprising rational functions, explicit or implicit algebraical functions, e^x, $\ln(x)$ and all finite combinations of the above. This is necessary because the rest of the book is "exclusively concerned with the question of the integration of elementary functions" [41, p.7]

Note the set-theoretic language used in the definition of an elementary function – set theory was still new at this time and a concurrent interest of Hardy's which is discussed in chapter 3. This is essentially the modern definition[7] of an elementary function.

Further distinctions in the types of elementary functions are made. First, the elementary functions that are not rational or algebraic are called elementary transcendental functions - this is adopted from Liouville.[8] The elementary transcendental functions are further classified into orders, again in a manner first indicated by Liouville. First order transcendental functions are functions in which all of the arguments of the exponential and logarithmic functions are rational or algebraic. Second order transcendental functions are functions in which the arguments of the exponential and logarithmic functions are first order transcendental functions and so on. Many functions of interest and importance are second order transcendental functions – for example e^{e^x} or $\ln(\ln(x))$. For Hardy, part of this interest was in these functions' application to classifying the orders of infinity which again is a reflection of Hardy's concurrent interest in set theory.

Hardy's *Integration* tract and his *Orders of Infinity* tract both owe something to Liouville's classification of transcendents[9]. In a 1912 paper titled *Properties of Logarithmico-Exponential Functions*, Hardy stated:

> in my opinion, the main interest of Liouville's classification lies in its application to two special problems – indefinite integration in finite terms on one hand, orders of infinity on the other. [48, p. 59].[10]

[7]For example, see [98] which states "A function built up of a finite combination of constant functions, field operations (addition, multiplication, division, and root extractions – the elementary operations) - and algebraic, exponential, and logarithmic functions and their inverses under repeated compositions"

[8]Following Liouville (1837, 1838, 1839), Watson (1966, p. 111) defines the elementary transcendental function as $l_1(z) \equiv l(z) \equiv \ln(z), e_1(z) \equiv e(z) \equiv e^z, \zeta_1 f(z) \equiv \zeta f(z) \equiv \int f(z)dz$ and lets $l_2 \equiv l(l(z))$, etc.

[9]After some criticism of his classification scheme, Liouville published a very thorough paper on it in 1837 titled *Mémoire sur la classification des transcendantes et sur l'impossibilité d'exprimer les raciness de certaines équations en fonction finie explicite des coefficients* in Liouville's Journal 2 (1838), 523-547

[10]Du Bois-Reymond, in a paper published in 1877, used his *Infinitärcalcül* to investigate the behavior for

It is important to note that functions defined by transcendental equations are not elementary. For example $y = x\ln(y)$ is not elementary, meaning that y cannot be explicitly expressed in a finite number of terms of x. Liouville showed this in a paper published in the *Journal de Mathématiques.*[11]

Hardy felt that the theorems stating that e^x and $\ln(x)$ are not algebraical, that $\ln(x)$ is not expressible as e^y where $y = f(x)$, f algebraic, and that transcendental functions actually exist to be the foundation of his analysis and, as such, these should be proved by writers on elementary analysis.

2.2.3 Integration of Elementary Functions – A Summary

For the purposes of this text, the fundamental question about integration is: given an elementary function $f(x)$ to first determine if its integral is elementary and, if so, to find it. A precise statement of the integration question requires that both the mode of expression to be used and the class of allowable functions be specified. Hardy specified that we are considering closed form solutions using elementary functions. Hardy stated that "complete answers to these questions have not and probably never will be given". [41, p.7]

This was the case in 1905 but by 1969[12], Robert H. Risch had discovered a decision procedure that determined whether or not a given integral is elementary. Further, if the integral is elementary, a closed-form formula is determined. Hardy's intuition was thus shown to be incorrect; however, his insistence on a methodical approach to integration was prescient since, with the advent of computer algebra systems, an algorithmic approach to integration is absolutely necessary. Although currently the interest in methodical, algorithmic integration techniques stems from computing applications, Hardy may have taken this approach in a search for a "complete" theory of integration – one that would methodically find the solution to all integration problems regarding elementary functions that could be asked.

The Risch algorithm builds a series of logarithmic and exponential functions with algebraic extensions as it solves the integration problem. As of 2008, the Risch algorithm has not been fully implemented in any computer algebra system but it forms the backbone of

large values of λ of Dirichlet integrals $\int_0^{\xi} \frac{sin\lambda x}{x} f(x)dx$ when $f(x)$ contains factors such as $\cos\psi(\frac{1}{x}$, where ψ is a function of the exponentio-logarithmic scale and $\psi(u) \to \infty$ as $u \to \infty$. In 1909, 4 and 1909, 6 Hardy re-examines some of Du Bois-Reymond's results, providing simpler proofs or obtaining more general theorems [92, p.3]

[11] Journal de Mathématiques, t. III p. 523.

[12] See Robert H. Risch's articles [79, 80, 78]

the method used by computer systems once simple and quick heuristic methods have been tried and found lacking.

Many introductory calculus courses teach and MAPLE – the computer algebra system that will be used here as an example – uses a number of heuristic approaches to the integration problem including substitution, trigonometric substitution, integration by parts and partial fractions. This is followed by a table lookup for approximately thirty-five simple functions including the trigonometric functions. Then a technique called the "derivative-divides" method is used which is a form of substitution aimed at determining whether or not the integrand has a composite function structure. If the integrand has a composite structure, an attempt is made to divide the integrand by the derivative of the composite function, $f(x)$, and produce an integrand that is independent of x after the substitution $u = f(x)$.

Finally, if all of the above methods fail to produce a result, the Risch algorithm for integrating elementary functions is used. The Risch algorithm is not directly employed at the outset for reasons of time. It is desirable to solve simple problems quickly and a "surprisingly large percentage of integral problems" [25, p. 474] are solved by MAPLE's heuristic methods.

Geddes et al [25, p. 512] in their text *Algorithms for Computer Algebra* provide a brief history of the problem of integration in finite terms. Here is a quote from their work:

> The problem of integration in finite terms (or integration in closed form) has a long history. It was studied extensively about 150 years ago by the French mathematician Joseph Liouville. The contribution of another nineteenth-century mathematician, Charles Hermite, to the case of rational function integration is reflected in computational methods used today. For the case of transcendental elementary functions, apart from the sketch of an integration algorithm presented in G.H. Hardy's 1928 treatise, the constructive (computational, algebraic) approach to the problem received little attention beyond Liouville's work until the 1940's when J.F. Ritt[13] started to develop the topic of Differential Algebra. With the advent of computer languages for symbolic mathematical computation, there has been renewed interest in the topic since 1960 and the mathematics of the indefinite integration problem has evolved significantly. The modern era takes as its starting point the fundamental work by Risch in 1968, where a complete

[13] J.F. Ritt, *Integration in Finite Terms*, Amer. Math Monthly, 79 pp. 963-927 (1972)

decision procedure was described for the first time.

In the above quote, Hardy's *Integration* is not mentioned by name but the second edition of *Integration* was published in 1916 and then reprinted in 1928. Since there are no other papers of Hardy's published in 1928 concerning integration in finite terms, it is safe to assume that Geddes et al are referring to *Integration* in the above quote.

A much longer, more thorough history of integration in finite terms is given by Lützen [69, p. 351-422], and by Kasper [65]. What follows is a summary of these works, highlighting the role of Hardy. Hardy's *Integration* was the first work after Liouville's to approach integration in finite terms in a general way.

Both Kasper and Lützen credit Liouville with founding the theory of integration in finite elementary terms – a theory that was created in a series of papers between 1833 and 1841. Liouville used his theory to prove that elliptic integrals cannot be represented by explicit finite elementary expressions. This was the first case of a proven, non-elementary integral. In 1923, Joseph Fels Ritt (1893-1951) proved elliptic integrals cannot be expressed as implicit finite functions.

Ritt was an American mathematician who spent most of his professional life at Columbia University. His study of elementary function integration culminated in a book, published in 1948, titled *Integration in Finite Terms; Liouville's Theory of Elementary Methods*. This book was reviewed by Laurence Young, son of William and Grace Chisholm Young, contemporaries of Hardy, who said "This little book is the first treatise to deal with the theory of elementary integrals according to Abel and Liouville".[14] It is worth noting that Ritt "took an unusual interest in reading the great mathematical works of his predecessors. He drew substantial inspiration from careful study of classic texts, especially those by the great figures of the eighteenth and nineteenth centuries. Among his heroes were Niels Henrik Abel, Augustin Louis Cauchy, David Hilbert, Carl G. J. Jacobi, Joseph-Louis Lagrange, the marquis Pierre-Simon de Laplace, Joseph Liouville and Jules Henri Poincaré"[15] It may be that the work of Hardy went unnoticed by Ritt .

During the time period between Liouville and Ritt, developments in integration in finite terms took place in Russia. In the second half of the nineteenth century, Ostrogradsky, Youschkevitch, and Chebyshev all published on this topic.

---

[14]http://www-groups.dcs.st-and.ac.uk/ history/Biographies/Ritt.html

[15]J W Dauben, Joseph Fels Ritt, American National Biography 18 (Oxford, 1999), 550-551.

Early in the twentieth century, Russian mathematician D.D. Mordukhai-Boltovskoi (1876-1952) wrote and contributed much to the Liouville theory [65, p. 198]. And in 1946, A. Ostrowski

> broadened Liouville's general theory (of 1835) and extended it to the wider class of meromorphic functions (single-valued and analytic, except possibly for poles) in regions of the complex plane. [65, p. 199].

This was done by using field extensions and was the genesis of the algebraic approach that led to the final solution of the problem of indefinite integration.

Ritt's 1948 monograph summarized the results to that date and led to the creation of the modern theory. "By his own account, Risch was introduced to the subject of finite integration by Ritt's 1948 monograph" [65, p. 200].

Mirroring Hardy's comment that the student of calculus does not view integration coherently and would benefit from seeing integration presented methodically, Risch

> makes the interesting suggestion that some features of his algorithm are suitable for presentation to calculus students. No calculus text at present provides this material, an omission that not only leaves the story of finite elementary integration incomplete, but deprives the calculus student of some valuable insights. [65, p 201]

Lützen considers the Ritt monograph to be the summary of Liouville's work that Liouville never wrote and claims that Ritt, like Hardy, stressed the functional-analytic aspects of the theory.

2.2.4 Integration of Rational Functions

Hardy's description of integration of rational functions is short (7 pages) and straightforward. Using results from algebra[16] that state that any polynomial of the form $Q(x) = x^n + b_1 x^{n-1} + \ldots + b_n$ can be expressed as a product of n linear factors of the form $(x - a)$, where a is a root, and that any rational function $R(x) = \frac{P(x)}{Q(x)}$ can be written in the form

$$A_0 x^p + A_1 x^{p-1} + \ldots + A_p + \sum_{s=1}^{r} \{ \frac{\beta_1}{(x-\alpha_s)} + \frac{\beta_2}{(x-\alpha_s)^2} + \ldots + \frac{\beta_m}{(x-\alpha_s)^m} \},$$

[16] Hardy references Chrystal's *Algebra*.

the integral in the general case is shown to be

$$\int R(x)dx = A_0\frac{x^{p+1}}{p+1} + A_1\frac{x^p}{p} + \ldots + A_p x + C + \sum_{s=1}^{r}\{\beta_1 \log(x-\alpha_s) - \frac{\beta_2}{x-\alpha_s} - \ldots - \frac{\beta_s}{(m-1)(x-\alpha_s)^{m-1}}\}$$

The solution of the general case allows Hardy to claim that the integral of any rational function is an elementary function. Integration of rational functions then is an exercise in algebra. MAPLE, for example, can solve this problem using polynomial division with remainder, GCD computation, polynomial factorization and equation solving.

However, there remains the problem of determining the constants in the above result, which cannot be expressed explicitly as functions of the constants of the integrand. Hardy describes Hermite's[17] method – a method that allows the integration of the rational part of the integral without needing to factor the denominator of the rational function integrand. Hermite's method reduces the problem of $\int \frac{P(x)}{Q(x)}dx$ to $\frac{C(x)}{D(x)} + \int \frac{A(x)}{B(x)}dx$ using only polynomial operations where the degree of $A(x)$ is less than the degree of $B(x)$ and $B(x)$ is monic and square-free. The remaining integral is solved using logarithms. Hardy takes two pages to go through an example using Hermite's method. He concludes that, for rational function integration, the complete integral can be found if it is possible to find the roots of $Q(x) = 0$ and that it is always possible to find the rational part of the integral.

Hardy was, however, unable to obtain a general method and so he describes "the maximum of information which can be obtained about the logarithmic part of the integral in the general case in which the factors of the denominator cannot be determined explicitly" [41, p.16]. Some of the techniques for this include partial factoring of the denominator until $Q(x)$ is irreducible by the adjunction of any algebraic irrationality (that is, with coefficents in $\mathbb{Q}(\sqrt{p})$), and the integration of rational functions when the result is only logarithmic. The above suggestions, however, only solve specific cases that depend on the value of the integrand.

The final part of Hardy's discussion of the logarithmic part of the integration result have been superseded by new techniques for integrating $\int \frac{A(x)}{B(x)}dx$ where the degree of $A(x)$ is less than the degree of $B(x)$ and $B(x)$ is monic and square-free. This method is called the

[17] Hermite, Charles. *Sur l'Integration des Fractions Rationelles*, Nouvelles Annales de Mathematiques, pp. 145-148 (1872)

Rothstein-Trager method discovered independently by M. Rothstein [82] and B. Trager[94] in 1976. It solves the problem completely with the minimum number of algebraic extensions.

2.2.5 Integration of Algebraic Functions

The integration of algebraic functions, explicit or implicit, is the longest section of *Integration* at 24 pages. It is far more difficult and Hardy gives only "a brief account of the most important results and of the most obvious of their applications" [41, p. 18].

Hardy first discusses integrands that can be reduced to rational functions by a substitution, either real or imaginary. He notes that integrals of functions of the type $\int R(x, y)dx$ in which the x and y are connected by a variable t, such that $x = R_1(t), y = R_2(t)$ with R_1, R_2 rational, can be evaluated in finite terms by means of elementary functions. A variety of specific substitutions are discussed for different integrands. Such a parameterization defines a curve, and curves of this type are called unicursal.

Hardy then defines the deficiency of a curve as the number of possible double points a curve can have minus the number of double points that it actually has.[18] If a curve has deficiency zero, then it is unicursal. For curves of deficiency zero, a general procedure is described to determine the substitution required in order to reduce the integrand to a rational function in the substituted variable.

When the curve is not unicursal, the integral is in general not an elementary function. If the deficiency is one, then the integral will be expressible in terms of elementary functions and elliptic integrals. As the deficiency increases, the complexity of required new transcendents increases. However, there are many particular cases of curves with non-zero deficiency that can be integrated in terms of elementary functions. Hardy attempts to classify which non-unicursal curves are expressible in terms of elementary functions but he is well aware of limitations saying, "It will be as well to say at once that this problem has not been completely solved" [41, p. 30].

A modern algorithmic treatment (the Risch algorithm) of this topic [25, p.511-573], handles the explicit algebraic function integration and transcendental function integration together under elementary function integration with the special case of rational function integration removed. Here the chief insight is that all of the algebraic or transcendental functions are expressed using exponentials and logarithms (of complex numbers if necessary)

[18]A double point is a point that is traced out twice as a closed curve is traversed.

so that special notation for trigonometric, inverse trigonometric, hyperbolic and inverse hyperbolic functions is discarded. This allows the finite decision procedure to be invoked and the integral to be expressed in exp-log notation. It is difficult to translate back from exp-log notation in order express results in the more familiar form – for this reason, a heuristic approach is applied first in an attempt to solve the integration problem with the more familiar notation.

Implicit algebraic functions alone can be handled by an algorithm developed by Trager [95] in 1984 and the case of an integrand consisting of mixed transcendental and algebraic integrands was completely solved in 1987 by Bronstein. [7]

2.2.6 Integration of Transcendental Functions

When defining the class of elementary functions early in the book, Hardy was careful to note that there is no general theory of transcendental functions in the way that there is a general theory of algebraic functions. He said

> The theory of integration of transcendental functions is naturally much less complete than that of the integration of rational or even of algebraical functions. It is obvious from the nature of the case that this must be so, as there is no general theorem concerning transcendental functions which in any way corresponds to the theorem that any combination of algebraical functions, explicit or implicit, may be regarded as a simple algebraical function, the root of an equation of a simple standard type [41, p.42]

By this he meant that there is no transcendental equivalent to the implicit definition of algebraic functions. So, for the example of $y = x\ln(y)$ given earlier, y is incapable of a finite explicit expression in terms of x.

The last chapter of the book, on the integration of elementary transcendental functions, is short – just shy of ten pages. Hardy claims that the theory of integration of transcendental functions is much less complete precisely because there is no general theory of transcendental functions. In fact there is, for Hardy, no general theory. The theory of integration of elementary transcendental functions is reduced to

> an enumeration of the few cases in which the integral may be transformed by an appropriate substitution into an integral of a rational or algebraical function.

> These few cases are however of immense importance in the applications of the general theory of integration [41, p.42].

Here Hardy, in a sense, contradicts his original aim – to show that integration is not a series of clever tricks.

This final section is just clever tricks that allow a very few of the possible integration questions that can be asked to be solved. Among these are integration of functions of the form $F(e^{ax}, e^{bx}, \ldots, e^{kx})$ where F is an algebraic function and the constants are commensurable – this includes the particular case of sin, cos, sinh, and cosh. Functions of the type $P(x, e^{ax}, e^{bx}, \ldots, e^{kx})$ where P is a polynomial and the constants are any numbers are also given a general treatment. But apart from these two types of functions, there are "no really general classes of transcendental functions which we can *always* integrate in finite terms" [41, p.45].

In the case of the innumerable other possible integrals where it is not possible to give a general theory, Hardy seeks to apply a systematic reduction theory that will split up the integral into a part that can be integrated and a part that cannot. The latter part is minimized and proved to be incapable of further reduction.

The prime example here is an integral of the form $\int e^x R(x)dx$ where $R(x)$ is a rational function of x. All integrals of this form can be made to depend on known functions and on the single transcendent $\int \frac{e^x}{x} dx$ normally denoted $li(e^x)$. Hardy justified this result.

The concluding page compares these new types of transcendent functions with those that arise from the integration of algebraic functions, notably elliptic integrals, and states that they are often of great interest and importance. For example $li(e^x)$ is important in describing the distribution of prime numbers. Often, these transcendents may be expressed using definite integrals or by means of an infinite series.

Chapter 3

G.H. Hardy and Set Theory in Britain, 1900-1910

In 1910, David Hilbert who was, at that time, arguably the world's most influential mathematician, claimed that set theory was

> that mathematical discipline which today occupies an outstanding role in our science, and radiates its powerful influence into all branches of mathematics. [59].

In this chapter I examine how this branch of mathematics, which originated in the work of Georg Cantor (1845-1918) and Richard Dedekind (1831-1916)[1] was introduced to English speaking mathematicians in Britain.

Specifically, I want to consider the earliest introduction of set theory into Britain during the years 1900 to 1910 looking mainly at the work of Hardy, Philip Jourdain (1879-1919), Bertrand Russell (1872-1970), and, to a lesser extent, William Henry (1863-1942) and Grace Chisholm Young (1868-1944). All of these people wrote on set theory in the earliest part of the 20th century in different ways and with differing impacts.

The Youngs, a married couple who frequently collaborated, wrote a textbook called *The Theory of Sets of Points*, published in 1906. Hardy wrote a series of 5 papers on set theory which were published between 1904 and 1910. Russell wrote *Principles of Mathematics* first

[1]Some historical accounts of set theory treat it as though it were the brainchild of Cantor alone whereas other accounts emphasize the roles played by others, most notably Dedekind and Weierstrass. For a discussion of this issue, see [21, p. xv-xvii].

published in 1903, in which the philosophical point of view of set theory is discussed. And the first volume of Russell and Whitehead's *Principia Mathematica* appeared in 1910.

During the time interval under consideration, Jourdain maintained a large correspondence with both Hardy and Russell and, in 1915, translated and wrote an 85-page introduction to Cantor's *Contributions to the Founding of the Theory of Transfinite Numbers*. How was this writing received and how did it influence the work of others?

Hardy went on to write *A Course of Pure Mathematics* in 1908, an enormously influential textbook on analysis which is discussed in chapter 4. For the most part, he chose not to introduce set theory in *A Course of Pure Mathematics* but instead[2] took special care to introduce the theory of real numbers developed by Dedekind. Ernest William Hobson (1856-1933) actively engaged with Hardy's papers, writing critically of Hardy's work and refuting some of Cantor's statements in 1905. Then, in 1907, Hobson wrote a book titled *The Theory of Functions of a Real Variable and the Theory of Fourier Series* which contained "copious references to the literature of set theory" [103, p. viii].

I also discuss from where and from whom these authors informed themselves of set theory and point to the common characteristic that all were fluent in German – a "skill [fluency in reading French and German] which Whittaker claimed was at that time [1893] almost unknown amongst mathematical lecturers, and which they showed little inclination to learn." [58, p.150]

Since it is mainly the work of Cantor and Dedekind that was brought to Britain at this time, I outline some features of this work in order to provide background information. We will see that the work of Schönflies is also cited by the British authors – this is work that summarizes Cantor's developments.

3.1 The Development of Set Theory in Late 19th Century Germany

3.1.1 Short Summary of Cantorian Set Theory

This summary is based mainly on four sources: the historical work of Ferreirós [21], and Dauben [17, 18] and the introduction that Jourdain provides in his translation of Cantor's *Contributions to the Founding of the Theory of Numbers* [64].

[2]beginning with the 2^{nd} edition

In 1869, Cantor, at the suggestion of Eduard Heine (1821-1881) began to consider if and under what conditions Fourier series representations of functions are unique. By 1872, he was able to prove that they were, first for trigonometric series that were everywhere convergent, then for series with a finite number of exceptions and finally for series even with an infinite number of exceptions if the exceptions were distributed in particular manner. In order to form these proofs and to describe the distribution of an infinite number of exceptions, Cantor developed the ideas of point-sets and their derived sets. In Cantor's words of November 1871:

> some explanations, or rather some simple indications, intended to put in a full light the different manners in which numerical magnitudes, in number finite or infinite, can behave [64, p. 25]

This in turn required him to develop a theory of real numbers, which he built from the rational numbers.

In developing a theory of real numbers, Cantor relied on the work of Karl Weierstrass (1815-1897) saying that:

> I believe, a propos of Weierstrass's theory, that this logical error[3] which was first avoided by Weierstrass, escaped notice almost universally in earlier time, and was not noticed on the grounds that it is one of the rare cases in which actual errors can lead to none of the more important mistakes in calculation [64, p. 18]

Jourdain describes Cantor's theory of irrational numbers as a "happy modification of Weierstrass" [64, p. 26]. Weierstrass developed his theory of real numbers, defined as the limits of convergent series, beginning in 1859, in order to make his lectures in analytic function theory systematic.

Cantor built the real numbers from the set A of the rationals, by associating real numbers with sequences of rational numbers that were constructed such that after some number of terms in the sequence, the difference between any two terms of the sequence remained arbitrarily small. All such numbers constructed this way formed a set of new numbers, B. Beyond this, a new domain was defined in a similar manner using sequences of numbers from set B, and similarly on to further domains C through L. At this point it was by use

[3] There was a vicious circle in the definition of a real number that was considered to be the limit of a convergent sequence since the limit itself involved the prior assumption of the existence of a real number.

of an axiom[4], "to every real number a definite point of the straight line corresponds, whose ordinate is equal to that number" [18, p.183] that B was associated with what Cantor termed the linear continuum.

Cantor then defined a limit-point as follows:

> Given a point set P, if an infinite numbers of points of the set P lie within every neighbourhood, however small, of a point p, then p is said to be a 'limit-point' of the set P. [18, p. 183 quoting Cantor]

The above definition is used to differentiate between point-sets of the first and second species given that the elements of the $(k+1)^{st}$ set are the limit points of the k^{th} set. Point-sets of the first species have the property that the n^{th}, with n finite, derived limit-points set is the empty set. Point-sets with nonempty n^{th} derived sets, for all finite n, are point-sets of the second species. The proof of the uniqueness of trigonometric function representation used point-sets of the first species whereas transfinite numbers were developed from point-sets of the second species.

Ferreirós [21] divides Cantor's work into four phases. The work on point-sets between 1870 and 1872 just described was the first phase. The second phase of Cantor's work, which lasted from 1873 to 1878, consisted of the work on sets of infinite cardinality. That the rationals were dense but don't form a linear continuum led Cantor to suspect that there were more irrationals than rationals. He corresponded with Dedekind in 1873 about whether it would be possible to have a one-to-one correspondence between the real numbers and the integers and then found in 1874, that it was not possible. Cantor proved this by contradiction. He assumed that a one-to-one correspondence was possible and then constructed a one-to-one correspondence between the real numbers and the integers such that at least one number was shown to be left out. Hence, the real numbers are non-denumerable - not able to be counted by a one-to-one correspondence with the infinite set of integers. Immediately then, given that the algebraic numbers are denumerable, transcendental numbers on any given interval are infinite. This argument is a now standard method for proving the existence of transcendental numbers. For example, it is used in Gelfond's *Transcendental and Algebraic Numbers* [26, p. 2].

Cantor then tried to generalize further, looking for distinct powers of infinity greater

[4]This axiom is now called the Cantor-Dedekind axiom - "the points on a line can be put into a one-to-one correspondence with the real numbers" [97].

than the power of the real numbers. Again, corresponding with Dedekind, in 1874, he asked whether or not a square could be mapped one-to-one onto a line. He answered his own question in the affirmative in 1877, but incredulously, saying "je le vois, mais je ne le crois pas" [21, p. 171]. Dedekind found a minor error in the mapping Cantor first used to show this, which Cantor subsequently fixed. This realization allowed Cantor to study continuity in the linear continuum of the real numbers.

The third period, covering the interval from 1879 to 1884 was "guided by the core objective of proving the Continuum Hypothesis"[5] [21, p. 257]. In 1880, Cantor introduced transfinite numbers (originally just symbols but soon to be named transfinite ordinal numbers) from derived sets of the second species and showed that there was an unending sequence of larger and larger transfinite symbols. It was here that Cantor moved from a potential to an actual infinity:

> I was logically forced, almost against my will, because in opposition to traditions which had become valued by me in the course of scientific researches extending over many years, to the thought of considering the infinitely great, not merely in the form of the unlimitedly increasing, and in the form, closely connected with this, of convergent infinite series, but also to fix it mathematically by numbers in the definite form of a completed infinite [64, p. 53].

Arithmetic with transfinite ordinals is noncommutative and unlike finite sets where ordinal and cardinal numbers coincide, Cantor said:

> The conception of number which, *in finito*, has only the background of enumeral [Jourdain's invented word used to translate "Anzahl", Dauben uses "numberings"] splits, in a manner of speaking, when we raise ourselves to the infinite, into the two conceptions of power... and enumeral ... ; and, when I again descend to the finite, I see just as clearly and beautifully these two conceptions again unite to form that of the finite integer [64, p. 52]

It is at this time that Cantor invented his famous ternary set which is everywhere dense but

[5] the proposal that there is no cardinal number between the cardinal number of the integers and the cardinal number of the reals. Symbolically this can be written as $c = \aleph_1$. This remains unproven and it has been shown that (Gödel/Cohen) no contraction will arise in Zermelo-Fraenkel set theory if either the Continuum Hypothesis or its negation is added as an axiom. Set theoreticians generally feel that the continuum hypothesis should be considered false [16, p. 282].

has measure zero.[6] By the end of this period, Cantor was having increasing mental health problems, had a falling out with Gösta Mittag-Leffler[7], and felt persecuted by Kronecker[8].

The fourth period, from 1885 to the end of Cantor's life, began with proof of the continuum hypothesis which was subsequently withdrawn. In an unpublished paper of 1885, Cantor equated pure mathematics with pure set theory and his interest shifted from point-sets to abstract set theory where the notions of cardinality and order are foremost. In the late 1880's, Cantor constructed the diagonalization proof, showing that the set of all subsets is of a greater power than the parent set, which Hardy responded to. In 1895, Cantor's last major publication introduced the alephs for the cardinal numbers. This widely read work was translated immediately into French and Italian, and defended by younger mathematicians. Perhaps because Cantor realized how directly confrontational his transfinite numbers were with regard to widely held beliefs about numbers and infinite, he was led to detailed philosophical and theological discussions about the meaning of his work.

3.1.2 Short Summary of Dedekind's contributions

To determine when a mathematician began using set theory, Ferreirós [21] has pointed to the necessity of clarifying exactly what is meant by set theory. Set theory can be considered in three different and somewhat overlapping ways: formally, as a foundation for mathematics, or less formally, as a language for mathematics as in the "set-theoretical approach", or as a separate branch of mathematics typically called abstract set theory. However, in all three cases, serious study of set theory appears to require the conception of the actual infinite. It wasn't until 1885 that Cantor began to distinguish abstract set theory from the theory of point-sets [21, p. xix] by which time the set-theoretical approach had been established in the work of Dedekind.

As with Weierstrass, Dedekind's teaching motivated his foundational work. His lectures on differential calculus made him keenly aware of "the lack of a really scientific foundation

[6]This set is sometimes called the Cantor set, the Cantor comb or the no middle third set. It is formed by taking the interval [0,1] and removing the middle third. Then the middle third is removed from the two remaining pieces and the process continues indefinitely. "The Cantor set is the only totally disconnected, perfect, compact metric space up to a homeomorphism" [3]

[7]Mittag-Leffler was responsible for translating all of Cantor's work into French and for providing a publication outlet for Cantor's work via *Acta Mathematica*.

[8]Leopold Kronecker (1823-1891), a Berlin mathematician who believed that mathematics was about finite numbers only and only finite numbers of operations were permissible. He was the editor of *Crelle's Journal* and tried to prevent publication of Cantor's work because he was mathematically opposed to it.

for arithmetic" [8, p. 221]. He wanted an arithmetical foundation for calculus founded on the operations of natural numbers. Like Weierstrass, he sought and found a method of constructing the real numbers from the rationals.[9] The idea of a set was absent in Dedekind's work prior to 1855[10] but present in the work on irrational numbers where sets were used to define new numbers, and in his algebraic work from 1856-58 [21, p. 77].

In this early work, Dedekind used the words 'System', 'Klasse' and 'Complex' for sets and freely makes use of infinite classes, "tracing an analogy between these infinite classes and the natural numbers, which were most concrete objects for a traditional mathematician" [21, p. 88]. During the 1870's and 1880's, Dedekind began using a set-theoretical approach in his writing. For example, in a work on Galois theory, where he used the word substitution as we would now use mapping, Dedekind said:

> By a substitution one understands, in general, any process by which certain elements $a, b, c \ldots$ are transformed into others $a', b', c' \ldots$, or are replaced by these; in what follows we shall consider only those substitutions in which the complex of replacing elements a', b', c' is identical with that of the replaced a, b, c. [21, p. 88]

Or, in 1879, Dedekind said:

> Upon this mental faculty of comparing a thing ω with a thing ω', or relating ω with ω', or making ω' correspond to ω, without which it is not at all possible to think, rests also the entire science of numbers [21, p. 89].

When Cantor writes of point-sets between 1879 and 1884, he uses Dedekind's terminology introduced in Dedekind's algebraic work. Cantor used divisor, greatest common divisor, and multiple for set inclusion, intersection and union. These words had been introduced by Dedekind to describe operations on a field, which was understood to be

> every set of infinitely many real or complex numbers, which is so closed and complete in itself, that addition, subtraction, multiplication, and division of any of those numbers yields always a number of the same set ... we call a field A

[9] Dedekind worked on this in 1858 but published it as *Stetigkeit und Irrationalen Zahlen* in 1872. Ferreirós claims that for Dedekind "pure mathematics was the science of number in all its extension and derivations, number systems being more basic than any possible abstract structure" [21, p. 81]

[10] For example, there is no mention of sets in his Habilitationsvortrag in 1854.

> divisor of field M, and this a multiple of that, if all the numbers contained in A are also found in M. [21, p. 92]

In 1888, Dedekind published his most important foundational work, *Was sind und was sollen die Zahlen*.[11] This work was started in 1872, revised in 1878, and brought into its final form in 1887. On the cover of the original German edition is the Greek phrase that translates to "man eternally arithmetizes". This is indicative of both the importance that Dedekind placed on the concept of number and of his philosophical stance that numbers are the creation of the human mind. From Plato's position of "God eternally geometrizes" to Gauss's "God eternally arithmetizes" to Dedekind's statement, a shift from geometry to arithmetic and from God to man is seen.

Here is a long quote from Dedekind that nicely outlines his philosophical position and demonstrates the fundamental position of the concept of a set:

> numbers are free creations of the human mind; they serve as a means of apprehending more easily and more sharply the difference of things. It is only through the purely logical process of building up the science of numbers and by thus acquiring the continuous number-domain that we are prepared accurately to investigate our notions of space and time by bringing them into relation with this number-domain created in our mind. If we scrutinize closely what is done in counting an aggregate or number of things, we are led to consider the ability of the mind to relate things to things, to let a thing correspond to a thing, or to represent a thing by a thing, an ability without which no thinking is possible. Upon this unique and therefore absolutely indispensable foundation, ... , must, in my judgment, the whole science of numbers be established [19, p. 32]

The work succinctly outlined the basic notions of set ('thing' or 'system'), subset ('part' or 'proper part'), union ('compounded system') and intersection ('community of systems'). This was followed by a definition of a mapping which appears quite modern.

> by a transformation [Abbildung[12]] ϕ of a system S we understand a law according to which to every determinate element s of S there belongs a determinate thing

[11] The Nature and Meaning of Numbers [19]

[12] This word can also be translated as 'representation' which then may better capture the idea of representing one thing by another (see [21, p. 229].)

> which is called the transform of s and denoted by $\phi(s)$; we say also that $\phi(s)$ corresponds to the element s, the $\phi(s)$ results or is produced from s by the transformation ϕ. [19, p. 50]

Given a particular transformation, ϕ, the question naturally arises as to whether the transformation of a system S, $\phi(S)$, into a system S', results in an S' such that S' is a subset of S or not. If S' is a subset of S, then S is a chain, with respect to the given transformation [19, p. 56]. Two pages later, this notion is generalized and a "chain of System A" [ibid, p.58] is defined as the intersection of all of the chains of A. Because this theory of chains is built solely on the concept of a transformation, Ferreirós feels that this is Dedekind's "most original contribution to abstract set theory" [21, p. 230]. It is this concept that allows construction of the set of natural numbers starting from 1 and a properly defined mapping[13].

Further, chains are then used to prove the mathematical validity of induction whether based on the natural numbers or on an arbitrarily defined chain. Dedekind stated and proved the

> *Theorem of complete induction.* In order to show that the chain A_0 is part of any system Σ - be this latter part of S or not – it is sufficient to show,
> ρ. that $A_0 \subset \Sigma$[14]
> σ. that the transform of every common element of A_0 and Σ is likewise an element of Σ. [19, p. 60-61].

The definition of an infinite set as a set which is "similar to a proper part of itself" [19, p. 63] followed. This is, in modern terms, a set which can be mapped one-to-one onto a subset of itself. Ferreirós claims that this idea was formed before any of Cantor's work on set theory and "can be seen as the first noteworthy and influential attempt to elaborate an abstract theory of finite and infinite sets" [21, p. 109]. Dedekind, in a footnote to the definition of an infinite set, pointed to the importance of this definition and to whom and when he communicated it.

> In this form I submitted the definition of the infinite which forms *the core of my whole investigation*[15] in September, 1882, to G. Cantor and several years earlier

[13] [21, p. 230-231]
[14] $\subset$ is not the symbol Dedekind uses for subset, he uses a $\ni$.
[15] my emphasis

> to Schwarz and Weber. All other attempts that have come to my knowledge to distinguish the infinite from the finite seem to me to have met with so little success that I think I may be permitted to forego any criticism of them [19, p. 63]

Dedekind, unlike Cantor, did not explicitly develop point-sets and there is no development of a transfinite number - instead set theory is used in a foundational manner and, importantly, as a language to describe other mathematics. Beginning in the 1850's, a set-theoretic approach enters into Dedekind's mathematics with sets and mappings becoming "the central notions for Dedekind's understanding of arithmetic, algebraic number theory, algebra and also, one may safely conjecture, analysis". [21, p. 90]

3.2 G.H. Hardy and Set Theory

It is during the early part of his career that Hardy published all five of his set theory papers. Richard Rankin (1915-2001), a member of the editorial board for Hardy's collected papers, wrote a brief introduction to Hardy's miscellaneous papers, which include the set theory papers. He said that "Hardy's direct contribution to set theory was not great" [77, p. 417]. But he also pointed out that Hardy, unlike many professional mathematicians at that time, "fully accepted Cantor's theories as a valid contribution to mathematics" [ibid].

In the winter of 1902-03, Hardy attended lectures on set theory given by Albert North Whitehead (1861-1947). He referred to these lectures in a footnote while discussing the construction of an aggregate[16] whose cardinal[17] is 2^{α_1} :

> Mr. Whitehead worked out some of the most interesting properties of such aggregates in his lectures on the application of symbolic logic to the theory of aggregates. [40, p. 94]

The first edition of Hardy's *A Course of Pure Mathematics*, in 1908, assumed the existence of the real numbers. However, in a very positive ("his book seems to us the best elementary treatise on the Calculus in the English language" [43, p. 308]) book review of *A First Course in the Differential and Integral Calculus* by W.H. Osgood in 1907, Hardy said

[16] Aggregate is an early English translation of the German work menge - set is used now.

[17] Hardy used a lower case Greek alpha, rather than the now standard aleph, to denote a cardinal number. Now we would see this as $2^{\aleph_1}$.

> we should have liked to have seen included at anyrate [sic] a short and popular sketch of Dedekind's theory of number. It is not very difficult to explain in general terms and enables a good many gaps to be filled in, notably in connection with the 'Fundamental Principle' that an increasing function approaches a limit or tends to infinity. [43, p. 308]

Similarly, in another book review in 1909, this time of *The Thirteen Books of Euclid's Elements* by T.L Heath, in a context where it may not be expected, Hardy commented again on the work of Cantor, Weierstrass and Dedekind. When talking about the theory of proportion from Book V, Hardy said:

> Rendered into a modern and a more convenient notation it may still stand, almost word for word, as the one and only possible theory of proportion, as irreproachable and in as little danger of supersession as when Barrow declared that "there is nothing in the whole body of the elements of a more subtile [sic] invention, nothing more solidly established and more accurately handled than the theory of proportionals." How subtle and accurate it is is clearer now, to the mathematician familiar with the work of Weierstrass, Cantor and Dedekind, than it can possibly have been to Barrow or even to De Morgan. There is, as Dr. Heath points out, "an exact correspondence, almost coincidence between Euclid's definition of equal ratios and the modern theory of irrationals due to Dedekind." [45, p .857][18]

By the second edition[19] of *A Course of Pure Mathematics* in 1914, Hardy used Dedekind's method of constructing the real numbers from sections of the rationals using cuts and introduced the theorem of Weierstrass[20] on accumulation points. The two quotes below from

[18]It is, of course, possible that Hardy and or Heath are reading into Euclid the more modern theories and that these ideas are not really present in the Greek work. Sabetai Unguru cautions against this in his article *On the Need to Rewrite the History of Greek Mathematics* [96], where he refutes the claim that the Greeks were, in essence, doing algebra when their work is translated into modern notation.

[19]In a 1916 article for *Mind*, Jourdain reviewed and compared the first and second editions of *A Course of Pure Mathematics*.

[20]This is now known as the Bolzano-Weierstrass theorem and Hardy stated it thus: "If a set S contains infinitely many points, and is entirely situated in an interval (α, β), then at least one point of the interval is a point of accumulation of S" [56, p. 32]. By accumulation point, Hardy meant what Cantor termed a limit point.

the second edition of the book serve to show how Hardy's perception of the importance of set theory had increased in the six years between the two editions.

> The idea of a 'section', first brought into prominence in Dedekind's famous pamphlet *Stetigkeit und irrationale Zahlen*[21] , is one which must be grasped by every reader of this book. [56, p. 28]

> The general theory of sets of points is of the utmost interest and importance in the higher branches of analysis; but it is for the most part too difficult to be included in a book such as this. [56, p. 31]

It is interesting that when Hardy chose to construct the real numbers in the second edition of *A Course of Pure Mathematics*, he chose the method of Dedekind. He had three possibilities to choose from: Dedekind's method using cuts, Cantor's method using Cauchy sequences, or Weierstrass's method using infinite series. Since Hardy introduced the construction of real numbers very early on in his textbook (by page 30), I believe that Dedekind's method was the only possibility as Hardy had yet to talk about sequences (functions of an integral variable in his text) or series; these were part of the material he was leading up to.

Hardy was not the first English author to choose this approach in a textbook. Thomas l'Anson Bromwich (1875-1929) wrote *An Introduction to the Theory of Infinite Series* in 1908. He did not introduce the arithmetic treatment of limits in the main body of the text but, in appendix 1, he defined the irrational numbers using Dedekind sections and stated in a footnote that Meray, Weierstrass, and Cantor had framed other definitions. This work was not done independently of Hardy. In the preface of the text, Bromwich credited Hardy in particular:

> [Hardy] has given me great help during the preparation of the book; he has read all the proofs, and also the manuscript of Chapter XI and the Appendices. I am deeply conscious that the value of the book has been much increased by Mr. Hardy's valuable suggestions and by his assistance in the selection and manufacture of the examples [6, p. viii]

Perhaps it can also be argued that the method of cuts is the simplest of the three methods.

---

[21] *Continuity and Irrational Numbers* by Richard Dedekind, 1872, translated into English in 1901. [19]

In this introductory section on Hardy and set theory, I have established that Hardy was one of a small group of British mathematicians[22] who were the first to engage with the new ideas of number, point-sets and transfinite numbers. This can be seen from the following facts: Hardy published papers on set theory, Hardy attended lectures on set theory, Hardy introduced the construction of the real numbers into his influential textbook and, as I will show below, Hardy corresponded with both Jourdain and Russell on set theory during this initial period of his career. So far I have not identified set theory being used as a language for mathematics (the set-theoretic approach) in England during the first decade of the 20th century in the way in which it was described by Hilbert in the passage cited at the beginning of the chapter. I would now like to look at the papers on set theory that Hardy published.

3.2.1 Hardy's five set theory papers, 1904-1910

3.2.1.1 The first and fifth paper on convergence

Hardy's first set theory paper was published in the Proceedings of the London Mathematical Society in 1904 and is titled *A General Theorem Concerning Absolutely Convergent Series.* Series convergence was a topic that was of considerable interest to him at this time as can be seen from the many papers published in the years between 1901 and 1904 on absolute and conditionally convergent series and integrals. In this paper, Hardy proved a general theorem:

> If a series is absolutely convergent in type β[23] , it remains absolutely convergent when its terms are rearranged in another type β', and its sums in the two types are the same [39, p .286]

This generalized more simply stated theorems about absolutely convergent series such as: the sum of the series is independent of the permutation of the terms, double series can be rearranged in any manner of simpler series, or double series may be summed by either rows or columns.

The thrust of the argument is that any enumerably infinite set (or ordinary simply infinite set with ordinal type $< \omega$) can be arranged in terms of type β so that, if a series is

[22] I will show that this group also includes Whitehead, Russell, Jourdain, the Youngs and to a lesser extent Hobson.

[23] Hardy denoted "by Greek letters $\alpha, \beta, \gamma, \ldots$ numbers of *Cantor's first and second classes* 0, 1, 2, ..., ω , $\omega + 1$, ..., $\omega * 2$, ..., ω^2 , ..., ω^ω , ... When it is necessary to distinguish specially the finite numbers (numbers of the *first* class) I shall use $m, n, p, \ldots$" [Har04a, p.286]

absolutely convergent in type β and then can be shown to be convergent in type β', then the results apply for all ordinary infinite series. First the fact that the series being considered are absolutely convergent for Cantor's first number class is generalized to Cantor's second number class using induction. Then the converse is also shown to be true, so that an absolutely convergent series of type β is shown to be convergent when arranged in type ω. But, any rearrangement of an ordinary series of type ω must lead to an arrangement in some type β. And so the desired result is obtained.

Lastly Hardy used a theorem of Cantor's, which he stated as "any set[24] of intervals on a straight line must be enumerable" to justify that his theorem is the most general possible theorem of its kind. For this theorem and its proof, Hardy referenced a work of Young where the theorem, "Every set of intervals on a straight line is countable, provided no two overlap" [101, p. 248], is stated. Young pointed out that he has stated and proved the theorem in precisely the same way as Cantor.

In this first foray into set theory, Hardy used some newly established results in set theory in order to prove a general theorem useful in analysis. Since the fifth and final set theory paper is of a similar type, it will be discussed now, out of chronological order.

In the last set theory paper, published in 1910, again in the Proceedings of the London Mathematical Society, titled *The Ordinal Relations of the Terms of a Convergent Sequence*, Hardy discussed the question:

> How far is it possible to discriminate, on the ground simply of the ordinal relations that hold between their terms, between sequences which converge to a limit and sequences which do not? And, in so far as this is possible, what is the simplest expression of the ordinal relations which characterize convergent sequences? [46, p. 295]

Hardy used the example of the two sequences $1 - \frac{1}{n}$ and n to show that it is not possible to discriminate between convergent and divergent sequences simply on ordinal grounds since both sequences have the ordinal relationship $a_p \geq a_q, p > q$ and the first sequence converges and the second diverges to infinity. However, if diverging to plus or minus infinity can be considered as a special type of convergence then ordinal relationships can be used to discriminate between convergent and oscillatory sequences.

The answer to the question asked is

[24]Note: set not aggregate, even though Hardy's still uses theory of aggregates in 1914.

> that the ordinal relations of convergent sequences do possess an exceedingly simple characteristic which distinguishes them from all other sequences except one very special type of oscillatory sequence; but that this special type of oscillatory sequence cannot be distinguished from a convergent sequence by any marks of its ordinal relations [46, p. 296].

and it results in the theorem

> the necessary and sufficient condition that a sequence should converge to a limit, greater than any of its terms, is that it should be quasi-monotonic (increasing) [46, p. 297]

This is a result Hardy had obtained previously, and it appeared as an example in Bromwich's *Infinite Series* text in the appendix on the *Arithmetic Theory of Irrational Numbers and Limits* discussed above. The novel idea in this paper is that the quasi-monotonic sequences are defined using three 'classes'[25], such that the first two must be finite. The proof that these quasi-monotonic sequences converge is based on the finiteness of the first two classes. Again, set theory here is being used to obtain a result in analysis.

3.2.1.2 Three papers on abstract set theory

The other three papers are of a different nature, dealing with abstract set theory. The first and third of these three remaining papers are related so I will discuss them as a pair after first discussing the 1904 paper, published in the Messenger of Mathematics and titled *The Cardinal Number of a Closed Set of Points.*

In this paper, Hardy noted that Cantor had proved that "the cardinal number of a closed set of points contained in the linear continuum (0,1) is α_0[26] or 2^{α_0}" [Har04c, p. 67], a result for which Hardy referenced Schönflies' *Bericht über die Mengenlehre.* Hardy proved the same result in a "simpler and more direct manner than Cantor" [38, p. 68].

Hardy's proof rests on representing each number in (0,1) in its binary decimal representation. Since the whole interval is non-enumerable, Hardy argues, at some point in the binary decimal expansion of a number, you can divide all of the decimal expansions into two sets

[25]Note: class, not set or aggregate. I suspect that this terminology is a result of the influence of Bertrand Russell.

[26]Hardy used α where $\aleph$ is now used

where the n^{th} number in one set is 1 and in the other set it is 0, such that both of these new sets are non-enumerable. Similarly at some point further along in the expansion in either of these two new sets, there can be a further division into two sets based on the same argument. This process can be repeated indefinitely resulting in 2^n "classes of sets"[27] constructed from the original interval after n steps. Now, since each one of the decimal expansions describes a different limit point, the original set has 2^{α_0} limit points and if it is closed, its own cardinal is 2^{α_0}.

The remaining two set theory papers were published in 1904 and 1907. The first of these two papers, titled *A Theorem Concerning the Infinite Cardinal Numbers*, was published in the Quarterly Journal of Mathematics, and the second titled, *The Continuum and the Second Number Class*, was published in the Proceedings of the London Mathematical Society. The second paper is a response to criticisms of the first paper published by Hobson in 1905 in a paper titled *On the General Theory of Transfinite Numbers and Order Types.*

The purpose of *A Theorem Concerning the Infinite Cardinal Numbers* is to state and prove a theorem concerning the infinite cardinal numbers - i.e.:

> to prove rigorously that the cardinal number of the continuum is greater than or equal to the cardinal numbers of Cantor's second number class; in symbols, that $2^{\alpha_0} \geq \alpha_1$, and more generally, that $2^{\alpha_\beta} \geq \alpha_{\beta+1}$ [40, p. 78].

Hardy showed that this follows from Cantor's theorem that $2^{\alpha_0} > \alpha_1$ but noted that he has not seen it explicitly stated, and in fact, Russell in his 1903 publication *Principles of Mathematics* stated that it is unknown whether or not

> that of any two different cardinal numbers one must be the greater, and it may be that 2^{α_0} is neither greater nor less than α_1 and α_2 and their successors [87, p. 323].

Hardy argued that his theorem concerning the infinite cardinal numbers is true based on an extension of Cantor's argument that every aleph is greater than or equal to aleph-zero. Care is taken to make the justification independent of whether or not the cardinal numbers can be well-ordered. Following this, Hardy constructed a set of points of α_1, a construction that he hoped would throw

[27]Now we would use 'set of sets'. Here it appears that a set contains numbers where a class can contain sets.

> some light (though of course a very partial one) on one of the most fundamentally important and apparently hopeless questions in the whole range of pure mathematics [40, p. 87].

The construction proceeded as follows. An infinite number of integer sequences are formed, each consisting of consecutive integers where the first sequence starts with 1, the second with 2, the third with 3, etc. Then, a new sequence is formed, consisting of diagonal terms from each of above sequences; this sequence is $\{1, 3, 5, 7, 9, \dots\}$ and it is sequence ω since it follows the infinite number of sequences listed above. Then the process starts again and an infinite number of sequences are formed from sequence ω by starting with the second, third, fourth, etc. number. From this second set of infinite sequences, another sequence is formed by again taking diagonals. This sequence is $\omega \cdot 2$. This process continues indefinitely such that an infinite number of sequences of type $\omega \cdot n$ are formed. When the diagonal procedure is applied to these sequences, sequence number ω^n is formed. This building process continues indefinitely and Hardy then set about proving that all of the sequences just constructed are distinct; the proof rests on the fact that each one of the sequences constructed are ordered. Hardy pointed to the importance of the freedom of choice required to make each sequence distinct. Once the distinctness of the sequences is proved, a collection of size α_1 has then been constructed.

3.2.1.3 Hobson's criticisms

In 1905, Hobson published a fairly lengthy paper titled *On the General Theory of Transfinite Numbers and Order Types* [60] which was an attempt to define a 'norm' that would force an ordering onto an aggregate at the time of its formation so that the Burali-Forti paradox[28] could be avoided. The paper also discussed Zermelo's introduction of the Axiom of Choice[29] It appears that Zermelo was unaware of the English debate about the Axiom of Choice by Hardy, Hobson, Jourdain, and Russell. [72, p. 144]

Hobson wanted to

> decide what limitations or qualifications must be imposed upon the nature of an aggregate, so that in the development of the theory [Cantor's general theory

[28] Since the set W of all ordinals is well-ordered, it has an order-type, say β, which must be the largest ordinal; but then the set $W \cup \{\beta\}$ is well-ordered and has as its ordinal $\beta + 1$, a contradiction [72, p. 110]

[29] If a set S is partitioned into a disjoint family A of non-empty sets, then there exists at least one subset T of S which has exactly one element in common with each member of A [72, p. 144].

> of ordinal and aleph-numbers], the possibility of being confronted by such a contradiction as that which was pointed out by Burali Forti [sic] may be removed at its source [60, p. 172].

The end of the paper contained a criticism of Hardy's construction of a set of size $\aleph_1$ outlined above. "This criticism is merely incidental to a much more comprehensive attack on the whole theory of Cantor's transfinite numbers" [42, p. 10]. It is however, the part of the paper I will focus on since it is to this part of the Hobson paper that Hardy responded, leaving the other criticisms to be rebutted by Russell and Jourdain.

Hobson first outlined the method Hardy used to construct the set of size $\aleph_1$ and then objected to Hardy's use of the freedom of choice[30] required to make each sequence in the construction unique. Hobson wanted Hardy to define "a definite norm, or finite set of rules" [60, p. 186] which would suffice to make each sequence unique. Hobson felt the definition of such a norm would be difficult, if not impossible and since Hardy had not provided one, each sequence that Hardy had constructed cannot be admitted as unique, and hence a set of the required size had not been constructed as claimed.

Hobson took care at the end of the paper to praise the theories of Cantor which "have rendered inestimable service in formulating the limitations to which many results in analysis, formerly supposed to be universally valid, are subject" [60, p. 188]. He also allowed that given the "great logical difficulties of the subject" [ibid] that his criticisms may not be fully valid and hoped that they would at least contribute "towards the discussion of questions of great interest which, at the present time, cannot be regarded as having been settled" [ibid].

3.2.1.4 Hardy, Russell and Jourdain Rebut Hobson

Hardy, Russell, and Jourdain all responded in print to the Hobson paper. Russell, at the end of 1905, in a paper titled *On Some Difficulties in the Theory of Transfinite Numbers and Order Types* [88] and Jourdain, early in 1906, with *On the Question of the Existence of Transfinite Numbers* [62], both published in the Proceedings of the LMS. Grattan-Guinness claims that it was this paper of Hobson's that motivated Russell's progress during this time:

[30] Jourdain describes this as "an infinite series of acts of arbitrary selection" [62, p. 266]

> Reformulating Hobson's norm as a propositional function, Russell separated issues surrounding the paradoxes from those related to Zermelo's axiom. He compared the latter with the multiplicative axiom (which he still thought less general), and gave theorems from set theory when needed. He also published for the first time his charming illustration about the need for infinite selections to show that $\aleph_0$ boots divide into pairs as any reasonable owner would desire [33, p. 357]

I focus on the response from Hardy who began his paper, *The Continuum and the Second Number Class* [42], of 1905 (published in 1907), by noting his intention to only discuss the mathematical disagreement between himself and Hobson and leave it to Russell to respond to the rest of the criticism. Hardy then simplified his construction of the set of size $\aleph_1$ without affecting Hobson's arguments and then proceeded to explain the validity of his construction and why Hobson's arguments did not invalidate it. In fact, Hardy simply denied Hobson's argument. "My answer to Dr. Hobson's argument consists simply of a denial of his major premiss" [42, p. 13].

A large part of Hobson's criticisms came from his attempt to deny "the postulate of the existence of the multiplicative class" [42, p. 14]. It is interesting that Russell was in agreement with Hobson about this. He wrote a letter to Jourdain detailing the use of infinite selections,

> I think Hardy's argument that an Nc is either equal to some Aleph or greater than all of them fallacious. It would involve, if written out formally, the assumption that, given a mutually exclusive class of classes, none of them null, it is possible to find at least one class composed of one term out of each class of the class of classes. A similar assumption is involved in Zermelo. I have worked long at assumptions of this kind, without seeing any reason to think them true [29, p. 48]

While agreeing that the allowing of infinite selection is unproved, Hardy pointed to several mathematical examples where it is tacitly assumed. In summary, Hardy was

> therefore, in default of proof, prepared to accept the multiplicative axiom provisionally on the grounds
> (i.) that to deny it appears to be paradoxical;
> (ii.) that no reason has been given for denying it;

> (iii.) that to deny it reduces to a state of chaos a great deal of very interesting mathematics [42, p. 17]

Further, Hardy did not want "to imply that I accept Zermelo's proof that every aggregate is well-ordered. I agree with Dr. Hobson in thinking it open to objection on other grounds" [ibid].

3.2.2 Interactions between Hardy, Russell and Jourdain

In the first decade of the 20th century, Hardy, Russell and Jourdain maintained a three way correspondence. There are letters between Hardy and Jourdain, Hardy and Russell, and Jourdain and Russell. Upon Jourdain's death in 1919, Mittag-Leffler acquired two notebooks of Jourdain's containing letters from a number of mathematicians as well as drafts of his replies. One can see Jourdain's intense interest in set theory from whom he corresponded with. The largest collection of letters in Jourdain's papers is with Russell (about 115 letters) and these letters form the basis of book called *Dear Russell-Dear Jourdain* by Ivor Grattan-Guinness [29]. The other two substantial collections of letters are with Hardy (about sixty-five letters)[31] and Cantor (nearly 50 letters). Jourdain also corresponded with Frege, Peano, and Zermelo. This collection of letters "form an outstandingly valuable collection for the history of mathematics of the period" [34, p. 368].

I have not yet ascertained how many letters where exchanged between Hardy and Russell in this period but the following quotes are used to establish that it was significant. For example, in an article (see [30]) discussing how Bertrand Russell discovered his paradox of the class of all classes which do not belong to themselves, Grattan-Guinness said:

> The most detailed account of which I know is in an exchange of letters with G.H. Hardy. It is largely forgotten now that Cantorian set theory was one of Hardy's early mathematical interests, but he and Russell discussed such questions in the 1900's both in correspondence and in conversation together at Trinity College where they were both fellows. Hardy kept no papers, but he would often send letters back to their writers together with his own replies. Thus there are some extensive letters to Hardy from Russell in Russell's manuscripts [30, p. 130-131].

[31] Interestingly Rankin, in his introduction to Hardy's set theory papers in Hardy's collected works, points specifically to just two letters between Hardy and Jourdain, one of 11 Dec 1903 and another of 04 Aug 1904 to "confirm Hardy's considerable interest at this time in the foundations of set theory" [77, p. 417].

Grattan-Guinness used the letters between Hardy and Russell to try and reconstruct Russell's line of thought that led to the discovery of Russell's paradox.

In addition to corresponding, Hardy, Russell and Jourdain met for discussions. Hardy and Russell were both fellows at Trinity College and members of the Apostles, a secret club that met to discuss questions, primarily of a philosophical nature. A July 5th, 1905 letter from Jourdain to Russell mentions a recent visit from Hardy at which they discussed the difficulties of the multiplicative class [29, p. 54]. Russell visited Jourdain in July of 1907 at Broadwindsor while on a bicycling trip [29, p. 108].

In the winter of 1901-1902 Jourdain attended the first course in mathematical logic given in Britain. Russell's audience for this course was small but included Whitehead as well as Jourdain [33, p. 331].

> The small amount of surviving material [from the lectures] suggests that in addition to the basic Peanist logic he seems to have covered quite a bit of set theory and some aspects of geometry [33, p. 331].

Jourdain, Russell and Hardy read each other's work and reviewed it. As seen above, Jourdain reviewed Hardy's *A Course of Pure Mathematics.*

In May of 1903, Russell's *Principles of Mathematics (PofM)* was published. Russell gave copies to Whitehead, Johnson, Moore, Bradley, Stout, Jourdain and probably Hardy. Overseas, copies went to Couturat, Frege, Peano, Vailati and Pieri. By 1909, fifty of the original print run of one thousand copies were left [33, p. 328]. Hardy reviewed *PofM* for the Times Literary Supplement in September of 1903[32] where he gave it a mixed review calling it "exceedingly difficult", more so than necessary, and claimed it was too short – in fact sometimes too short to be able to follow.

Hardy quoted Russell on the existence of number – "We no more create numbers than Columbus created America" [37, p. 851] and pointed to Russell's definition of pure mathematics as all that can be deduced from twenty premisses[33]. This is the book that introduced Hardy to Frege, "of whom we must confess that we had never heard until Mr. Russell introduced him to our notice" [ibid, p.852].

The problem of Russell's paradox was mentioned,

---

[32]see [37]

[33]"By the help of ten principles of deduction and ten other premisses of a general logical nature . . . , all mathematics can be strictly and formally deduced; and all the entities that occur in mathematics can be defined in terms of those that occur in the above twenty premisses" [87, p. 4]

> Part I raises many difficult questions, especially in the case of the notion of class in which Mr. Russell has discovered a strange contradiction hitherto unresolved [37, p. 852].

And, perhaps most interesting in this context, Hardy stated that most of the readers of *PofM* will probably turn with interest to the fifth part of the book on infinity and continuity. Here, Hardy claimed, if the reader was familiar with modern mathematics, he will not find much that is new but

> the truth is that the supposed contradictions of the infinite have been scattered once and for all by the illustrious Cantor. All that has remained for Mr. Russell is to restate Cantor's doctrines with a philosophical clearness not always to be found in the writings of their originator. [37, p. 853]

Finally, the last sentence of the review was a comment on mathematics at Cambridge and Hardy's own feelings about pure mathematics. This was just before the publication of his first set theory paper.

> Cambridge is generally supposed to care much for the physical and little for the abstract side of mathematics; this book should do much to dispel so unfortunate an impression [37, p. 854].

3.2.3 Hardy's continued interest in set theory and the foundations of mathematics following 1910

Hardy's last paper on set theory was published in 1910 and Hardy's collaboration with Littlewood started the following year – a remarkably productive collaboration which produced a large volume of results, primarily in analysis. But, I will argue here, that while it was undoubtedly a minor interest of Hardy's, he did not lose interest in the development of set theory or in the foundations of mathematics in the decades following 1910.

Hardy, Jourdain, and Hobson were given a copy of the first volume of Russell and Whitehead's *Principia Mathematica (PM)* at end of 1910.[34] Jourdain did seven comprehensive reviews. Hardy reviewed the work in September of 1911 for the Times Literary Supplement in a piece titled *The New Symbolic Logic.*

[34] as were the Royal Society (who gave money to help cover the printing costs), Trinity College, Berry, Couturat, Forsyth, Frege, Hawtry, Johnson (the Press's reader), Peano and Royce. [33, p. 385]

The review is generally positive but pragmatic stating that while "all professional mathematicians, for example, can and ought to read it" [47, p. 859], it can't be expected that more than "twenty or thirty people in England may be expected to" [ibid]. Further - "it is a strange and disconcerting fact that mathematicians as a class are utterly impatient of inquires into the foundations of their own subject" [ibid].

Hardy claimed the authors of *PM* were not widely respected:

> In England we find the authors regarded by mathematicians as amusing cranks. In France, we find the great Poincaré, who has a weakness for philosophy to which we owe several most entertaining volumes, pouring contempt on la Logistique, and preaching a form of pragmatism as hazy and elusive as any philosopher's. Even in Germany, the home of mathematical precision, we find the successors of Cantor and Weierstrass protesting angrily that to ask really fundamental questions is an indecency and an insult to mathematics [47, p. 859].

Early in 1914, Hardy, having made the acquaintance of the American Norbert Wiener, presented Wiener's paper on *A Simplification of the Logic of Relations* to the Cambridge Philosophical Society. Later in 1914 and then continuing for the next five years, Hardy offered a free course in the Easter term on *Elements of Mathematics (for non-mathematical students).* Lecture notes, written by influential philosopher G.E. Moore, show that Hardy was teaching

> a course of set theory influenced by *PM*, and including variables, finite and transfinite cardinal and ordinal arithmetic, mathematical induction, the multiplicative class, continuity, and some elementary geometry [33, p. 421].

In 1927, Russell's book *The Analysis of Matter* was published. In the prefatory material, the publisher noted a forthcoming volume from Hardy on "mathematics for philosophers" [33, p. 451]. The book never appeared and no trace of it has been found.

In 1928, Hardy gave the Rouse Ball lecture at Cambridge. The title of the lecture was *Mathematical Proof* and its subject matter was "from the doubtful ground disputed by mathematics, logic and philosophy" [52, p. 1]. Hardy claimed expertise in mathematics, apologized for intruding as an amateur into mathematical logic and, regarding philosophy, said, "philosophy proper is a subject, on the one hand so hopelessly obscure, and on the other so astonishingly elementary, that there knowledge hardly counts" [52, p. 2]. Examples of mathematical, logical and philosophical questions were given and the continuum hypothesis[35] was used as an illustration of a mathematical question which lies near the border with logic, and "a mathematician interested in the problem is likely to hold logical and even philosophical views of his own" [52, p. 3].

For a mathematician, Hardy said, an acceptable philosophy of mathematics must rationally account for both propositions (mathematical theorems) and proof (a collection of propositions). The paper continued by outlining the objective reality and the validity of a mathematical theorem and insisted on the admittance of the infinite into mathematics.

Hardy divided the existing schools of mathematical logic three: the logicists (Russell, Whitehead, Wittgenstein, and Ramsey[36]), the finitists (Brouwer and Weyl) and the formalists (Hilbert) and stated that there had not been enough attention paid to formalism in England. A summary of the main points of each system followed, with particular reference to the writings of Frank Ramsay (1903-1930), work that Hardy appears to have read with approval – "I cannot hope to find popular language clearer than Ramsay's, and I shall follow him very closely" [52, p. 7].

The theory of aggregates of Cantor and Dedekind[37], the paradoxes, Russell's theory of types, Dedekind's theorem, Russell's Axiom of Reducibility were covered in some detail, the finitists were summarily dismissed and of Hilbert, Hardy separated the logic of Hilbert from the philosophy of Hilbert, which he disliked. Hardy quoted Hilbert on formalism, "the axioms and demonstrable theorems which arise in our formalistic game, are the images of the ideas which form the subject-matter of the ordinary mathematics" [52, p. 11]. Hilbert's system was then described in some detail based on the work of Hilbert's postdoctoral pupil John von Neumann[38], "whose statement I [Hardy] find sharper and more sympathetic than Hilbert's own" [52, p. 14]. Hilbert's formalism was contrasted with Russell's logicism.

Metamathematics was introduced and explained, "and of course it is the metamathematics which is the exciting subject and affords the real justification for our interest in this particular sort of mathematics" [52, p. 16]. This allowed Hardy to distinguish between

[35]stated in this context by Hardy as "Is the cardinal number of the continuum the same as that of Cantor's second number class?" [52, p. 2]

[36]Recall Jourdain died in 1919 or, presumably, he would be part of this list. Frank Ramsay (1903-1930), fellow of Kings College, Cambridge, published *The Foundations of Mathematics* in 1925, and *Mathematical Logic* in 1926.

[37]Note here that, like Ferreirós, Hardy credits both Cantor and Dedekind.

[38]John von Neumann (1903-1957), doctorate in mathematics (1926) with a thesis on set theory.

proof "inside the system", "formal, mathematical, official proof" [52, p. 16] and proof outside the system, "informal, unofficial, significant proofs" [52, p. 17] whose object is to "produce conviction, unofficial conviction of the absence of official contradiction – which is what we want". [52, p. 17]

Thus, given the above - Hardy's book reviews, his Cambridge lecture series and his choice of topic for the Rouse Ball lecture of 1928 - it is evident that Hardy followed developments after 1910.

3.3 Other British Interactions with Set Theory

I have so far discussed primarily Hardy, Russell and Jourdain, who I believe were the main English mathematicians interacting with the new disciple of set theory in the first decade of the 20th century. Because of my interest in Hardy, I have taken particular care to emphasize his role, possibly overemphasizing it. So far there has been one obvious omission – William Henry Young and Grace Chisholm Young. First I intend to give reasons why, despite a large volume of work, they do not appear to have had a commensurate impact and then I will give a brief summary of their work.

William Henry Young (1863-1942), entered Cambridge in 1881, was 4th wrangler in the mathematical tripos of 1884 and was a fellow of Peterhouse College from 1886-1892. Prior to 1898, he did no research, preferring to work long hours as a mathematical coach and to accumulate a large savings. In 1896, he married Grace Chisholm, who had done well in the mathematical tripos exams, but was unranked because of her gender, and had a Ph.D. from Göttingen under Felix Klein. They chose[39] to leave England in 1897, for the continent, settling first in Göttingen and then in Switzerland; a decision taken because Cambridge, they felt, was mathematically dead. "There was no mathematician – or more properly no mathematical thinker – in the place"[40] [99, p.129].

This decision is perhaps one of the reasons for their lack of influence. "His continental associations disenchanted the English, whereas his English background alienated continental

[39]Felix Klein received an honorary doctorate from Cambridge in 1897 and, while visiting, suggested the move to the Youngs. [81, p. 92]. It was also Klein who suggested the new field of Mengenlehre (set theory) to the Youngs, as presented by Schönflies [33, p. 132]

[40]This quote is from Grace's autobiographical notes. Note that Sylvia Wiegand is the granddaughter of Will and Grace, being the daughter of their son Laurence, so her analysis is not necessarily dispassionate.

universities"[41] [99, p. 131]. As well, neither William nor Grace held a traditional academic post; Grace not at all, and William several temporary positions at minor universities. William's lack of a position partially stemmed from his following an unconventional career path. He did not earn a formal degree until 1903, he did not compete for a Smith's prize, and he did no major research until mid-life.

William also had an abrasive personal manner[42], and was disliked by Grace's family. He was also disliked by Russell, who, in a letter to Jourdain, wrote, "I am amused at the thought of your having the Youngs as neighbours. I should suppose that he might possibly become a little trying in the long run" [29, p. 111]. "Will's letters indicate that he was well-intentioned but demanding, outspoken, and critical as well as overly sensitive and especially paranoid about his difficulties in finding a good position" [99, p.131]. "More than once electors to a chair passed him over in favor of men less powerful as mathematicians but less exacting as colleagues" [11, p. 573]

The Youngs often collaborated but consciously chose to publish most of their work under William's name alone, in order to establish a career for him. In a letter from Will to Grace in 1902, Will said

> The fact is our papers ought to be published under our joint names, but if this were done neither of us get the benefit of it. No. Mine the laurels now and the knowledge. Yours the knowledge only [81, p. 94]

The Youngs published 214 papers between 1900 and 1924, and authored three books. Their second book, *The Theory of Sets of Points*, was published in Cambridge in 1906. The Youngs' publications were not of uniformly high quality which likely restricted their influence. Hardy, in his obituary of William in 1942, said, "His style is better in his books than in his papers, which are sometimes rather rambling and diffuse" [55, p. 223]. The Youngs appear to have made a conscious choice of quantity over quality. In the same 1902 letter quoted above, Will said,

> But we must flood the societies with papers. They need not all be up to the continental standard, but they must show knowledge that the others have not got and they must be numerous [81, p. 95]

[41] There is also the possibility of a religious/class divide. Young was originally a Baptist who was baptized into the Church of England while at Cambridge. [11]

[42] See for example [99, p. 131] and [99, p. 314] where family members characterize him this way.

Despite the above, the Youngs were recognized for their work. William was "one of the most profound and original of the English mathematicians of the last fifty years" [55, p. 218] and

> he was awarded the De Morgan medal of the London Mathematical Society in 1917 and was its President during 1922-24. In 1928 the Royal Society awarded him the Sylvester medal in recognition of "a life of invincible mathematical activity" [81, p. 97].
>
> Dame Mary Cartwright (F.R.S. 1947) wrote that G.H. Hardy (F.R.S. 1910) considered that it was through contacts with the Youngs and their continental standards of rigour that E.W. Hobson (F.R.S. 1893) commenced his work on functions of a real variable [81, p. 93].

Rothman (see [81]) claims that this is how Grace's years at Göttingen greatly influenced pure mathematics at Cambridge. But Hobson is generally not seen as a first rate player, and even though he published in set theory, he was slower to grasp the new ideas and use them creatively. Hardy reviewed the second edition of Hobson's *The Theory of Functions of a Real Variable and the Theory of Fourier's Series* in 1922. Of the chapters on number, transfinite numbers and order type, Hardy said:

> Prof. Hobson often allows himself to use language which suggests the Oxford philosopher rather than the Cambridge mathematician. ... We have an uneasy feeling that if one scratched the mathematician one might find the idealist, and that all these discussions, and especially those which concern the 'principle of Zermelo', ought to be stated in a sharper and clearer form [50, p. 436]

There is no doubt that the Youngs felt that set theory was a profoundly important area of study, "we are entering in this subject with the holy of holies of mathematical thought" [103] and that they were the first to publish a book length, systematic treatment of the "Georg Cantor's magnificent theory"[43] [103, p. ix] in English. In fact after mentioning Schönflies *Bericht über die Mengenlehre*, and Russells's *Principles of Mathematics*, the Youngs felt confident enough to claim that their work was "the first attempt at a systematic exposition

[43] Unlike Hardy, the Youngs focused solely on the work of Cantor and viewed set theory as his creation alone.

of the subject as a whole" [103, p. ix]. They sent Cantor a copy of their book, corresponded with him about it, and he came to visit them. Cantor was warmly receptive of their work, indeed of his reception in Britain in general, writing in 1908 that

> My greatest wish is to be able to see that country [Britain], with whose high-minded inhabitants I feel myself as one; quite otherwise is it with the Germans, who do not know me, although I have lived among them fifty-two years [102, p. 423]

However their book was out of print by 1912. This may in part be because it was a pioneering work and "the mathematical community was not yet receptive to this theory" [99, p. 134]. As well, the set theory work was not considered Young's best. Hardy felt Young's best work was the "on the theory of Fourier and other orthogonal series, the differential calculus, and on certain parts of the theory of integration" [55, p. 224] and when summarizing Young's work in his obituary, concluded that it was possible that his work on integration, impeded his recognition. The work was not popular in England or France and Young was dismissed as "the man who was anticipated by Lebesgue" [55, p. 225]. Of the three books the Youngs wrote, Hardy summarized them thus:

> It is curious that Young should never have written a really successful book. He wrote three, alone or in collaboration with his wife; the *Sets of Points*, this tract [*The fundamental theorems of the differential calculus*] (both "classics" which somehow hung fire) [55, p. 234]

The Young's work has likewise attracted rather little attention from historians, perhaps a reflection of its limited reception. For example, the index of Grattan-Guinness's book, *The Search for Mathematical Roots 1870-1940, Logics, Set Theories, and the Foundations of Mathematics from Cantor through Russell to Gödel*, lists all of the pages on which Hardy, Young, Jourdain, Hobson, and Russell's names appear. See table 3.1 below.

Table 3.1: Relative author mention in *The Search for Mathematical Roots*

Name	Total Pages Mentioned	Pages mentioning a relationship with Russell
Hardy	18	10
Hobson	4	0
Jourdain	51	34
Russell	307	NA
Young	5	0

My conclusion is that the Youngs, despite a lot of work, did not have as significant an impact in the early transmission of set theory to Britain as did Hardy, Jourdain and Russell.

Chapter 4

A Course of Pure Mathematics

The purpose of this chapter is to examine the impact of Hardy and *A Course of Pure Mathematics* on British mathematics by considering the following questions: why did Hardy write *A Course in Pure Mathematics*, what did it replace, how is it different from what it replaced, how is it different from textbooks of today, and from where did the "new" mathematics come to Britain?

In considering these questions, I hope to establish that Hardy himself and this textbook in particular exerted a large influence on British mathematics by bringing new standards of rigour to Britain, rigour that was first established in the mathematical work of both French and German mathematicians. Hardy was not the only one who did this - in the previous chapter I discussed the more minor roles played by the Youngs and Hobson for example - but he played an important part, perhaps the most important part.

The process of rigourisation of analysis during the 19^{th} century was motivated by a variety of factors. New technical developments, of which Fourier series is a particularly important example, made it necessary to examine the concepts of limit, function, convergence, and continuity more closely. The separation of mathematics from physics, and the separation of analysis from geometry, removed two previous foundational justifications for analysis, which then needed replacement.

Teaching also formed a main motivating factor for clarification of the foundations of analysis; Cauchy (1789-1857), Weierstrass (1815-1897), and Dedekind (1831-1916) were all motivated to examine foundational issues while preparing to lecture or while authoring textbooks. Hardy, in the preface to the first edition of his book, captured this motivation:

> It has been my good fortune during the last eight or nine years to have a share in the teaching of a good many of the ablest candidates for the Mathematical Tripos[1]; and it is very rarely indeed that I have encountered a pupil who could face the simplest problem involving the ideas of infinity, limit, or continuity, with a vestige of the confidence with which he would deal with questions of a different character and of far greater intrinsic difficulty. ... The fault is not that of the subject or of the student, but of the text-book and the teacher. It is not enough for the latter, if he wishes to drive sound ideas on these points well into the mind of his pupils to be careful and exact himself. He must be prepared not merely to tell the truth, but to tell it elaborately and ostentatiously. [44, p. vii]

The rigorous foundations of analysis recognized today developed primarily in two places and its development was dominated by two people. Cauchy, in France, played the major role in the first half of the 19^{th} century and Weierstrass, in Germany, played the major role in the second half of the 19^{th} century resulting in a satisfactory foundation for analysis by the beginnings of the 20^{th} century. These developments in analysis are well explained in [70, p.155-195] and [2, p.1-13].

By comparing the textbooks used at Cambridge from which Hardy himself was likely to have been taught with *A Course of Pure Mathematics*, I will show what new ideas Hardy felt important to introduce to "first year students at the Universities whose abilities reach or approach something like what is usually described as 'scholarship' standard" [44, p. v].

In particular, the way in which Hardy treats the properties of the real number system and the definition of a function, how he introduces and defines the notion of a limit, and how logarithmic and exponential functions are introduced will be examined in detail. I will show how, or if, these concepts are handled differently from what came before and will comment briefly on how Hardy's methods compare with a modern first year calculus textbook of today. I will also provide some of the historical context in which the changes that occurred in British analysis at the beginning of the 20^{th} century happened.

---

[1] Mathematical examinations at Cambridge University explained below

4.1 The Impact of *A Course of Pure Mathematics*

The importance of a book can be judged in a variety of ways. One simple measure is the length of time a book remains in print and how many editions are printed. The year 2008 marked the publication of the 11^{th} edition of *A Course of Pure Mathematics* and the 100^{th} anniversary of the book. The centenary edition of the book was published with a new forward by T.W. Körner who stated:

> One hundred years after it was first published, CUP (Cambridge University Press) is issuing this 11^{th} edition, not as act of piety, but because *A Course In [sic] Pure Mathematics* remains an excellent seller bought and read by every generation of mathematicians [66, p. 1].

By this measure, this book has influenced mathematics for 100 years.

Furthermore, the impact of this book has not been limited to Britain and other English speaking countries. *A Course of Pure Mathematics* has been translated into Spanish, Chinese, Polish and Russian.

Another measure of a book's importance is the frequency and way in which it is referred to or referenced by mathematicians and others who have used it either for teaching purposes or in their own training. The following quotations illustrate some specific examples of the book's impact in this regard.

Hardy's student, E.C. Titchmarsh, wrote the obituary that was published in the Obituary Notices of Fellows of the Royal Society in 1949. When speaking of the early part of Hardy's career, he said

> To this period belongs his well-known book *A Course of Pure Mathematics*, first published in 1908, which has since gone through several editions and been translated into several languages. The standard of mathematical rigour in England at that time was not high, and Hardy set himself to give the ordinary student a course in which elementary analysis was for the first time done properly. *A Course of Pure Mathematics* is hardly a *Cours d'analyse* in the sense of the great French treatises, but so far as it goes it serves a similar purpose. It is to Hardy and his book that the outlook of present-day English analysts is very largely due. [93, p. 3]

Burkhill, who spent most of his career at Cambridge, wrote several textbooks, one of which, *The Lebesgue Integral*, was published in 1951. In the introduction to this text – which aims to give a straightforward introduction to the theory of integration due to Lebesgue – the reader is told of the assumed background:

> The groundwork in analysis and calculus with which the reader is assumed to be acquainted is, roughly, what is in Hardy's *A course of pure mathematics (1908)* [9].

Similarly, E.H Neville[2] in the *Correspondence of The Mathematical Gazette* wrote in 1941 that

> it would be an impertinence to present a sentence from this source [*A Course of Pure Mathematics*] as if it could be unfamiliar ... the slovenly teaching in elementary analysis was all but universal in England until Professor Hardy directed his expository genius to its eradication [73, p. 217].

Clearly these quotations indicate widespread use and familiarity with Hardy's book nearly fifty years after its publication.

Finally, Ivor Grattan-Guinness wrote in his paper *The Emergence of Mathematical Analysis and its Foundational Progress, 1780-1880* that

> The Continental Analysis did not make much impact in Britain until the early 1840s, when William Thompson, later to become Lord Kelvin but then still a teenager, began to study Fourier series and integrals. Even then British interest lay chiefly in applications to mathematical physics, where the achievements were very brilliant, rather than in foundations ... Not until the work in the early years of this century [the 20^{th}] by G.H. Hardy and W.H. Young were foundational studies brought fully into British education and research [31, p. 98].

---

[2] E.H. Neville, an English mathematician who played a large role in helping Hardy bring the Indian genius Ramanujan to Cambridge in 1914.

4.2 The Motivation for *A Course of Pure Mathematics*

4.2.1 Mathematics training at Cambridge prior to 1907

Mathematics training at Cambridge prior to 1907 was centered on preparing students to sit examinations called the mathematical Tripos. These exams were divided into two parts, of which some students wrote only the first part. Highly competitive, students were ranked by their performance, and their future career prospects depended on their results. Top students were called wranglers; the first wrangler was the top scoring student. Hardy was 4^{th} Wrangler in Part 1 of the mathematical Tripos in 1898 and was placed in the first division of the first class in Part II in 1900. This led to a junior fellowship in 1900, which lasted until 1906 when he obtained a lectureship. Dissatisfaction with this system – particularly with the order of merit – led to a sweeping reform in 1910, a reform in which Hardy played a major role.

Cambridge students of the late 19^{th} and very early 20^{th} century did not attend lectures in order to learn the material required to perform well on the critically important Tripos examinations and so tended not to attend lectures at all.

> It was impossible for undergraduates, whose future career depended on their positions in the order of merit in a highly competitive examination, rigidly confined to a stereotyped syllabus, to 'waste their time' with professors who were eagerly extending the bounds of knowledge, and seeking after new truths generally too complicated to be dealt with in a three hours' examination. Thus arose the strange paradox that Cambridge possessed a number of eminent professors whose lectures had little (if any) influence on even the best students, and with whom most of the undergraduates were wholly unacquainted [76, p. 461].

Rather, students employed, at their own expense, a private coach who lectured, set, and helped students solve typical Tripos questions with the explicit intent of training a student to perform well on the examinations. Mathematical training to equip a student to become a research mathematician was completely secondary to examination preparation.

One of the most famous and successful coaches was Edward John Routh (1831-1907) who coached over 600 students [85, p. 320]. Late in Routh's career he was the coach of Andrew Russell Forsyth (1858-1942) and of Robert Rumsey Webb (1850-1936). Webb was subsequently the coach of both Bertrand Russell and Hardy. Hardy did not approve of Webb's views of mathematics.

> Webb was not interested in the subject of mathematics, only in the tricks of examinations. [74]

Hardy was later coached by A. E. Love. The negative commentary of Forsyth, Hardy and Russell below clearly demonstrate their belief in the inadequacy of mathematical training based on examination preparation.

This type of education impeded the production of textbooks in Britain since coaches had a financial incentive to preserve their teaching methods and materials for the exclusive use of their paying students. Forsyth commented that a coach "codified mathematical knowledge into small tracts or pamphlets, kept in manuscript as his own private prescription for his own set of students. Thus it came about that there were relatively few books" [23, p. 167]. And when he commented on Routh's teaching methods, he said "It was superbly direct for the purpose in view: and it was strong in the measured completeness with which he covered the whole ground for the Tripos. Independence on the part of the student was not encouraged; for independence would rarely, if ever, be justified by the event. Foreign books were seldom mentioned: Routh himself had summarized from them all that could be deemed useful for the examination" [22, p. xvi][3] In these two comments, one can see the disincentive for British textbook production as well as the disincentive to study texts from foreign countries if the material contained in those books did not form part of the Tripos examination material.

Indeed the latter in a sense followed from the former as subjects from the Tripos examination were selected from textbooks that were available and suitable to Cambridge students. Walter William Rouse Ball[4] (1850-1925), best known as a historian of mathematics, explains:

> The character of the instruction in mathematics at the university [Cambridge] has at all times largely depended on the text-books then in use. The importance of good books of this class has been emphasized by a traditional rule that questions should not be set on a new subject in the Tripos unless it had been discussed in

---

[3]It should be noted that not all who studied under Routh had the same opinion as Forsyth about the Tripos system. Karl Pearson (1857-1936), also a student of Routh's, said

> Every bit of mathematical research is really a 'problem', or can be thrown into the form of one, and in post-Cambridge days in Heidelberg and Berlin I found this power of problem-solving gave one advantages in research over German students, who had been taught mathematics in theory, but not in 'problems'. The problem-experience at Cambridge has been of the greatest service to me in life, and I am grateful indeed for it [75, p. 27].

[4]see his book, *The History of Mathematics at Cambridge*

> some treatise suitable and available for Cambridge students [84, p. 128].

Russell's poor opinion of his Cambridge mathematics training is seen in the following quotes.

> The mathematical teaching at Cambridge when I was an undergraduate was definitely bad. Its badness was partly due to the order of merit of the Tripos ... The 'proofs' that were offered of mathematical theorems were an insult to the logical intelligence. Indeed the whole subject of mathematics was presented as a set of clever tricks by which to pile up marks for the Tripos. The effect of all this upon me was to make me think mathematics disgusting. When I had finished my Tripos, I sold all my mathematical books and made a vow that I would never look at a mathematical book again [90, p. 37-38].

Or, when discussing the motivation for studying philosophy, Russell stated:

> My teacher offered me proofs which I felt to be fallacious and which, as I learnt later, had been recognized as fallacious. I did not know then, or for some time after I left Cambridge, that better proofs had been found by German mathematicians. ... I was encouraged in my transition to philosophy by a certain disgust with mathematics, resulting from too much concentration and too much absorption in the sort of skill that is needed in examinations. [89, p. 15-6].

And of the instructors at Cambridge:

> The men who taught me at Cambridge where almost wholly untouched by the Continental mathematics of the previous twenty or thirty years; throughout my undergraduate time, I never heard the name of Weierstrass. It was only by subsequent travel that I came in contact with modern mathematics [91, p. 166].

Hardy was just 5 years younger than Russell and was trained in the same system.

In fact, it has been argued[5] that the situation was worse than the above quotes imply. In analyzing an 1896 unpublished article of Russell's, Griffin and Lewis concluded that not only was Russell ignorant of the concepts of Weierstrass but that he didn't have "any appreciation

[5]see Nicholas Griffin and Albert C Lewis, *Bertrand Russell's Mathematical Education* in the Notes and Records of the Royal Society of London.

of even early 19^{th} century work on limits" [35, p. 64]. At the tail end of the 19^{th} century, Russell became aware of his lack of knowledge of mathematical foundations and rectified it.

In 1926, Hardy gave the Presidential Address to the Mathematical Association, titled *The Case Against the Mathematical Tripos.* Keeping in mind that this was well after the reforms of the Tripos system that occurred in the first decade of the 20^{th} century, Hardy was still a vehement opponent of the Tripos system claiming that historically it had impeded mathematical progress in England and that, to a lesser extent, that it was still doing so in 1926. It is the historical effect of the Tripos, prior to 1907, that is of most interest here. After justifying England's and in particular, Cambridge's stature as a place where first-rate mathematics could be expected to develop, Hardy stated:

> Since Newton, England has produced no mathematician of the very highest rank. There have been English mathematicians, for example Cayley, who stood well in the front rank of the mathematicians of their time, but their number has been quite extraordinarily small; where France or Germany produces twenty or thirty, England produces two or three. There has been no country, of first-rate status and high intellectual tradition, whose standard has been so low; and no first-rate subject, except music, in which England has occupied so consistently humiliating a position. And what have been the peculiar characteristics of such English mathematics as there has been? Occasional flashes of insight, isolated achievements sufficient to show that the ability is really there, but, for the most part, amateurism, ignorance, incompetence, and triviality [51, p. 63].

Hardy continued by noting that the quality of pure mathematics at Cambridge was negatively correlated with the strength of the Tripos system,

> When, in the years perhaps between 1880 and 1890, the Tripos stood, in difficulty, complexity, and notoriety, at the zenith of its reputation, English mathematics was somewhere near it lowest ebb [51, p. 63].

Hardy even questioned the quality of education of mathematical physicists in the 19^{th} century under the Tripos system which is widely thought to have been excellent, producing such well known figures as James Clerk Maxwell, John William Strutt (Lord Rayleigh), William Thomson (Lord Kelvin) and Peter Guthrie Tait. He said:

> A mathematical physicist, I may be told, would on the contrary have received an appropriate and an excellent education. It is possible; it would no doubt be

> very impertinent for me to deny it. Yet I do remember Mr. Bertrand Russell telling me that he studied electricity at Trinity for three years, and that at the end of them he had never heard of Maxwell's equations; and I have been told by friends whom I believe to be competent that Maxwell's equations are really rather important in physics. And when I think of this I begin to wonder whether the teaching of applied mathematics was really quite so perfect as I have sometimes been led to suppose [51, p. 65].

This rather stunning admission was corroborated by Littlewood who said, of the time between 1903 and 1905 when he studied for part 1 of the Tripos examination, that,

> Electricity was completely scrappy and I never saw Maxwell's equations [68, p. 83].

It was not just the opinion of British mathematicians that their training and research abilities were second rate. American mathematician G.D. Birkhoff opinion is evident here:

> We[6] talked in the Prologue[7] about that "third place" which Italian mathematical research was given in the international ranking at the beginning of the century. Italy still appears in third place, at the beginning of the 1920's, in the notes of a US mathematician, G.D. Birkhoff, who was especially interested in the European reality [36, p. 284].

4.2.2 Other cultural changes influencing British mathematical training early in the 20^{th} century

The Tripos training educational system at Cambridge is intimately tied to cultural and educational philosophy issues at Cambridge and in England in general and has a long history; the first 'Mathematical Tripos List' appeared in 1747. The Tripos examinations still take place today; however they in no way dominate the mathematical landscape as they did during the 19^{th} century. The abolition of the order of merit in 1910, the production of a series of rigorous, up to date English language textbooks in the late 19^{th} and early 20^{th}

[6] Guerraggio and Nastasi in their book *Italian Mathematics Between the Two World Wars.*

[7] "The virtues of the Risorgimento generation are to be seen, however, in terms of the creation of the conditions which made possible the second generation to transform Italian mathematics into a great power, second only to France and Germany." [36, p. 10]

century, the importation of latest continental mathematics, and the people at Cambridge in the early 20^{th} century all played a role in vastly improving the quality of mathematics at Cambridge so that by the third decade of the 20^{th} century, British analysis, centered at Cambridge was world class. The question of how and why this happened when it did is complicated and difficult to answer.

There are many issues that affected the rigourisation of British analysis, some of which pertain to mathematics in general and some of which pertain in particular to British mathematics. One of the general issues that affected the rigourisation of mathematics was technical; Fourier series is a good example. It was already discovered by Abel that it was possible to sum a convergent infinite Fourier series of continuous functions to a discontinuous function. This also meant that term by term differentiation had to be more cautiously approached and ideas of continuity and convergence had to be clarified. Other technical issues – where it was possible to get wrong results because of unclear foundations – arose in differential equations and elliptic functions.

Another issue affecting the rigourisation of analysis was pedagogical. Several instructors of mathematics found that in order to clearly explain analysis, they themselves needed to formulate the basic ideas and definitions clearly. This was the case for Cauchy, Weierstrass and Dedekind.

The last two general points involve disciplinary separations. First, mathematics separated from physics during the 19^{th} century. Prior to this separation, physical results could be used to corroborate mathematical results; in particular, the existence of a solution was demonstrated by the physical situation.[8] This meant that the correctness of a mathematical result now needed to be provided by mathematics itself. Secondly, analysis separated from geometry; this meant the correctness of analysis required demonstration by methods other than geometric argument. In fact, with the construction of new geometries and the gaps discovered in Euclid's proofs, recourse to geometrical argument itself was not sufficient. So, there were several reasons why the development of rigour was pursued in the 19^{th} century and it happened primarily in France and Germany[9].

In Britain, in the period following Newton particularly after the Newton/Leibniz controversy over precedent in the invention of calculus, English (which in effect was Cambridge) mathematicians isolated themselves from others, feeling a strong sense of superiority [35, p.

[8]see [70]

[9]see [70] or [2]

56]. Tripos examinations even expected students to know lemmas from Newton's work by number alone. The first attempt at reform happened early in the 19^{th} century.

Robert Woodhouse[10] (1773-1827) introduced analytic techniques to Cambridge when he wrote *Principles of Analytical Calculation* in 1803. In this work, he "explained the differential notation and strongly pressed the employment of it" [84] while at the same time "he exposed the unsoundness of some of the usual methods of establishing it" [84]. Rouse Ball notes that Woodhouse was critical of the foundations of this new analysis and that this is "not infrequently neglected in modern [Ball was writing in 1908] textbooks" [84]. The most important effect of Woodhouse's efforts was the formation of the short-lived *Analytical Society* formed by three undergraduates: George Peacock (1791-1858), Charles Babbage (1792-1871) and Sir John Frederick William Herschel (1792-1871).

In the context of reforming Cambridge mathematics, Peacock was the most important of the founders. As a newly appointed lecturer, with the responsibility for setting Tripos examination papers, he switched, without warning, to Continental notation.[11] "Even then we took examinations seriously; from that moment Peacock had won and Cambridge was confirmed as the home of English mathematics" [100, p. 278]. These reforms influenced several generations of undergraduates including Augustus De Morgan, William Whewell, Arthur Cayley, J.J. Sylvester, William Thomson and George Gabriel Stokes. It was during the 1830's and 1840's that analysis became a Tripos examination subject and the Tripos examination became famous.[12] Griffin argues that the result of this reform was that the system

> produced during those years some of Cambridge's best pure mathematicians of the 19^{th} century: J.J. Sylvester and Arthur Cayley [35, p. 56].

This first reform came to an end, largely and ironically, given that he had benefited from it, by the actions of Whewell. He decided that the role of mathematics education was primarily as a foundation for a liberal education and as such emphasis was to be placed on permanent (geometry) rather than progressive (contemporary developments in algebra) ideas. The Tripos was reformed to reflect this in 1848; at which time Whewell also established

---

[10] A Cambridge trained fellow of Caius College who wrote several textbooks in the first two decades of the 19^{th} century.

[11] Joseph Louis Lagrange's algebraic notation

[12] see [35] and [20]

a preference for "physical sciences with their practical utility to pure mathematics" [20, p. 156].

Isaac Todhunter (1820-1884) (a student of De Morgan who was a student of Peacock's) is best known for his textbooks and was a fellow and lecturer at Cambridge on geometrical optics, an important Tripos topic. Given the emphasis on physical sciences instituted by Whewell, here is another quote, again involving Maxwell, which indicates that the physical science instruction was not excellent, as Hardy has suggested. Maxwell invited Todhunter to observe his experiments on conical refraction at Cavendish laboratory; a invitation that Todhunter refused

> since he thought it might upset him. "Then", Maxwell asked, "allow your students to come." "Sir" said Todhunter, "If a young man will not believe his Tutor, a gentleman and often in Holy Orders, I fail to see what can be gained by practical demonstration." [100, p. 280]

This was approximately the system in place when Hardy arrived at Cambridge in 1896 with the exception that in the 1880's, the Tripos exam was divided into two parts; elementary and advanced. This division may have been a further detriment since only a small number of students studied for part two of the mathematical Tripos.

The reform needed for Cambridge to become world class began late in the 19^{th} century with the 1893 publication of Forsyth's text titled *Theory of Functions*. At least one person[13] felt that this was a tremendously important publication and called it "one of the most influential British mathematics books after the Principia" [100, p. 281]. Hardy joined the reform movement in 1898, and his text, "another book designed to liberate Cambridge mathematics" [100, p. 281], was published in 1908.

4.2.3 Hardy's informal mathematical training

Given that the Tripos based system at Cambridge for undergraduate mathematical training was lacking, particularly in pure mathematics, how did Hardy educated himself in the methods he was later to promote? Since Bertrand Russell also successfully undertook this task, his method will also be discussed.

Hardy, in an often-quoted passage, has stated where his first real understanding of mathematics came from:

[13] H.H. Williams [100]

> My eyes were first opened by Professor Love, who first taught me a few terms and gave me my first serious conception of analysis. But the great debt which I owe to him was his advice to read Jordan's *Cours d'analyse*; and I shall never forget the astonishment with which I read that remarkable work, the first inspiration for so many mathematicians of my generation, and learnt for the first time as I read it what mathematics really meant [54, p. 147].

In 1922, Hardy wrote an obituary of Camille Jordan (1838-1922) in which he calls Jordan's *Cours*

> the first systematic treatise on analysis in which the fundamental problems of the theory of functions were envisaged from a really modern point of view, and it has accordingly played a great part in the education of most of the leading analysts of the day.[14] [49, p. 721]

There are three editions of the three volume *Cours d'analyse*. The first edition appeared between 1882 and 1887, the second between 1893 and 1896 and third between 1909 and 1915. It is in the second edition that the point-set theory of Cantor is detailed in the opening pages of the first volume. It is probable that Hardy is referring to the second edition in the above quotations since he refers to the second edition of Jordan's work in Jordan's obituary[15], written in 1922. More compellingly, Hardy, in his first tract *The Integration of Functions of a Single Variable*, written in 1905, used an example from Jordan that he credits as "Cf. Jordan, *Cours d'analyse*, ed. 2, vol 2, p.21" [41, p. 26].

Hardy has also stated that he

> was really quite ignorant, even when I took the Tripos, of the subjects on which I have spent the rest of my life [54, p. 147]

and depending on which part of the Tripos he is referring to, it is possible to conclude then that he read the second edition of Jordan's *Cours d'analyse* between 1898 or 1900 and 1905.

[14]In a biography of Julian Schwinger (see [71]), there is a note on Freeman Dyson who attended Winchester College before going to Cambridge, as did Hardy. While at Winchester, Dyson "worked through the three volumes of Camille Jordan's *Cours d'analyse* which they found on the upper shelves of the library. Jordan's *Cours* had probably been donated by the Cambridge mathematician G.H. Hardy." [71, p. 238] This is further evidence of how important this book was to Hardy as well as the strength of his convictions that this material should be made available to future mathematicians.

[15]This is the conclusion drawn by Ivor Grattan-Guinness in his article *Russell and G.H. Hardy: A Study of their Relationship*.

Unlike *A Course of Pure Mathematics*, *The Integration of Functions of a Single Variable* has a bibliography; works Hardy must have read prior to 1905. Of course these referenced works are the ones that deal with integration of a function of a single variable and with one exception, all twenty-five works are written in German or French, with the works of Abel, Liouville and Chebyshev the most prominent. I conclude that Hardy had, by 1905 and probably quite a bit sooner, a full appreciation of the continental mathematics that he was not aware of prior to 1898.

Bertrand Russell returned to mathematics in 1894[16] to write a fellowship dissertation on the foundations of geometry. It is not known exactly when – although it would have been after his unpublished 1896 article – or how he realized that his mathematical knowledge was out of date. It is known however, that he first heard of Weierstrass when visiting the United States in 1896 and that he read *Introduction to the theory of analytic functions*[17] in 1899. For Russell, mathematics was to be the foundation of a rational belief system, one secure against skeptical attack and as such, the foundations of mathematics were extraordinarily important to him. Mathematics had "to be placed on a rational foundation" because "only when it were done would mathematics be fit to stand as the cornerstone for the whole edifice of human knowledge" [35, p. 67].

4.3 The Mathematics in *A Course of Pure Mathematics*

4.3.1 Overview of the Text

A Course of Pure Mathematics, still in print, was written in 1908 and is currently in its 11^{th} edition. New editions appeared roughly every 5 years from 1908 to 1944 when the 9^{th} edition was printed and then reprinted in 1945. The 10^{th} edition was published in 1963 and the 11^{th} edition in 2008. Major revisions occurred in the 2^{nd} (1914) and the 7^{th} (1938) edition. The first edition has a comprehensive index whereas, inexplicably, there is no index by the 9^{th} edition.

The book has always consisted of ten chapters with the same or very similar chapter titles. The first three chapters introduce real numbers, functions of integer and real variables

---

[16]See [35] for a detailed explanation on which this summary is based.

[17]A monograph authored by James Harkness, professor of mathematics at Bryn Mawr, and Frank Morley, professor of mathematics at Haverford College in Pennsylvania. Both of these professors attended Russell's lectures at Bryn Mawr in 1896.

and complex numbers. Chapters four and five define the limits of functions of a positive integral variable and follow with the limits of functions of a continuous variable. Chapters six and seven cover derivatives and integrals and chapter eight details convergence of infinite series and infinite integrals. The last two chapters, nine and ten, define the logarithmic and exponential functions, starting from the definition of the logarithm as an integral and culminate in the general theory of the logarithmic, exponential and circular functions.

It is, in fact, the material in the last two chapters that provided the impetus for writing the book and directed the material in the earlier chapters. According to Hardy:

> It was the desire to write an elementary treatise of this theory that originally led me to begin the book, and I have generally decided my choice of what was to be included in the earlier chapters by a considerations of what theorems would be wanted in the last two [44, p. vii].

I found this surprising and will comment on it below.

When Hardy revised the book in 1937 in preparation for the 7^{th} edition, he reread the book in detail for the first time in twenty years. In the preface of the 7^{th} edition, his writing gives a clue both to his original motivations when writing the text and to how much things had changed at Cambridge in the preceding twenty years. He said

> It (*A Course of Pure Mathematics*) was written when analysis was neglected at Cambridge, and with an emphasis and enthusiasm which seem rather ridiculous now. If I were to rewrite it now I should not write (to use Prof. Littlewood's simile) like "a missionary talking to cannibals", but with decent terseness and restraint and, writing more shortly, I should be able to include a great deal more. . . . It is perhaps fortunate that I have no time for such an undertaking, since I should probably end by writing a much better but much less individual book, and one less useful as an introduction to the books on analysis of which, even in England, there is now no lack [56, p. v].

4.3.2 The Real Number Line

Hardy's book first introduces the properties of the various classes of numbers in the real number system, which he calls the arithmetical continuum. This is handled very differently in the first and second editions. In the first edition, Hardy does not:

> attempt to include any account of any purely arithmetical theory of irrational number, since I believe all such theories to be entirely unsuitable for elementary teaching [44, p. v].

The linear continuum (Cantor's term) is taken for granted and it is assumed that there exists "a definite number corresponding to each of its points" [44, p. v].

Chapter one, twenty-four pages in length in the first edition, consists of an introduction to number via line length and ratios of line lengths, leading to a definition of a rational number. There is a justification that there are infinitely many rational numbers in every segment of the number line. Hardy then states

> There is, however, good reason for supposing that *there are on the line points which are not rational points*[18] [44, p. 4].

This is justified by appealing to the length of the hypotenuse of a right-angled isosceles triangle with equal sides of 1 and to Euclid's construction for a mean proportional between 1 and 2. After proving that there is no rational number whose square is m/n, where m/n is any positive fraction in its lowest terms, unless m and n are both perfect squares, Hardy defines a number that is not rational to be irrational.

Next, quadratic surds (numbers of the form $a + \sqrt{b}$, where a and b are rational) are introduced with the idea that the existence of quadratic surds was suggested by geometrical considerations. But a more general form of an irrational number can be thought of, one with no geometrical significance. It is shown that infinitely many irrational numbers can be found by considering divisions of rational numbers into two classes, the lower class T and the upper class U. This is proven and then

> our common-sense notion of the attributes of a straight line demands *the existence of a number x and a corresponding point P such that P divides the class T from the class U*. [Har08, p.13]

and this number is not rational. The example of $\sqrt{2}$ is justified but there is no general proof. So, despite Hardy's quote in the preface of the first edition that he does not include any account of any purely arithmetical theory of irrational number, he does go part way down the path of defining an irrational number arithmetically.

---

[18] Emphasis is in the original.

Finally, at the end of the first chapter, the linear continuum is defined as the "aggregate of points contained in a straight line L" which includes the rational points, the irrational points and "all other points of the line, if any such there may be" [44, p. 15]. Following from this, real numbers are the signed measures of the lengths of finite portions of the linear continuum, and the arithmetic continuum is the aggregate of all real numbers. Again Hardy states that this is not a rigorous exposition and indeed provides examples using a quintic equation and pi to show that all irrational numbers are not found using combinations of surds. Interestingly, the terms algebraic and transcendental are not introduced in the main text but are defined in the exercises at the end of the chapter.

In contrast, in later editions of the book, starting in 1914, Hardy defines a rational number as a ratio of two integers and then follows with an explanation of the representation of rational numbers by points on a line. He then takes care to mention that analysis in no way depends on geometry and that geometry is "employed merely for the sake of clearness of exposition" [56, p. 2]. The same two geometrical examples are used to motivate the discussion of an irrational number but only to "suggest the desirability of enlarging our conception of 'number' by the introduction of further numbers of a new kind" [56, p. 7]. Irrational numbers are then introduced as necessary for the solution of equations of the form $x^2 = 2$. The rational numbers are, as before, divided into two classes, which are now R and L rather than T and U.

Now the reader is invited to accept the existence of these "irrational numbers" and skip to the section on quadratic surds – if he wishes to avoid abstract discussion. This results in the discussion flowing much as it did in the first edition. Otherwise, for someone interested in a rigorous definition of an irrational number, there is a more extensive explanation of irrational numbers. Irrational numbers are defined using sections such that an irrational number is a real number that is not rational where a real number is "a section of the rational numbers, in which both classes exist and the lower class has no greatest number" [56, p.14]. Then, rather than assuming the laws of arithmetic apply to these newly defined numbers, Hardy expends considerable effort to show the validity of the relational operators between real numbers and to show that all of the algebraic operations, such as addition, subtraction, multiplication and division still hold. The discussion then continues with quadratic surds as in the first edition.

Next, as in the first edition, this discussion culminates in a description of the continuum. This section is greatly enlarged and the reader is not invited to skip any of this material;

it is considered of utmost importance. First the arithmetical continuum is defined as the aggregate of all real numbers, rational and irrational, and the linear continuum is only mentioned to supply the reader with a convenient image of the arithmetic continuum. Surds, a small subclass of irrational numbers are roots of algebraic equations, which are in turn a small subclass of all of the irrational numbers.

Now, the idea of sections that was applied to the discussion of rational numbers in order to define irrational number is applied to real numbers.

> The idea of a 'section', first brought into prominence in Dedekind's famous pamphlet *Stetigkeit und irrationale Zahlen*[19] , is one which must be grasped by every reader of this book, even if he be one of those who prefer to omit the discussion of the notion of an irrational number [56, p. 28].

Sections are used to prove that, unlike the aggregate of rational numbers, the aggregate of real numbers is complete[20] – there is no necessity for a further generalization of number. As such the continuum is closed. The results proved using sections of real numbers are stated as Dedekind's thereom. Next, an accumulation point is defined and finally Weierstrass's theorem that

> If a set S contains infinitely many points, and is entirely situated in an interval (α, β), then at least one point of the interval is a point of accumulation of S [56, p. 32]

is stated.

This is the point where Hardy steps back from things he feels to be too difficult for an elementary text.

---

[19] *Continuity and Irrational Numbers* by Richard Dedekind, 1872, translated into English in 1901.

[20] In [56, p. 29-30], completeness is introduced in making the point that, while sections of rational numbers lead to a new, more general conception of number, sections of real numbers do not lead to a further generalization of number such that "the aggregate of real numbers, or the continuum has a kind of completeness which the aggregate of the rational numbers lacked, a completeness which is expressed in technical language by saying that the continuum is closed". The words complete or closed do not appear in Dedekind but the same idea is expressed in two theorems. The first identifies that the domain $\Re$ possesses *continuity* and states that "if the system $\Re$ of all real numbers breaks up into two classes, $U1$ and $U2$ such that every number α_1 of the class $U1$ is less than every number α_2 of the class $U2$ then there exists one and only one number α by which this separation is produced" [19, p. 20]. The second theorem is "if a magnitude x grows continually but not beyond all limits it approaches a limiting value" [19, p. 25]. Dedekind then states that this second thereom is equivalent to the first and both lose their validity when a single real number is not contained in the domain $\Re$.

> The general theory of sets of points is of the utmost interest and importance in the higher branches of analysis; but it is for the most part too difficult to be included in a book such as this [56, p. 31].

From a modern perspective, it is apparent that set theory is of the utmost importance:

> set theory plays an organizing role in the polis of mainstream modern mathematics and represents one of the highest achievements of mathematical wisdom [21, p. xv].

Hardy used some set notation when dealing with real numbers but, as will be shown below, used none when dealing with functions; something that is not seen today.

How is Hardy's treatment of real numbers different from what we now expect to see in an advanced first year introductory textbook? Hardy provides a more advanced and rigorous introduction to the continuum than we would expect in a modern course that "would either leave the construction to much later or omit it altogether" [66, p. 4]. For example, in Adams[21] , real numbers are introduced with a series of examples which include integers, rational numbers, and irrational numbers, both algebraic and transcendental. The real numbers are geometrically identified with the real number line and the algebraic, order and completeness properties of the real number line are stated as axioms. Completeness is identified as a subtle and difficult concept and is stated as:

> If A is any set of real numbers having at least on number in it, and if there exists a real number y with the property $x \leq y$ for every x in A, then there exists a smallest number y with the same property. Roughly speaking, this says that there can be no holes or gaps on the real line – every point corresponds to a real number. [1, p. 3]

The interested reader is referred to an appendix that contains $\epsilon - \delta$ proofs of some of the fundamental theorems of calculus; however completeness of the real number line is stated as an axiom; Dedekind cuts are not mentioned.

Similarly, Clark[22] takes an axiomatic approach,

> Only in the 19^{th} century was a successful theory developed, in which the entire real number system could be defined in terms of the system of positive integers.

[21] *Single Variable Calculus* by Robert A Adams, a standard first year university calculus textbook

[22] *Elementary Mathematical Analysis* by Colin Clark used in second year introductory analysis courses.

> Since this theory is quite complicated, it is not convenient to present it in an elementary book. Instead, we will treat the properties of the real number system as axioms. [15, p. 6]

It is not until an advanced text such as Rudin[23] – "intended to serve as a text for the course in analysis that is usually taken by advanced undergraduates or by first-year graduate students who study mathematics" [86, p. v] – that the real numbers are constructed from the rationals, Dedekind cuts are introduced and the completeness property is proved.

Rudin's text in this section is astonishingly similar to Hardy's. After stating Theorem 1.32 (Dedekind), Rudin states that because gaps were found in the rational number system and filled and that

> if we tried to repeat the process which lead us from the rationals to the reals, by constructing cuts ... whose members are real numbers, every cut would have a smallest upper number, we could immediately identify every cut with its smallest upper number, and nothing new would be obtained. For this reason, Theorem 1.32 is sometimes called the completeness theorem for the real numbers [86, p. 10].

Both Hardy and Rudin make the same argument that cuts of rational numbers identify gaps and that cuts of real numbers do not, that this property of real numbers is called completeness and that Dedekind's theorem is a statement of this completeness. The only real difference is that Hardy makes the completeness argument before stating Dedekind's theorem[24] and Rudin does it after. Their statements of Dedekind's theorem are different, however, in that Rudin states Dedekind's theorem using set notation and Hardy does not.

Of the sixteen books that Rudin lists in his bibliography, two are authored by Hardy. The 9^{th} edition of *A Course of Pure Mathematics* is one and Hardy's book, coauthored with Rogosinski, on Fourier series is the other. A third book in the bibliography, *The Theory*

---

[23] *Principles of Mathematical Analysis* by Walter Rudin

[24] Hardy's statement of Dedekind theorem is "If the real numbers are divided into two classes L and R in such a way that i) every number belongs to one or other of the two classes, ii) each class contains at least one number, iii) any member of L is less than any member of R, then there is a number α, which has the property that all the numbers less than it belong to L and all the numbers greater than it to R. The number α itself may belong to either class" [56, p. 30]. Rudin's statement of the thereom is "Let A and B be sets of real numbers such that (a) every real number is either in A or in B; (b) no real number is in A and in B; (c) neither A nor B is empty; (d) if $\alpha \in A$, and $\beta \in B$, then $\alpha < \beta$. Then there is one (and only one) real number γ such that $\alpha \leq \gamma$ for all $\alpha \in A$, and $\gamma \leq \beta$? for all $\beta \in B$" [86, p. 9].

of Functions, was written by Hardy's student Titchmarsh. Through Rudin's work, Hardy's influence on the way to define real numbers continues to this day.

When we look at English textbooks that are contemporary to Hardy's or that Hardy may have studied himself, the picture is not so straightforward. For example, Edouard Goursat's *A Course in Mathematical Analysis*, written in 1902, was translated and published in English in 1904 in order to "fill the need so generally felt throughout the American mathematical world" because the "lack of standard texts on mathematical subjects in the English language is too well known to require insistence" [28, p. v]. The English version was intended for use as a text for a second course in calculus.

This book starts with functions of a single variable, without mention of real numbers. Goursat states in his preface that

> Since mathematical analysis is essentially the science of the continuum, it would seem that every course in analysis should begin, logically, with the study of irrational numbers. I have supposed, however, that the student is already familiar with that subject. The theory of incommensurable numbers is treated in so many excellent well-known works that I have thought it useless to enter upon such a discussion [28, p. iii].

This remark forced the book's translator, E.R. Hedrick, to comment in a footnote

> such books are *not* common in English. The reader is referred to Pierpont, *Theory of Functions of Real Variables*, 1905... Tannery, *Lecons d'arithmetique*, 1900 and other foreign works on arithmetic and on real functions [28, p. iii].

Chrystal's[25] *Algebra*, first published in 1889 and in its 5^{th} edition by 1904, is another example of an analysis book contemporary to Hardy's. It was written for higher classes of secondary schools and for colleges. The second volume of this book was identified by Littlewood as one of the more advanced books he read while attending St. Paul's School prior to taking the Entrance Scholarship Examination for Cambridge. Littlewood felt his education was good considering the system in which it occurred and that

[25] George Chrystal(1851-1911) was a graduate of Cambridge University (1877) who spent most of his career as the chair of mathematics at the University of Edinburgh where he was known an an excellent teacher. *Algebra* is his most famous publication. The first volume was published in 1886 and the second volume in 1889.

> Ideally I should have learnt analysis from a French *Cours d'analyse* instead of from Chrystal and Hobson[26], but this would have been utterly unconventional. I did not see myself as a pure mathematician (still less as an analyst) until after my Tripos Part I, but I had enough instinctive interest in rigour to make me master the chapters of Chrystal on limits and convergence. The work is rigorous (within reasonable limits), and I really did understand, for instance, uniform convergence, but it is appallingly heavy going [68, p. 81-82].

This level of mathematics is so far beyond a high school student of today that it barely seems possible.

It appears that Chrystal, like Hardy, believed that examination preparation was causing the fundamental notions of limits and infinite series to be rushed so that the machinery of calculus could be learned in order to solve problems.

> Besides being to a large extent an educational sham, this course is a sin against the spirit of mathematical progress [14, p. vi].

Therefore, in his chapters on inequalities, limits and convergence of series, Chrystal attempted to

> avoid trenching on the ground already occupied by standard treatises: the subjects taken up, although they are all important, are either not treated at all or else treated very perfunctorily in other English text-books [14, p. vii].

This motivated the discussion of the work of Weierstrass, Dedekind, and Cantor, all well referenced to French or German works, and the decision to base the theory of irrationals on a mixture of Dedekind and Cantor – a method "best suited to bring the issues clearly before the mind of a beginner" [14, p. 98]. This is followed by the Dedekind's theory of sections in a presentation that is more difficult to follow than the one given by Hardy.

In summary, with regard to the continuum, Hardy's book appears to be one of a very few English books published before 1908 in which the real numbers are constructed and are shown to be complete. Goursat's book treats this as known to the reader but that is because

[26]Hobson published *Treatise on Plane Trigonometry* in 1891. Although the title would not suggest this book's subject to be analysis, "the later portions of this book were for many years the only place (with the exception of Chrystal's *Algebra*) where could be found an accurate account in English of complex numbers and of infinite series. In 1907 the fame of his *Trigonometry* was eclipsed by that of his *Treatise on the Functions of a Real Variable and the Theory of Fourier's Series*" [76, p. 463]

Goursat wrote in French and assumed familiarity with other French work or perhaps he thought this was not very important. And then, when compared with textbooks of today, Hardy's, and even Chrystal's work, are more rigorous than most of the textbooks designed for students of a similar age. Today, textbooks written for high school and first and second year university mathematics favour the axiomatic approach over the constructivist approach – a method which has "all of the advantages of theft over honest toil"[27].

4.3.3 The Definition of a Function

Hardy motivates the concept of a function geometrically by considering two straight lines on which two distances are measured, respectively x and y, which are continuous real variables. Exactly what is meant by variable is not further analyzed and there is no attempt to link a variable to a section or cut. If there is a relationship between x and y such that when x is known, y is known then y is a function of x. The first examples of functions are explicit formula and geometrical constructions. However, "all that is essential is that there should be some relation between x and y such that to some values of x at any rate correspond values of y" [44, p. 26].

Following this definition, the graphical representation of functions in Cartesian coordinates is explained, with a definition of terms such as coordinates, abscissa, ordinate, graph, and locus given in the context of graphing a straight line. The exposition in the 1^{st} edition on graphing in Cartesian coordinates is longer and more detailed and followed with more examples than in subsequent editions. Polar coordinates are introduced as an alternate way of specifying the location of a point but there is no explanation of why they may be useful or needed.

The remainder of the chapter outlines various types of functions and their graphs including, in this order, polynomials, rational functions, explicit algebraical functions, implicit algebraical functions, and transcendental functions including the direct and indirect trigonometric functions or circular functions. The exponential and logarithmic functions are mentioned but are discussed separately in later chapters. As well, other transcendental functions

[27]Russell coined this phrase and used it to compare the axiomatic and the constructivist approach in defining the real numbers. He also used it in the same sense in *Introduction to Mathematical Philosophy* when he said "The method of 'postulating' what we want has many advantages; they are the same as the advantages of theft over honest toil."

such as elliptic functions, Bessel's and Legendre's functions and gamma functions are mentioned but lie beyond the scope of the book. The final functions mentioned serve to illustrate the possible variety of functional relationships and include step and sawtooth functions or unusual functions such as "y is the largest prime factor of x" or "y is the denominator of x" among others.

This function definition remained constant throughout all editions of the text and this is interesting because there is no mention of sets. In modern textbooks, even elementary ones, a function is typically defined as a mapping of one set onto another even when, like Hardy, the concept is illustrated graphically. So, for example, in Adams

> A function f is a rule that assigns to each real number x in some set $D(f)$ (called the *domain* of f) a unique real number $f(x)$ called the value of f at x

and

> The range of a function f is the set of all real numbers y that are obtained as values of the function; that is, it is the set of all numbers $y = f(x)$ corresponding to all numbers x in the domain of f

and then for graphing purposes

> the domain of the function f can be represented as the set of points on the x-axis; the range is a set on the y-axis. [1, p. 14-15]

Similarly in Clark, the modern set-theoretic definition of function is given as

> Let A and B be given (nonempty) sets. A set f of ordered pairs (a, b) with $a \in A$ and $b \in B$, is called a function from A to B, provided that for every $a \in A$ there exists a unique $b \in B$ such that $(a, b) \in f$. In case $(a, b) \in f$, we write $f(a) = b$. [15, p. 234]

However, Clark gives this full definition in an appendix. The word function is used freely in the main body of the text, on limits, convergence, continuity etc., without prior definition. No geometrical discussion is given.

In Rudin, set notation is used and more logically than in Clark, the definition of a function is given following the chapter on real and complex number systems before the word is used. The definition is:

> Consider two sets A and B, whose elements may be any objects whatsoever, and suppose that with each element x of A there is associated, in some manner, an element of B, which we denote by $f(x)$. Then f is said to be a function from A to B (or a mapping of A into B). The set A is called the domain of definition of f (we also say f is defined on A), and the elements $f(s)$ are called the values of F. The set of all values of f is called the range of f. [86, p. 21]

This is the most abstract of the three definitions and it is very clear.

It is surprising that Hardy did not ever change his function definition; he also chose not to introduce countable sets or any discussion of infinity as both Clark and Rudin do. In the first edition of *A Course of Pure Mathematics*, Hardy does not introduce the Heine-Borel theorem; by the 9^{th} edition he does – in the context of oscillations of a function on an interval. To do this, he introduces sets of intervals on a line. However, Hardy's statement of the Heine-Borel theorem[28] is quite different from what is in Rudin[29]. Maybe the set theory material is some of what Hardy alludes to in the quote above (section 4.1) when he says that if he were to rewrite the book some twenty years later, he would be briefer and then able to include a great deal more. "The book would then be much more like a *Traite d'analyse* of the standard pattern" [56, p. v].

As with the real number definition, Hardy's treatment of a function is similar to that of Chrystal who defines a function using the word quantity, which given its usage, appears to be interchangeable with variable. "There are an infinite number of ways in which we may conceive one quantity y to depend upon, be calculable from, or, in technical mathematical language, be a function of, another quantity x" [14, p. 273]. Following several simple algebraic function examples of the type $y = f(x)$, Chrystal continues with "For convenience x is called the independent variable, and y the dependent variable; because we imagine that any value we please is given to x, and the corresponding value of y derived from it by means of the functional relation. All the other symbols of quantity that occur in the above relations, such as 3, 17, a, b, c, 2, &c., are supposed to remain fixed, and are therefore called

[28] Suppose that we are given an interval (a, b), and a set of intervals I each of whose members is included in (a, b). Suppose further that I possesses the following properties: (i) every point of (a, b), other than a or b, lies inside at least one interval of I; (ii) a is the left-hand end point, and b the right-hand end point, of at least one interval of I. Then it is possible to choose a finite number of intervals from the set I which form a set of intervals possessing the properties (i) and (ii). [56, p. 197-198]

[29] If a set E in R^k has one of the following three properties, then it has the other two: (a) E is closed and bounded. (b) E is compact. (c) Every infinite subset of E has a limit point in E. [86, p. 35]

constants" [14, p. 273-274].

Again, in Goursat, a function definition based on relationship between variables is given. "When two variable quantities are so related that the value of one of them depends upon the value of the other, they are said to be functions of each other" and "In short, any absolutely arbitrary law may be assumed for finding the value of y from that of x. The word function, in its most general sense, means nothing more nor less than this: to every value of x corresponds a value of y" [28, p. 2].

In summary, Hardy and the two books (Chrystal and Goursat) roughly contemporary to his, all treat functions in approximately the same manner; as the relationship between two variables such that when one variable is known, the other can then be determined. This is different from all of the modern books examined (Adams, Clark and Rudin) that treat functions as a mapping between one set and another. The set-theoretic definition of a function is not something that was present in mathematical textbooks like *A Course of Pure Mathematics*.

4.3.4 The Introduction and Definition of a Limit

Hardy, after the chapters on real numbers, functions, and complex variables,[30] starts his discussion of limits with an entire chapter devoted to limits of functions of a positive integral[31] variable. Right away this is different from a modern text – the word sequence is not used. Functions of a positive integral variable can be functions of a continuous real variable taken just at positive integral values n or they can be functions that are only defined for positive integral values. Then, without motivation, a class is defined and it is "roughly speaking the aggregate or collection of all the entities or objects which possess a certain property" [56, p. 112]. A finite class has a finite number of members that can be ascertained by counting; an infinite class has not a finite number of members. Both types of integral functions above are such that "the values of the variables for which they are defined form an infinite class" [56, p. 113]. This definition is not used later in the book and it is not clear exactly why it was introduced; perhaps just to establish that the integral functions on which limits are being introduced are defined for an infinite number of points.

Next the phrase "tends to infinity" is carefully defined, with the idea of changing over

[30] the chapter on complex variables will not be discussed

[31] This is Hardy's terminology. We would perhaps expect to see limits of a positive integer variable

time being used in the same way that graphs were used when defining functions – as a matter of convenience. The statement n tends to infinity means then that "n is supposed to assume a series of values which increase beyond all limit" [56, p. 114]. After a discussion designed to convince the reader that $1/n$ gets close to 0 as n gets large, the formal definition of a limit of a function of an integer variable $\phi(n)$ is stated "The function $\phi(n)$ is said to tend to the limit l as n tends to ∞, if, however small be the positive number δ, $\phi(n)$ differs from l by less than δ for sufficiently large values of n; or if, however small be the positive number δ, we can determine a value n_0 corresponding to δ and such that $\phi(n)$ differs from l by less than δ for all values of n greater than or equal to n_0" [44, p. 118]. This discussion is leisurely, taking 10 pages to arrive at this definition of the limit of a function defined on the positive integers.

In the modern texts under consideration there are three approaches. First, Adams does not bother with sequences right away and after defining a function, proceeds to give a definition of a limit of a real valued function using standard $\epsilon\delta$ notation under the assumption the function is defined on a given interval. Much later, in chapter 8, an infinite sequence is taken as "a special case of the concept of function. The sequence $\{a_1, a_2, a_3, a_4, \cdots\}$ can be regarded as a function a whose domain is the set of positive integers and that takes the value $a(n) = a_n$ at each integer in the domain" [1, p. 442]. The limit of the sequence is then defined and it is very similar to Hardy's. The difference here is the order of presentation.

Second, Clark takes an approach very similar to Hardy. He introduce limits of sequences earlier, after the properties of real numbers, by giving an introduction to the idea, followed by numerical examples which results in the definition "Let $\{x_n\}$ be a given sequence of real numbers, and let a be a given real number. Then $\lim_{n\to\infty} x_n = a$ means that: for any given $\epsilon > 0$, there is a corresponding integer N (which may depend on ϵ) such that $|x_n - a| < \epsilon$ for every $n >= N$" [15, p. 17]. This is virtually the same definition and presentation as in Hardy except that sequences (using set notation) are used rather than functions of positive integral values.

Finally, Rudin's approach is to follow the real and complex number systems with a chapter on set theory and then to state the definition of convergent sequence in a metric space.[32] "A sequence $\{p_n\}$ in a metric space X is said to converge if there is an integer N such that $n \geq N$ implies that $d(p_n, p) < \epsilon$. (Here d denotes the distance in X). In this case

[32]A set possessing a distance function that satisfies 4 properties: $d(x, y) + d(y, z) \geq d(x, z) d(x, y) = d(y, x)$ $d(x, x) = 0 d(x, y) = 0 \Rightarrow x = y$

we also say the $\{p_n\}$ converges to p, or that p is the limit of $\{p_n\}$" [86, p. 41]. This is a much terser and more comprehensive definition than in Hardy, as befits a text aimed at a more advanced audience.

Chrystal introduces the idea of a sequence in the formation of the real numbers using "Cantor's theorem". He defines a sequence of rational numbers, given an arbitrary a_0, so that $a_n = a_0+p_1/10+\ldots+p_n/10^n, b_n = a_n+1/10^n$ where the $p_n s$ are chosen so that a_n and b_n are separated by a section (A, B). Depending on the $p_n s$, either a rational or irrational number is determined by finding enough $p_n s$ so that a_n is the greatest possible number in A. This sequence of $a_n s$, satisfies the following property (Cantor's theorem): "Given any positive rational number ϵ, however small, we can always find an integer ν such that $u_n - u_{n+r} < \epsilon$ when $n \geq \nu$, r being any positive integer whatsoever" [14, p. 102]. Such a sequence is a convergent sequence. This notion is generalized several pages later by removing the restriction that the numbers considered need to be rational and this results in the following definition: "the limit of the infinite sequence of real quantities $u_1, u_2, \ldots, u_n, \ldots (\Sigma)$, as a quantity u such that, if ϵ be any real quantity however small, then there exists always a positive integer ν such that $u_n - u < \epsilon$ when $n \geq \nu$ " [14, , p.108]. Chrystal ties together the convergence of a sequence with the construction of the real numbers making it difficult to understand what he means by a convergent sequence. Unlike Hardy, he provides no graphical aid to visualize what is happening and he provides no examples before moving on to the limit of a function of x where the definition of x is not specified.

Goursat reviews the definition of a convergent sequence in the chapter on infinite series: "In order that a sequence should be convergent, it is necessary and sufficient that, corresponding to any preassigned positive number ϵ, a positive integer n should exist such that the difference $s_{n+p} - s_n$ is less than ϵ in absolute value for any positive integer p" [28, p. 327]. He then provides the example of $1/n$.

In summary, Hardy's exposition is clearer than Chrystal's, and quite similar to a modern presentation with the exception that he presents the material more gradually and places more emphasis on functions of positive integral variables rather than using the term sequence with the $\{\}$ notation.

4.3.5 The Introduction of the Logarithmic and Exponential Function

As previously mentioned, Hardy was motivated to write *A Course of Pure Mathematics* in order to present an elementary account of the theory of the logarithm and exponential. So, how does Hardy introduce this theory? First, he draws an analogy between the necessity of introducing new functions to provide solutions to problems which cannot be solved with existing functions and the expansion of number from rational numbers to irrational and complex numbers. Integration has provided this type of motivation in the discovery of new functions since it was found that it was impossible to integrate some functions in terms of already known functions. The new function, then, is defined by the property that its derivative is given, such that the function itself is the integral of the given derivative.

After giving reasons why differentiating any rational, irrational, or trigonometric function will fail to produce $1/x$, Hardy refers the reader to his tract *The Integration of Functions of a Single Variable* and, in the first edition but not the ninth, the French version of Goursat's *Cours d'analyse* for proof that the integral of $1/x$ is a new function. Then, the new function, $\log x$, the logarithm of x, is defined by the equation

$$\log x = \int_1^x \frac{dt}{t}$$

After defining $\log x$, the graph of $\log x$ is shown, it is shown that $\log x$ is the only non-trivial solution of the functional equation $f(xy) = f(x) + f(y)$, and the manner in which $\log x$ tends to infinity is discussed.

Then e, a number of immense importance in analysis, is defined by the equation

$$1 = \int_1^e \frac{dt}{t}$$

This leads to the definition of the exponential function as the inverse of the logarithmic function. Then this e, defined here, is identified with a previously discussed limit (from chapter 4 on limits) that was provisionally denoted by e. The limit of

$$\lim_{n\to\infty} (1 + \frac{1}{n})^n$$

was shown to be greater than 2 and less than 3 and set equal to e such that

$$2 < e \leq 3.$$

More generally,

$$\lim_{n\to\infty} (1 + \frac{y}{n})^n = \lim_{n\to\infty} (1 - \frac{y}{n})^{-n} = e^y$$

That the function defined as the inverse of the logarithm is the same as the function resulting from the above limit "is of very great importance" and is proved twice.

The above is a summary of most of the content of the ninth chapter of *A Course of Pure Mathematics*. Chapter ten generalizes logarithmic, exponential, and circular functions to complex variables.

This is completely different from Chrystal who has, I feel, an incoherent description that covers more material which is interspersed between two volumes. For example, Chrystal introduces logarithms in chapter 21 of the first volume in a discussion of the exponential function $y = a^x$ where a is real and greater than 1. In a very confusing demonstration, Chrystal states that x, while not necessarily rational, can be considered rational of the form $\frac{m}{n}$ because "we can always find two commensurable values, $\frac{m}{n}$ and $\frac{m+1}{n}$ (where m and n are positive integers), between which x lies, and which differ from one another as little as we please" [14, p. 509]. This approximation allows him to claim that $y = a^x$ is a "continuous function of x susceptible of all positive values between 0 and $+\infty$" [ibid]. Under this assumption and motivated by the graph of the exponential function which "if we look at the matter from a graphical point of view, we see that the continuity of the graph means the continuity of y as a function of x, and also the continuity of x as a function of y" [14, p. 511], Chrystal claims "when we determine x as a function of y by means of the equation $y = a^x$, we obviously introduce a new kind of transcendental function into algebra ... the two equations, $y = a^x$, $x = \log_a y$ are thus merely different ways of writing the same functional relation" [14, p. 511]. This is followed by logarithm tables and instructions on their use!

The waters are further muddied in volume 2 in the chapter on exponential and logarithmic series where it is assumed that a convergent series expansion of a^x in ascending powers of x exists. The coefficients of these series are used to show that $L_{n=\infty}(1 + \frac{1}{n})^n$ is Napier's base or e. In terms of this notation, Hardy takes pains to point out the one should not write anything equal to infinity since that is meaningless and liable to confuse someone learning about limits and convergence. The next mention of the logarithm is to expand $\log(1 + x)$ in a series of ascending powers of x with the base understood to be e.

In Goursat, the situation is also very different from Hardy and is such that I presume that Goursat assumes the reader to be already familiar with exponential and logarithmic functions. The logarithm and its derivative are stated in a table on page 15 and at next mention, a theorem using Jacobians is used to show that the logarithm satisfies the functional equation $f(x) + f(y) = f(xy)$ and that this definition "might have led to the discovery of the fundamental properties of the logarithm had they not been known before the integral calculus" [28, p. 57]. Here, there is much more knowledge assumed (for example, partial differentiation) than in Hardy.

Unlike the situation for real numbers and for functions, Hardy's treatment of the logarithmic and exponential functions are very different from both Chrystal and Goursat. Particularly the treatment by Goursat could be seen as a reason why Hardy felt that a text was needed that would handle the theory of logarithms and exponential theory in an elementary and rigorous way.

To compare with the modern textbooks, Adams introduces the natural logarithm exactly as Hardy did, by giving examples to show that, so far, a function with derivative of $\frac{1}{x}$ has not been encountered. The presentation and justification in Hardy is much more thorough but the result is the same. The logarithmic function is introduced as a function whose derivative is $\frac{1}{x}$, its properties detailed and "since ln is increasing on its domain $(0, \infty)$, it is one-to-one there and so has an inverse function. For the moment, let us call this inverse function exp" [1, p.149] which gives

$$y = exp(x) \Leftrightarrow x = ln(y).$$

The level of detail that Hardy uses is absent, as are the proofs, but the basic idea of how to introduce the two functions is similar. In Clark, the number e is introduced as a limit and the derivative of the exponential and the logarithmic function are stated in the chapter on sequences, limits, and real numbers. Later, in chapter 5, Clark states that "a second method of defining e^x is to define $\log x$ first by means of an integral; then e^x can be defined as the inverse function of $\log x$. This method is quite simple and elegant, and is used in many modern calculus textbooks". The reader is referred to Apostol[33]. Then, the method Clark uses is to "due to Weierstrass, and is based on the elementary theory of power series. Thus we will *define*

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Historically this is backward, but logically it is acceptable and logical" [15, p. 182]. $Log(x)$ is defined as the inverse of the exponential function. So, the method ultimately chosen is different from that of Hardy but Hardy's method is outlined and, Clark says, is the one used

[33] T.M Apostol , Calculus, Vol. I, Blaisdell Publishing Company (1967).

in many modern texts. Rudin's approach is to define

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}$$

and then state and prove as a theorem the limit definition of e. Then, the exponential function is defined in the complex plane as

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

and it is shown that for all real x,

$$E(x) = e^x.$$

The function E is strictly increasing and differentiable, it has an inverse function, L, defined by

$$E(L(y)) = y, (y > 0).$$

This implies

$$L(y) = \int_1^y \frac{dx}{x}.$$

Rudin, like Clark, states that this last result is "quite frequently taken as the starting point of the theory of the logarithm and exponential function" [86, p. 166].

In summary, with the caveat that not all of the English elementary analysis texts that existed at Hardy's time have been examined, it is apparent from reading Goursat and Chrystal that either: 1) the theory of logarithms and exponentials were assumed to be already understood at an elementary level and the discussion starts at a higher level or, 2) the presentation is disjointed, not rigorous and difficult to understand. Hardy presents a very rigorous picture of the theory built from first principles based on defining the logarithm as the integral of $\frac{1}{x}$. This is a presentation that can still be seen in textbooks today. It is the method followed by Adams, with much less detail, and while not exactly what Clark and Rudin do, both refer to Hardy's method and state that it is a common starting point. More so here than in the discussion of real numbers or functions or limits, we see Hardy breaking with the past and setting the tone for the modern way of introducing the theory of logarithms and exponentials.

Chapter 5

Conclusion

The first decade of the twentieth century saw rapid change in British mathematics. New standards of rigour, new conceptions of the infinite and its place in mathematics, and a changing educational climate all impacted the type of mathematics studied and researched at Cambridge. G.H. Hardy, through his research, his writing and his activism, played a large role in effecting this change.

Hardy was educated at Cambridge at the end of the 19th century in a system dominated by the mathematical Tripos examinations. The impact of these examinations on textbooks, lecture topics and attendance, and research are widely thought to have been the cause of the poor performance of British mathematicians relative to their continental counterparts at that time.

Hardy educated himself in the latest continental mathematics citing Jordan's *Cours d'analyse* as a particularly important influence. Hardy then actively sought to abolish the Tripos examinations; he was unsuccessful but important reforms that severely restricted its effect were implemented in 1910.

Hardy's self-education using French and German textbooks and journal articles is seen in the three books he wrote between 1905 and 1910. Two of these books, *The Integration of Functions of a Single Variable* and *A Course of Pure Mathematics*, have been discussed in considerable detail and the third, *Orders of Infinity: The Infinitärcalcül of Paul du Bois-Reymond*, has been briefly mentioned. All of these books have the common feature that they include mathematical content developed in France or Germany - content which was often entirely new to Hardy's English speaking audience.

Hardy, in his first monograph, *The Integration of Functions of a Single Variable* chose an interesting approach. First he chose not to present indefinite integration in the standard way, which was and still is, as a collection of clever tricks. To do this, he relied on the earlier work of Liouville, work that had largely been ignored for the 70 years prior to Hardy's 1905 monograph. This brought the work of Liouville to the attention of others.

Hardy's methodical approach to integration may have been motivated by a desire to provide a complete theory, or at least as complete a theory as possible. It is clear that he thought a complete solution to integration of elementary functions in finite terms was unlikely to be found. Here he was wrong, as the subsequent work of Ritt, Trager, Rothstein, and Risch have shown, but the method with which he approached the problem turned out to be exactly what was required for the development of computer algebra systems in the second half of the twentieth century. For a *pamphlet*, written early in a career, at the beginning of a new series of tracts, *Integration* was a remarkable piece of work.

Hardy wrote *A Course of Pure Mathematics* in 1908; a textbook that has gone through eleven editions and is still in print 100 years later, a textbook to which other authors referred to during the fifty years following its publication with the assumption that others would have read it or have been familiar with it, a textbook that defined a first analysis course in Britain for seventy years following its publication [66], and a textbook that has influenced the presentation of analysis in textbooks used today.

A detailed examination of Hardy's presentation of real numbers, functions, limits of sequences, logarithmic and exponential theory in *A Course of Pure Mathematics*, and a comparison with contemporary and modern authors provides a more nuanced picture of what Hardy did. He most definitely, as he set out to do, provided a clear, rigorous introduction to the theory of logarithms and exponentials far superior to that of the contemporaneous textbooks I compared his to and provided what is still a perfectly reasonable introduction to the theory.

Hardy provided a rigorous, comprehensive, constructivist definition of the real numbers, which was available in other books of his time but the clarity of his prose makes his presentation very accessible. This material is typically no longer presented to students of the level Hardy was aiming at, but when it is included in an advanced analysis book, it is often done in a manner similar to Hardy's presentation. The main difference between Hardy and a modern presentation is the modern incorporation of set theory.

In fact, it is the lack of set theory in the definitions in *A Course of Pure Mathematics* that most sharply divides it from a more modern text. This is particularly evident in his definition of a function, which is a relationship between two continuous real variables rather than a mapping from one set to another. A first year university student of today who mastered the material in *A Course of Pure Mathematics* would, with small changes involving set theory, be well prepared for further study in analysis.

However, despite the lack of set theory in *A Course of Pure Mathematics* relative to textbooks of today, Hardy was one of a small group of British mathematicians researching and writing about set theory between 1900 and 1910. Set theory was developed in Germany in the second half of the 19^{th} century primarily by Georg Cantor and Richard Dedekind with Schönflies providing an early summary of Cantor's work.

The first British mathematicians to engage with this material were the Youngs, Hardy, Russell and Jourdain. The Youngs appear to have concentrated more on the point-set aspects of Cantor's theory and, for a variety of reason, failed to significantly impact developments in Britain.

Hardy, Russell and Jourdain were, I feel, the major figures in British set theory between 1900 and 1910 and each one was affected by the work of the other two through meetings, correspondence, and reading, reviewing and responding to one another's work. Hobson provided critical commentary and stimulated discussion particularly surrounding the Zermelo's Axiom of Choice and Russell's multiplicative axiom.

Hardy's five set theory papers, all published in the first decade of the 20^{th} century, have been examined in detail. Two of the papers show Hardy linking set theory with his early interest in the topics of analysis. Hardy used newly established results from set theory to prove general results in analysis regarding series convergence. The three other papers, on abstract set theory, demonstrate Hardy's interest in this foundational topic as an independent topic.

By 1910, there was not yet consensus on many aspects of set theory – for example, the well-ordering principle or the continuum hypothesis. Also, set theory as a language of mathematics was not yet present in mathematical textbooks. But this does not lessen the magnitude of the shift that occurred to the foundations of the mathematics during the previous ten years. In surprisingly warlike language (given that Hardy was a near-pacifist), Hardy captured this sentiment in 1929.

> The history of mathematics shows conclusively that mathematicians do not evacuate permanently ground which they have conquered once. There have been

> many temporary retirements and shortenings of the line, but never a general retreat on a broad front. We may be confident that, whatever the precise issue of current controversies, there will be no general surrender of the ground which Weierstrass and his followers have won. 'No one', as Hilbert says himself, 'shall chase us from the paradise that Cantor has created': the worst that can happen to us is that we shall have to be a little more particular about our clothes. [52, p. 5]

Hardy's career spanned a period of great change in British mathematics, a period of time during which British mathematics modernized. Hardy recognized this and, in 1934, wrote

> The most commonplace Cambridge mathematician now has forgotten the superstition that it is impossible to be "rigorous" without being dull, and that there is some mysterious terror in exact thought: now we go to the opposite extreme, and say that rigour is of secondary importance in analysis because it can be supplied, granted the right idea, by any competent professional. [53, p. 236]

This is a testament to the changes that Hardy wrought.

Bibliography

[1] Robert A. Adams. *Single-Variable Calculus.* Addison-Wesley: Ontario, 1983.

[2] Tom Archibald. Rigour in analysis. In Timothy Gowers, editor, *Princeton Companion to Mathematics*, pages 117–129. Princeton University Press: Princeton, 2007.

[3] Margherita Barile and Eric W. Weisstein. *Cantor Set.* From MathWorld - A Wolfram Web Resource http://mathworld.wolfram.com/CantorSet.html.

[4] B. Bollabás, A. Katok W. Fulton, F. Kirwan, P. Sarnak, B. Simon, and B. Totaro. *Cambridge Tracts in Mathematics.* From http://www.cambridge.org/uk/series/sSeries.asp?code=CTM.

[5] B. Bollobás. Forward. In J.E. Littlewood, editor, *Littlewood's Miscellany.* Cambridge University Press: Cambridge, 1986.

[6] Thomas John l'Anson Bromwich. *An Introduction to the Theory of Infinite Series.* Macmillan and Co. Limited: London, 1908.

[7] Manuel Bronstein. Integration of Elementary Functions. *J. Symbolic Comp.*, 9(2):117–173, 1990.

[8] R. Bunn. *Developments in the Foundations of Mathematics from 1870 to 1910*, chapter 6, pages 220–255. In Grattan-Guinness [32], 2000.

[9] John Burkhill. *The Lebesgue Integral.* Cambridge University Press, 1951.

[10] John Burkhill. *Hardy, Godfrey Harold*, pages 113–114. Volume 6 of Gillispie [27], 2008.

[11] John Burkhill. *Young, William Henry*, pages 572–573. Volume 14 of Gillispie [27], 2008.

[12] Georg Cantor. *Contributions to the Founding of Transfinite Numbers, article 1, 1895.* Dover Publications: New York, 1955.

[13] Georg Cantor. *Contributions to the Founding of Transfinite Numbers, article 2, 1897.* Dover Publications: New York, 1955.

[14] G. Chrystal. *Algebra, an Elementary Text-book. 7th ed.* Chelsea Publishing Company: New York, 1964.

[15] Colin W. Clark. *Elementary Mathematical Analysis*. Wadsworth: Belmont, California, 1982.

[16] John H. Conway and Richard Guy. *The Book of Numbers*. Springer-Verlag: New York, 1996.

[17] Joseph W. Dauben. *Georg Cantor. His Mathematics and Philosophy of the Infinite.* Princeton University Press: Princeton, New Jersey, 1979.

[18] Joseph W. Dauben. *The Development of Cantorian Set Theory*, chapter 5, pages 181–219. In Grattan-Guinness [32], 2000.

[19] Richard Dedekind. *Essays on the theories of numbers*. Dover Publications: New York, 1963.

[20] Janet Delve. The College of Preceptors and the Educational Times: Changes for British mathematics in the mid-nineteenth century. *Historia Mathematica*, 30:140–172, 2003.

[21] Jose Ferreirós. *Labyrinth of Thought, A History of Set Theory and Its Role in Modern Mathematics*. Birkhauser Verlag: Basel-Boston-Berlin, 2007.

[22] A.R. Forsyth. Edward John Routh. *Proceedings of the London Mathematical Society*, 5, 1906.

[23] A.R. Forsyth. Old Tripos Days at Cambridge. *The Mathematical Gazette*, 19(234):162–179, 1935.

[24] Lars Garding. Marcel Reisz in memoriam. *Acta Mathematica*, 124:I–IX, 1970.

[25] Keith O. Geddes, Stephen R. Czapor, and George Labahn. *Algorithms for Computer Algebra*. Kluwer Academic Publishers: Boston, 1992.

[26] A. O. Gelfond. *Transcendental and Algebraic Numbers*. Dover Publications: New York, 1960.

[27] Charles Coulston Gillispie, editor. *Complete Dictionary of Scientific Biography.* Charles Scribner's Sons, Detroit, 2008.

[28] Edouard Goursat. *A Course in Mathematical Analysis*. Dover Publications: New York, 1904.

[29] Ivor Grattan-Guinness. *Dear Russell-Dear Jourdain*. Duckworth: London, 1977.

[30] Ivor Grattan-Guinness. How Bertrand Russell Discovered his Paradox. *Historia Mathematica*, 5:127–137, 1978.

[31] Ivor Grattan-Guinness. *The Emergence of Mathematical Analysis and its Foundational Progress, 1780-1880*, chapter 3, pages 94–145. In [32], 2000.

[32] Ivor Grattan-Guinness, editor. *From the Calculus to Set Theory, 1630-1910. An Introductory History.* Princeton University Press, 2000.

[33] Ivor Grattan-Guinness. *The Search for Mathematical Roots 1870-1940 Logics, Set Theories and the Foundations of Mathematics from Cantor through Russell to Gödel.* Princeton University Press: Princeton and Oxford, 2000.

[34] Ivor Grattan-Guinness and Magnus Gustav Mittag-Leffler. Materials for the History of Mathematics in the Institut Mittag-Leffler. *Isis*, 62(3):363–374, 1971.

[35] Nicholas Griffin and Albert C. Lewis. Bertrand Russell's Mathematical Education. *Notes and Records of the Royal Society of London*, 44(1):51–71, 1990.

[36] Angelo Guerraggio and Pietro Nastasi. *Italian Mathematics Between the Two World Wars.* Birkhauser Verlag: Basel-Boston-Berlin, 2005.

[37] G.H. Hardy. The Philosophy of Mathematics, a review of The Principles of Mathematics by Bertrand Russell. *Times Literary Supplement*, 88:263, 1903.

[38] G.H. Hardy. The cardinal number of a closed set of points. *Messenger of Mathematics*, 33:67–69, 1904.

[39] G.H. Hardy. A general theorem concerning absolutely convergent series. *Proceedings of the London Mathematical Society*, 1(2):285–290, 1904.

[40] G.H. Hardy. A theorem concerning the infinite cardinal numbers. *Quarterly Journal of Mathematics*, 35:87–94, 1904.

[41] G.H. Hardy. *The Integration of Functions of a Single Variable.* Cambridge Tracts in Mathematics and Mathematical Physics. Cambridge University Press: Cambridge, 1905.

[42] G.H. Hardy. The Continuum and the Second Class Number. *Proceedings of the London Mathematical Society*, 4(2):438–445, 1907.

[43] G.H. Hardy. Review of A First Course in the Differential and Integral Calculus. by W.F. Osgood. *Mathematical Gazette*, 4:307–309, 1907.

[44] G.H. Hardy. *A course of pure mathematics 1st ed.* Cambridge University Press: Cambridge, 1908.

[45] G.H. Hardy. The "Fall" of Euclid, a review of The Thirteen Books of Euclid's Elements by T.L. Heath. *Times Literary Supplement*, 409:420–421, 1909.

[46] G.H. Hardy. The ordinal relations of the terms of a convergent sequence. *Proceedings of the London Mathematical Society*, 8(2):295–300, 1910.

[47] G.H. Hardy. The New Symbolic Logic, a review of Principia Mathematica by A.N. Whitehead and Bertrand Russell, Vol. I. *Times Literary Supplement*, 504:321–322, 1911.

[48] G.H. Hardy. Properties of logarithmico-exponential functions. *Proceedings of the London Mathematical Society*, 2(10):54–90, 1912.

[49] G.H. Hardy. Camille Jordan, 1838-1922. *Proceedings of the Royal Society*, 104(A), 1922.

[50] G.H. Hardy. Review of 'The Theory of Functions of a Real Variable and the Theory of Fourier's Series' by Prof. E.W. Hobson. *Nature*, 109:435–437, 1922.

[51] G.H. Hardy. The Case Against the Mathematical Tripos. *Mathematical Gazette*, 13:61–71, 1926.

[52] G.H. Hardy. Mathematical Proof. *Mind*, 38:1–25, 1929.

[53] G.H. Hardy. Ernest William Hobson. *Journal of the London Mathematical Society*, 9:225–237, 1934.

[54] G.H. Hardy. *A Mathematicians Apology.* Cambridge University Press: Cambridge, 1940.

[55] G.H. Hardy. William Henry Young. *Journal of the London Mathematical Society*, 17:218–237, 1942.

[56] G.H. Hardy. *A course of pure mathematics 9th ed.* Cambridge University Press: Cambridge, 1945.

[57] G.H. Hardy. *A course of pure mathematics 11th ed.* Cambridge University Press: Cambridge, 2008.

[58] John Michael Heard. *The Evolution of the Pure Mathematician in England, 1850-1920.* PhD thesis, Imperial College, London, 2004.

[59] David Hilbert. Gedachtnissrede auf H. Minkowski. *Math. Annalen*, 68:445–471, 1910.

[60] E.W. Hobson. On the General Theory of Transfinite Numbers and Order Types. *Proceedings of the London Mathematical Society*, s2-3:170–188, 1905.

[61] Wallie Abraham Hurwitz. Review of G.H. Hardy, Orders of Infinity. The Infinitärcalcul of Paul Du Bois-Reymond. *Bulletin of the American Mathematical Society*, 21(4):201–202, 1915.

[62] Philip E. B. Jourdain. On the Question of the Existence of Transfinite Numbers. *Proceedings of the London Mathematical Society*, pages 266–283, 1906.

[63] Philip E. B. Jourdain. Review of 'A Course of Pure Mathematics', first and second edition by G.H Hardy and 'A Course of Modern Analysis' by E.T. Whittaker, first and second edition. *Mind, New Series*, 25(100):525–533, 1916.

[64] Philip E. B. Jourdain. *Introduction to Contributions to the Founding of the Theory of Transfinite Numbers*. Dover Publications: New York, 1955.

[65] Toni Kasper. Integration in Finite Terms: The Liouville Theory. *Mathematics Magazine*, 53:195–201, 1980.

[66] Tom Körner. *Forward*, pages v–xi. In [57], 2008.

[67] J. L. Dr. J.G. Leathem. *Nature*, 111:437–437, 1923.

[68] J.E. Littlewood. *Littlewood's Miscellany*. University Press: Cambridge, 1986.

[69] Jesper Lützen. *Joseph Liouville 1809-1882 Master of Pure and Applied Mathematics*. Springer-Verlag: New York, 1990.

[70] Jesper Lützen. The foundations of analysis in the 19th century. In H. Jahnke, editor, *A History of Analysis*. American Mathematical Society: USA, 2003.

[71] Jagdish Mehra and Milton A. Kimball. *Climbing the Mountain: The Scientific Biography of Julian Schwinger*. Oxford University Press: USA, 2003.

[72] Gregory H.. Moore. *Zermelo's Axiom of Choice Its Origins, Development and Influence*. Springer-Verlag: New York, Heidelberg, Berlin, 1982.

[73] E.H. Neville. Correspondence. *The Mathematical Gazette*, 25(266):217, 1941.

[74] J.J O'Connor and E.F. Robertson. G.H. Hardy. *MacTutor History of Mathemtics*, 2008.

[75] Karl Pearson. Old Tripos Days at Cambridge, as Seen from Another Viewpoint. *The Mathematical Gazette*, 20(237):27–36, 1936.

[76] H.T.H. Piaggio. Three Sadleirian Professors: A.R. Forsyth, E.W. Hobson and G.H. Hardy. *The Mathematical Gazette*, 15(215):461–465, 1931.

[77] Richard Rankin. Miscellaneous papers, introduction. In A Committee appointed by the London Mathematical Soceity, editor, *Collected Papers of G.H. Hardy, Volume 7*, pages 417–418. Clarendon Press: Oxford, 1979.

[78] Robert H. Risch. On the Integration of Elementary Functions which are built up using Algebraic Operations. *Sys. Dev. Corp.*, Report SP-2801/002/00, 1968.

[79] Robert H. Risch. The Problem of Integration in Finite Terms. *Transactions of the American Mathematical Society*, 139:162–189, 1969.

[80] Robert H. Risch. The solution of the problem of integration in finite terms. *Bulletin of the American Mathematical Society*, 76:605–608, 1970.

[81] Patricia Rothman. Grace Chisholm Young and the Division of Laurels. *Notes and Records of the Royal Society of London*, 50(1):89–100, 1996.

[82] M. Rothstein. *Aspects of Symbolic Integration and Simplification of Exponential and Primitive Functions.* PhD thesis, Univ. of Wisconsin, Madison, 1976.

[83] W.W. Rouse Ball. *The History of Mathematics at Cambridge.* Cambridge University Press, 1889.

[84] W.W. Rouse Ball. *A Short Account of the History of Mathematics, 4th edition.* MacMillan and Co: London, 1908.

[85] W.W. Rouse Ball. The Cambridge School of Mathematics. *The Mathematical Gazette*, 6(99):311–323, 1912.

[86] Walter Rudin. *Principles of Mathematical Analysis, 2nd ed.* McGraw Hill: New York, 1964.

[87] Bertrand Russell. *Principles of Mathematics.* Merchant Books: London, 1903.

[88] Bertrand Russell. On Some Difficulties in the Theory of Transfinite Numbers and Order Types. *Proceedings of the London Mathematical Society*, pages 29–53, 1904.

[89] Bertrand Russell. *Portraits from Memory and other essays.* George Allen and Unwin: London, 1958.

[90] Bertrand Russell. *My Philosophical Development.* Simon and Schuster: New York, 1959.

[91] Bertrand Russell. *On Education, especially in early childhood, 2nd ed.* Unwin Books: London, 1960.

[92] F. Smithies. Introduction to papers on the integral calculus. In *Collected Papers of G.H. Hardy*, volume 5. Oxford University Press: London, 1972.

[93] E.C. Titchmarsh. Godfrey Harold Hardy 1877-1947. *Obituary Notices of the Fellows of the Royal Society*, 6:447–458, 1949.

[94] B. Trager. Algebraic Factoring and Rational Function Integration. In R.D. Jenks, editor, *Proc. SYMSAC '76*, pages 219–226. ACM Press, 1976.

[95] B. Trager. *Integration of Algebraic Functions.* PhD thesis, Dept. of EECS, M.I.T., 1984.

[96] Sabetai Unguru. On the Need to Rewrite the History of Greek Mathematics. *Archive for the History of the Exact Sciences*, 15:67–114, 1975.

[97] Eric W. Weisstein. *Cantor-Dedekind Axiom*. From MathWorld - A Wolfram Web Resource http://mathworld.wolfram.com/Cantor-DedekindAxiom.html.

[98] Eric W. Weisstein. *Elementary Function*. From MathWorld - A Wolfram Web Resource http://mathworld.wolfram.com/ElementaryFunction.html.

[99] Sylvia Wiegand. Grace Chisholm Young and William Henry Young: A Partnership of Itinerant British Mathematicians. In Helena Pycior, Nancy Slack, and Pnina Abir-Am, editors, *Creative Couples in the Sciences*. Rutgers University Press: New Brunswick, New Jersey, 1996.

[100] H.H. William. A History of Teachers of Mathematics. *The Mathematical Gazette*, 55(393):272–283, 1971.

[101] W.H. Young. Sets of Intervals on the Straight Line. *Proceedings of the London Mathematical Society*, s1-35:245–268, 1902.

[102] W.H. Young. The Progress of Mathematical Analysis in the Twentieth Century. *Proceedings of the London Mathematical Society*, s2-24:421–434, 1926.

[103] W.H. Young. *The Theory of Sets of Points, 2nd edition*. Chelsea Publishing Company: New York, 1972.