133 II

MATHEMATICAL
Surveys and Monographs

Volume 37

# Structural Properties of Polylogarithms

Leonard Lewin
Editor

American Mathematical Society
Providence, Rhode Island

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 39B50, 33A70, 30D05, 19F27; Secondary 11F67, 39B70, 51M20, 57R20.

**Library of Congress Cataloging-in-Publication Data**

Structural properties of polylogarithms/Leonard Lewin, editor.
p. cm.—(Mathematical surveys and monographs, ISSN 0076-5376; v. 37)
Includes bibliographical references and index.
ISBN 0-8218-1634-9
1. Logarithmic functions. I. Lewin, Leonard, 1919– . II. Series.
QA342.S77 1991 91-18172
512.9'22—dc20 CIP



The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability. ♾
This publication was typeset using $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX,
the American Mathematical Society's TEX macro system.
10 9 8 7 6 5 4 3 2 1 95 94 93 92 91

# Contents

# Preface

As editor of this monograph on polylogarithms I would like to take the liberty of commencing with a few personal reminiscences. I first encountered the dilogarithm function many years ago in high school; it was a fascinating discovery for me, and it initiated a romance that has lasted almost sixty years. For the dilogarithm, the transition from its standing as a curious mathematical oddity to its current status as an important element in the fabric of modern mathematical structure began about fifteen years ago with Bloch's studies on its applications in algebraic $K$-theory and algebraic geometry. Since then, the pace of discovery has quickened dramatically. In 1980, when I was in the throes of completing my "Polylogarithms and Associated Functions," I became dimly aware that the handful of peculiar numerical identities that had been known since the time of Euler and Landen were, in fact, just the tip of an iceberg of unlimited extent. Thus emerged the new discoveries on cyclotomic equations and their related polylogarithmic "ladders"—a nomenclature that came to me in a dream, after much chewing over of other, more artificial, verbal constructs. Ten years of development in this arena, conducted mostly by the methods of classical analysis with the help of number-crunching computers, ran parallel with other, and more important, discoveries in diverse branches of abstract algebra and algebraic geometry. The confluence of these two streams of thought in the last few years, due to the work of several mathematicians, but particularly to studies of Browkin in Poland and Zagier in Germany, has lead to the present synthesis which I have tried to present in this timely, I hope, monograph.

One of the biggest problems has been the pace of new research. It is obviously extremely difficult to produce a book that is current when new discoveries are taking place all the time and making already-written material partially outdated—though it is also a sign of a very flourishing field when things go this way. During the approximately twelve months that the book has been in active preparation many new discoveries were made. I have endeavored to keep the material up-to-date by the last minute inclusion of two appendices: one on a special workshop on polylogarithms held in November 1990, and one on very recent discoveries on the relation of functional equations to polylogarithmic ladders, Dedekind's zeta function; and including the remarkable discovery by D. Zagier and H. Gangl at the Max-Planck-Institut für Mathematik of a two-variable functional equation for the hexalogarithm—the first significant advance in this area since Kummer's work of 150 years ago. In my earlier (1958) book on dilogarithms, talking about the difficulty of making much further progress in this area, I had written "But the complexity of the present results makes a completely new approach imperative if much progress is to be made." It is now clear what this new approach is entailing: on the one hand the structural analysis arising from algebraic $K$-theory and related fields; and on the other the extensive use of computers, both for high precision numerical work and also for machine computation using symbolic logic. It is doubtful that many of the new and interesting formulas could have been found by hand alone; powerful computer programs are becoming almost as important as mathematical skills and the ability to generate new constructive conjectures.

This book could not have been written without the splendid help and cooperation of the several contributors who gave generously of their time and effort. Many helpful suggestions and contacts were made. I would particularly like to thank Richard Hain for his assistance in the compilation of the bibliography, Don Zagier for his extensive up-to-date appendix, and Han Sah and Robert MacPherson for their report on the recent polylogarithm workshop.

Authors have very individual styles of writing and it is not practical, for the purpose of uniform presentation, to constrain them into one common pattern of text organization. Even so, I think the overall volume has not suffered from any ensuing tendency to be "patchy," and I hope that, the disparate contributions notwithstanding, the material as a whole is sufficiently coherent to give the entire work the integrity that I, as editor, have sought.

Most authors have written their chapters in the absence of knowing in detail what others were writing. This has given rise to a small amount of redundancy which I have not thought fit to try to remove; I do not think the work has suffered in any way from this. Rather, it has been interesting to see how similar ideas have arisen independently and received corresponding treatment. The whole subject is now in a state of rapid transition; even as I write, new discoveries vie for admission. With reluctance I have had to call a halt to the inclusion of a flood of new material. It will be fascinating to see what further developments the coming decade will bring. Don Zagier once wrote that "the dilogarithm is the only mathematical function with a sense of humor." As this subject matures and gets more important, and more serious, I hope it manages to retain its once light-hearted beginnings. Its ability over the years to attract and hold the interest of so many mathematicians, many of them of the finest caliber, has been outstanding. I hope that its capacity for fruitful exploration will continue unabated for a long time to come.

Leonard Lewin

January 1991

# Acknowledgments

Much credit for the preparation of this volume is due the various contributors, who, together with their affiliations, are listed on the following pages.

Some of the formulas on ladders had already appeared earlier in the literature, and acknowledgment is gratefully made to the publishers for permission to use material from some of their publications of the past decade. In particular, credit is due to the Academic Press, publisher of the *Journal of Number Theory*, for permission to use material from: *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory, **19** (1984), 345–373 and *The polylogarithm in algebraic number fields*, J. Number Theory **21** (1985), 214–244. Credit is also due to Birkhäuser Verlag AG, publisher of *Aequationes Mathematicae*, for permission to use material from: *The order independence of the polylogarithmic ladder structure*, Aequationes Math. **30** (1986), 1–20; *Polylogarithmic functional equations*, Aequationes Math. **31** (1986), 223–242; *The polylogarithm in the field of two irreducible quintics*, Aequationes Math. **31** (1986), 315–321; *Polylogarithms in the field of omega*, Aequationes Math. **33** (1987), 23–45; and *Supernumary polylogarithmic ladders and related functional equations*, Aequationes Math. **39** (1990), 210–253.

Much gratitude is also due to the reviewers of the original proposal for this book for making helpful suggestions, many of which have been incorporated into the current version.

# List of Contributors

Mohamed D. Abouzahra, Ph.D.
MIT Lincoln Laboratory, Lexington, MA 02173, USA
Spencer Bloch, Ph.D.
Department of Mathematics, University of Chicago, Chicago, IL 60637, USA
Jerzy Browkin, Ph.D.
Institute of Mathematics, Warsaw University, ul. Banacha 2, PL-00-913, Warsaw 59, Poland
Richard M. Hain, Ph.D.
Department of Mathematics, University of Washington, Seattle, WA 98195, USA
Ruth Kellerhals, Ph.D.
Max-Planck-Institut für Mathematik, Gottfried-Claren-Straße 26, 5300 Bonn 3, Germany
Leonard Lewin, D.Sc.
Professor Emeritus, Campus Box 425, University of Colorado, Boulder, CO 80309, USA
John H. Loxton, Ph.D.
Head of School of Mathematics, Macquarie University, NSW 2109, Australia
Robert MacPherson, Ph.D.
Department of Mathematics, MIT, Cambridge, MA 02139, USA
Gary A. Ray, Ph.D.
University of Washington, Seattle, WA 98195, USA (currently at Boeing High Technology Center, Seattle, WA 98124, USA)
C. Han Sah, Ph.D.
Department of Mathematics, State University of New York at Stony Brook, Stony Brook, NY 11794, USA
Gerd Wechsung, Ph.D.
Prorektor für Naturwissenschaft und Technik, Friedrich-Schiller-Universität, 6900 Jena, Germany

Zdzislaw Wojtkowiak, Ph.D.
Institut des Hautes Etudes Scientifiques, 35 Route de Chartres, 91440 Bures-Sur-Yvette, France. (Formerly at Max-Planck-Institut für Mathematik, Gottfried-Claren-Straße 26, 5300 Bonn 3, Germany)
Don Zagier, Ph.D.
Max-Planck-Institut für Mathematik, Gottfried-Claren-Straße 26, 5300 Bonn 3, Germany

CHAPTER 1

# The Evolution of the Ladder Concept

L. LEWIN

## 1.1 Early history.

1.1.1. *The dilogarithm.* This function was first studied by Euler [1] and Landen [2] in the second half of the eighteenth century. Known as the *Euler dilogarithm*, it is defined by the series

$$\mathrm{Li}_2(z) = \sum_{r=1}^{\infty} z^r/r^2, \qquad |z| \le 1. \tag{1.1}$$

The range can be extended outside the unit circle by using the integral formulation

$$\mathrm{Li}_2(z) = -\int_0^z \log(1-t)\,dt/t. \tag{1.2}$$

For $z$ real and greater than unity the dilogarithm is complex. The earlier notation $\mathrm{dl}(z)$, which is not capable of extension to higher orders, is no longer in use. The function is transcendental, and Chudnovsky [3] has recently proved that $\mathrm{Li}_2(z)$ is irrational when $z$ is rational.

1.1.2. *Simple functional equations.* By using (1.2) Euler readily obtained the simple single-variable functional equations

$$\mathrm{Li}_2(z) + \mathrm{Li}_2(-z) = \tfrac{1}{2}\mathrm{Li}_2(z^2), \tag{1.3}$$

$$\mathrm{Li}_2(-z) + \mathrm{Li}_2(-1/z) = 2\,\mathrm{Li}_2(-1) - \tfrac{1}{2}\log^2(z), \tag{1.4}$$

$$\mathrm{Li}_2(z) + \mathrm{Li}_2(1-z) = \mathrm{Li}_2(1) - \log z \log(1-z). \tag{1.5}$$

In the above, $\log^2(z)$ means $[\log_e(z)]^2$. The first two are known, respectively, as the duplication and inversion relations. From them one obtains, by taking $z = e^{i\pi}$,

$$\mathrm{Li}_2(1) = \pi^2/6 = \zeta(2); \qquad \mathrm{Li}_2(-1) = -\pi^2/12, \tag{1.6}$$

where the $\pi^2$ comes from $-\log^2(e^{i\pi})$. These are, of course, well-known results. Lesser known is Euler's formula obtained by taking $z = 1/2$ in (1.5),

$$\mathrm{Li}_2(\tfrac{1}{2}) = \pi^2/12 - \tfrac{1}{2}\log^2(\tfrac{1}{2}). \tag{1.7}$$

This is the forerunner of the ladder results developed later.

1.1.3. *Landen's functional equation.* A companion result to (1.5), due to Landen, is

$$\mathrm{Li}_2(z) + \mathrm{Li}_2[-z/(1-z)] = -\tfrac{1}{2}\log^2(1-z). \tag{1.8}$$

Defining the quantity $\rho = (\sqrt{5}-1)/2$, (the inverse of the golden ratio), which is the solution in $(0, 1)$ of

$$u^2 + u = 1, \tag{1.9}$$

Landen, by taking $z = \rho^2$, $1 - z = \rho$, and using (1.3), (1.5) and (1.8), was able to show

$$\mathrm{Li}_2(\rho) = \pi^2/10 - \log^2\rho, \tag{1.10}$$

$$\mathrm{Li}_2(\rho^2) = \pi^2/15 - \log^2\rho. \tag{1.11}$$

These results are the only ones known (apart from their inversions using (1.4), and $\mathrm{Li}_2(-\rho)$ from (1.3)) for which the dilogarithm can be expressed directly in terms of simpler functions, in closed form.

1.1.4. *Landen's trilogarithm results.* Extending the definition to the third-order function, Landen examined

$$\mathrm{Li}_3(z) = \sum_{r=1}^{\infty} z^r/r^3 = \int_0^z \mathrm{Li}_2(t)\,dt/t. \tag{1.12}$$

Utilizing both (1.5) and (1.8) he showed that

(1.13)
$$\mathrm{Li}_3\left(\frac{-z}{1-z}\right) + \mathrm{Li}_3(z) + \mathrm{Li}_3(1-z) = \zeta(3) + \zeta(2)\log(1-z) - \tfrac{1}{2}\log(z)\log^2(1-z) + \tfrac{1}{2}\log^3(1-z).$$

From this result he obtained

$$\mathrm{Li}_3(\tfrac{1}{2}) = \tfrac{7}{8}\zeta(3) + \tfrac{1}{2}\zeta(2)\log(\tfrac{1}{2}) - \tfrac{1}{6}\log^3(\tfrac{1}{2}) \tag{1.14}$$

and

$$\mathrm{Li}_3(\rho^2) = \tfrac{4}{5}\zeta(3) + \tfrac{4}{5}\zeta(2)\log\rho - \tfrac{2}{3}\log^3\rho. \tag{1.15}$$

These are the only known formulas of this type for $\mathrm{Li}_3$, and none are known for the higher orders. The way in which two functional equations at the second order come together to yield one at the third order is noteworthy. Here it is just a simple integrability requirement, but we shall later encounter it in a much wider context.

**1.2. Functional equations.**

1.2.1. *Abel's equation.* A two-variable equation, usually attributed to Abel [4], and easily verified by differentiation, is

$$\mathrm{Li}_2\left[\frac{x}{1-x}\cdot\frac{y}{1-y}\right] = \mathrm{Li}_2\left[\frac{x}{1-y}\right] + \mathrm{Li}_2\left[\frac{y}{1-x}\right] - \mathrm{Li}_2(x) - \mathrm{Li}_2(y) - \log(1-x)\log(1-y). \tag{1.16}$$



An equivalent formula, published by Spence [5] some twenty years earlier, seems to be relatively poorly known. A modified form, found by utilizing (1.8) in (1.16) is

$$\mathrm{Li}_2\left[\frac{x}{1-x}\cdot\frac{y}{1-y}\right] = \mathrm{Li}_2\left[\frac{x}{1-y}\right] + \mathrm{Li}_2\left[\frac{y}{1-x}\right] + \mathrm{Li}_2\left[\frac{-x}{1-x}\right] + \mathrm{Li}_2\left[\frac{-y}{1-y}\right] + \frac{1}{2}\log^2\left[\frac{1-x}{1-y}\right]. \tag{1.17}$$

1.2.2. *Alternative formulations.* There are a number of equations, reducible to Abel's, associated with different researchers of the nineteenth century.

From (1.17), putting $x$ for $x/(1-y)$ and $y$ for $y/(1-x)$, we obtain a form due to Hill [6],

$$\mathrm{Li}_2(xy) = \mathrm{Li}_2(x) + \mathrm{Li}_2(y) + \mathrm{Li}_2\left[-x\cdot\frac{1-y}{1-x}\right] + \mathrm{Li}_2\left[-y\cdot\frac{1-x}{1-y}\right] + \frac{1}{2}\log^2\left[\frac{1-x}{1-y}\right]. \tag{1.18}$$

From (1.16), with $1-x$ for $x$, and using (1.5), we obtain a form due to Schaeffer [7],

$$\mathrm{Li}_2\left[\frac{y(1-x)}{x(1-y)}\right] = \mathrm{Li}_2(x) - \mathrm{Li}_2(y) + \mathrm{Li}_2(y/x) + \mathrm{Li}_2\left[\frac{1-x}{1-y}\right] - \zeta(2) + \log x\log\left[\frac{1-x}{1-y}\right]. \tag{1.19}$$

Finally, if in Hill's equation we take $-x(1-y)/(1-x)$ and $-(1-y)/y(1-x)$ for $x$ and $y$ we get a structure due to Kummer [8],

$$\mathrm{Li}_2\left[\frac{x(1-y)^2}{y(1-x)^2}\right] = \mathrm{Li}_2\left[-x\cdot\frac{1-y}{1-x}\right] + \mathrm{Li}_2\left[-\frac{1}{y}\cdot\frac{1-y}{1-x}\right] + \mathrm{Li}_2\left[\frac{x(1-y)}{y(1-x)}\right] + \mathrm{Li}_2\left[\frac{1-y}{1-x}\right] + \frac{1}{2}\log^2(y). \tag{1.20}$$

The structure of Kummer's form seems to be the most versatile, and he succeeded in extending his method of developing functional equations as far as the fifth order. For this reason we shall here, as a convenience, refer to all such equations as Kummer's, the assignation of different variants to other authors notwithstanding.

1.2.3. *Rogers' multi-variable equation.* Using the notation $\mathrm{L}(z)$, Rogers [9] defined a function closely related to the dilogarithm. We shall later generalize his function to higher orders, and, in anticipation, include a subscript 2 in his notation. Rogers' function (and its generalization) has the property of removing the logarithmic terms [10] from functional equations like Kummer's. Define

$$L_2(z) = -\frac{1}{2}\int_0^z \left[\frac{\log(1-t)}{t} + \frac{\log(t)}{1-t}\right] dt = \mathrm{Li}_2(z) + \frac{1}{2}\log(z)\log(1-z). \tag{1.21}$$

Rogers generated for this function (and therefore equivalently for the dilogarithm) a functional equation with $m$ variables and involving $(m^2+1)$ transcendental terms. With $m=2$ the equation is reducible to Kummer's, but for larger $m$ the variables involve the roots of a polynomial of degree $m$, and seemingly the general form, as was pointed out by Rogers, cannot be recovered from this special case. Recently, however, research by Dupont **[11]** indicates that this conclusion is not correct, though the route to develop the construction does not appear from his analysis. This aspect, of *accessibility* of certain formulas from Kummer's, appears again and again in various guises. In some instances the inaccessible nature of a particular result may turn out to be an artifact; in other cases it may be genuine. The matter is currently unresolved.

However, recently, G. Ray (see Chapter 7), using generalizations of Rogers' multi-variable equation, and H. Gangl, using multiple applications of Abel's equation, have succeeded in proving analytically some of the heretofore inaccessible results.

**1.3. More recent numerical results.**

1.3.1. *Formulas of Coxeter.* The results of Euler and Landen in equations (1.7), (1.10), (1.11), (1.14), and (1.15) were the only such expressions recorded until Coxeter **[12]**, in a 1935 paper, obtained equations in which dilogarithms of *several* arguments, each a power of $\rho=(\sqrt{5}-1)/2$, were developed from the properties of a certain infinite series. Three of his results can be put in the form

$$\mathrm{Li}_2(\rho^6)=4\,\mathrm{Li}_2(\rho^3)+3\,\mathrm{Li}_2(\rho^2)-6\,\mathrm{Li}_2(\rho)+7\pi^2/30, \tag{1.22}$$

$$\begin{aligned}\mathrm{Li}_2(\rho^{12})=2\,\mathrm{Li}_2(\rho^6)&+3\,\mathrm{Li}_2(\rho^4)+4\,\mathrm{Li}_2(\rho^3)\\ &-6\,\mathrm{Li}_2(\rho^2)+\pi^2/10,\end{aligned} \tag{1.23}$$

$$\mathrm{Li}_2(\rho^{20})=2\,\mathrm{Li}_2(\rho^{10})+15\,\mathrm{Li}_2(\rho^4)-10\,\mathrm{Li}_2(\rho^2)+\pi^2/5. \tag{1.24}$$

Although one cannot isolate the dilogarithm of a single power of $\rho$ from these expressions, they can be seen as a natural generalization of (1.10) and (1.11): these latter can be considered as degenerate special cases of multi-term formulas in which the number of transcendental terms reduces to unity.

The above expressions of Coxeter are the earliest published results of a structure that was later to be designated a *ladder.* A number of properties appear on inspection of the above forms. Calling the highest powers, appearing on the left, the *index*, it is apparent that

(i) the powers of $\rho$ on the right run through the factors of the index,
(ii) the coefficients associated with these powers appear closely related to the index divided by the corresponding factors,
(iii) the coefficient of $\pi^2$ is a simple rational.

We shall see later in what way these observations will need to be modified



in more general cases, but they provided excellent guidelines for the initial research on ladders.

1.3.2. *Formulas of Watson.* In a 1937 paper, Watson **[13]** examined the dilogarithm of arguments involving the roots of the cubic

$$x^3+2x^2-x-1=0. \tag{1.25}$$

Denoting the roots by $\alpha$, $-\beta$ and $-1/\gamma$, (with $\alpha=\frac{1}{2}\sec(2\pi/7)$), then all three quantities $\alpha$, $\beta$ and $\gamma$ lie in $(0, 1)$. By applying Euler's and Landen's functional equations a number of times over, Watson deduced a relation in which $\mathrm{Li}_2(\alpha)$ and $\mathrm{Li}_2(\alpha^2)$ were both accumulated seven times, resulting in

$$\mathrm{Li}_2(\alpha)-\mathrm{Li}_2(\alpha^2)=\pi^2/42+\log^2\alpha. \tag{1.26}$$

Related equations for the other two roots were similarly derived.

Although these results are accessible from Kummer's functional equation, they are not *obviously* so. Watson indicated that he had long suspected the existence of certain results, and although his eventual proof is easy enough to follow, it is clear that it was not all that easy to come by. This comment is germane to the many later results that have been found without, to-date, any analytic proof. Designating them as "inaccessible" may turn out to be only an interim labeling. In fact, an interesting counter-example is Coxeter's (1.24) of index 20. This had long been believed to be inaccessible (on the basis of many failed tries), but has now been shown, using directions furnished by Dupont **[11]**, to be capable of being derived from Hill's equation, (1.18). An outline of the proof is given in Chapter 4.

It may be noted that when the relationship (1.25) involving powers of $x$ is rearranged, for example as $1-x=x^2/(1+x)^2$, and put into the functional equations, the process does not distinguish between the various roots of the cubic. Although manipulation of the final result to put it in convenient form may be necessary, it is clear that if a ladder result exists for one root, it can be expected for all of them. This feature is exemplified by Watson's trio of results, and is confirmed further in many other cases where derivation via a functional equation has not yet been demonstrated. No exceptions are known. This may be taken as indirect support for a conjecture that all ladder results may be assumed to be, in principle, derivable from a functional equation; though there is no current proof of this. Its correctness has in fact been questioned, particularly for those ladders of order greater than seven, for which no nontrivial functional equations of any sort are currently known.

1.3.3. *Formulas of Loxton.* In the early 1980's Loxton **[14]**, using an analysis based on properties of Rogers-Ramanujan partition functions, deduced a formula, which can be presented in ladder form, for the root in $(0, 1)$ of the cubic $x^3+3x^2=1$. His analysis is presented in Chapter 13. Based on an analogy with Watson's results, Lewin **[15]** suggested formulas involving the other two roots, and was able to structure the ladder expressions so that the otherwise undetermined coefficients of $\pi^2$ could be easily found by numerical computation. They turned out to be simple rationals (to the accuracy of the calculations), and there can be little doubt as to the correctness of the identifications. Loxton was able to come up with a partition-function proof for one of the pair, but not, for technical reasons, for the other. All three results were, until quite recently, thought to be "inaccessible" in the sense that no one seems to have been able to derive any of them from Kummer's equation. However, H. Gangl (1989) has succeeded in finding a derivation by multiple uses of Abel's equation.

**1.4. Current developments.**

1.4.1. *Ladders unlimited.* In a 1982 paper Lewin **[15]** showed that the previous few results represented the tip of an iceberg of unlimited extent. By using Kummer's equation, and some others derived from it, it was possible to generate ladders for a large number of algebraic bases. Virtually all of them exhibited the properties listed in 1.3.1, and the odd exception (a power appearing that was not a factor of the index) was considered, for the time being, an anomaly. However, the polynomial equations defining the base belonged to a rather restricted family of equations. This group of equations was expanded somewhat in a 1984 paper **[10]**, but the usable bases are nevertheless rather severely circumscribed. As will be explained shortly, a necessary, though by no means sufficient, condition is that the equation defining the base be capable of rearrangement into a form resembling a cyclotomic factorization, and called a *cyclotomic equation.*

1.4.2. *Definitions and notation.* The actual form representing the ladder, and its notation, have evolved somewhat since the earlier work. Dictated partly by convenience and partly by the discovery of the ladder properties, the preferred definition and notation for a ladder of index $N$ and order $n$, for the base $u$, is

$$L_n(N, u) = \frac{\mathrm{Li}_n(u^N)}{N^{n-1}} - \left\{\sum_r \frac{A_r \mathrm{Li}_n(u^r)}{r^{n-1}} + \frac{A_0 \log^n(u)}{n!}\right\}, \tag{1.27}$$

where the $A_r$, $0 \le r < N$, are rational coefficients defined implicitly by the rational polynomial giving rise to $u$, and $\mathrm{Li}_n$ is the $n$th order polylogarithm given by

$$\begin{aligned}\mathrm{Li}_n(z) &= \sum_{r=1}^{\infty} z^r/r^n, \qquad |z| \le 1, \\ &= \int_0^z \mathrm{Li}_{n-1}(t)\,dt/t.\end{aligned} \tag{1.28}$$

In practice, the powers $r$ in (1.27) are contained in a rather sparse group of integers most, but not necessarily all, of which may be factors of $N$.

Associated with (1.27) is a group of logarithmic terms of the form

$$L_n = \sum_{m=2}^{n} \frac{D_m \zeta(m) \log^{n-m}(u)}{(n-m)!}, \tag{1.29}$$



to be subtracted from (1.27). According to convenience some or all of these terms may or may not be explicitly included in the definition of the ladder. When $L_n(N, u)$ is equated to $L_n$, starting at $n = 2$, the resulting equations become a set for determining the $D_m$, which are, in general, irrational. But there is an important subset of $u$ for which $D_2$, and in many cases some of the other $D_m$, are rational. When this is the case the ladder is said to be *valid*, and it is only these cases which currently are of any interest.

The disadvantage of (1.27) as a definition is the presence of factors $N^{n-1}$ and $r^{n-1}$ in the denominators, and one can define a *modified ladder* $M_n(N, u)$ by the relation

$$M_n(N, u) = N^{n-1} L_n(N, u). \tag{1.30}$$

This removes most of the fractional coefficients from (1.27), but it has the disadvantage that it no longer possesses certain valuable order-independent features exhibited by $L_n(N, u)$. There seems to be no way round this particular difficulty; as a convenience, both forms are retained. It will be recognized that, with these definitions, equations like (1.22) or (1.26) are really modified ladders.

1.4.3. *The order-independence property and the cyclotomic equation.* It will be shown that if an expression like (1.27), coupled with (1.29), is valid for a certain value of $n$, it will also be valid for all lower values down to $n = 1$, where it takes the structure, in logarithmic form, of a cyclotomic equation:

$$(1 - u^N) = u^{-A_0} \prod_r (1 - u^r)^{A_r}. \tag{1.31}$$

This result was first discovered empirically by observation of the properties of a considerable number of ladders. No exceptions are known. A partial proof, dependent on the supposition that the ladder can, in principle, be found from a polylogarithmic functional equation from which the use of Rogers' function removes the logarithmic terms, will be presented in Chapters 2 and 3. This process works for Kummer's equations, and for certain other single-variable equations. The cyclotomic equation, which seems to be a *sine qua non* for the validity of its corresponding ladder, can be written down by inspection from (1.27); and, conversely, (1.27) can be written down from (1.31). But the coefficients $D_m$ in $L_n$ cannot be determined in this way, and depend either on numerical computation or an appropriate analytical development.

The order-reduction property can be exhibited from (1.27) by treating $u$ *as if* it were a variable, and differentiating with respect to $u$. We call this process *pseudodifferentiation*. The inverse process, which can be designated *pseudointegration*, is very useful for generating expressions at higher orders, and which can then be investigated for validity. All of the currently-known ladders for $6 \le n \le 9$ were found in this way.

1.4.4. *Component-ladders and transparency.* Although a valid ladder always yields a related cyclotomic equation, it is not true, in reverse, that an arbitrary cyclotomic equation corresponds to a valid ladder; counter examples are easy to construct. Rather, the set of terms corresponding to a given cyclotomic equation can be called a *component-ladder*. If it is indeed valid, the pseudointegration process is called *transparent*. Sometimes there are several component-ladders of different indices for the same base. These can be thought of as pieces of a ladder to be formed by combining them linearly with certain rational coefficients.

$$L_n^{(1)}(N, u) = \sum_{m=1}^{N} C_m L_n(m, u). \tag{1.32}$$

The resulting ladder of index $N$ can then be tested for validity at $n = 2$. If valid, further combinations can be sought to give valid ladders for larger values of $n$. In practice the set of coefficients $C_m$ is quite sparse, often reducing to two or, rarely, three. Unless these coefficients $C_m$ can be determined analytically they must be found by numerical search, a time-consuming effort involving multiprecision calculations to prevent the generation of artifacts. Nevertheless, a very large number of new results have been found in this way, and the vast majority of them currently lack any analytical derivation.

Because of the order-independent property of $L_n(m, u)$ the coefficients $C_m$ in (1.32) are *independent of order*. This is a quite remarkable, and by no means obvious, feature of these coefficients. When a valid result is obtained via a functional equation the coefficients involved are usually integral, and the result expressible in terms of the modified ladders $M_n(m, u)$. It is therefore necessary to use (1.30) to put the result in terms of $L_n(m, u)$ before the $C_m$ coefficients can be extracted.

**1.5. Base on the unit circle and Clausen function ladders.**

1.5.1. *Clausen's function.* This function, related to the dilogarithm of complex argument, was studied by Clausen [16] in the 1830's; and by others, particularly Kummer and Rogers, who both gave a functional equation involving five variables with a single constraint between them. On the unit circle we have

$$\mathrm{Li}_2(e^{i\theta}) = [\zeta(2) - \theta(2\pi - \theta)/4] + i\,\mathrm{Cl}_2(\theta), \qquad 0 \le \theta \le 2\pi, \tag{1.33}$$

where

$$\mathrm{Cl}_2(\theta) = \sum_{r=1}^{\infty} \sin(r\theta)/r^2 \tag{1.34}$$

$$= -\int_0^{\theta} \log|2\sin(t/2)|\,dt \tag{1.35}$$

is Clausen's function.

In recent studies in $K$-theory *Bloch's dilogarithm function* $D(z)$ has emerged as an important aspect. It can be related to Clausen's function via the formula



$$D(z) = \arg(1 - z)\log|z| - \mathrm{Im}\int_0^z \log(1 - t)\,dt/t \tag{1.36}$$

$$= \tfrac{1}{2}\{\mathrm{Cl}_2(2\theta) + \mathrm{Cl}_2(2\omega) - \mathrm{Cl}_2(2\theta + 2\omega)\}, \tag{1.37}$$

with

$$\theta = \arg(z), \qquad \omega = \arg(1 - z^*). \tag{1.38}$$

1.5.2. *Clausen function ladders.* The constraint equations of Kummer or Rogers, relating to five Clausen function variables, of which four are independent, can be put in either exponent or trigonometric form [15]. When this is done it becomes apparent that there are two distinct groups of relations. The first gives rise to real variables in $(0, 1)$ and consequent dilogarithmic ladders; the second to variables on the unit circle and consequent Clausen function relatic ns, involving various multiples of an angle as arguments. These are really ladders with the base on the unit circle, and can be considered *Clausen function ladders*. A few analytical results have been found, deduced directly from the Clausen functional equation, but the most interesting of such formulas, given by Browkin in Chapter 11, *appear* to be of the inaccessible sort, and are the consequence of the application of several unproven conjectures.

1.5.3. *The form of the constant term.* Whereas, in the dilogarithmic ladders, the constant term appearing in the ladders is a simple rational multiple of $\mathrm{Li}_2(1)$, the corresponding constants in the Clausen function ladders are more involved. The simplest involves Catalan's constant $G = \mathrm{Cl}_2(\pi/2)$, but others involve further submultiples of $\pi$, as called for in a conjecture of Lichtenbaum.

These new ladders appear to be the counterpart of some of the *supernumary ladders* discussed in Chapter 5.

## References


1. L. Euler, *Institutiones calculi integralis*, **1** (1768), 110–113.
2. J. Landen, *Mathematical memoirs*, **1** (1780), 112.
3. G. V. Chudnovsky, *Padé approximations to the generalized hypergeometric functions*, J. Math. Pure Appl. **58** (1979), 445–476.
4. N. H. Abel, *Oeuvres complètes*, **2**, 189–192.
5. W. Spence, *An essay on logarithmic transcendents*, London and Edinburgh, 1809.
6. C. J. Hill, *Specimen exercitii analytici, functionum integralum* $\int_0^x \frac{dx}{x}\log(1 + 2x\cos\alpha + x^2)$ *tum secundum amplitudinem, tum secundum modulum comparandi exhibentis*, Lund (1830), 9.
7. W. Schaeffer, *De integrali* $-\int_0^x \frac{\log(1-\alpha)}{\alpha}\,d\alpha$, J. für Math. (Crelle) **30** (1846), 277–295.
8. E. E. Kummer, *Ueber die Transzendenten welche aus wiederholten Integrationen rationaler Funktionen entstehen*, J. für Math. (Crelle) **21** (1840), 74–90.
9. L. J. Rogers, *On function sum theorems connected with the series* $\sum_1^\infty x^n/n^2$, Proc. London Math. Soc. **4** (1907), 169–189.
10. L. Lewin, *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), 345–373.

11. J. Dupont, private correspondence, 1989.
12. H. S. M. Coxeter, *The functions of Schläfli and Lobatschefsky*, Quart. J. Math. Oxford Ser. **6** (1935), 13–29.
13. G. N. Watson, *A note on Spence's logarithmic transcendent*, Quart. J. Math. Oxford Ser. **8** (1937), 39–42.
14. J. H. Loxton, *Special values of the dilogarithm function*, Acta Arith. **43** (1984), 155–166.
15. L. Lewin, *The dilogarithm in algebraic fields*, J. Austral. Math. Soc. 33 (1952), 302–330.
16. T. Claussen, *Über die Function* $\sin\phi + \frac{1}{2^2}\sin 2\phi + \frac{1}{3^2}\sin 3\phi +$ etc., J. für Math. (Crelle) **8** (1832), 298–300.

CHAPTER 2

# Dilogarithmic Ladders

L. LEWIN

**2.1. Derivation from Kummer's functional equation.**

2.1.1. *The role of Kummer's equation.* Apart from the (analytic) derivations of Coxeter and Loxton, referred to in Chapter 1, all the early ladder results stemmed from Kummer's functional equation. There are three useful forms of it, due, respectively, to Hill, Kummer and Schaeffer, and they are given in (1.18), (1.19) and (1.20). In these expressions the arguments appear in the form of simple rational functions; in what may be called the *algebraic form* of these equations. They look so different from the ladder arguments, powers of a base, that it was not obvious how to go from the one to the other. In fact there are so many unexplained current ladder results, for the time being labelled "inaccessible", that it is far from clear whether they are all obtainable from Kummer's formula. It has been a working hypothesis that they may be obtainable from *some* functional equation, and this guide has enabled, in a rather small number of cases, the construction of *ad hoc* single-variable equations to explain the related ladders. These single-variable equations share a set of significant arguments with Kummer's, but also contain other terms that seemingly cannot be obtained therefrom. So the question of accessibility is again raised. Whatever the eventual outcome, the expression of Kummer's equation in what might be called *exponent form* is the first step in the generation of a class of ladders from which an unlimited number of special cases can be deduced.

2.1.2. *The exponent form, two-term base equation.* All the arguments in (1.18) to (1.20), and also in Kummer's higher-order equations, involve variables $x$, $y$, $(1-x)$ and $(1-y)$ raised to various powers. All the arguments can be converted to powers of a base $u$ satisfying the rather limited *base equation*

$$u^p + u^q = 1, \tag{2.1}$$

by taking $x = u^p$, $1 - x = 1 - u^p = u^q$, and similar expressions for $y$. In fact the choice is wider than this. Kummer's equations are based on the harmonic group of automorphic functions

$$\{z;\ 1-z;\ -z/(1-z);\ 1/z;\ 1/(1-z);\ -(1-z)/z\}, \tag{2.2}$$

so if $x$ and $y$ are taken as any of these with $z = u^p$, a net power of $u$ will always result. With the form $-z/(1-z)$, for example, one gets $-u^{p-q}$. The negative terms are removed via the duplication formula (1.3), or can be temporarily left as is, and the functional equation is transformed into a sum of dilogarithms of powers of $u$. There are six independent results obtainable in this way, and including expressions from Euler's and Landen's equations (1.5) and (1.8), which give the first two, they are

$$\mathrm{Li}_2(u^p) + \mathrm{Li}_2(u^q) = \zeta(2) - pq\log^2 u, \tag{2.3}$$

$$L_2(u^p) + \mathrm{Li}_2(-u^{p-q}) = -\tfrac{1}{2}q^2\log^2 u, \tag{2.4}$$

$$\mathrm{Li}_2(u^{p+q}) - \mathrm{Li}_2(-u^{2p-q}) + \mathrm{Li}_2(-u^{p-2q}) = -(3q^2/2)\log^2 u, \tag{2.5}$$

(2.6)
$$\mathrm{Li}_2(u^{3p-3q}) - [\mathrm{Li}_2(u^{2p-2q}) + \mathrm{Li}_2(u^{p-q}) + \mathrm{Li}_2(-u^{2p-q}) + \mathrm{Li}_2(-u^{p-2q})] = (q^2/2)\log^2 u,$$

$$\mathrm{Li}_2(-u^{3p}) - [\mathrm{Li}_2(u^{2p}) + \mathrm{Li}_2(-u^p) + \mathrm{Li}_2(u^{p+q}) + \mathrm{Li}_2(-u^{2p-q})] = (q^2/2)\log^2 u \tag{2.7}$$

$$\mathrm{Li}_2(-u^{3q}) - [\mathrm{Li}_2(u^{2q}) + \mathrm{Li}_2(-u^q) + \mathrm{Li}_2(u^{p+q}) + \mathrm{Li}_2(-u^{2q-p})] = (p^2/2)\log^2 u. \tag{2.8}$$

To get some of the arguments into the range $(-1, 1)$ it may be necessary to use the inversion relation (1.4). Apart from such rearrangements, which are considered "trivial", the above six are the *only* results known for *general* $p$ and $q$. There are, however, *additional* results occurring for particular values; they are called *supernumary*, and are discussed at length in Chapter 5. The majority of them, at this time, exist purely by numerical confirmation—analytical proofs seem to be very hard to come by. However, G. Ray has recently succeeded in proving some of them (see Chapter 7).

The equations (2.3) to (2.8) are not, as they stand, in ladder form. They can also be linked in arbitrary ways by forming linear combinations. Preferred expressions, in ladder structure, are given in Chapter 3, both for the dilogarithm, and for the higher-order functions up to the fifth order, beyond which no equations of Kummer's type exist. A proof of this latter theorem, due to Wechsung, is given in Chapter 9.

Many special cases can now be found by giving $p$ and $q$ particular integer values. A selection is given in §2.6.

2.1.3. *The exponent form, three-term base equation.* For the dilogarithm, and also for the trilogarithm, though not for any of the higher-order functions,



Kummer's equation involves the factors $(1-x)$ and $(1-y)$ in only their ratio, $(1-x)/(1-y)$. This can be clearly seen from (1.18) to (1.20). Hence slightly less restrictive equations than (2.1) can be found by taking

$$x = \pm u^m; \qquad y = -u^{r-s}; \qquad (1-x)/(1-y) = u^s. \tag{2.9}$$

Eliminating $x$ and $y$ gives the pair of three-term equations for the base:

$$u^r + u^s \pm u^m = 1. \tag{2.10}$$

When the negative sign is used, one should take $m > r + s$ to ensure a real root for $u$ in $(0, 1)$.

Using the plus sign, insertion into (1.18) to (1.20) gives

(2.11)
$$\mathrm{Li}_2(-u^{m+s-r}) - [\mathrm{Li}_2(u^m) + \mathrm{Li}_2(-u^{s-r}) + \mathrm{Li}_2(-u^{m-r}) + \mathrm{Li}_2(u^s)] = \tfrac{1}{2}r^2\log^2 u,$$

(2.12)
$$\mathrm{Li}_2(-u^{r+s-m}) - [\mathrm{Li}_2(u^s) + \mathrm{Li}_2(-u^{r-m}) + \mathrm{Li}_2(-u^{s-m}) + \mathrm{Li}_2(u^r)] = \tfrac{1}{2}m^2\log^2 u,$$

(2.13)
$$\mathrm{Li}_2(-u^{s-m}) - [\mathrm{Li}_2(u^m) - \mathrm{Li}_2(-u^{s-r}) + \mathrm{Li}_2(-u^{s-r-m}) + \mathrm{Li}_2(u^r)] = -\zeta(2) + mr\log^2 u.$$

Apart from minor rearrangements based on the inversion relation (1.4) they may be inferred one from another by interchanging $r$, $s$ and $m$. Indices involving all three exponents occur once in each equation.

Using the minus sign in (2.10) gives a related triad:

(2.14)
$$\mathrm{Li}_2(u^{m+r-s}) - [\mathrm{Li}_2(-u^m) + \mathrm{Li}_2(-u^{s-r}) + \mathrm{Li}_2(u^{m-r}) + \mathrm{Li}_2(u^s)] = \tfrac{1}{2}r^2\log^2 u,$$

$$\mathrm{Li}_2(u^{m-r-s}) - [\mathrm{Li}_2(u^{m-r}) + \mathrm{Li}_2(u^{m-s}) - \mathrm{Li}_2(u^r) - \mathrm{Li}_2(u^s)] = \zeta(2) - rs\log^2 u, \tag{2.15}$$

(2.16)
$$\mathrm{Li}_s(u^{m+r-s}) - [\mathrm{Li}_2(-u^m) + \mathrm{Li}_2(-u^{r-s}) + \mathrm{Li}_2(u^{m-s}) + \mathrm{Li}_2(u^r)] = \tfrac{1}{2}s^2\log^2 u.$$

Equation (2.1) can be re-expressed in the form of the three-term equations by rearrangement:

$$u^p + u^{p+q} + u^{2q} = 1, \tag{2.17}$$

$$u^q + u^{p+q} + u^{2p} = 1, \tag{2.18}$$

$$u^q + u^{p-q} - u^{2p-q} = 1. \tag{2.19}$$

These are the forms suitable for $p > q$; no additional results come from using them in the previous sets of equations.

A somewhat more prolific result ensues if the base simultaneously satisfies two different equations of the type (2.10). For example, let $u$ satisfy

$$u^m + u^{m+k} + u^{m+2k} = 1. \tag{2.20}$$

The series on the left can be summed to $u^m(1-u^{3k})/(1-u^k)$, so the equation can be rewritten as

$$u^m + u^k - u^{m+3k} = 1. \tag{2.21}$$

This leads to six relations. However, one of them must be considered redundant, since, with the help of the others, it is reducible to an identity through the duplication relation (1.3). The six are

(2.22)
$$\mathrm{Li}_2(u^{2m+2k}) - [\mathrm{Li}_2(-u^{m+3k}) + \mathrm{Li}_2(-u^{m-k}) + \mathrm{Li}_2(u^{m+2k}) + \mathrm{Li}_2(u^m)] = \tfrac{1}{2}k^2\log^2 u,$$

(2.23)
$$\mathrm{Li}_2(u^{4k}) - [\mathrm{Li}_2(-u^{m+3k}) + \mathrm{Li}_2(-u^{k-m}) + \mathrm{Li}_2(u^{3k}) - \mathrm{Li}_2(u^k)] = \tfrac{1}{2}m^2\log^2 u,$$

(2.24)
$$\mathrm{Li}_2(u^{2k}) - [\mathrm{Li}_2(u^{m+2k}) + \mathrm{Li}_2(u^{3k}) - \mathrm{Li}_2(u^k) - \mathrm{Li}_2(u^m)] = \zeta(2) - mk\log^2 u,$$

(2.25)
$$\mathrm{Li}_2(-u^{m+k}) - [\mathrm{Li}_2(u^m) + \mathrm{Li}_2(u^{m+2k})] = -\zeta(2) + m(k+m/2)\log^2 u,$$

(2.26)
$$\mathrm{Li}_2(-u^{2k}) - [\mathrm{Li}_2(u^m) - \mathrm{Li}_2(-u^k) + \mathrm{Li}_2(-u^{k-m}) + \mathrm{Li}_2(u^{m+k})] = -\zeta(2) + m(k+m)\log^2 u,$$

(2.27)
$$\mathrm{Li}_2(-u^{m+3k}) - [\mathrm{Li}_2(u^{m+2k}) + \mathrm{Li}_2(-u^k) + \mathrm{Li}_2(u^{m+k}) + \mathrm{Li}_2(-u^{2k})] = \tfrac{1}{2}m^2\log^2 u.$$



A further relation comes by taking $y = -x$ in (1.18) to (1.20), and suggests the equation

$$u^r = (1-u^s)/(1+u^s) \quad \text{or} \quad u^r + u^s + u^{r+s} = 1. \tag{2.28}$$

This is a variant of (2.10), and leads to three independent relations:

(2.29)
$$\mathrm{Li}_2(u^{4r}) - [3\,\mathrm{Li}_2(u^{2r}) + 2\,\mathrm{Li}_2(u^{r+s}) + \mathrm{Li}_2(u^{2r-2s}) - 2\,\mathrm{Li}_2(u^{r-s})] = s^2\log^2 u,$$

(2.30)
$$\mathrm{Li}_2(u^{4s}) - [3\,\mathrm{Li}_2(u^{2s}) + 2\,\mathrm{Li}_2(u^{r+s}) - \mathrm{Li}_2(u^{2r-2s}) + 2\,\mathrm{Li}_2(u^{r-s})] = -2\zeta(2) + s(2r-s)\log^2 u,$$

(2.31)
$$\mathrm{Li}_2(u^{2r}) + \mathrm{Li}_2(u^{2s}) - 4[\mathrm{Li}_2(u^r) + \mathrm{Li}_2(u^s)] = -3\zeta(2) + 2rs\log^2 u.$$

This seems about as far as one can go in obtaining results of a general character from Kummer's equation, at least in its *basic* forms.

The above results can be expressed in ladder form but it is much easier to do so when the exponents are given explicitly. Some examples are discussed in §2.6.

**2.2. Relation to Clausen's function.**

2.2.1. *The imaginary part of the dilogarithm.* A very important result, due to Kummer [1], expresses the imaginary part of the dilogarithm of complex argument in terms of Clausen's function. From

$$\mathrm{Li}_2(re^{i\theta}) = -\int_0^{re^{i\theta}} \log(1-z)\,dz/z = -\int_0^r \log(1-ye^{i\theta})\,dy/y$$

one obtains, on taking the imaginary part (with $r$ and $\theta$ real),

$$\mathrm{Im}\,\mathrm{Li}_2(re^{i\theta}) = \int_0^r \tan^{-1}\left[\frac{y\sin\theta}{1-y\cos\theta}\right]\frac{dy}{y}. \tag{2.32}$$

This integral can be evaluated by integration by parts and using a subsidiary quantity $\omega$ defined by

$$\tan\omega = r\frac{\sin\theta}{1-r\cos\theta} \quad \text{or} \quad r = \frac{\sin\omega}{\sin(\omega+\theta)}. \tag{2.33}$$

The result can be written

(2.34)
$$\mathrm{Im}\,\mathrm{Li}_2(re^{i\theta}) = \omega\log r + \tfrac{1}{2}[\mathrm{Cl}_2(2\theta) + \mathrm{Cl}_2(2\omega) - \mathrm{Cl}_2(2\theta+2\omega)],$$

and is closely related to Bloch's definition (1.36).

2.2.2. *Reversion to Kummer's equation.* Using (1.32) this formula can be converted back into an equation for the dilogarithm only, by eliminating the three Clausen functions:

(2.35)
$$\begin{aligned}\mathrm{Li}_2(re^{i\theta}) - \mathrm{Li}_2(re^{-i\theta}) = 2i\omega\log r + 2\omega\theta - \zeta(2) + \mathrm{Li}_2(e^{i2\omega}) \\ + \mathrm{Li}_2(e^{i2\theta}) - \mathrm{Li}_2[e^{i2(\omega+\theta)}].\end{aligned}$$

Rather surprisingly, this is but a variant of Kummer's functional equation in the form (2.15), as may be seen by taking $e^{i2\omega} = u^s$, $e^{i2\theta} = u^{m-r-s}$ and $r = u^{(m+r-s)/2}$. With hindsight one can therefore get the expression for the imaginary part of the dilogarithm directly from the functional equation. This process may be repeated by taking the variables $x$ and $y$ in Kummer's equation as complex quantities, putting the imaginary part in terms of Clausen's function and then reverting again to the dilogarithm form. This operation was carried out in [2] and led to an equation between five variables $u^{q_m}$, $m = 1$ to 5, with a single constraint relation between them.

2.2.3. *A fifteen-term functional equation.* The derivation uses the fifteen-term functional equation for Clausen's function due to both Kummer [1] and Rogers [3]. The result can be written

(2.36)
$$\begin{aligned}\frac{1}{2}\left\{\sum_{m=1}^{5}\sum_{n=1}^{5}\mathrm{Li}_2(-u^{s-q_n-q_m}) - \sum_{m=1}^{5}\mathrm{Li}_2(-u^{s-2q_m})\right\} - \sum_{m=1}^{5}\mathrm{Li}_2(u^{q_m}) \\ = \tfrac{1}{2}\left(s^2 - \sum_{m=1}^{5} q_m^2\right)\log^2 u - 4\zeta(2),\end{aligned}$$

where

(2.37)
$$2s = \sum_{m=1}^{5} q_m,$$

and the constraint relation is

(2.38)
$$u^s + 1 = \sum_{m=1}^{5}(u^{q_m} + u^{s-q_m}).$$

This equation, for given (real) $q_m$, can give both real values of $u$, and also values on the unit circle. This latter can be seen by taking $u$ in the form $e^{i2\theta}$, when (2.38) goes over into

(2.39)
$$\cos(s\theta) = \sum_{m=1}^{5}\cos[(\tfrac{1}{2}s - q_m)\theta],$$

which will, in general, possess some real roots in $\theta$.

This important link connects the ladders in $(0, 1)$ with related Clausen-function formulas.



**2.3. A three-variable dilogarithmic functional equation.**

2.3.1. *Derivation from rational function arguments.* Wechsung [4] indicated a method of derivation of functional equations which, although limited to order $n \le 5$, might possibly provide results not accessible from Kummer's original equation. The simplest, for the dilogarithm, commences with some such algebraic argument as $\prod_{m=1}^{M}[(z-a_m)/(z-b_m)]$, and proceeds by developing the logarithm thereof. A similar idea was presented earlier by Mantel [5], and in the case of the biquadratic form, $M = 2$, leads to an equation with nine transcendental terms. The quantities $a_m$ and $b_m$ are treated as constant parameters, and $z$ as a variable; but in the ultimate equation this distinction is no longer necessary. If we take $(z-a_1)$, $(z-a_2)$, $(z-b_1)$, and $(z-b_2)$ as four variables $u_1$ to $u_4$, we apparently obtain a relation with four independent quantities. However, they occur in such combinations that only three of them are really independent.

2.3.2. *A nine-term, three-variable functional equation.* If the quantities $u_m$ are normalized to $(u_1 + u_2 - u_3 - u_4)/(u_1u_2 - u_3u_4)$ the equation, with a suitable change of variables, becomes

(2.40)
$$\begin{aligned}\mathrm{Li}_2\left(\frac{vw}{xy}\right) = \mathrm{Li}_2(\tfrac{v}{x}) + \mathrm{Li}_2(\tfrac{w}{y}) + \mathrm{Li}_2(\tfrac{v}{y}) + \mathrm{Li}_2(\tfrac{w}{x}) + \mathrm{Li}_2(x) + \mathrm{Li}_2(y) \\ - \mathrm{Li}_2(v) - \mathrm{Li}_2(w) + \tfrac{1}{2}\log^2(-x/y),\end{aligned}$$

subject to the constraint

(2.41) $\quad (1-v)(1-w) = (1-x)(1-y) \quad$ or $\quad x + y - v - w = xy - vw.$

This result can be contrasted with the apparently distinct 5-term form of Kummer's equation.

A ladder-type relation, in which the base satisfies a 5-term equation (in contrast to Kummer's 3-term equation) can be found from (2.40) by taking $x = -u^m$, $y = u^{n+m}$, $v = -u^{p+m}$ and $w = u^{q+m}$. Then the constraint relation (2.41) becomes

(2.42)
$$u^p + u^n - u^q + u^{n+m} - u^{q+p+m} = 1,$$

where one needs $q > n$ to give a real root in $(0, 1)$. (This relation is slightly different from equation (28) of [6], where this result was first reported, in order to give real roots for $u$.)

In exponent form, (2.40) becomes

(2.43)
$$\begin{aligned}\mathrm{Li}_2(u^{p+q-n}) = \mathrm{Li}_2(u^p) + \mathrm{Li}_2(u^{q-n}) + \mathrm{Li}_n(-u^{p-n}) + \mathrm{Li}_2(-u^q) + \mathrm{Li}_2(-u^m) \\ + \mathrm{Li}_2(u^{n+m}) - \mathrm{Li}_2(-u^{p+m}) - \mathrm{Li}_2(u^{q+m}) + \tfrac{1}{2}n^2\log^2 u.\end{aligned}$$

2.3.3. *Accessibility from Kummer's equation.* The above result can be obtained indirectly from Kummer's 5-term equation. The most suitable variant to demonstrate this is Hill's form (1.18) in which one takes, in turn, for $x$ and $y$ in that equation, (i) $1/y, v$; (ii) $w, 1/x$; (iii) $w/y, v/x$. The three results are combined, and (2.41) is utilized (this is quite essential). Finally, with three applications of the inversion relation (1.4), (2.40) reduces to an identity. It is therefore algebraically accessible from Kummer's equation, the dissimilarity between the two notwithstanding.

**2.4. Functional equations in the complex plane.**

2.4.1. *The real part of the dilogarithm.* Unlike the imaginary part, which can be expressed in terms of a single-variable function via (2.34), the real part defines a new, two-variable function:

$$\mathrm{Li}_2(r, \theta) = \mathrm{Re}\,\mathrm{Li}_2(re^{i\theta}). \tag{2.44}$$

It is discussed extensively in [7], where many useful formulas are developed. Although it is a "new" function, it can be expressed in terms of the dilogarithm of a real variable for certain combinations of $r$ and $\theta$; and in particular when $\theta/\pi$ is rational. The following four equations are typical:

$$\mathrm{Li}_2(\tan x, \tfrac{\pi}{2} - 2x) = x^2 + \tfrac{3}{4}\mathrm{Li}_2(\tan^2 x) - \tfrac{1}{8}\mathrm{Li}_2(\tan^4 x), \tag{2.45}$$

$$\mathrm{Li}_2(r, \tfrac{\pi}{3}) = \tfrac{1}{6}\mathrm{Li}_2(-r^3) - \tfrac{1}{2}\mathrm{Li}_2(-r), \tag{2.46}$$

(2.47)

$$\mathrm{Li}_2(r, \pi/4) = \tfrac{1}{4}\mathrm{Li}_2(r\sqrt{2} - r^2) - \tfrac{1}{2}\mathrm{Li}_2\left(\frac{-r}{\sqrt{2} - r}\right) + \tfrac{1}{8}\mathrm{Li}_2\left(\frac{-r^2}{(\sqrt{2} - r)^2}\right),$$

(2.48)

$$\mathrm{Li}_2(r, \pi/6) = \tfrac{1}{12}\mathrm{Li}_2\left[-\left(\frac{r}{(\sqrt{3} - r)}\right)^3\right] - \tfrac{3}{4}\mathrm{Li}_2\left[\frac{-r}{\sqrt{3} - r}\right] + \tfrac{1}{4}\mathrm{Li}_2[r(\sqrt{3} - r)].$$

Moreover, from (1.32), with $0 \le \theta \le \pi$, we have

$$\begin{aligned}\mathrm{Li}_2(\pm 1, \theta) &= \mathrm{Re}\{\mathrm{Li}_2(e^{i\theta}), \mathrm{Li}_2(e^{i(\theta+\pi)})\} \\ &= \zeta(2) - \theta(2\pi - \theta)/4,\ \zeta(2) - (\pi^2 - \theta^2)/4,\end{aligned} \tag{2.49}$$

and corresponding variants for other ranges of $\theta$.

These will all be utilized to provide specific results in §2.6.

2.4.2. *Multiple-angle formulas.* There are many such; the following is typical:

(2.50)

$$\mathrm{Li}_2\left[-\frac{\sin[(M-1)a]}{\sin a}, Ma\right] = \tfrac{1}{2}a^2 M(M-1) - \tfrac{1}{2}\sum_{m=1}^{M-1}\mathrm{Li}_2\left[\frac{\sin^2 a}{\sin^2 ma}\right] + \sum_{m=2}^{M-1}\log\left[\frac{\sin ma}{\sin a}\right]\log\left[\frac{\sin[(m-1)a]}{\sin ma}\right].$$



The particular case $a = \pi/M$ is

(2.51)

$$\sum_{m=2}^{M-1}\mathrm{Li}_2\left[\frac{\sin^2(\pi/M)}{\sin^2(m\pi/M)}\right] = \frac{M-2}{2M}\pi^2 + 2\sum_{m=2}^{M-1}\log\left[\frac{\sin(m\pi/M)}{\sin(\pi/M)}\right]\log\left[\frac{\sin[(m-1)\pi/M]}{\sin(m\pi/M)}\right].$$

One of Watson's results can be found from (2.51) by taking $M = 7$.

Taking the limit $a \to 0$ and writing $M + 1$ for $M$ gives

$$\mathrm{Li}_2(-M) = \frac{1}{2}\sum_{m=1}^{M}\mathrm{Li}_2(1/m^2) + \sum_{m=2}^{M}\log(m)\log(1 - 1/m), \tag{2.52}$$

a form which can also be generated directly from Kummer's equation. Taking the limit as $M \to \infty$ gives

$$\sum_{m=2}^{\infty}\mathrm{Li}_2(1/m^2) = \zeta(2) + \sum_{m=2}^{\infty}\log(m)\log(1 - 1/m^2). \tag{2.53}$$

This result is due to Richmond and Szekeres [8] from evaluating the coefficients of certain Rogers-Ramanujan partition identities (due to George Andrews). Szekeres also put this result in terms of the Riemann zeta-function:

$$\sum_{m=1}^{\infty}\left[\frac{1}{m^2}\zeta(2m) - \frac{1}{m}\zeta'(2m)\right] = 2\zeta(2). \tag{2.54}$$

2.4.3. *A formula of Richmond and Szekeres.* A formula that was, at first, thought to be of a quite different character from (2.51) can be written

$$\sum_{i=1}^{r}\mathrm{Li}_2(d_i) = \frac{\pi^3}{3}\frac{r}{2r+3} - \frac{1}{2}\sum_{i=1}^{r}\log(d_i)\log(1 - d_i), \tag{2.55}$$

where

$$d_1 = 2\left[1 - \cos\left(\frac{\pi}{2r+3}\right)\right] \tag{2.56}$$

$$d_{j+1} = d_j\prod_{i=1}^{j}(1 - d_i)^{-2}, \qquad j = 1, 2, \ldots, (r-1). \tag{2.57}$$

Taking the limit $r \to \infty$ gives (2.53).

Appearances notwithstanding, (2.55) can be shown equivalent to (2.51). A brief outline of the proof follows.

Writing $d_j = 1/D_j^2$ in (2.57), inverting and taking the square root gives

$$D_{j+1} = D_j\prod_{i=1}^{j}(1 - 1/D_i^2). \tag{2.58}$$

Putting $j+1$ for $j$ and dividing the two-equations gives

$$D_j D_{j+2} = D_{j+1}^2 - 1. \tag{2.59}$$

Recalling the identity $\sin^2 x - \sin^2 y = \sin(x+y)\sin(x-y)$ enables (2.59) to be solved by inspection in the form

$$D_j = \sin(jB + C)/\sin B, \tag{2.60}$$

with $B$ and $C$ constants to be determined by (2.56) and (2.57). Hence we find that

$$d_j = \sin^2\left(\frac{\pi}{2r+3}\right)\Big/\sin^2\left(\pi\frac{r+2-j}{2r+3}\right), \tag{2.61}$$

so (2.55) goes over to the form (2.51) with $M = 2r+3$.

**2.5. Cyclotomic equations and Rogers' function.**

2.5.1. *The logarithm-removal property of Rogers' function.* As may be readily verified, the use of (1.21) removes the logarithmic terms from Euler's (1.5). It does the same for the other equations of §1.1.2, and also the two-variable formulas of §1.2, provided the minus signs in front of some of the arguments are ignored. This has led to a need for a slight redefinition of Rogers' function in the form

$$L_2(z) = \mathrm{Li}_2(z) + \tfrac{1}{2}\log|z|\log(1-z), \tag{2.62}$$

where the use of $|z|$ is meant primarily to remove any initial minus sign in the argument. It could, perhaps, be reinterpreted as $\log|z| = \frac{1}{2}\log(z^2)$, with the principal value of the latter implied; but in fact the situation for complex $z$ is more involved than this. For example, the multiplication theorem, of which (1.3) is the special case $M = 2$, is

$$\left(\frac{1}{M}\right)\mathrm{Li}_2(z^M) = \sum_{m=0}^{M-1}\mathrm{Li}_2(ze^{i2m\pi/M}), \tag{2.63}$$

and this requires (2.62) as it stands if a logarithm-free equivalent for Rogers' function is to be formed.

Complex $z$ is not an issue in the ladder-related use made here of (2.62), so simply the minus-suppressing role of $|z|$ is needed for now. With this proviso, all the known second-order functional equations for $\mathrm{Li}_2$ can be checked to go over into a corresponding form for $L_2$, but absent any logarithmic terms. It will be shown later that (2.62), and its generalization (3.14) to higher-orders, is the only possible definition of a function that can accomplish this, though this still does not prove that this function does in fact operate in this way; but no exceptions are known.

2.5.2. *Application to the ladder structure.* Since, with the use of $L_2(z)$, the logarithmic terms are removed from Kummer's equation, *any* result that can be obtained therefrom in a finite number of steps must also be free of logarithms. Thus, if a valid ladder is, in principle, at least, known to



be obtainable in this way, the replacement of $\mathrm{Li}_2$ by $L_2$ in the ladders must result in a ladder expression for $L_2$ free of logarithms. Thus, (1.26) combined with (1.29) at $n = 2$, takes the form

$$\frac{\mathrm{Li}_2(u^N)}{N} - \left\{\sum_r \frac{A_r\,\mathrm{Li}_2(u^r)}{r} + \frac{A_0\log^2(u)}{2}\right\} = D_2\zeta(2), \tag{2.64}$$

and if the above operation is carried out the logarithmic terms that accumulate take the form $L = -\frac{1}{2}\log(u)\log U$ where

$$U = (1-u^N)u^{A_0}\Big/\prod_r (1-u^r)^{A_r}. \tag{2.65}$$

But there is no *net* logarithmic term in the modified equation, so $L = 0$, $U = 1$, and (2.65) becomes the *cyclotomic equation* (1.31) associated with the ladder. This one-to-one relation works only in the direction (2.64) to (2.65), and no exceptions are known. In reverse, an equation of the form (2.65) does not *necessarily* lead to a valid ladder, though if it does the structure has to be of the form (2.64).

**2.6. Accessible and analytic ladders.**

2.6.1. *Ladders from functional equations.* We will retain here the previous use of certain lower-case Greek symbols for some designated bases. Results obtained by the use of the two-term or three-term base equations of §2.1 extend in all cases to the third, or occasionally higher orders, and are considered in more detail in Chapter 3. Many novel results fall into the supernumary category of Chapter 5. Of those remaining, Watson's formulas [9] stem directly from multiple uses of Kummer's equation. His results in ladder form are

$$\frac{\mathrm{Li}_2(\alpha^2)}{2} - \tfrac{1}{2}\mathrm{Li}_2(\alpha) + \tfrac{1}{2}\log^2\alpha = \frac{-\zeta(2)}{14}, \tag{2.66}$$

$$\frac{\mathrm{Li}_2(\beta^2)}{2} + \mathrm{Li}_2(\beta) + \log^2\beta = \frac{5\zeta(2)}{7}, \tag{2.67}$$

$$\frac{\mathrm{Li}_2(\gamma^2)}{2} + \mathrm{Li}_2(\gamma) + \tfrac{1}{2}\log^2\gamma = \frac{4\zeta(2)}{7}, \tag{2.68}$$

where

$$\alpha = \frac{1}{2}\sec(2\pi/7) \tag{2.69}$$

$$\beta = \frac{1}{2}\sec(\pi/7) = 1/(1+\alpha) \tag{2.70}$$

$$\gamma = 2\cos(3\pi/7) = \alpha/(1+\alpha). \tag{2.71}$$

They are obtained via the cubic $x^3 + 2x^2 - x - 1 = 0$, whose roots are $\alpha$, $-\beta$, and $-1/\gamma$. It is straightforward to obtain the cyclotomic equations corresponding to (2.66) to (2.68) and to verify that they are obtainable by simple rearrangement of the defining base cubic equation.

Loxton, using the methods of Chapter 13 on partition identities was the first to obtain results for the quantity $a = (3^{1/2} - 1)/2$. This is a root of the equation

$$2u^2 + 2u - 1 = 0. \tag{2.72}$$

It has a second root $b = -1/(3^{1/2} - 1)$. If we write $c = -1/b = (3^{1/2} - 1)$, then $0 < c < 1$ and it is easy to verify that

$$1 - c^6 = (1 - c^3)(1 - c^2)^3(1 - c)^{-2}. \tag{2.73}$$

From this a ladder can be written down by inspection, and the coefficient $D_2$ obtained by computation. The result was first reported in **[10]** and the two ladders are

$$\frac{\mathrm{Li}_2(a^3)}{3} - \left\{\frac{\mathrm{Li}_2(a^2)}{2} + 2\,\mathrm{Li}_2(a) + \frac{1}{2}\log^2 a\right\} + 5\zeta(2)/6 = 0, \tag{2.74}$$

$$\frac{\mathrm{Li}_2(c^6)}{6} - \left\{\frac{\mathrm{Li}_2(c^3)}{3} + \frac{3\,\mathrm{Li}_2(c^2)}{2} - 2\,\mathrm{Li}_2(c)\right\} - \zeta(2)/2 = 0. \tag{2.75}$$

Subsequently, these results have been proved from (2.48) by taking $r = 1$ and $-1$, and using (2.49) with $\theta = \pi/6$.

2.6.2. *Ladders from other analytical formulas.* The only results originally fitting this category were Loxton's formula and ladders for base $\rho$ of index 20 and 24. The latter pair have since been proven by the development of *ad hoc* functional equations. Loxton's base equation, treated in detail in Chapter 13, is

$$u^3 + 3u^2 - 1 = 0. \tag{2.76}$$

If the three roots are designated $\kappa$, $-\lambda$, $-1/\mu$ then all three quantities $\kappa$, $\lambda$ and $\mu$ lie in $(0, 1)$, and, as in the case of Watson's cubic, they are similarly related:

$$\kappa = \tfrac{1}{2}\sec(\pi/9), \tag{2.77}$$

$$\lambda = \tfrac{1}{2}\sec(2\pi/9) = 1/(1+\kappa), \tag{2.78}$$

$$\mu = 2\cos(4\pi/9) = \kappa/(1+\kappa). \tag{2.79}$$

From the corresponding easily-developed cyclotomic equations one finds, by



computation of the constant term,

$$\frac{\mathrm{Li}_2(\kappa^3)}{3} - \{\mathrm{Li}_2(\kappa^2) + \mathrm{Li}_2(\kappa) + \log^2(\kappa)\} + \frac{7\zeta(2)}{9} = 0, \tag{2.80}$$

$$\frac{\mathrm{Li}_2(\lambda^6)}{6} - \left\{\frac{\mathrm{Li}_2(\lambda^3)}{3} + \frac{3\,\mathrm{Li}_2(\lambda^2)}{2} - \mathrm{Li}_2(\lambda) + \log^2(\lambda)\right\} + \frac{\zeta(2)}{9} = 0, \tag{2.81}$$

$$\frac{\mathrm{Li}_2(\mu^6)}{6} - \left\{\frac{\mathrm{Li}_2(\mu^3)}{3} + \frac{3\,\mathrm{Li}_2(\mu^2)}{2} - \mathrm{Li}_2(\mu)\right\} - \frac{\zeta(2)}{9} = 0. \tag{2.82}$$

The first two are proved in Chapter 13 from the partition-function identities, and the last has recently been proved by H. Gangl. It is included here with the other results because of their very close affinity.

### 2.7. Inaccessible ladders.

2.7.1. *Ladders from quadratic equations.* The equation

$$u^2 + 3u - 1 = 0 \tag{2.83}$$

is the particular case $m = 3$ of $u^2 + mu - 1 = 0$ discussed in **[11]**. The cases for other usable values of $m$ are analytic, and are dealt with in Chapters 3 and 4. But (2.83) does not give a ladder with a known analytic derivation, though a valid result has been found by computation. Details are given in the reference, where the cyclotomic equation

$$1 - u^6 = (1 - u^3)^2(1 - u^2)^2(1 - u)^{-4}u \tag{2.84}$$

was found. From it the ladder structure can be written down and the constant $D_2$ determined by computation, leading to

$$\frac{\mathrm{Li}_2(S^6)}{6} - \left\{\frac{2\,\mathrm{Li}_2(S^3)}{3} + \mathrm{Li}_2(S^2) - 4\,\mathrm{Li}_2(S) - \tfrac{1}{2}\log^2(S)\right\} = \frac{7\zeta(2)}{6}, \tag{2.85}$$

where $S = (\sqrt{13} - 3)/2$.

The only other quadratic equation leading to real roots and a valid ladder, not already examined in the literature and discussed in later chapters, is one resulting from Browkin's research on cyclotomic equations **[12]**. It is almost a "companion" equation to (2.83), and takes the form

$$3u^2 + u - 1 = 0, \tag{2.86}$$

which has two roots $u = U, -V$ where both

$$U = (\sqrt{13} - 1)/6 \quad \text{and} \quad V = (\sqrt{13} + 1)/6 \tag{2.87}$$

are real and lie in $(0, 1)$.

A cyclotomic equation derived from (2.86) is

$$1 - u^6 = (1 - u^3)^3(1 - u^2)^{-1}(1 - u)^{-3}u^2. \tag{2.88}$$

From this the ladder for $U$ can be constructed, and a simple computation gives

(2.89)
$$\frac{\mathrm{Li}_2(U^6)}{6} - \left\{\mathrm{Li}_2(U^3) - \frac{\mathrm{Li}_2(U^2)}{2} - 3\,\mathrm{Li}_2(U) - \log^2(U)\right\} = \frac{4\zeta(2)}{3}.$$

The corresponding structure for $V$ comes by replacing $u$ by $-u$ in (2.88), followed by minor rearrangements, to give

(2.90)
$$\frac{\mathrm{Li}_2(V^6)}{3} - \{\mathrm{Li}_2(V^3) + 2\,\mathrm{Li}_2(V^2) - 3\,\mathrm{Li}_2(V) + \log^2(V)\} = \frac{2\zeta(2)}{3}$$

2.7.2. *Ladders from quintic equations.* P. Barrucand [13] has questioned whether certain irreducible quintics may have roots which could give rise to valid ladders. In particular he specified the pair

$$u^5 - u^3 + u^2 + u - 1 = 0, \tag{2.91}$$
$$u^5 + u^4 - u^3 + u^2 - 1 = 0. \tag{2.92}$$

Neither of these fits the forms of the base equations of §§2.1 to 2.3. Although the two equations appear superficially quite similar, in fact (2.91) alone can be maneuvered into a three-term equation, from which an extraordinarily rich sequence of ladders is found. It was first discussed in [14], and is reviewed in Chapter 5 on supernumary results.

It seems impossible to put (2.92) into a similar form to fit any of the accessible results. However, it can be rearranged as two cyclotomic equations:

$$1 - u^6 = (1-u^3)(1-u^2)^{-1}u^4, \tag{2.93}$$
$$1 - u^5 = (1-u^2)^{-2}(1-u)u^6. \tag{2.94}$$

The corresponding component-ladders are

$$L_2(6, u) = \frac{\mathrm{Li}_2(u^6)}{6} - \frac{\mathrm{Li}_2(u^3)}{3} + \frac{\mathrm{Li}_2(u^2)}{2} + 2\log^2 u, \tag{2.95}$$
$$L_2(5, u) = \frac{\mathrm{Li}_2(u^5)}{5} + \mathrm{Li}_2(u^2) - \mathrm{Li}_2(u) + 3\log^2 u. \tag{2.96}$$

Neither is individually valid, but it is found numerically that

$$5L_2(5, u) + 6L_2(6, u) = \zeta(2). \tag{2.97}$$

More details are given in the reference. There is no currently known prescription about (2.92) that determines that it should possess a valid ladder.



## References


1. E. E. Kummer, *Ueber die Transzendenten, welche aus wiederholten Intergrationen rationaler Funktionen entstehen*, J. für Math. (Crelle), **21** (1840), 74–90.
2. L. Lewin, *The dilogarithm in algebraic fields*, J. Austral. Math. Soc. **33** (1982), 303–330.
3. L. J. Rogers, *On function sum theorems connected with the series* $\sum_1^\infty x^n/n^2$, Proc. London Math. Soc. **4** (1907), 169–189.
4. G. Wechsung, *Lineare Funktionalgleichungen von Polylogarithmen*, Wiss. Z. Friedrich-Schiller-Univ. Jena-Thuringen **14** (1965), 401–408.
5. W. Mantel, *Dilogarithmen*, Nieuw Archief 3 (1898), 292–320.
6. L. Lewin, *The order-independence of the polylogarithmic ladder structure—implications for a new category of functional equations*, Aequationes Math. **30** (1986), 1–20.
7. ——, *Polylogarithms and associated functions*, Elsevier-North-Holland, New York and Oxford, 1981, (See Chapter V).
8. R. Richmond and G. Szekeres, *Some formulas related to dilogarithms, the zeta-function and the Andrews-Gordon identities*, J. Austral. Math. Soc. **31** (1981), 362–373.
9. G. N. Watson, *A note on Spence's logarithmic transcendent*, Quart. J. Math. Oxford Ser. **8** (1937), 39–42.
10. J. H. Loxton, *Special values of the dilogarithm function*, Acta Arith. **43** (1984), 155–166.
11. L. Lewin, *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), 345–373.
12. J. Browkin, private correspondence.
13. P. Barrucand, private correspondence.
14. M. Abouzahra and L. Lewin, *The polylogarithm in the field of two irreducible quintics*, Aequationes Math. **31** (1986), 315–321.

CHAPTER 3

# Polylogarithmic Ladders

M. ABOUZAHRA AND L. LEWIN

## 3.1. Kummer's function and its relation to the polylogarithm.

3.1.1. *Definition.* Kummer utilized the function $\Lambda_n(z)$ defined by

$$\Lambda_n(z) = \int_0^z \frac{\log^{n-1}|t|}{1+t}\,dt. \tag{3.1}$$

For real values of $z$, positive or negative, the use of $|z|$, exactly as in the definition of Rogers' function, was meant to suppress possible negative signs in the logarithm, to ensure $\Lambda_n(z)$ real when $z$ is real. In the subsequent development for complex $z$ the use of the modulus symbol takes its usual meaning, and again the comparison with Rogers' function is useful. In the present development only real $z$ is of concern, so $\log|z|$ may be interpreted as $\frac{1}{2}\log(z^2)$. The relevance here is that differentiation with respect to $z$ then gives $1/z$, whereas differentiation of $\log|z|$ with respect to $z$ is not defined. However, a useful procedure is to differentiate in the radial direction, that is, with respect to $re^{i\theta}$ holding $\theta$ constant—but this will not be needed here.

Kummer used simple functional equations for $(\log)^{n-1}$ for $2 \leq n \leq 5$ to produce functional equations for $\Lambda_n$ of two variables $x$ and $y$. He denoted $(1-x)$ and $(1-y)$ by $\xi$ and $\eta$ respectively, and this notation will be retained in this section.

3.1.2. *Relation to the polylogarithm.* Repeated integration by parts of (3.1) gives

$$\begin{aligned}\operatorname{Li}_n(z) = \operatorname{Li}_n(1) &+ \sum_{r=1}^{n-1}(-1)^{r-1}\frac{\log^r|z|}{r!}\operatorname{Li}_{n-r}(z)\\ &+ \frac{(-1)^{n-1}}{(n-1)!}[\Lambda_n(-1) - \Lambda_n(-z)].\end{aligned} \tag{3.2}$$

Thus, any functional equation for $\Lambda_n(z)$ gives rise to an equation of like form for $\operatorname{Li}_n(-z)$, together with a string of logarithmic and lower-order polylogarithmic terms. It is rather laborious to do this, but when done it is found that none of the lower-order polylogarithms survive—on using their own functional equations they are all seen to vanish identically. Thus the resulting equations for $\mathrm{Li}_n$ are of *pure* form, involving only $\mathrm{Li}_n$ and ordinary logarithms. No exceptions to this observation are known, and it would be valuable to have a proof that was more direct than one that simply relied on verification in each case. The comparison with the logarithm-removal property of Rogers' function and its generalization is quite close in this regards. Be that as it may, this purity property, at least for $n \leq 5$, is enough to show the order-independence feature of accessible ladders as far as the transcendental terms are concerned; the logarithmic part is treated similarly through the use of Rogers' function.

The constant term in (3.2) can be removed by noting that when $z = 0$ the equation reduces to

$$0 = \mathrm{Li}_n(1) + (-1)^{n-1}\Lambda_n(-1)/(n-1)!, \tag{3.3}$$

so they mutually cancel.

**3.2. Functional equations for the polylogarithm.**

3.2.1. *Results by integration.* At the second order the equations of Hill, Kummer, and Schaeffer, (1.18) to (1.20), exemplify forms for the dilogarithm. Many different variants can be obtained by making harmonic group transformations therein for either variable. None of these equations integrates directly to the third order, but by combining three of them [1] the integration can be done. Similarly five variants of the third-order functional equation are needed to obtain an integration to the fourth order. Apart from a special case when $x = y$, when only two forms are needed, this process has not been attempted to the fifth order; rather, Kummer's equation for $\Lambda_5$ has been converted to one for $\mathrm{Li}_5$ through the use of (3.2) [2]. The transformation is straightforward but extremely laborious because of the many terms involved and the extensive use needed of the lower-order functional equations to simplify and, in fact, to completely eliminate them; thus confirming the purity feature whereby only $\mathrm{Li}_5$ terms are needed. It may be noted that the $\mathrm{Li}_3$ formula involves $\zeta(3)$, but the $\mathrm{Li}_4$ and $\mathrm{Li}_5$ equations are absent this term; this exemplifies the conjecture that $\zeta(n)$, for $n$ odd, is essentially absent from functional equations of degree greater than $n$.

In Chapter 6 three separate families of single-variable equations are integrated, through to the fifth order in two cases. Because of the large number of terms involved, the MACSYMA [3] computer program for handling symbolic forms is ideal, once the equations have been but in appropriate form.

3.2.2. *Functional equations for the polylogarithm.* There are six groups of equations that are relevant to the present studies. The first two are single-variable and are relatively simple to prove. They exist for all orders, and are almost trivial in comparison; variations of ladder results depending on them are termed trivial variants. The third group consists of Kummer's equations for $2 \leq n \leq 5$ in polylogarithmic form. The fourth consists of multivariable equations for the second and third orders, and due, respectively to Rogers and Sandham [4]. A generalization, due to Ray, is given in Chapter 7. The fifth, considered in detail in Chapter 9, extends Kummer's formulas for automorphic functions other than the harmonic group, though the latter appears to be the only set to give real arguments when the variables are real. The sixth group consists of the three families of single-variable equations, referred to above, developed as *ad hoc* formulas to provide analytic derivations for some otherwise numerically-determined ladders. None of the equations in the last two groups are capable of extension to the sixth order, and proofs of this are developed in Chapter 9. Single-variable and two-variable equations to the sixth order, and single-variable equations to the seventh order, recently obtained by H. Gangl, are explained in Appendix A and in Chapter 16.

The formulas of the first three groups follow.

3.2.2.1. *Duplication formula.*

$$\mathrm{Li}_n(z) + \mathrm{Li}_n(-z) = 2^{1-n}\,\mathrm{Li}_n(z^2). \tag{3.4}$$

The corresponding formula for $\Lambda_n$ is almost the same;

$$\Lambda_n(z) + \Lambda_n(-z) = 2^{1-n}\Lambda_n(-z^2). \tag{3.5}$$

3.2.2.2. *Inversion formula.*

$$\mathrm{Li}_n(-z)+(-1)^n\,\mathrm{Li}_n(-1/z) = -\frac{1}{n!}\log^n(z)+2\sum_{r=1}^{[n/2]}\frac{\log^{n-2r}(z)}{(n-2r)!}\,\mathrm{Li}_{2r}(-1), \tag{3.6}$$

where $[n/2]$ is the greatest integer contained in $n/2$.

In (3.6) the constants $\mathrm{Li}_{2r}(-1)$ can be put in terms of the Bernoulli numbers through

$$\mathrm{Li}_{2r}(-1) = -(2^{2r-1}-1)B_r\pi^{2r}/(2r)!. \tag{3.7}$$

The corresponding formula for $\Lambda_n$ is somewhat simpler:

$$\Lambda_n(z) + (-1)^n\Lambda_n(1/z) = \text{constant} + (1/n)\log^n|z|, \tag{3.8}$$

where the constant depends on the sign of $z$. (This feature is a characteristic of many of Kummer's equations.)

3.2.2.3. *Two-variable functional equations of Kummer.* Using Kummer's notation $(1-x) = \xi$ and $(1-y) = \eta$, Kummer's equations, in polylogarithmic form, are:

(a) $n = 2$.

(3.9)

$$\mathrm{Li}_2\left(\frac{x\eta^2}{y\xi^2}\right) = \mathrm{Li}_2\left(\frac{-x\eta}{\xi}\right) + \mathrm{Li}_2\left(\frac{-\eta}{y\xi}\right) + \mathrm{Li}_2\left(\frac{x\eta}{y\xi}\right) + \mathrm{Li}_2\left(\frac{\eta}{\xi}\right) + \frac{1}{2}\log^2 y.$$

(b) $n = 3$.

$$
\begin{aligned}
&\operatorname{Li}_3\left(\frac{x\eta^2}{y\xi^2}\right) + \operatorname{Li}_3(xy) + \operatorname{Li}_3(x/y) = 2\operatorname{Li}_3\left(\frac{x\eta}{y\xi}\right) + 2\operatorname{Li}_3\left(\frac{-x\eta}{\xi}\right) \\
&\quad + 2\operatorname{Li}_3\left(\frac{\eta}{\xi}\right) + 2\operatorname{Li}_3\left(\frac{-\eta}{y\xi}\right) + 2\operatorname{Li}_3(x) + 2\operatorname{Li}_3(y) - 2\zeta(3) \\
&\quad + \log^2 y \log\eta - 2\zeta(2)\log y - \frac{1}{3}\log^3 y .
\end{aligned}
\tag{3.10}
$$

(c) $n = 4$.

$$
\begin{aligned}
&\operatorname{Li}_4\left(\frac{-x^2y\eta}{\xi}\right) + \operatorname{Li}_4\left(\frac{-y^2x\xi}{\eta}\right) + \operatorname{Li}_4\left(\frac{x^2y}{\eta^2\xi}\right) + \operatorname{Li}_4\left(\frac{y^2x}{\xi^2\eta}\right) \\
&= 6\operatorname{Li}_4(xy) + 6\operatorname{Li}_4\left(\frac{xy}{\xi\eta}\right) + 6\operatorname{Li}_4\left(\frac{-xy}{\eta}\right) + 6\operatorname{Li}_4\left(\frac{-xy}{\xi}\right) \\
&\quad + 3\operatorname{Li}_4(x\eta) + 3\operatorname{Li}_4(y\xi) + 3\operatorname{Li}_4\left(\frac{x}{\eta}\right) + 3\operatorname{Li}_4\left(\frac{y}{\xi}\right) + 3\operatorname{Li}_4\left(\frac{-x\eta}{\xi}\right) \\
&\quad + 3\operatorname{Li}_4\left(\frac{-y\xi}{\eta}\right) + 3\operatorname{Li}_4\left(\frac{-x}{\eta\xi}\right) + 3\operatorname{Li}_4\left(\frac{-y}{\eta\xi}\right) \\
&\quad - 6\operatorname{Li}_4(x) - 6\operatorname{Li}_4(y) - 6\operatorname{Li}_4(-x/\xi) - 6\operatorname{Li}_4(-y/\eta) + \frac{3}{2}\log^2\xi\log^2\eta .
\end{aligned}
\tag{3.11}
$$

(d) $n = 5$.

$$
\begin{aligned}
&\operatorname{Li}_5\left(\frac{-x^2y\eta}{\xi}\right) + \operatorname{Li}_5\left(\frac{-y^2x\xi}{\eta}\right) + \operatorname{Li}_5\left(\frac{x^2y}{\eta^2\xi}\right) + \operatorname{Li}_5\left(\frac{y^2x}{\xi^2\eta}\right) + \operatorname{Li}_5\left(\frac{x^2\eta}{y^2\xi}\right) \\
&\quad + \operatorname{Li}_5\left(\frac{x\xi}{y\eta}\right) + \operatorname{Li}_5\left(\frac{-x\xi\eta^2}{y}\right) + \operatorname{Li}_5\left(\frac{-x}{\xi^2y\eta}\right) + \operatorname{Li}_5\left(\frac{x\eta^2}{\xi^2y}\right) \\
&= 9\operatorname{Li}_5(xy) + 9\operatorname{Li}_5\left(\frac{x}{y}\right) + 9\operatorname{Li}_5(x\eta) + 9\operatorname{Li}_5\left(\frac{-xy}{\eta}\right) + 9\operatorname{Li}_5\left(\frac{-x\eta}{y}\right) \\
&\quad + 9\operatorname{Li}_5\left(\frac{-xy}{\xi}\right) + 9\operatorname{Li}_5\left(\frac{-x}{y\xi}\right) \\
&\quad + 9\operatorname{Li}_5\left(\frac{-x\eta}{\xi}\right) + 9\operatorname{Li}_5\left(\frac{-x}{\eta\xi}\right) + 9\operatorname{Li}_5\left(\frac{xy}{\xi\eta}\right) \\
&\quad + 9\operatorname{Li}_5\left(\frac{x\eta}{\xi y}\right) + 9\operatorname{Li}_5(\xi y) + 9\operatorname{Li}_5(\xi\eta) + 9\operatorname{Li}_5(-\xi y/\eta) + 9\operatorname{Li}_5(y/\xi) \\
&\quad + 9\operatorname{Li}_5\left(\frac{\eta}{\xi}\right) + 9\operatorname{Li}_5\left(\frac{-y}{\eta\xi}\right) + 9\operatorname{Li}_5\left(\frac{x}{\eta}\right) - 18\operatorname{Li}_5\left(\frac{-x}{\xi}\right) - 18\operatorname{Li}_5(x) \\
&\quad - 18\operatorname{Li}_5(\xi) - 18\operatorname{Li}_5(y) - 18\operatorname{Li}_5(\eta) - 18\operatorname{Li}_5(-y/\eta) + 18\zeta(5) \\
&\quad + 18\zeta(4)\log\xi + 3\zeta(2)\log^2\xi\log(\xi/\eta^3) + \tfrac{3}{2}\log^2\xi\log^2\eta\log(y^3/\eta) \\
&\quad + \tfrac{3}{20}\log^4\xi\log(\xi^2y^5/x^5) .
\end{aligned}
\tag{3.12}
$$



There are variants of these equations obtained by utilizing harmonic group forms for $x$ and $y$, but the number decreases as the order increases, and at the fifth order there is only the single form (3.12), since the equation itself utilizes all the harmonic-group structures, as discussed in relation to (3.70).

### 3.3. A generalization of Rogers' function to the $n$th order.

3.3.1. *A definition of* $L_n(z)$. In reference [5] it was shown that the transformation

$$\operatorname{Li}_3(z) = L_3(z) + \log|z| L_2(z) - \frac{1}{6}\log^2|z|\log(1-z) \tag{3.13}$$

removes the logarithmic terms from Landen's third-order functional equation (1.13), and that it could be further generalized by defining $L_n(z)$ through

$$\operatorname{Li}_n(z) = L_n(z) + \sum_{r=1}^{n-2} \frac{\log^r|z| L_{n-r}(z)}{r!} - \frac{1}{n!}\log^{n-1}|z|\log(1-z) . \tag{3.14}$$

However, this is a *definition* only; a *direct proof* that this leads to logarithm-free equations for $L_n$ is currently lacking. What can be shown is (i) $L_n(z)$ as so defined achieves the claimed property if used on the fifth-order formula (3.12), and (ii) it is the only possible (differentiable) function that *can* do this for any order. This leaves open the question of whether it in fact always *does* so, though no exceptions to particular cases are known; the success in developing the known ladder results into the trans-Kummer range ($6 \le n \le 9$) depends on it. This amounts to a substantial verification.

3.3.2. *A formula for* $L_n(z)$ *by induction.* To proceed, first differentiate (3.14) to obtain

$$
\begin{aligned}
\frac{\operatorname{Li}_{n-1}(z)}{z} &= L_n'(z) + \sum_{r=1}^{n-2} \frac{\log^r|z| L_{n-r}'(z)}{r!} - \frac{n-1}{n!}\frac{\log^{n-2}|z|\log(1-z)}{z} \\
&\quad + \frac{1}{n!}\frac{\log^{n-1}|z|}{1-z} + \frac{1}{z}\left\{L_{n-1}(z) + \sum_{r=1}^{n-3}\frac{\log^r|z| L_{n-r-1}(z)}{r!}\right\} ,
\end{aligned}
\tag{3.15}
$$

(Use has been made of the previously-discussed interpretation $(d/dz)\log|z| = 1/z$.)

In (3.15) almost all the terms with $1/z$ as a factor cancel out on using (3.14) with $n-1$ for $n$, and the equation simplifies to

$$L_n'(z) + \sum_{r=1}^{n-2}\frac{\log^r|z| L_{n-r}'(z)}{r!} + \frac{\log^{n-2}|z|}{n!}\left[\frac{\log|z|}{1-z} + \frac{\log(1-z)}{z}\right] = 0 . \tag{3.16}$$

This can be solved as a difference equation in $L_n'$ and gives simply

$$L_n'(z) = (-1)^{n-1}\frac{(n-1)}{n!}\left[\frac{\log^{n-1}|z|}{1-z} + \frac{\log^{n-2}|z|\log(1-z)}{z}\right] . \tag{3.17}$$

On integration this becomes

$$L_n(z) = \frac{(-1)^{n-1}}{n!}[\log^{n-1}|z|\log(1-z) - n\Lambda_n(-z)], \tag{3.18}$$

and this could be taken as an alternative definition of $L_n$. It has been tested for Kummer's two-variable functional equation for $\Lambda_5$. The process is quite tedious, and the seven formulas of (7.102) and (7.103) of [1, Chapter 7] are all needed to handle the very many terms that are generated. The ultimate effect, however, is to demonstrate the claimed feature of self-cancellation of all the logarithmic terms, confirming (3.18) or (3.14) as the needed relation to achieve this, at least for $n \le 5$.

Using (3.18) and (3.2) the relation in (3.14) can now easily be inverted to give $L_n$ in terms of polylogarithms:

$$\begin{aligned} L_n(z) = \mathrm{Li}_n(z) + \sum_{r=1}^{n-2}(-1)^r \frac{\log^r|z|\,\mathrm{Li}_{n-r}(z)}{r!} \\ + (-1)^n\frac{(n-1)}{n!}\log^{n-1}|z|\log(1-z). \end{aligned} \tag{3.19}$$

As far as $n > 5$ is concerned the only analytic relations are the duplication and inversion formulas (3.5) and (3.8), and it is readily shown that (3.18) removes the logarithmic terms therein. And as will now be demonstrated, (3.18) is the *only* appropriate relation that can achieve this.

3.3.3. *The uniqueness of the definition of* $L_n(z)$. Putting (3.18) into (3.5) and (3.8) leads to a pair of functional equations developed and solved as follows: inserting (3.18) into (3.5), and into (3.8) and removing the logarithmic terms, since they self-cancel, leads to

$$L_n(z) + L_n(-z) = 2^{1-n}L_n(z^2), \tag{3.20}$$

$$L_n(-z) + (-1)^n L_n(-1/z) = C_n, \tag{3.21}$$

where $C_n$ is a constant depending only on $n$. Consider the pair of functional equations for an unknown $F_n(z)$;

$$F_n(z) + F_n(-z) = 2^{1-n}F_n(z^2), \tag{3.22}$$

and

$$F_n(-z) + (-1)^n F_n(-1/z) = C_n. \tag{3.23}$$

Put $F_n(z) = L_n(z) + g_n(z)$ so that this pair of equations reduces to

$$g_n(z) + g_n(-z) = 2^{1-n}g_n(z^2), \tag{3.24}$$

and

$$g_n(-z) + (-1)^n g_n(-1/z) = 0. \tag{3.25}$$



Equation (3.25) essentially says that $g_n$ is an even function in $(\log|z|)$, apart from an initial factor to give $(-1)^{n-1}$ on inversion of $z$, whence

$$g_n(z) = \log^{n-1}|z|G_n[(\log|z|)^2], \tag{3.26}$$

for some function $G_n(w)$ where $w = (\log|z|)^2$. Substitution of this into (3.24) gives

$$2\log^{n-1}|z|G_n(w) = 2^{1-n}\cdot 2^{n-1}\log^{n-1}|z|G_n(4w), \tag{3.27}$$

or

$$G_n(4w) = 2G_n(w). \tag{3.28}$$

This has the solution *

$$G_n(w) = K_n \cdot |w|^{1/2} = K_n \cdot |\log|z||, \tag{3.29}$$

whence

$$g_n(z) = K_n \cdot (\log|z|)^{n-1}|\log|z||, \tag{3.30}$$

for some constant $K_n$ that possibly depends on $n$. This function satisfies (3.24) and (3.25) for any $K_n$, but insertion into any of the Kummer equations for $n \le 5$, or the numerically-determined trans-Kummer ladders for $6 \le n \le 9$ shows that $K_n$ must be zero. Apart from a very dubious possible nonzero value for $n > 9$, for which no data exists, this determines $K_n$ and hence $g_n(z)$ to be zero. Accordingly, $F_n(z) = L_n(z)$, so that $L_n(z)$ is the unique function that could, in principle, possess the desired logarithm-suppressing feature.

**3.4. Ladder order-independence on reduction of order.**

3.4.1. *The transcendental terms.* According to (3.2) any $\Lambda_n$ ladder obtainable from the functional equation for $\Lambda_n$ would translate into a comparable ladder for $\mathrm{Li}_n$ with some extra polylogarithmic (and logarithmic) terms of lower order. Corresponding to a term $A_m\,\mathrm{Li}_n(u^m)/m^{n-1}$ there will be terms, summed over $r$, of the form

$$\frac{1}{r!}\left[\frac{A_m}{m^{n-1}}\log^r(u^m)\,\mathrm{Li}_{n-r}(u^m)\right] = \frac{\log^r(u)}{r!}\left[A_m\frac{\mathrm{Li}_{n-r}(u^m)}{m^{n-r-1}}\right]. \tag{3.31}$$

Apart from the initial factor $\log^r(u)/r!$, which is common to all terms involving $\mathrm{Li}_{n-r}$, this is of the form required to ensure that a ladder of degree $n$ is associated with sums of polylogarithmic terms of lower order, each having the identical ladder structure. However, the purity property dictates that there shall be no such terms surviving. Hence, apart from apportioning the ordinary logarithmic terms in an appropriate way, these lower-order polylogarithmic ladders must each be zero. This confirms the invariance of the structure (1.26) in so far as the transcendental part is concerned.

* D. Zagier has pointed out that (3.29) could be multiplied by any function periodic in $\log w$ of period $\log 4$.

3.4.2. *The logarithmic terms.* Suppose that known analytical equations, for which the generalized Rogers' function operates according to the dictates of §3.3, lead to a valid ladder in the form

$$\frac{\mathrm{Li}_n(u^N)}{N^{n-1}} - \left\{\sum_r \frac{A_r \,\mathrm{Li}_n(u^r)}{r^{n-1}} + \frac{A_0 \log^n(u)}{n!}\right\} = D_n\zeta(n) + L_n\,, \tag{3.32}$$

where $L_n$ is an expression, as in (1.29), involving only logarithms. The transformation (3.14) will give rise to a parallel relation for the generalized Rogers' function, but omitting the logarithmic terms:

$$\frac{L_n(u^N)}{N^{n-1}} - \left\{\sum_r \frac{A_r L_n(u^r)}{r^{n-1}}\right\} = D_n\zeta(n)\,. \tag{3.33}$$

However, if the transformation is actually performed, it takes the form

(3.34)
$$\begin{aligned}\frac{1}{N^{n-1}}&\left[L_n(u^N) + \sum_{m=1}^{n-2} \frac{\log^m |u^N| L_{n-m}(u^N)}{m!} - \frac{1}{n!}\log^{n-1}(u^N)\log(1-u^N)\right]\\ &= \sum_r \frac{A_r}{r^{n-1}}\left[L_n(u^r) + \sum_{m=1}^{n-2} \frac{\log^m |u^r| L_{n-m}(u^r)}{m!} - \frac{1}{n!}\log^{n-1}(u^r)\log(1-u^r)\right]\\ &\quad + D_n\zeta(n)A_0\log^n(u)/n! + L_n\,.\end{aligned}$$

On using (3.33) with orders from 2 to $n$, together with (2.65), the cyclotomic equation, in logarithmic form, (3.34) reduces to

$$L_n = \sum_{m=1}^{n-2} \frac{\log^m(u)}{m!} D_{n-m}\zeta(n-m)\,. \tag{3.35}$$

Together with the term $D_n\zeta(n)$, and putting $n-m$ for $m$, this becomes (1.29), confirming the claimed invariance property of the ladder logarithmic component. Incidentally it defines the constants $D_n$ through (3.33).

**3.5. Generic ladders for the base equation $u^p + u^q = 1$.**

3.5.1. *Cyclotomic equations.* This base equation was introduced in §2.1.2, and is the only base equation giving ladders up to $n = 5$ via Kummer's equations. It possesses eight cyclotomic equations, obtained by simple manipulation of the base equation. As far as is known there are no others, for general $p$ and $q$, though there is currently no proof of this. These equations



are

$$1 - u^q = u^p\,, \tag{3.36}$$
$$1 - u^p = u^q\,, \tag{3.37}$$
$$1 - u^{2(p-q)} = (1-u^{p-q})u^{-q}\,, \tag{3.38}$$
$$1 - u^{2(2p-q)} = (1-u^{2p-q})(1-u^{2(p-2q)})(1-u^{p-2q})^{-1}u^q\,, \tag{3.39}$$
$$1 - u^{3(p-q)} = (1-u^{p-q})(1-u^{2(p-2q)})(1-u^{p-2q})^{-1}\,, \tag{3.40}$$
$$1 - u^{p+q} = (1-u^{2(p-2q)})(1-u^{p-2q})^{-1}u^{2q}\,, \tag{3.41}$$
$$1 - u^{6q} = (1-u^{3q})(1-u^{2q})(1-u^{2(p-2q)})(1-u^{p-2q})^{-1}u^{2q-p}\,, \tag{3.42}$$
$$1 - u^{6p} = (1-u^{3p})(1-u^{2p})(1-u^{2(p-2q)})(1-u^{p-2q})^{-1}u^q\,. \tag{3.43}$$

It will be noticed that $(p-2q)$ occurs in many places in the above formulas. For the smaller values of $p$ and $q$ this will often be a factor of the indices appearing on the left-hand sides of these equations, giving rise to the observed result that, in many cases, all the powers $r$ in (1.31) are factors of the index. For the record, these generic indices are:

$$N = \{q\,, p\,, 2(p-q)\,, 2(2p-q)\,, 3(p-q)\,, p+q\,, 6q\,, 6p\}\,. \tag{3.44}$$

Apart from the special case $p = q = 1$, $(u = 1/2)$, we take $p$ as the greater of $p$ and $q$; and for the interim, $p > 2q$ also. This ensures that all the powers in (3.36) to (3.43) are positive. Modifications for the ladders with $q < p < 2q$ are found by inverting those arguments with negative exponent. The special case $p = 2q$ is discussed in more detail in §§4.1.2 and 5.2.1.

3.5.2. *Component-ladders.* Corresponding to the eight cyclotomic equations, eight ladders can be written down. Thus, from (3.36), one can define, at $n = 2$, the ladder

$$L_2(q\,, u) = \frac{\mathrm{Li}_2(u^q)}{q} + \frac{p}{2}\log^2 u\,. \tag{3.45}$$

It is only in certain very special cases that the ladders constructed directly in this way are valid, and it may be preferable to think of them more as *component-ladders*, i.e., pieces of what will eventually be a valid ladder. The two known exceptions in this family are (i) $p = q = 1$; $u = 1/2$, for which the indices degenerate into only two surviving values, namely $N = 1$ and 6; and (ii) $q = 1$, $p = 2$; $u \equiv \rho = (5^{1/2}-1)/2$, for which the indices degenerate to $N = 1$, 2, 6, and 12. In both these cases the ladders pseudo-integrate directly [2] to give valid results as far as $n = 5$ (for additional results on $\rho$, see §§4.1.2 and 4.1.3), and this property is called *transparency*. The only other known transparent results stem from the family $u^2 - mu + 1 = 0$, $2 \le m \le 10$ $(m \ne 9)$ as discussed in references [2] and [6]; and a new result, stemming from research on cyclotomic equations generated by quadratic equations, and given in §2.7.1.

3.5.3. *Euler's and Landen's forms.* The first three cyclotomic equations, (3.36) to (3.38) can be grouped together. Algebraically they are almost trivial rearrangements of the base equation and correspond to three component-ladders which, at $n = 2$, combine in pairs to yield two valid ladders. These two then combine to give a single valid ladder at $n = 3$, and which involves $\zeta(3)$ via $D_3$ in (1.29). These are the well-established results of Euler and Landen, and are given in (3.54), (3.55), and (3.64).

This tendency for ladders to combine in pairs, resulting in one fewer at the next higher order, in going from even to odd orders, is very marked.

3.5.4. *The five main component-ladders.* The remaining five cyclotomic equations give five component-ladders at $n = 2$ which combine in pairs to give four valid ladders at $n = 2$ and three at $n = 3$. At this juncture we encounter the effect of the arbitrariness of the combinations used to construct the cyclotomic equations. Thus, in (3.43), if the final factor $u^q$ is replaced by $(1 - u^p)$ from (3.37), the effect is to add a component-ladder in $L_2(p, u)$ into the definition of $L_2(6p, u)$. This affects the ensuing ladder at $n = 3$, altering the multiplier of $\zeta(3)$ therein. Now the valid ladders at $n = 4$ contain no net $\zeta(3)$ term. It is a conjecture that, quite generally, zeta functions of odd argument less than $n$ do not appear in ladders of order $n$; and this hypothesis was critical in enabling the trans-Kummer results to $n = 9$ to be generated. Thus we have two options: either to use Landen's result at $n = 3$ to cancel out any $\zeta(3)$ term before proceeding to $n = 4$; or, with some hindsight, to choose combinations of component-ladders so that the $\zeta(3)$ term is absent from the beginning. The outcome is the same, of course, but the *form* taken by the results is simpler if the second alternative is taken. Now it happens that a very convenient functional equation, (6.108) of [1], is absent any $\zeta(3)$ term, and if the valid ladders at $n = 3$ are generated from it, combinations of component-ladders free of $\zeta(3)$ are produced quite naturally. Thus the component-ladder from (3.40) less two thirds of that from (3.38) is used; that from (3.42) plus one third that from (3.36); and (3.43) plus one third of (3.37). The resulting component-ladders, for general $n$, are given in (3.46) to (3.53) and the valid formulas at $n = 2$ in (3.54) to (3.59).

*Component-ladders*

(3.46) $$L_n(q, u) = \frac{\mathrm{Li}_n(u^q)}{q^{n-1}} + \frac{p \log^n(u)}{n!},$$

(3.47) $$L_n(p, u) = \frac{\mathrm{Li}_n(u^p)}{p^{n-1}} + \frac{q \log^n(u)}{n!},$$

(3.48) $$L_n[2(p-q), u] = \frac{\mathrm{Li}_n(u^{2(p-q)})}{[2(p-q)]^{n-1}} - \left\{\frac{\mathrm{Li}_n(u^{p-q})}{(p-q)^{n-1}} + \frac{q \log^n(u)}{n!}\right\},$$



(3.49) $$L_n[2(2p-q), u] = \frac{\mathrm{Li}_n(u^{2(2p-q)})}{[2(2p-q)]^{n-1}} - \left\{\frac{\mathrm{Li}_n(u^{2p-q})}{(2p-q)^{n-1}} + \frac{\mathrm{Li}_n(u^{2(p-2q)})}{[2(p-2q)]^{n-1}} - \frac{\mathrm{Li}_n(u^{p-2q})}{(p-2q)^{n-1}} - \frac{q \log^n(u)}{n!}\right\},$$

(3.50) $$L_n[3(p-q), u] = \frac{\mathrm{Li}_n(u^{3(p-q)})}{[3(p-q)]^{n-1}} - \left\{\frac{2}{3}\frac{\mathrm{Li}_n(u^{2(p-q)})}{[2(p-q)]^{n-1}} + \frac{\mathrm{Li}_n(u^{p-q})}{3(p-q)^{n-1}} + \frac{\mathrm{Li}_n(u^{2(p-2q)})}{[2(p-2q)]^{n-1}} - \frac{\mathrm{Li}_n(u^{p-2q})}{(p-2q)^{n-1}} - \frac{2q}{3}\frac{\log^n(u)}{n!}\right\},$$

(3.51) $$L_n(p+q, u) = \frac{\mathrm{Li}_n(u^{p+q})}{(p+q)^{n-1}} - \left\{\frac{\mathrm{Li}_n(u^{2(p-2q)})}{[2(p-2q)]^{n-1}} - \frac{\mathrm{Li}_n(u^{p-2q})}{(p-2q)^{n-1}} - \frac{2q \log^n(u)}{n!}\right\},$$

(3.52) $$L_n(6q, u) = \frac{\mathrm{Li}_n(u^{6q})}{(6q)^{n-1}} - \left\{\frac{\mathrm{Li}_n(u^{3q})}{(3q)^{n-1}} + \frac{\mathrm{Li}_n(u^{2q})}{(2q)^{n-1}} - \frac{\mathrm{Li}_n(u^{q})}{3q^{n-1}} + \frac{\mathrm{Li}_n(u^{2(p-2q)})}{[2(p-2q)]^{n-1}} - \frac{\mathrm{Li}_n(u^{p-2q})}{(p-2q)^{n-1}} + \frac{2(p-3q)}{3}\frac{\log^n(u)}{n!}\right\},$$

(3.53) $$L_n(6p, u) = \frac{\mathrm{Li}_n(u^{6p})}{(6p)^{n-1}} - \left\{\frac{\mathrm{Li}_n(u^{3p})}{(3p)^{n-1}} + \frac{\mathrm{Li}_n(u^{2p})}{(2p)^{n-1}} - \frac{\mathrm{Li}_n(u^{p})}{3p^{n-1}} + \frac{\mathrm{Li}_n(u^{2(p-2q)})}{[2(p-2q)]^{n-1}} - \frac{\mathrm{Li}_n(u^{p-2q})}{(p-2q)^{n-1}} - \frac{4q}{3}\frac{\log^n(u)}{n!}\right\}.$$

*Valid ladders at $n = 2$.* From Kummer's functional equations it is found that:

(3.54) $qL_2(q, u) + pL_2(p, u) = \zeta(2)$,

(3.55) $pL_2(p, u) + (p-q)L_2[2(p-q), u] = 0$,

(3.56) $3(p-q)L_2[3(p-q), u] - (2p-q)L_2[2(2p-q), u] = 0$,

(3.57) $(p+q)L_2(p+q, u) - (2p-q)L_2[2(2p-q), u] = 0$,

(3.58) $3qL_2(6q, u) - (2p-q)L_2[2(2p-q), u] = -\zeta(2)$,

(3.59) $3pL_2(6p, u) - 2(2p-q)L_2[2(2p-q), u] = 0$.

As already mentioned, (3.54) and (3.55) are essentially due to Euler and Landen.

As can be readily seen, there are six relations between the eight component-ladders, so two of them are independent; they are conveniently taken to be $L_2(q,u)$ or $L_2(p,u)$, and $L_2[2(2p-q),u]$; and these are candidates for combination with the supernumary component-ladders considered in Chapter 5.

3.5.5. *Trilogarithmic ladders.* At $n=3$ certain combinations of component-ladders give rise to valid structures, as found from the functional equations. Again there is a certain amount of arbitrariness, but the following second-degree ladders are convenient to define:

$$L_n^{(2)}(p,u) = p^2L_n(p,u) + q^2L_n(q,u) + (p-q)^2L_n[2(p-q),u], \tag{3.60}$$

$$L_n^{(2)}[3(p-q),u] = 9(p-q)^2L_n[3(p-q),u] + (p+q)^2L_n(p+q,u) - 2(2p-q)^2L_n[2(2p-q),u], \tag{3.61}$$

$$L_n^{(2)}(6q,u) = 9q^2L_n(6q,u) - 2(p+q)^2L_n(p+q,u) + (2p-q)^2L_n[2(2p-q),u], \tag{3.62}$$

$$L_n^{(2)}(6p,u) = 9p^2L_n(6p,u) - 2(p+q)^2L_n(p+q,u) - 2(2p-q)^2L_n[2(2p-q),u]. \tag{3.63}$$

Then at $n=3$ we have

$$L_3^{(2)}(p,u) = \zeta(3) + \zeta(2)q\log(u), \tag{3.64}$$

$$L_3^{(2)}[3(p-q),u] = 0, \tag{3.65}$$

$$L_3^{(2)}(6p,u) = 0, \tag{3.66}$$

$$L_3^{(2)}(6q,u) = -3\zeta(2)q\log(u). \tag{3.67}$$

Landen's formula reduces to (3.64) and, as already discussed, the remaining results are free of $\zeta(3)$. In extending results to $n=4$, (3.64) is therefore not needed, in the general case, and only (3.65) to (3.67) are used.

3.5.6. *Fourth- and fifth-order ladders.* At $n=4$ these remaining three equations combine in pairs to give

$$pL_4^{(2)}(6p,u) + (p-q)L_4^{(2)}[3(p-q),u] = 0, \tag{3.68}$$

$$qL_4^{(2)}(6q,u) - (p-q)L_4^{(2)}[3(p-q),u] = \tfrac{19}{12}\zeta(4) - \tfrac{3}{2}q^2\zeta(2)\log^2(u). \tag{3.69}$$

Were it not that (3.65) to (3.67) are already free of $\zeta(3)$, (3.64) would be needed to eliminate it first, so the equations really combine in threes to produce two results from four initial ladders. In the supernumary cases at higher



orders we see this type of combining more clearly in effect: ladders combine in threes when going from odd to even orders, resulting in a loss of two results as the order is increased.

At $n=5$ there is only one valid ladder coming from combining the forms in (3.68) and (3.69):

$$3p^2L_5^{(2)}(6p,u) + 3q^2L_5^{(2)}(6q,u) + 3(p-q)^2L_5^{(2)}[3(p-q),u] = 7\zeta(5) + \tfrac{19}{4}\zeta(4)q\log(u) - \tfrac{3}{2}\zeta(2)q^3\log^3(u). \tag{3.70}$$

All these results come from Kummer's functional equations, or by pseudo-integration of the earlier results. However, the latter method does not give the $D_m$ coefficients, the multipliers of $\zeta(m)$, and these have to found numerically or by other means.

This is as far as one can go in the general case. One can hypothesize that *at least* one more result containing $\zeta(5)$, and, more likely, two such, would be required to reach a result at $n=6$, where, as previously stated, no equations of Kummer's type exist. Two nontrivial single-variable functional equations at $n=6$ and 7 have recently been found by H. Gangl.

3.5.7. *Flow chart.* Figure 3.1 is a flow chart showing how the ladders at one order combine to give results at the next. The three elementary cyclotomic equations (order 1) give two valid ladders at order two, and one at order 3, where this sequence stops. The remaining five cyclotomic equations

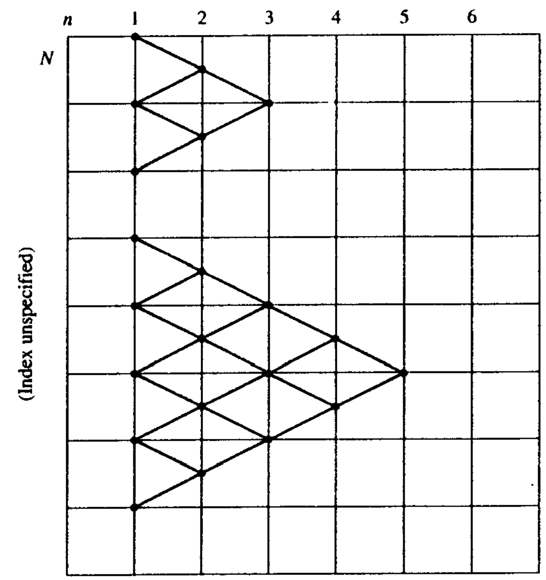



FIGURE 3.1. Flow chart for general case

give four results at the second order, decreasing by one each time the order increases, and terminating with a single result at $n = 5$. It is believed that this is as far as the sequence goes in the general case (arbitrary $p$ and $q$).

One interesting consequence that can be drawn from these results concerns the special case $p = q = 1$, $u = 1/2$. Euler's and Landen's formulas give $\mathrm{Li}_2(1/2)$ and $\mathrm{Li}_3(1/2)$ in terms of $\log(1/2)$, $\zeta(2)$ and $\zeta(3)$. But according to Figure 3.1 the sequence stops there (because of the term in $\zeta(3)$), so there is no corresponding formula for $\mathrm{Li}_4(1/2)$ in terms of, say, $\log(1/2)$, $\zeta(2)$ and $\zeta(4)$. Some researchers have sought (in vain) to find such a result via numerical computations. According to Figure 3.1, no such result is to be expected, though this demonstration falls short of constituting a rigorous proof.

**3.6. Examples of ladders for $n \le 3$.**

3.6.1. *Accessible ladders.* Many of the ladders extending beyond $n = 2$ go to $n = 5$. This is true for the two-term base equation (2.1) but not for the three-term relation (2.10). In this section results are given only for ladders that extend to $n = 3$ but no further. Both Euler's and Landen's results fall into this category and they can be put in the form

(3.71)
$$L_n(2, \rho) = \frac{\mathrm{Li}_n(\rho^2)}{2^{n-1}} - \left\{ \frac{-\log^n \rho}{n!} + \frac{\zeta(2)}{5}\frac{\log^{n-2}\rho}{(n-2)!} + \frac{\zeta(3)}{5}\frac{\log^{n-3}\rho}{(n-3)!} \right\} = 0,$$

and

(3.72)
$$L_n(1, \delta) = \mathrm{Li}_n(\delta) - \left\{ -\frac{\log^n \delta}{n!} + \frac{\zeta(2)}{2}\frac{\log^{n-2}\delta}{(n-2)!} + \frac{7}{8}\zeta(3)\frac{\log^{n-3}\delta}{(n-3)!} \right\} = 0.$$

3.6.1.1. The base $\varepsilon = 1/3$ satisfies (2.10) with $r = s = m = 1$, with the plus sign for the third term. The cyclotomic equation is $(1-\varepsilon^2) = (1-\varepsilon)^3\varepsilon^{-1}$ and the corresponding ladder, valid for $1 \le n \le 3$, is

(3.73)
$$L_n(2, \varepsilon) = \frac{\mathrm{Li}_n(\varepsilon^2)}{2^{n-1}} - \left\{ 3\,\mathrm{Li}_n(\varepsilon) + \frac{\log^n \varepsilon}{n!} - \zeta(2)\frac{\log^{n-2}\varepsilon}{(n-2)!} - \frac{13}{6}\zeta(3)\frac{\log^{n-3}\varepsilon}{(n-3)!} \right\} = 0.$$

3.6.1.2. The base $\tau = \tan(\pi/8) = \sqrt{2} - 1$ satisfies $u^2 + u + u = 1$ and possesses the cyclotomic equation $(1 - \tau^4)^2 = (1 - \tau^2)^5\tau^{-1}$, for which the corresponding ladder is

(3.74)
$$L_n(4, \tau) = \frac{\mathrm{Li}_n(\tau^4)}{4^{n-1}} - \left\{ \frac{5}{2}\frac{\mathrm{Li}_n(\tau^2)}{2^{n-1}} + \frac{1}{2}\frac{\log^n \tau}{n!} - \frac{\zeta(2)}{4}\frac{\log^{n-2}\tau}{(n-2)!} - \frac{11}{32}\zeta(3)\frac{\log^{n-3}\tau}{(n-3)!} \right\} = 0.$$



This can also be presented as a formula for $\tau_2 = \tau^2 = 3 - 2\sqrt{2}$, which is the case $m = 6$ of the equation $u^2 - mu + 1 = 0$ considered in §3.6.2. The resulting ladder of index 2 is

(3.75)
$$L_n(2, \tau_2) = \frac{\mathrm{Li}_n(\tau_2^2)}{2^{n-1}} - \left\{ \frac{5}{2}\mathrm{Li}_n(\tau_2) + \frac{1}{4}\frac{\log^n \tau_2}{n!} - \frac{\zeta(2)}{2}\frac{\log^{n-2}\tau_2}{(n-2)!} - \frac{11}{8}\zeta(3)\frac{\log^{n-3}\tau_2}{(n-3)!} \right\} = 0.$$

3.6.1.3. The base $\sigma$ is the real root of $\sigma^3 + \sigma^2 + \sigma = 1$ and satisfies at least four cyclotomic equations

(3.76) $$1 - \sigma^2 = (1-\sigma)^2\sigma^{-2},$$

(3.77) $$1 - \sigma^3 = (1-\sigma)\sigma^{-1},$$

(3.78) $$(1-\sigma^4)^2 = (1-\sigma^2)^3(1-\sigma)^2\sigma^{-4},$$

(3.79) $$(1-\sigma^8)^8 = (1-\sigma^4)^{13}(1-\sigma^2)^{5/2}(1-\sigma)^{-1}\sigma^{-1}.$$

The corresponding component-ladders are

(3.80)
$$L_n(2, \sigma) = \frac{\mathrm{Li}_n(\sigma^2)}{2^{n-1}} - \left\{ 2\,\mathrm{Li}_n(\sigma) + 2\frac{\log^n \sigma}{n!} \right\},$$

(3.81)
$$L_n(3, \sigma) = \frac{\mathrm{Li}_n(\sigma^3)}{3^{n-1}} - \left\{ \mathrm{Li}_n(\sigma) + \frac{\log^n \sigma}{n!} \right\},$$

(3.82)
$$L_n(4, \sigma) = \frac{\mathrm{Li}_n(\sigma^4)}{4^{n-1}} - \left\{ \frac{3}{2}\frac{\mathrm{Li}_n(\sigma^2)}{2^{n-1}} + \mathrm{Li}_n(\sigma) + 2\frac{\log^n \sigma}{n!} - \frac{3}{4}\zeta(2)\frac{\log^{n-2}\sigma}{(n-2)!} \right\},$$

(3.83)
$$L_n(8, \sigma) = \frac{\mathrm{Li}_n(\sigma^8)}{8^{n-1}} - \left\{ \frac{13}{8}\frac{\mathrm{Li}_n(\sigma^4)}{4^{n-1}} + \frac{5}{16}\frac{\mathrm{Li}_n(\sigma^2)}{2^{n-1}} - \frac{\mathrm{Li}_n(\sigma)}{8} + \frac{1}{8}\frac{\log^n \sigma}{n!} - \frac{\zeta(2)}{32}\frac{\log^{n-2}\sigma}{(n-2)!} + \frac{\zeta(3)}{64}\frac{\log^{n-3}\sigma}{(n-3)!} \right\}.$$

A relevant second-degree ladder stemming from these structures is

(3.84) $$L_n^{(2)}(3, \sigma) = 3L_n(3, \sigma) - L_n(2, \sigma) + \frac{\zeta(2)}{2}\frac{\log^{n-2}\sigma}{(n-2)!},$$

and for $1 \le n \le 2$ we have

(3.85) $$\begin{cases} L_n^{(2)}(3, \sigma) = 0, \\ L_n(4, \sigma) = 0. \end{cases}$$

A further relevant second-degree ladder is

$$L_n^{(2)}(4, \sigma) = L_n(4, \sigma) - L_n^{(2)}(3, \sigma) + \frac{\zeta(3)}{4}\frac{\log^{n-3}\sigma}{(n-3)!}, \tag{3.86}$$

and for $1 \le n \le 3$ we have

$$\begin{cases} L_n^{(2)}(4, \sigma) = 0, \\ L_n(8, \sigma) = 0. \end{cases} \tag{3.87}$$

3.6.1.4. The base $\zeta$ is the real root of the cubic $\zeta^3 + 2\zeta = 1$. There are three known cyclotomic equations:

$$1 - \zeta^2 = (1-\zeta)^3\zeta^{-2}, \tag{3.88}$$

$$1 - \zeta^4 = (1-\zeta^2)(1-\zeta)\zeta^{-1}, \tag{3.89}$$

$$1 - \zeta^6 = (1-\zeta^3)^2(1-\zeta)\zeta^{-1}. \tag{3.90}$$

The corresponding component-ladders are

$$L_n(2, \zeta) = \frac{\mathrm{Li}_n(\zeta^2)}{2^{n-1}} - \left\{3\,\mathrm{Li}_n(\zeta) + \frac{2\log^n\zeta}{n!}\right\}, \tag{3.91}$$

$$L_n(4, \zeta) = \frac{\mathrm{Li}_n(\zeta^4)}{4^{n-1}} - \left\{\frac{\mathrm{Li}_n(\zeta^2)}{2^{n-1}} + \mathrm{Li}_n(\zeta) + \frac{\log^n\zeta}{n!}\right\}, \tag{3.92}$$

$$L_n(6, \zeta) = \frac{\mathrm{Li}_n(\zeta^6)}{6^{n-1}} - \left\{\frac{2\,\mathrm{Li}_n(\zeta^3)}{3^{n-1}} + \mathrm{Li}_n(\zeta) + \frac{\log^n\zeta}{n!}\right\}. \tag{3.93}$$

From these one can form the valid second-degree ladders:

$$L_n^{(2)}(4, \zeta) = L_n(4, \zeta) - \frac{L_n(2, \zeta)}{4} + \frac{\zeta(2)}{4}\frac{\log^{n-2}\zeta}{(n-2)!}, \tag{3.94}$$

$$L_n^{(2)}(6, \zeta) = L_n(6, \zeta) - \frac{L_n(2, \zeta)}{6} + \frac{\zeta(2)}{6}\frac{\log^{n-2}\zeta}{(n-2)!}. \tag{3.95}$$

Then $L_n^{(2)}(4, \zeta) = L_n^{(2)}(6, \zeta) = 0$ for $n = 1$ and 2. For $n = 3$ a third-degree ladder is formed:

$$L_n^{(3)}(6, \zeta) = 9L_n^{(2)}(6, \zeta) - 8L_n^{(2)}(4, \zeta) + \frac{7}{4}\zeta(3)\frac{\log^{n-3}\zeta}{(n-3)!}. \tag{3.96}$$

Then $L_n^{(3)}(6, \zeta) = 0$ for $1 \le n \le 3$.

3.6.2. *Inaccessible ladders.* In reference [6] the quadratic equation

$$u^2 - mu + 1 = 0 \tag{3.97}$$

was examined. For $m = 2, 3, 4, 6$, and 7 the bases are, respectively,

$$u = 1, \quad \rho_2 = \rho^2, \quad \chi = 2 - \sqrt{3}, \quad \tau_2 = \tau^2 = 3 - 2\sqrt{2}, \quad \rho_4 = \rho^4.$$

Of these the base $\chi$ is treated in §3.7 and $\rho_4$ in Chapter 4. For $m = 5, 8$, and 10 (but not 9) it was shown in the reference how cyclotomic equations could be generated and ladders constructed to correspond to them.



3.6.2.1. As an example, for $m = 5$, the base equation can be arranged as

$$\begin{cases} (1-u)^2 = 3u, \\ 1 - u + u^2 = 4u, \\ 1 + u + u^2 = 6u. \end{cases} \tag{3.98}$$

The elimination of factors 2 and 3 enables (3.98) to be put in the form

$$1 - u^6 = (1-u^3)^3(1-u^2)(1-u)^{-7}u. \tag{3.99}$$

The corresponding dilogarithmic ladder is constructed and found valid. It pseudointegrates transparently to $n = 3$ and the resulting ladder, numerically determined, is

(3.100)

$$\frac{\mathrm{Li}_3(P^6)}{6^2} - \left\{\frac{3\,\mathrm{Li}_3(P^3)}{3^2} + \frac{\mathrm{Li}_3(P^2)}{2^2} - 7\,\mathrm{Li}_3(P) - \frac{\log^3 P}{3!} + \frac{5}{3}\zeta(2)\log P\right\} = \frac{77}{18}\zeta(3),$$

where $P = (5 - \sqrt{21})/2$.

3.6.2.2. For $m = 8$ the cyclotomic equation becomes

$$(1-u^4)^2 = (1-u^3)^{-3}(1-u^2)^2(1-u)^{15}u^{-1}. \tag{3.101}$$

Again the related dilogarithmic ladder is found to be valid, and pseudointegrates transparently to $n = 3$, where it is found numerically that

$$\begin{aligned} \frac{\mathrm{Li}_3(Q^4)}{4^2} - \Bigg\{ -\frac{3}{2}\frac{\mathrm{Li}_3(Q^3)}{3^2} + \frac{\mathrm{Li}_3(Q^2)}{2^2} + \frac{15}{2}\mathrm{Li}_3(Q) \\ + \frac{(1/2)\log^3 Q}{3!} - \frac{5}{4}\zeta(2)\log Q\Bigg\} \\ = -\frac{179}{48}\zeta(3), \end{aligned} \tag{3.102}$$

where $Q = 4 - \sqrt{15}$.

3.6.2.3. The remaining example, for $m = 10$ (the case $m = 9$ has not yielded any cyclotomic equations), gives the structure

$$(1-u^6)^3 = (1-u^3)^3(1-u^2)^{15}(1-u)^{-23}u. \tag{3.103}$$

The related dilogarithmic ladder is found to be valid, and it pseudointegrates transparently to give

(3.104)

$$\frac{\mathrm{Li}_3(R^6)}{6^2} - \left\{\frac{\mathrm{Li}_3(R^3)}{3^2} + 5\frac{\mathrm{Li}_3(R^2)}{2^2} - \frac{23}{3}\mathrm{Li}_3(R) - \frac{1}{3}\frac{\log^3 R}{3!} + \zeta(2)\log R\right\} = \frac{29}{9}\zeta(3),$$

where $R = 5 - 2\sqrt{6}$. As with the others, this result was dependent on numerical determination.

**3.7. Examples of ladders for $n \le 4$.**

3.7.1. *Accessible ladders.* There are relatively few bases that give ladders valid to $n = 4$ but not $n = 5$. The bases $\psi$ and $\phi$ are handled in the wider context of supernumary ladders in Chapter 5, leaving only the one known accessible base, $\rho$. From Kummer's equations an index-6 ladder can be derived and takes the form

(3.105)
$$L_n(6, \rho) = \frac{\mathrm{Li}_n(\rho^6)}{6^{n-1}} - \left\{ \frac{2\,\mathrm{Li}_n(\rho^3)}{3^{n-1}} + \frac{1}{12}\frac{\mathrm{Li}_n(\rho^2)}{2^{n-1}} - \frac{1}{6}\mathrm{Li}_n(\rho) + \frac{3}{4}\frac{\log^n \rho}{n!} - \frac{\zeta(2)}{12}\frac{\log^{n-2}\rho}{(n-2)!} + \frac{5}{54}\zeta(4)\frac{\log^{n-4}\rho}{(n-4)!} \right\} = 0.$$

This equation is valid for $1 \le n \le 4$, and is a good example of the transparency property, as shown in the flow chart of Figure 4.1.

3.7.2. *Inaccessible ladders.* The base $\chi = \tan(\pi/12) = 2 - \sqrt{3}$ satisfies (3.97) with $m = 4$. Corresponding to the cyclotomic equation

$$1 - \chi^4 = (1-\chi^2)(1-\chi)^4\chi^{-1}, \tag{3.106}$$

there is obtained the ladder,

(3.107)
$$L_n(4, \chi) = \frac{\mathrm{Li}_n(\chi^4)}{4^{n-1}} - \left\{ \frac{\mathrm{Li}_n(\chi^2)}{2^{n-1}} + 4\,\mathrm{Li}_n(\chi) + \frac{\log^n \chi}{n!} - \frac{5}{4}\zeta(2)\frac{\log^{n-2}\chi}{(n-2)!} \right\}.$$

For $1 \le n \le 3$ we then have

$$L_n(4, \chi) = \frac{-23}{8}\zeta(3)\frac{\log^{n-3}\chi}{(n-3)!}. \tag{3.108}$$

This formula, as it happens, can be found analytically from Kummer's equations (3.9) and (3.10) by taking $x = u$, $y = -u$, and $(1-u)/(1+u) = iu$. The arguments involve powers of $iu^2 = \chi$, and the above results follow. Not so, however, for the further cyclotomic equation

$$1 - \chi^6 = (1-\chi^3)(1-\chi^2)^3(1-\chi)^{-5}\chi. \tag{3.109}$$

A valid dilogarithmic ladder corresponding to this can be found analytically, but the method, involving a cut in the complex plane, does not apply to $n = 3$. However, the ladder can be extended and the constant evaluated numerically. The ensuing ladder is

$$L_n(6, \chi) = \frac{\mathrm{Li}_n(\chi^6)}{6^{n-1}} - \left\{ \frac{\mathrm{Li}_n(\chi^3)}{3^{n-1}} + \frac{3\,\mathrm{Li}_n(\chi^2)}{2^{n-1}} - 5\,\mathrm{Li}_n(\chi) - \frac{\log^n \chi}{n!} + \frac{4}{3}\zeta(2)\frac{\log^{n-2}\chi}{(n-2)!} \right\}. \tag{3.110}$$



For $1 \le n \le 3$ we then find

$$L_n(6, \chi) = \frac{-115}{36}\zeta(3)\frac{\log^{n-3}\chi}{(n-3)!}, \tag{3.111}$$

with the coefficient $\frac{115}{36}$ found by computation.

Neither (3.108) nor (3.111) individually extends to $n = 4$ because of the presence of the term in $\zeta(3)$, which is conjectured always to be absent from a $\mathrm{Li}_4$ ladder. However, by combining these two formulas in the ratio 10 to 9 one obtains

$$10L_n(4, \chi) + 9L_n(6, \chi) = 0, \qquad 1 \le n \le 3. \tag{3.112}$$

An attempt can now be made to extend this to $n = 4$ by seeking a simple rational multiple of $\zeta(4) = \pi^4/90$. The resulting equation is found to be

$$10L_n(4, \chi) + 9L_n(6, \chi) = \frac{143}{96}\zeta(4)\frac{\log^{n-4}\chi}{(n-4)!}, \tag{3.113}$$

valid for $1 \le n \le 4$. This formula apparently does not extend beyond $n = 4$.

**3.8. Examples of ladders for $n \le 5$.**

3.8.1. *Results for the base $\delta = 1/2$.* Most of the results reaching to $n = 5$ are either associated with supernumary ladders or extend still further into the trans-Kummer region, some as far as $n = 9$. Of course, all the generic bases from $u^p + u^q = 1$ attain a single result at $n = 5$ via (3.70). A rather unusual case is $u = \delta = 1/2$ because the generic indices degenerate into only $N = 1$ and 6. The former is exhibited in (3.72) and the latter gives

(3.114)
$$\frac{\mathrm{Li}_n(\delta^6)}{6^{n-1}} - \left\{ \frac{\mathrm{Li}_n(\delta^3)}{3^{n-1}} + \frac{2\,\mathrm{Li}_n(\delta^2)}{2^{n-1}} - \frac{4}{9}\mathrm{Li}_n(\delta) + \frac{5}{9}\frac{\log^n \delta}{n!} - \frac{\zeta(2)}{9}\frac{\log^{n-2}\delta}{(n-2)!} + \frac{5}{27}\zeta(4)\frac{\log^{n-4}\delta}{(n-4)!} + \frac{403}{1296}\zeta(5)\frac{\log^{n-5}\delta}{(n-5)!} \right\} = 0.$$

This result is valid for $1 \le n \le 5$.

3.8.2. *Flow chart for $\delta$.* The data contained in (3.72) and (3.114) can be represented in the flow chart of Figure 3.2. Because of the transparency of

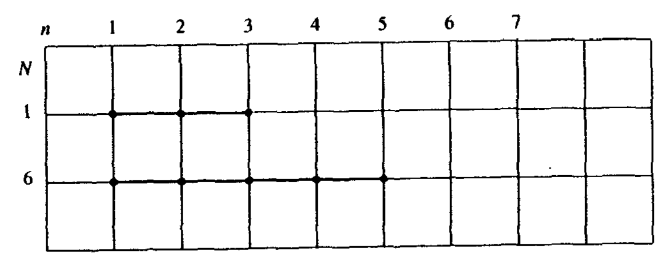



FIGURE 3.2 Flow chart for $\delta$-ladders

both ladders the chart consists merely of two separate lines, one from $n = 1$ to 3 and the other from $n = 1$ to 5. This may be compared with the $\rho$ flow chart of Figure 4.1, where the index-12 line similarly proceeds from $n = 1$ to 5, at which juncture it combines with other ladders to eliminate first the $\zeta(5)$ term and then the $\zeta(7)$ term, to pseudointegrate as far as $n = 9$. Nothing comparable occurs for the $\delta$-ladder because of the absence of any further ladders available to combine at $n = 5$.

**3.9. Polynomial relations for ladders.**

3.9.1. *The nonalgebraic character of the polylogarithm.* As far as can be surmised, the polylogarithm of any algebraic base is irrational, and almost certainly transcendental. Chudnovsky [7, 8] has given a proof of the irrationality of the dilogarithm for rational bases less than $1/2$. Apéry's recent proof [9] of the irrationality of $\zeta(3) = \mathrm{Li}_3(1)$ is the only such result published to date, for any base, for $n > 2$.

Although the polylogarithms themselves are transcendental, one can ask if there are any algebraic relations between them. Eastham [10] showed that there are no linear relations with algebraic coefficients, whilst Bombieri [11] showed that if $\xi$ is an element of a number field $K$, then no polynomial relations of polylogarithms of $\xi$, with coefficients in $K$, exist if $\xi$ lies between certain determinable upper and lower bounds in the interval (0, 1). However for $\xi = \pm 1$ the even-order polylogarithms can be expressed as a power of $\pi^2$ with rational coefficients, so that elimination of $\pi^2$ gives rise to such relations as [11]

$$7[\mathrm{Li}_2(-1)]^2 = 10\,\mathrm{Li}_4(-1)\,\mathrm{Li}_0(-1), \tag{3.115}$$

and many others; but these formulas are almost trivial.

3.9.2. *Examples of ladder relations.* Since, as far as is known, the $\zeta(n)$ for $n$ odd are algebraically unrelated to each other, and to powers of $\pi$, there results from this a sort of "decoupling" of odd-order ladders both from each other and from the even orders. It is only if two or more related results exist, enabling the $\zeta(n)$ terms to be eliminated between them, that this hiatus can be overcome and a bridge created between consecutive orders. An example of this process is given in the construction of (3.112). Following the method of Bombieri, a slightly more involved elimination of both $\pi^2$ and the logarithm of the base enables the formulation of a polynomial relation with rational coefficients for the polylogarithmic parts of ladders. The method requires ladders with $\zeta(2n+1)$ terms absent. An illustration comes from (3.105) for $L_n(6, \rho)$. If we denote the polylogarithmic parts of this ladder by $\overline{L}_n(\rho)$, eliminate $\pi^2$ between the ladders for $n = 2, 3, 4$, and write $y = \log^2\rho/8\overline{L}_2(\rho)$ then it is found that

$$\begin{cases} 4y^3 - 4y^2 + y - A = 0, & A = \overline{L}_3^2(\rho)/8\overline{L}_2^3(\rho), \\ 19y^2 - 14y - B = 0, & B = [\overline{L}_4(\rho)/2\overline{L}_2^2(\rho)] - 8/3. \end{cases} \tag{3.116}$$



The eliminant, which can be further simplified if desired, can be put in determinant form as

$$\begin{vmatrix} 0 & 0 & 19 & -14 & -B \\ 0 & 19 & -14 & -B & 0 \\ 19 & -14 & -B & 0 & 0 \\ 0 & 4 & -4 & 1 & -A \\ 4 & -4 & 1 & -A & 0 \end{vmatrix} = 0. \tag{3.117}$$

A closely similar calculation for the base $\delta = 1/2$ comes from (3.114), and gives

$$\begin{vmatrix} 0 & 0 & 225 & -51 & -B' \\ 0 & 225 & -51 & -B' & 0 \\ 225 & 51 & -B' & 0 & 0 \\ 0 & 200 & -60 & 9/2 & -A' \\ 200 & -60 & 9/2 & -A' & 0 \end{vmatrix} = 0. \tag{3.118}$$

where $A' = \overline{L}_3^2(\delta)/8\overline{L}_2^3(\delta)$, $B' = [\overline{L}_4(\delta)/2\overline{L}_2^2(\delta)] - 3$. It is surprising how close the expressions are for $A$, $B$ and $A'$, $B'$. Many other relations for other bases can be obtained in the same way.

## References


1. L. Lewin, *Polylogarithms and associated functions*, Elsevier-North Holland, New York and Oxford, 1981; see Chapter 6.
2. M. Abouzahra and L. Lewin, *The polylogarithm in algebraic number fields*, J. Number Theory **21** (1985), 214–244.
3. MACSYMA, A large symbolic manipulation program developed at the MIT Laboratory for Computer Science. MACSYMA is a trademark of Symbolics Inc.
4. H. F. Sandham, *A logarithmic transcendent*, J. London Math. Soc. **24** (1949), 83–91.
5. L. Lewin, *The order-independance of the polylogarithmic ladder structure—implications for a new category of functional equations*, Aequationes Math. **30** (1986), 1–20.
6. ——, *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), 345–373.
7. G. V. Chudnovsky, *Padé approximations to the generalized hypergeometric functions*, J. Math. Pure Appl. **58** (1979), 445–476.
8. ——, *Transcendental numbers*, Number Theory, Carbondale 1979 Proceedings, Lecture Notes in Math., Vol. 751, Springer-Verlag, New York and Berlin, 1980, pp. 45–69.
9. R. Apéry, *Irrationalité de $\zeta(2)$ et $\zeta(3)$*, Journées Arith. de Luminy, *Astérisque* **61** (1979), 11–13.
10. M. S. P. Eastham, *On polylogarithms*, Proc. Glasgow Math. Assoc. **6** (1961), 169–171.
11. E. Bombieri, *On G-functions*, Recent Progress in Analytic Number Theory, Vol. 2, Academic Press, New York, 1981, pp. 1–67.

CHAPTER 4

# Ladders in the Trans-Kummer Region

M. ABOUZAHRA AND L. LEWIN

**4.1. Ladder results to $n = 9$ for the base $\rho$.**

4.1.1. *Nonexistence of nontrivial functional equations for* $n > 5$. Apart from the relatively trivial duplication and inversion formulas, no functional equations were known for the polylogarithm of order greater than five until the generation of two single-variable equations at $n = 6$ by H. Gangl in 1990. In Chapter 9 a non-existence theorem is proven for equations of Kummer's type; and non-existence is shown for two of the three families of single-variable equations discussed in chapter 6; the third has not yet been taken beyond $n = 3$. It is conjectured that the currently-known 19 ladder results for $n > 5$ stem from functional equations of some sort, but, apart from Gangl's results at $n = 6$ and 7, there has been no success in finding them. Consequently, all the trans-Kummer results in this chapter were found by numerical computation after the structure of the presumptive ladders had been determined. Multiple precision calculations are needed both to avoid misleading artifacts and to avoid missing valid results.

4.1.2. *Analytic results to* $n = 5$. Since $\rho$ satisfies the two-term base equation (2.1) with $p = 2$, $q = 1$, the generic formulas of §3.5 apply. The index 6 results of (3.105) do not extend beyond $n = 4$ and are therefore irrelevant in the trans-Kummer region. Only the index 12 gives valid results to $n = 5$, and it is readily shown that

(4.1)
$$\begin{aligned} L_n(12, \rho) = \frac{\operatorname{Li}_n(\rho^{12})}{12^{n-1}} - \Bigg\{ &\frac{3}{2}\frac{\operatorname{Li}_n(\rho^6)}{6^{n-1}} + \frac{\operatorname{Li}_n(\rho^4)}{4^{n-1}} - \frac{11}{48}\frac{\operatorname{Li}_n(\rho^2)}{2^{n-1}} + \frac{13}{48}\frac{\log^n \rho}{n!} \\ &- \frac{\zeta(2)}{48}\frac{\log^{n-2}\rho}{(n-2)!} + \frac{19\zeta(4)}{1728}\frac{\log^{n-4}\rho}{(n-4)!} \Bigg\} \\ = \frac{67\zeta(5)}{6912}\frac{\log^{n-5}\rho}{(n-5)!}, &\qquad 1 \le n \le 5. \end{aligned}$$

A dilogarithmic ladder of index 20, long believed inaccessible from Kummer's equations, was found by Coxeter [1] from consideration of the properties of a certain infinite series. His method apparently does not apply for $n > 2$, and, initially, index-20 ladders to $n = 5$ were found numerically. Later an *ad hoc* single-variable equation was constructed from which these results could all be obtained as far as $n = 5$. And quite recently Dupont has outlined how Coxeter's formula can indeed be found from Kummer's functional equation at $n = 2$; the method is shown in §4.1.5.

The direct extension of Coxeter's ladder to $n = 3$ requires, as determined by numerical computation, a term $\zeta(3)/10$; but this can be eliminated on using (3.71) which contains a like term $\zeta(3)/5$. The resulting equation then extends directly to $n = 4$, with a determination of $178/20^3$ as the required coefficient of $\zeta(4)$. The ensuing ladder can be written

(4.2)
$$L_n(20,\rho) = \frac{\mathrm{Li}_n(\rho^{20})}{20^{n-1}} - \left\{\frac{\mathrm{Li}_n(\rho^{10})}{10^{n-1}} + 3\frac{\mathrm{Li}_n(\rho^4)}{4^{n-1}} - \frac{1}{2}\frac{\mathrm{Li}_n(\rho^2)}{2^{n-1}} + \frac{1}{2}\frac{\log^n\rho}{n!} - \frac{\zeta(2)}{25}\frac{\log^{n-2}\rho}{(n-2)!} + \frac{178}{20^3}\zeta(4)\frac{\log^{n-4}\rho}{(n-4)!}\right\},$$

for which

$$(4.3)\qquad \begin{aligned} L_n(20,\rho) &= 0, \qquad 1 \le n \le 4;\\ L_5(20,\rho) &= (201/10^4)\zeta(5).\end{aligned}$$

These results have now been found analytically from the equations of Chapter 6.

4.1.3. *Extension to the ninth order.* The fact that the factor 201 of $\zeta(5)$ is three times the factor 67 appearing in (4.1) enables a direct elimination of the $\zeta(5)$ term. Defining a second-degree ladder by

$$(4.4)\qquad L_n^{(2)}(20,\rho) = 3\cdot 20^5 L_n(20,\rho) - 80\cdot 12^5 L_n(12,\rho),$$

leads to the numerical determinations

$$(4.5)\qquad \begin{aligned} L_n^{(2)}(20,\rho) &= 0, && 1 \le n \le 5,\\ &= \frac{11606}{5}\zeta(6), && n = 6,\\ &= \frac{11606}{5}\zeta(6)\log\rho + 2166\zeta(7), && n = 7.\end{aligned}$$

In anticipation, it may be noted that the factor 2166 of $\zeta(7)$ is equal to $6\times 19^2$. A further ladder, of index 24, comes from formulas on multiple integration due to Phillips [2]. Although seemingly inaccessible from Kummer's equations these ladders can now be obtained analytically from the formulas of Chapter 6. Define



(4.6)
$$L_n(24,\rho) = \frac{\mathrm{Li}_n(\rho^{24})}{24^{n-1}} - \left\{\frac{4}{3}\frac{\mathrm{Li}_n(\rho^{12})}{12^{n-1}} + \frac{2\,\mathrm{Li}_n(\rho^8)}{8^{n-1}} - \frac{7}{3}\frac{\mathrm{Li}_n(\rho^4)}{4^{n-1}} + \frac{205}{576}\frac{\mathrm{Li}_n(\rho^2)}{2^{n-1}} - \frac{179}{24^2}\frac{\log^n\rho}{n!} + \frac{5\zeta(2)}{192}\frac{\log^{n-2}\rho}{(n-2)!} - \frac{629}{3\cdot 24^3}\zeta(4)\frac{\log^{n-4}\rho}{(n-4)!}\right\}.$$

Then

$$(4.7)\qquad \begin{aligned} L_n(24,\rho) &= 0, && 1 \le n \le 4,\\ &= -\frac{4623}{24^4}\zeta(5), && n = 5.\end{aligned}$$

The numerically determined 4623 equals $67\times 69$, suggesting elimination of $\zeta(5)$ via (4.1), where the numeric 67 also occurs.

Defining a second-degree ladder of index 24 by

$$(4.8)\qquad L_n^{(2)}(24,\rho) = \tfrac{1}{2}\cdot 24^5 L_n(24,\rho) + 23\cdot 12^5 L_n(12,\rho),$$

leads to the numerical determinations

$$(4.9)\qquad \begin{aligned} L_n^{(2)}(24,\rho) &= 0, && 1 \le n \le 5,\\ &= \frac{-6431}{6}\zeta(6), && n = 6,\\ &= \frac{-6431}{6}\zeta(6)\log\rho - \frac{47291}{48}\zeta(7), && n = 7.\end{aligned}$$

It may be remarked that $47291 = 131\times 19^2$, easing elimination of $\zeta(7)$ with (4.5).

The final step, after the elimination of $\zeta(7)$, requires defining

$$(4.10)\qquad \begin{aligned} L_n^{(3)}(24,\rho) = 131&\left[L_n^{(2)}(20,\rho) - \frac{11606}{5}\frac{\log^{n-6}\rho}{(n-6)!}\zeta(6)\right]\\ &+ 288\left[L_n^{(2)}(24,\rho) + \frac{6431}{6}\frac{\log^{n-6}\rho}{(n-6)!}\zeta(6)\right].\end{aligned}$$

Then it is found by numerical computation that

(4.11)
$$\begin{aligned} L_n^{(3)}(24,\rho) &= 0, && 1 \le n \le 7,\\ &= \frac{3,537,283}{1,800}\zeta(8), && n = 8,\\ &= \frac{3,537,283\zeta(8)\log\rho}{1,800} + \frac{66,452,911\zeta(9)}{36,000}, && n = 9.\end{aligned}$$

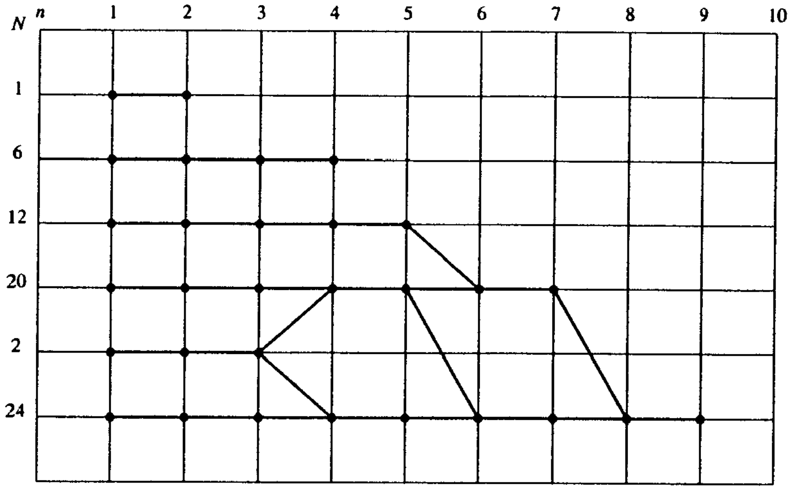



FIGURE 4.1. Flow chart for $\rho$-ladders

It may be noted that (4.11) and the preceding equations involve polylogarithms of $\rho^2 = \frac{1}{2}(3-\sqrt{5})$ only, so that the results can be expressed as ladders of $\rho_2 = \rho^2$, and halved indices.

4.1.4. *Flow chart for $\rho$-ladders.* A flow chart for the $\rho$-ladders is given in Figure 4.1. It may be seen that all the ladders pseudo-integrate transparently, except for the obligatory combinations needed to cancel out the terms in $\zeta(n)$ when $n$ is odd.

4.1.5. *Accessibility of Coxeter's formula at $n = 2$.* The proof makes use of Schaeffer's equation (1.19) in the form

$$\begin{aligned}\mathrm{Li}_2(x) - \mathrm{Li}_2(y) + \mathrm{Li}_2(y/x) + \mathrm{Li}_2\left(\frac{1-x}{1-y}\right) - \mathrm{Li}_2\left(\frac{y(1-x)}{x(1-y)}\right)&\\ = \zeta(2) - \log x \log\left(\frac{1-x}{1-y}\right),&\end{aligned} \tag{4.12}$$

and Landen's relation (1.8),

$$\mathrm{Li}_2(x) + \mathrm{Li}_2[-x/(1-x)] = -\tfrac{1}{2}\log^2(1-x). \tag{4.13}$$

Putting $g = e^{i2\pi/5}$, $\rho = g + 1/g = 2\cos(2\pi/5)$, then $\rho^2 + \rho = 1$. Take $x = -\rho^2/g$, $y = \rho^4$ in (4.12). Then it can be shown that

$$X = \frac{(1-x)}{(1-y)} = \frac{-g(1+g)}{(1-g)} \quad \text{and} \quad -\frac{X}{(1-X)} = \frac{-1}{g^2\rho^2}.$$

Similarly, if $Y = y(1-x)/x(1-y) = -\rho/(1-g)$ then $-Y/(1-Y) = -\rho^2/g^2$. Using the inversion relation on the term coming from $-X/(1-X)$, (4.12) takes the form

$$\begin{aligned}\mathrm{Li}_2(-g\rho^2) + \mathrm{Li}_2(-g^2\rho^2) + \mathrm{Li}_2(-g^3\rho^2) + \mathrm{Li}_2(-g^4\rho^2)&\\ = \mathrm{Li}_2(\rho^4) + L,&\end{aligned} \tag{4.14}$$

where $L$ is a logarithmic term given by

$$\begin{aligned}L = \frac{1}{2}\log^2\left(\frac{1+g^2}{1-g}\right) - 2\log^2(g\rho) - \frac{1}{2}\log^2\left(\frac{1+g}{g(1-g)}\right)&\\ - \log(-\rho^2/g)\log\left(\frac{-g(1+g)}{1-g}\right).&\end{aligned} \tag{4.15}$$

Now the multiplication formula (2.63) with $M = 5$ gives

$$\begin{aligned}(\tfrac{1}{5})\,\mathrm{Li}_2(-\rho^{10}) = \mathrm{Li}_2(-\rho^2) + \mathrm{Li}_2(-\rho^2 g) + \mathrm{Li}_2(-\rho^2 g^2)&\\ + \mathrm{Li}_2(-\rho^2 g^3) + \mathrm{Li}_2(-\rho^2 g^4).&\end{aligned} \tag{4.16}$$

Combining these, simplifying (4.15), and using the duplication formula to remove terms with negative arguments, one obtains Coxeter's formula (1.24). This proof is essentially due to Dupont.

**4.2. Ladder results to $n = 9$ for the base $\omega$.**

4.2.1. *The base $\omega$.* The quantity $\omega$ is the solution in $(0, 1)$ of the two-term base equation

$$u^3 + u^2 - 1 = 0. \tag{4.17}$$

If this is multiplied by $(u^2 - u + 1)$ the equation becomes

$$u^5 + u - 1 = 0. \tag{4.18}$$

Thus $\omega$ satisfies *two* equations of the form $u^p + u^q - 1 = 0$. It is the only base to do so with $(p, q) = 1$. From the studies of Ljunggren [3] and Tverberg [4], the polynomial $u^p + u^q - 1 = 0$ with any cyclotomic factors removed is irreducible. If $(p, q) = 1$ the only possible roots which are roots of unity are $e^{\pm i\pi/3}$, so the sole admissible cyclotomic factor is $u^2 - u + 1$. Thus the only way an algebraic number $\omega$ which is not a root of unity can be a root of two different equations $u^a + u^b = 1$ and $u^c + u^d = 1$ with $(a, b) = (c, d) = 1$ is if $(u^a + u^b - 1) = (u^2 - u + 1)(u^c + u^d - 1)$, and this reduces to (4.17) and (4.18). This demonstration is due to a referee of the 1987 paper [5] where the trans-Kummer results for $\omega$ were first reported. There can be no doubt that the extraordinarily rich sequence of results for $\omega$ comes from this unique property.

The solution to the cubic (4.17) is

$$\omega = \left[\frac{(1+\sqrt{23/27})}{2}\right]^{2/3} + \left[\frac{(1-\sqrt{23/27})}{2}\right]^{2/3} - \frac{1}{3}. \tag{4.19}$$

4.2.2. *Dilogarithmic ladders.* Various manipulations of the base equations (4.17) and (4.18) lead to a total of 12 different cyclotomic equations, with different indices $N$, comprised of

$$N = \{1, 2, 3, 5, 8, 12, 14, 18, 20, 28, 30, 42\}. \tag{4.20}$$

All of these, except for $\{14, 20, 28, 42\}$, constitute the generic set, but 14, 20, and 28 are all accessible from Kummer's equation. Only $N = 42$ is apparently excluded. It was first discovered by constructing a cyclotomic equation and numerically verifying the ensuing ladder. Subsequently, index-42 ladders up to the fifth order have been produced analytically.

Since the initial construction of the index-42 cyclotomic equation was not all that obvious, a brief outline of a proof follows:

$$1 - \omega^7 = 1 - \omega^2(1-\omega) = 2\omega^3,$$

so

$$1 - \omega^7 + \omega^{14} = (1-\omega^7)^2 + \omega^7 = \omega^3(4\omega^3 + \omega^4).$$

Moreover,

$$(1+\omega^3)^2 = (2-\omega^2)^2 = 4 - 4\omega^2 + \omega^4 = 4\omega^3 + \omega^4.$$

Combining these gives

$$1 - \omega^{42} = (1-\omega^{21})(1-\omega^{14})(1-\omega^7)^{-1}(1-\omega^6)^2\omega^{-1}, \tag{4.21}$$

from which the corresponding component-ladder can be written down.

Unlike the base $\rho$, the ladders for $\omega$ are not transparent. The index-1 ladder, for instance, does not, of itself, equal a rational multiple of $\zeta(2)$ but needs to be combined with the other component-ladders to give a valid result. There are thus eleven resulting valid $\omega$-ladders and, of the following equations, all except the first (index-1), are zero for $n = 1$ and 2. Using the convenient shorthand $L_n(N)$ for $L_n(N, \omega)$, they are

$$L_n(1) = \mathrm{Li}_n(\omega) + \frac{5}{n!}\log^n\omega, \tag{4.22}$$

$$L_n(2) = \frac{\mathrm{Li}_n(\omega^2)}{2^{n-1}} - \left\{-\mathrm{Li}_n(\omega) + \frac{\zeta(2)\log^{n-2}\omega}{(n-2)!} - \frac{8\log^n\omega}{n!}\right\}, \tag{4.23}$$

$$L_n(3) = \frac{\mathrm{Li}_n(\omega^3)}{3^{n-1}} - \left\{\frac{2}{3}\mathrm{Li}_n(\omega) - \frac{\zeta(2)}{3}\frac{\log^{n-2}\omega}{(n-2)!} + \frac{4}{3}\frac{\log^n\omega}{n!}\right\}, \tag{4.24}$$

$$L_n(5) = \frac{\mathrm{Li}_n(\omega^5)}{5^{n-1}} - \left\{-\frac{\mathrm{Li}_n(\omega)}{5} + \frac{\zeta(2)}{5}\frac{\log^{n-2}\omega}{(n-2)!} - \frac{2\log^n\omega}{n!}\right\}, \tag{4.25}$$

$$L_n(8) = \frac{\mathrm{Li}_n(\omega^8)}{8^{n-1}} - \left\{\frac{\mathrm{Li}_n(\omega^4)}{4^{n-1}} + \frac{\mathrm{Li}_n(\omega)}{4} - \frac{\zeta(2)}{4}\frac{\log^{n-2}\omega}{(n-2)!} + \frac{9}{4}\frac{\log^n\omega}{n!}\right\}, \tag{4.26}$$



$$\begin{aligned} L_n(12) = \frac{\mathrm{Li}_n(\omega^{12})}{12^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\omega^6)}{6^{n-1}} + \frac{\mathrm{Li}_n(\omega^4)}{4^{n-1}} - \frac{1}{4}\frac{\mathrm{Li}_n(\omega^3)}{3^{n-1}} \\ &- \frac{\zeta(2)}{12}\frac{\log^{n-2}\omega}{(n-2)!} + \frac{3}{2}\frac{\log^n\omega}{n!}\Bigg\}, \end{aligned} \tag{4.27}$$

$$\begin{aligned} L_n(14) = \frac{\mathrm{Li}_n(\omega^{14})}{14^{n-1}} - \Bigg\{&\frac{2\,\mathrm{Li}_n(\omega^7)}{7^{n-1}} - \frac{3}{14}\frac{\mathrm{Li}_n(\omega^2)}{2^{n-1}} + \frac{3}{14}\mathrm{Li}_n(\omega) \\ &- \frac{\zeta(2)}{7}\frac{\log^{n-2}\omega}{(n-2)!} + \frac{10}{7}\frac{\log^n\omega}{n!}\Bigg\}, \end{aligned} \tag{4.28}$$

$$L_n(18) = \frac{\mathrm{Li}_n(\omega^{18})}{18^{n-1}} - \left\{\frac{\mathrm{Li}_n(\omega^9)}{9^{n-1}} + \frac{\mathrm{Li}_n(\omega^6)}{6^{n-1}} - \frac{1}{3}\frac{\mathrm{Li}_n(\omega^3)}{3^{n-1}} + \frac{1}{3}\frac{\log^n\omega}{n!}\right\}, \tag{4.29}$$

$$\begin{aligned} L_n(20) = \frac{\mathrm{Li}_n(\omega^{20})}{20^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\omega^{10})}{10^{n-1}} + \frac{2\,\mathrm{Li}_n(\omega^4)}{4^{n-1}} - \frac{3}{50}\frac{\mathrm{Li}_n(\omega^2)}{2^{n-1}} \\ &+ \frac{6}{25}\mathrm{Li}_n(\omega) - \frac{17}{50}\zeta(2)\frac{\log^{n-2}\omega}{(n-2)!} + \frac{201}{50}\frac{\log^n\omega}{n!}\Bigg\}, \end{aligned} \tag{4.30}$$

$$\begin{aligned} L_n(28) = \frac{\mathrm{Li}_n(\omega^{28})}{28^{n-1}} - \Bigg\{&\frac{3}{2}\frac{\mathrm{Li}_n(\omega^{14})}{14^{n-1}} + \frac{\mathrm{Li}_n(\omega^4)}{4^{n-1}} \\ &- \frac{1}{14}\frac{\mathrm{Li}_n(\omega^2)}{2^{n-1}} - \frac{\zeta(2)}{14}\frac{\log^{n-2}\omega}{(n-2)!} + \frac{9}{7}\frac{\log^n\omega}{n!}\Bigg\}, \end{aligned} \tag{4.31}$$

$$\begin{aligned} L_n(30) = \frac{\mathrm{Li}_n(\omega^{30})}{30^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\omega^{15})}{15^{n-1}} + \frac{\mathrm{Li}_n(\omega^{10})}{10^{n-1}} + \frac{\mathrm{Li}_n(\omega^6)}{6^{n-1}} \\ &- \frac{1}{3}\frac{\mathrm{Li}_n(\omega^5)}{5^{n-1}} - \frac{1}{5}\frac{\mathrm{Li}_n(\omega^3)}{3^{n-1}} + \frac{4}{15}\frac{\log^n\omega}{n!}\Bigg\}, \end{aligned} \tag{4.32}$$

$$\begin{aligned} L_n(42) = \frac{\mathrm{Li}_n(\omega^{42})}{42^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\omega^{21})}{21^{n-1}} + \frac{\mathrm{Li}_n(\omega^7)}{7^{n-1}} + \frac{2\,\mathrm{Li}_n(\omega^6)}{6^{n-1}} \\ &- \frac{5}{14}\frac{\mathrm{Li}_n(\omega^3)}{3^{n-1}} - \frac{\zeta(2)}{21}\frac{\log^{n-2}\omega}{(n-2)!} + \frac{9}{7}\frac{\log^n\omega}{n!}\Bigg\}. \end{aligned} \tag{4.33}$$

All except the last are known to be accessible from Kummer's dilogarithmic functional equation.

4.2.3. *Third-order ladders.* From Kummer's equation at the third order, simple ladder combinations equaling a rational multiple of $\zeta(3)$ can be found. Only the index-42 result was initially dependent on numerical computation. Convenient combinations, prior to the cancellation of the $\zeta(3)$ term, are

(4.34) $$3^2L_3(3)+5L_3(2)=\zeta(3),$$

(4.35) $$3\cdot 5^2L_3(5)-13L_3(2)=\zeta(3),$$

(4.36) $$3\cdot 8^2L_3(8)+13\cdot 2^2L_3(2)=8\zeta(3),$$

(4.37) $$12^2L_3(12)+2\cdot 2^2L_3(2)=4\zeta(3),$$

(4.38) $$3\cdot 14^2L_3(14)+26\cdot 2^2L_3(2)=19\zeta(3),$$

(4.39) $$18^2L_3(18)+2^2L_3(2)=8\zeta(3),$$

(4.40) $$3\cdot 20^2L_3(20)+91\cdot 2^2L_3(2)=56\zeta(3),$$

(4.41) $$3\cdot 28^2L_3(28)+26\cdot 2^2L_3(2)=49\zeta(3),$$

(4.42) $$30^2L_3(30)+2\cdot 2^2L_3(2)=16\zeta(3),$$

(4.43) $$42^2L_3(42)+12\cdot 2^2L_3(2)=39\zeta(3).$$

At this point the association of factors $N^2$ with $L_3(N)$ is clearly seen, pointing to the desirability of working with the modified ladders $M_n(N)$ from here on, where $M_n(N)=N^{n-1}L_n(N)$.

4.2.4. *Fourth-order ladders.* Since these do not involve $\zeta(3)$ the first step is to eliminate it by the use of, for example, (4.34). The resulting equations can be pseudo-integrated and then combinations sought to give a rational multiple of $\zeta(4)$. In this way the first two ladders (4.34) and (4.35) become incorporated into the others. The results so generated can, except for the last, be checked against Kummer's equation at the fourth order, which, of course, is already absent any $\zeta(3)$ term.

In terms of modified ladders $M_n(N)$, define second-degree (modified) ladders $M_n^{(2)}(N)$. Then for $n=1,2,3$, and 4 each of the following expressions are found to be zero.

(4.44) $$M_n^{(2)}(8)=M_n(8)-\frac{8^{n-4}}{2}\left\{\frac{4M_n(5)}{5^{n-4}}+\frac{12M_n(3)}{3^{n-4}}-\frac{23M_n(2)}{2^{n-4}}+38\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\},$$

(4.45) $$M_n^{(2)}(12)=M_n(12)-\frac{12^{n-4}}{2}\left\{\frac{36M_n(5)}{5^{n-4}}+\frac{12M_n(3)}{3^{n-4}}-\frac{99M_n(2)}{2^{n-4}}+130\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\},$$



(4.46) $$M_n^{(2)}(14)=M_n(14)-\frac{14^{n-4}}{2}\left\{\frac{-34M_n(5)}{5^{n-4}}+\frac{78M_n(3)}{3^{n-4}}+\frac{117M_n(2)}{2^{n-4}}+102\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\},$$

(4.47) $$M_n^{(2)}(18)=M_n(18)-18^{n-4}\left\{\frac{36M_n(5)}{5^{n-4}}+\frac{28M_n(3)}{3^{n-4}}-\frac{54M_n(2)}{2^{n-4}}+72\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\},$$

(4.48) $$M_n^{(2)}(20)=M_n(20)-20^{n-4}\left\{\frac{62M_n(5)}{5^{n-4}}+\frac{90M_n(3)}{3^{n-4}}-\frac{605}{2}\frac{M_n(2)}{2^{n-4}}+547\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\}$$

(4.49) $$M_n^{(2)}(28)=M_n(28)-28^{n-4}\left\{\frac{134M_n(5)}{5^{n-4}}+\frac{78M_n(3)}{3^{n-4}}-\frac{338M_n(2)}{2^{n-4}}+765\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\},$$

(4.50) $$M_n^{(2)}(30)=M_n(30)-30^{n-4}\left\{108\frac{M_n(5)}{5^{n-4}}+100\frac{M_n(3)}{3^{n-4}}-135\frac{M_n(2)}{2^{n-4}}+254\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\},$$

(4.51) $$M_n^{(2)}(42)=M_n(42)-9\cdot 42^{n-4}\left\{57\frac{M_n(5)}{5^{n-4}}+\frac{29M_n(3)}{3^{n-4}}-\frac{128M_n(2)}{2^{n-4}}+\frac{1717}{9}\zeta(4)\frac{\log^{n-4}\omega}{(n-4)!}\right\}.$$

The index-42 result was initially dependent on numerical computation but can now be found from the formulas of Chapter 6.

4.2.5. *Fifth-order ladders.* The combination rule that seems to be emerging from these and other results is that ladders combine in pairs as the order increases from even to odd, and in threes, involving *inter-alia* the cancellation of the $\zeta$-function of odd order, in going from odd to even order. This would suggest the existence of seven ladders of order 5 from the eight results of order 4. However, Kummer's equations yield only three results. The "missing" four were first discovered by numerical computation, and only now, with the formulas of Chapter 6, have they been confirmed analytically. The needed third-degree (modified) ladders are found to be

(4.52) $$M_n^{(3)}(12) = M_n^{(2)}(12) - 4(12/8)^{n-5}M_n^{(2)}(8),$$

(4.53) $$M_n^{(3)}(14) = M_n^{(2)}(14) - \frac{11}{2}(14/8)^{n-5}M_n^{(2)}(8),$$

(4.54) $$M_n^{(3)}(18) = M_n^{(2)}(18) - 5(18/8)^{n-5}M_n^{(2)}(8),$$

(4.55) $$M_n^{(3)}(20) = M_n^{(2)}(20) - 66(20/8)^{n-5}M_n^{(2)}(8),$$

(4.56) $$M_n^{(3)}(28) = M_n^{(2)}(28) - 125(28/8)^{n-5}M_n^{(2)}(8),$$

(4.57) $$M_n^{(3)}(30) = M_n^{(2)}(30) - 35(30/8)^{n-5}M_n^{(2)}(8),$$

(4.58) $$M_n^{(3)}(42) = M_n^{(2)}(42) - \frac{633}{2}(42/8)^{n-5}M_n^{(2)}(8).$$

They are equal to $\zeta(5)$ multiplied, respectively, by 96, $-99$, 240, 528, 1353, 960, and 10,461 when $n = 5$.

4.2.6. *Ladders to the ninth order.*

4.2.6.1. From the seven formulas of order 5, six remain after eliminating $\zeta(5)$; this is conveniently done by using (4.52), except that (4.53) is used, for preference, with (4.56) and (4.58). Define

(4.59) $$\overline{M}_n(14) = 64M_n^{(3)}(14) + 77(14/12)^{n-6}M_n^{(3)}(12),$$

(4.60) $$\overline{M}_n(18) = 4M_n^{(3)}(18) - 15(18/12)^{n-6}M_n^{(3)}(12),$$

(4.61) $$\overline{M}_n(20) = 6M_n^{(3)}(20) - 55(20/12)^{n-6}M_n^{(3)}(12),$$

(4.62) $$\overline{M}_n(28) = 3M_n^{(3)}(28) + 82(28/14)^{n-6}M_n^{(3)}(14),$$

(4.63) $$\overline{M}_n(30) = M_n^{(3)}(30) - 25(30/12)^{n-6}M_n^{(3)}(12),$$

(4.64) $$\overline{M}_n(42) = M_n^{(3)}(42) + 317(42/14)^{n-6}M_n^{(3)}(14).$$

At the time of this writing, no analytical results exist at greater than the seventh order. A computer search was made for integers $p$ and $q$ such that $[p\overline{M}_6(N_1) + q\overline{M}_6(N_2)]/\zeta(6)$ should be integral, $N_1$ and $N_2$ being any pair of 14, 18, 20, 28, 30, 42. This leads to defining

(4.65)
$$M_n^{(4)}(18) = 17\overline{M}_n(18) - 18^{n-6}\left\{6\frac{\overline{M}_n(14)}{14^{n-6}} + 68,922\zeta(6)\frac{\log^{n-6}\omega}{(n-6)!}\right\},$$

(4.66)
$$M_n^{(4)}(20) = 17\overline{M}_n(20) + 20^{n-6}\left\{15\frac{\overline{M}_n(14)}{14^{n-6}} + 270,004\zeta(6)\frac{\log^{n-6}\omega}{(n-6)!}\right\},$$

(4.67)
$$M_n^{(4)}(28) = 34\overline{M}_n(28) - 28^{n-6}\left\{\frac{155}{2}\frac{\overline{M}_n(14)}{14^{n-6}} + 869,273\zeta(6)\frac{\log^{n-6}\omega}{(n-6)!}\right\},$$



(4.68)
$$M_n^{(4)}(30) = 68\overline{M}_n(30) + 30^{n-6}\left\{75\frac{\overline{M}_n(14)}{14^{n-6}} + 2,316,606\zeta(6)\frac{\log^{n-6}\omega}{(n-6)!}\right\},$$

(4.69)
$$M_n^{(4)}(42) = 68\overline{M}_n(42) - 42^{n-6}\left\{679\frac{\overline{M}_n(14)}{14^{n-6}} + 20,422\zeta(6)\frac{\log^{n-6}\omega}{(n-6)!}\right\}.$$

Then each of the above five expressions is zero for $n = 1$ through 6. At $n = 6$ they constitute five of the nineteen trans-Kummer results alluded to earlier.

It may be noted that the factor 17 occurs in all the initial terms on the right in these equations. The significance of this is not known, but this factor reappears again in the later equations developed to the ninth order.

4.2.6.2. At the seventh order the findings involve the fifth-degree (modified) ladders, obtained by numerical computation,

(4.70) $$M_n^{(5)}(20) = 15M_n^{(4)}(20) + 68(20/18)^{n-7}M_n^{(4)}(18),$$

(4.71) $$M_n^{(5)}(28) = 15M_n^{(4)}(28) - 371(28/18)^{n-7}M_n^{(4)}(18),$$

(4.72) $$M_n^{(5)}(30) = 5M_n^{(4)}(30) + 341(30/18)^{n-7}M_n^{(4)}(18),$$

(4.73) $$M_n^{(5)}(42) = M_n^{(4)}(42) + 37(42/18)^{n-7}M_n^{(4)}(18).$$

Then, for $n = 7$, we have

(4.74) $$M_7^{(5)}(20) = 2,998,800\zeta(7),$$

(4.75) $$M_7^{(5)}(28) = -60,286,590\zeta(7),$$

(4.76) $$M_7^{(5)}(30) = 50,649,120\zeta(7),$$

(4.77) $$M_7^{(5)}(42) = 30,208,320\zeta(7).$$

Double precision was necessary to achieve these results reliably. The factor of $\zeta(7)$ in (4.77), for example, was computed as

$$30,208,320.00000000000000011169\ldots,$$

an accuracy adequate to ensure against any erroneous interpretation.

4.2.6.3. To obtain results at the eighth order the first step is to eliminate $\zeta(7)$ from (4.74) through (4.77). The needed coefficients would be in the trillions were it not for very substantial common factors to the above multipliers of $\zeta(7)$. A prime factor devolution gives

(4.78) $$2,998,800 = 2^4\cdot 3^2\cdot 5^2\cdot 7^2\cdot 17,$$

(4.79) $$60,286,590 = 2\cdot 3^2\cdot 5\cdot 7\cdot 13\cdot 17\cdot 433,$$

(4.80) $$50,649,120 = 2^5\cdot 3^2\cdot 5\cdot 17\cdot 2069,$$

(4.81) $$30,208,320 = 2^6\cdot 3^2\cdot 5\cdot 17\cdot 617.$$

As mentioned previously, each of these has a factor 17. The elimination of $\zeta(7)$ between (4.74) to (4.77) is straightforward, and leads to defining

$$M_n^{(6)}(28) = 200M_n^{(5)}(28) + 5629\left(\frac{28}{20}\right)^{n-8} M_n^{(5)}(20), \tag{4.82}$$

$$M_n^{(6)}(30) = 245M_n^{(5)}(30) - 6207\left(\frac{30}{20}\right)^{n-8} M_n^{(5)}(20), \tag{4.83}$$

$$M_n^{(6)}(42) = 175M_n^{(5)}(42) - 3702\left(\frac{42}{20}\right)^{n-8} M_n^{(5)}(20). \tag{4.84}$$

These are all zero for $n = 1$ through 7.

In expectation that the previous pattern of combination in pairs would continue, integer combinations of (4.82) to (4.84) in pairs were constructed and rational multiples of $\zeta(8)$ were sought. Somewhat surprisingly, none were found (initially) so an integer combination of all three was examined. This came up with a valid solution, and also a second one! When $M_8^{(6)}(42)$ was eliminated between them it became apparent why the initial search had failed: the large integers involved were well outside the search range. In contrast, the integers involved in the triplets were much more modest, suggesting that, perhaps, the triplet combinations stemmed from a more fundamental source. The results are therefore retained in the triplet form, and are as follows.

$$\begin{aligned} &61\frac{M_n^{(6)}(28)}{28^{n-8}} + 23\frac{M_n^{(6)}(30)}{30^{n-8}} + 89\frac{M_n^{(6)}(42)}{42^{n-8}} \\ &\qquad + 1,234,036,640,200\frac{\zeta(8)\log^{n-8}\omega}{(n-8)!} = A_n, \end{aligned} \tag{4.85}$$

$$\begin{aligned} &11\frac{M_n^{(6)}(28)}{28^{n-8}} + 179\frac{M_n^{(6)}(30)}{30^{n-8}} - 292\frac{M_n^{(6)}(42)}{42^{n-8}} \\ &\qquad - 2,618,318,229,800\frac{\zeta(8)\log^{n-8}\omega}{(n-8)!} = B_n, \end{aligned} \tag{4.86}$$

with $A_n = B_n = 0$ for $n = 1$ through 8. The multipliers of $\zeta(8)$ factorize into $2^3{\cdot}5^2{\cdot}7{\cdot}13{\cdot}17{\cdot}1303{\cdot}3061$ and $2^3{\cdot}5^2{\cdot}7{\cdot}13{\cdot}17{\cdot}8,462,567$ respectively. They therefore contain a substantial common factor; but the significance of this, if any, is not currently known. There is not, for example, any call to cancel out the $\zeta(8)$ term as there was for the $\zeta(7)$, so no such use can be made of this information.

4.2.6.4. Due to the relatively inefficient search algorithm originally used, no result at $n = 9$ could be initially found, though the existence of one was strongly suspected. Recently, using a more powerful algorithm of Szekeres [**6**], researchers at the University of New South Wales [**7**] were able to come up with the ninth-order ladder result

$$7055A_9 + 3307B_9 + 50,284,867,249,200\zeta(9) = 0. \tag{4.87}$$



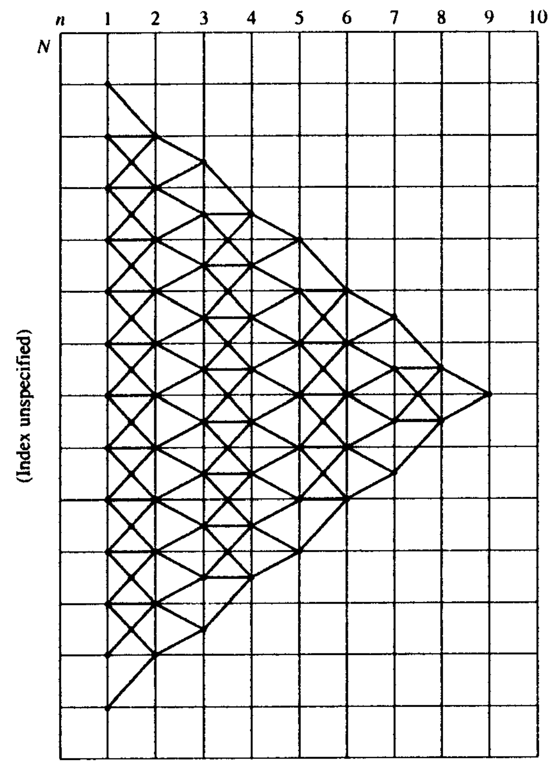



FIGURE 4.2. Flow chart for $\omega$-ladders

It may be noticed that the peculiar and unexplained appearance of the factor 17 in the earlier formulas, and in the prior coefficients of $\zeta(7)$ and $\zeta(8)$, continues here in both the coefficients of $\zeta(9)$ and also of $A_9$, though not in 3307, the (prime) coefficient of $B_9$.

D. Zagier (1989), using the complex roots conjugate to $\omega$, has obtained ladder-type expressions involving the Dedekind zeta-function. By eliminating it, in the same way as was the Riemann zeta-function earlier, he was able to arrive analytically at the above coefficients of $A_9$ and $B_9$ in (4.87).

4.2.7. *Flow chart for $\omega$-ladders.* The nature of the results of the preceding section is portrayed in the flow chart of Figure 4.2. Equations combine in pairs in going from even to odd orders. From odd to even orders the $\zeta$-functions of odd argument must first be eliminated, and then the eliminant equations are combined in pairs: in other words the ladders combine in threes. The trans-Kummer results give five equations at the sixth order, four at the seventh, two at the eighth, and a single one at the ninth; a total of twelve for $n \geq 6$. (Note that Figure 1 of [5] is in error because the ladder of index 1 was absorbed in the other ladders, but should have been shown separately in the chart.)

**4.3. Ladder results to $n = 6$ for the base $\theta$.**

4.3.1. *The base $\theta$.* The quantity $\theta$ is the solution in (0, 1) of the cubic

$$u^3 + u = 1, \tag{4.88}$$

and is given by

$$\theta = \left[\frac{(\sqrt{31/27}+1)}{2}\right]^{1/3} - \left[\frac{(\sqrt{31/27}-1)}{2}\right]^{1/3}. \tag{4.89}$$

It satisfies the two equations

$$1 - \theta = \theta^3; \qquad (1+\theta^2)^{-1} = \theta. \tag{4.90}$$

The pair of equations

$$1 - u = u^m; \qquad (1+u^2)^{-1} = u^n, \tag{4.91}$$

for $m$ and $n$ integral, $(m, n) = 1$, has the sole solution (4.90), since the equation $(1+u^2)^{-1} = u^n$ can be written $u^n + u^{n+2} = 1$; and from the discussion on the uniqueness of $\omega$ in §4.2.1 it is clear that the two equations can take only one form, which must, therefore, be (4.90). Thus, $\theta$ is the unique solution to (4.91), and requires $m = 3$, $n = 1$. This is relevant to the use of the single-variable functional equations of §6.4 of Chapter 6 to generate ladders. Designed specifically for the base $\theta$, it therefore does not apply to any other bases.

4.3.2. *Analytic results at $n = 2$.* There are six surviving generic indices

$$N = \{1, 3, 4, 6, 10, 18\}, \tag{4.92}$$

the indices 4 and 6 being repeated. There are also two supernumary indices $N = 14$ and 24, the former giving a valid ladder from Kummer's equation and the latter an analytically derived ladder from the functional equations of Chapter 6. The index-14 ladder comes from the easily-verified relation $\theta^7 = 1 - 2\theta^2$ which can be presented as the three-term equation $\theta^7 + \theta^2 + \theta^2 = 1$. It therefore gives accessible results from Kummer's equations at $n = 2$ and 3, but no further. Nevertheless the ladder can be pseudo-integrated and turns out to have valid extensions to higher orders. Although accessible at $n = 2$ and 3 this ladder is not part of the generic sequence.

It is convenient to retain the $\zeta(2)$ terms in the ladder definitions and to write the forms

$$L_n(3, \theta) = \frac{\mathrm{Li}_n(\theta^3)}{3^{n-1}} - \left\{\frac{-1}{3}\mathrm{Li}_n(\theta) + \frac{\zeta(2)}{3}\frac{\log^{n-2}\theta}{(n-2)!} - 2\frac{\log^n\theta}{n!}\right\}, \tag{4.93}$$



$$\begin{aligned} L_n(4, \theta) = \frac{\mathrm{Li}_n(\theta^4)}{4^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\theta^2)}{2^{n-1}} + \frac{1}{2}\mathrm{Li}_n(\theta) \\ &-\frac{1}{2}\zeta(2)\frac{\log^{n-2}\theta}{(n-2)!} + \frac{5}{2}\frac{\log^n\theta}{n!}\Bigg\}, \end{aligned} \tag{4.94}$$

$$\begin{aligned} L_n(6, \theta) = \frac{\mathrm{Li}_n(\theta^6)}{6^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\theta^3)}{3^{n-1}} + \frac{2\,\mathrm{Li}_n(\theta^2)}{2^{n-1}} + \frac{2}{3}\mathrm{Li}_n(\theta) \\ &-\zeta(2)\frac{\log^{n-2}\theta}{(n-2)!} + 6\frac{\log^n\theta}{n!}\Bigg\}, \end{aligned} \tag{4.95}$$

$$\begin{aligned} L_n(10, \theta) = \frac{\mathrm{Li}_n(\theta^{10})}{10^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\theta^5)}{5^{n-1}} + \frac{\mathrm{Li}_n(\theta^2)}{2^{n-1}} + \frac{1}{5}\mathrm{Li}_n(\theta) \\ &-\frac{2}{5}\zeta(2)\frac{\log^{n-2}\theta}{(n-2)!} + \frac{13}{5}\frac{\log^n\theta}{n!}\Bigg\}, \end{aligned} \tag{4.96}$$

$$\begin{aligned} L_n(14, \theta) = \frac{\mathrm{Li}_n(\theta^{14})}{14^{n-1}} - \Bigg\{&\frac{2\,\mathrm{Li}_n(\theta^7)}{7^{n-1}} + \frac{\mathrm{Li}_n(\theta^2)}{2^{n-1}} + \frac{\mathrm{Li}_n(\theta)}{7} \\ &-\frac{5}{14}\zeta(2)\frac{\log^{n-2}\theta}{(n-2)!} + \frac{17}{7}\frac{\log^n\theta}{n!}\Bigg\}, \end{aligned} \tag{4.97}$$

$$\begin{aligned} L_n(18, \theta) = \frac{\mathrm{Li}_n(\theta^{18})}{18^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\theta^9)}{9^{n-1}} + \frac{\mathrm{Li}_n(\theta^6)}{6^{n-1}} + \frac{2}{3}\frac{\mathrm{Li}_n(\theta^3)}{3^{n-1}} + \frac{\mathrm{Li}_n(\theta^2)}{2^{n-1}} \\ &+\frac{2}{3}\mathrm{Li}_n(\theta) - \frac{7}{9}\zeta(2)\frac{\log^{n-2}\theta}{(n-2)!} + \frac{14}{3}\frac{\log^n\theta}{n!}\Bigg\}, \end{aligned} \tag{4.98}$$

$$\begin{aligned} L_n(24, \theta) = \frac{\mathrm{Li}_n(\theta^{24})}{24^{n-1}} - \Bigg\{&\frac{\mathrm{Li}_n(\theta^{12})}{12^{n-1}} + \frac{2\,\mathrm{Li}_n(\theta^8)}{8^{n-1}} - \frac{\mathrm{Li}_n(\theta^6)}{6^{n-1}} \\ &-\frac{1}{4}\mathrm{Li}_n(\theta) - \frac{3}{4}\frac{\log^n\theta}{n!} + \frac{1}{6}\zeta(2)\frac{\log^{n-2}\theta}{(n-2)!}\Bigg\}. \end{aligned} \tag{4.99}$$

All of these are zero at $n = 1$ and 2.

4.3.3. *Analytic results at $n = 3$.* For $n = 3$, second-degree ladders, whose structure is dictated by Kummer's equations, are formed as follows:

$$L_n^{(2)}(4, \theta) = 4L_n(4, \theta) + 9L_n(3, \theta), \tag{4.100}$$

$$L_n^{(2)}(6, \theta) = 3L_n(6, \theta) + 11L_n(3, \theta), \tag{4.101}$$

$$L_n^{(2)}(10, \theta) = 25L_n(10, \theta) + 39L_n(3, \theta), \tag{4.102}$$

$$L_n^{(2)}(14, \theta) = 98L_n(14, \theta) + 123L_n(3, \theta), \tag{4.103}$$

$$L_n^{(2)}(18, \theta) = 27L_n(18, \theta) + 77L_n(3, \theta), \tag{4.104}$$

$$L_n^{(2)}(24, \theta) = 144L_n(24, \theta) - 99L_n(3, \theta). \tag{4.105}$$

At $n = 3$ these ladders are equal, respectively, to $\zeta(3)$ multiplied by 1, 1, 5, 27/2, 6, and $-5$. Equation (4.105) is apparently not directly accessible from Kummer's equation but comes from the functional equations of Chapter 6.

4.3.4. *Results at* $n = 4$. Both the index-14 and -24 ladders have no analytic proofs at this time, and the ladders have been determined numerically. All the others are part of the generic sequence. The $\zeta(3)$ terms associated with (4.100) to (4.105) can be eliminated by, for example, subtracting an appropriate amount of $L_n^{(2)}(4, \theta)$ from each of the other ladders. They are then combined with integer multipliers to get a rational multiple of $\zeta(4)$ at $n = 4$. The analysis is conveniently done numerically with guidance from Kummer's equation, and leads to defining

(4.106)
$$L_n^{(3)}(10, \theta) = 10[L_n^{(2)}(10, \theta) - 5L_n^{(2)}(4, \theta)] - 42[L_n^{(2)}(6, \theta) - L_n^{(2)}(4, \theta)] - \frac{19}{6}\zeta(4)\frac{\log^{n-4}\theta}{(n-4)!},$$

(4.107)
$$L_n^{(3)}(14, \theta) = 14[L_n^{(2)}(14, \theta) - \frac{27}{2}L_n^{(2)}(4, \theta)] - 285[L_n^{(2)}(6, \theta) - L_n^{(2)}(4, \theta)] - \frac{17}{12}\zeta(4)\frac{\log^{n-4}\theta}{(n-4)!},$$

(4.108)
$$L_n^{(3)}(18, \theta) = 18[L_n^{(2)}(18, \theta) - 6L_n^{(2)}(4, \theta)] - 36[L_n^{(2)}(6, \theta) - L_n^{(2)}(4, \theta)] - \frac{19}{3}\zeta(4)\frac{\log^{n-4}\theta}{(n-4)!},$$

(4.109)
$$L_n^{(3)}(24, \theta) = 24[L_n^{(2)}(24, \theta) + 5L_n^{(2)}(4, \theta)] - 324[L_n^{(2)}(6, \theta) - L_n^{(2)}(4, \theta)] + \frac{25}{2}\zeta(4)\frac{\log^{n-4}\theta}{(n-4)!}.$$

All of these are zero for $1 \le n \le 4$. Only the first and third have been obtained analytically, being accessible from Kummer's equation at $n = 4$.

4.3.5. *Results at* $n = 5$. In proceeding to $n = 5$ the above ladders are combined in pairs, leading to three valid results. Only that coming from (4.106) and (4.108) is accessible, as far as is known, from Kummer's fifth-order equation. Define

(4.110) $$L_n^{(4)}(14, \theta) = 14L_n^{(3)}(14, \theta) + 13L_n^{(3)}(10, \theta),$$

(4.111) $$L_n^{(4)}(18, \theta) = 3L_n^{(3)}(18, \theta) - 5L_n^{(3)}(10, \theta),$$

(4.112) $$L_n^{(4)}(24, \theta) = 4L_n^{(3)}(24, \theta) + 9L_n^{(3)}(10, \theta).$$

Then at $n = 5$ it is found that these are equal, respectively, to $\zeta(5)$ multiplied by 81/2, 14/3, and $-130/3$. Only (4.111) is accessible, being the generic result for $p = 3$, $q = 1$.



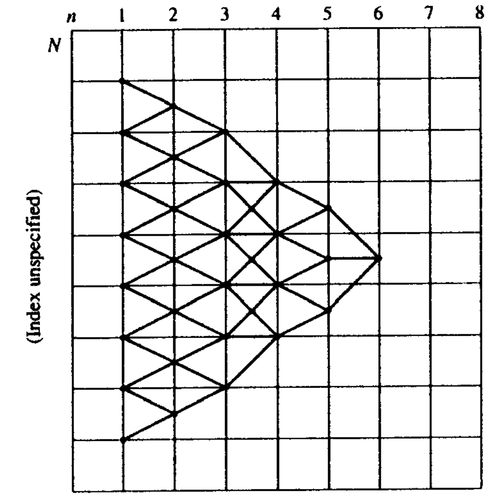



FIGURE 4.3. Flow chart for $\theta$-ladders

4.3.6. *The trans-Kummer result at* $n = 6$. The final step in going to $n = 6$ is to eliminate $\zeta(5)$ and then combine the resulting forms. Define

(4.113) $$L_n^{(5)}(18, \theta) = 243L_n^{(4)}(18, \theta) - 28L_n^{(4)}(14, \theta),$$

(4.114) $$L_n^{(5)}(24, \theta) = 7L_n^{(4)}(24, \theta) + 65L_n^{(4)}(18, \theta).$$

These are clearly zero at $n = 5$. At $n = 6$ it is then found numerically that

(4.115) $$324L_6^{(5)}(24, \theta) + 3L_6^{(5)}(18, \theta) + 5965\zeta(6) = 0.$$

This is the sole trans-Kummer result for $\theta$-ladders, and no others are anticipated.

4.3.7. *Flow chart for* $\theta$*-ladders.* Figure 4.3 shows the flow chart. As with other bases, the ladders combine in threes in going from odd to even orders, and in pairs from even to odd orders. Three results are thus required at the fifth order for the single valid ladder at $n = 6$.

**4.4. The nonexistence of functional equations at** $n = 6$ **with arguments limited to** $\pm z^m(1-z)^r(1+z)^s$.

4.4.1. *Limitations from the ladder structures; location of permissible forms.* The success reported in Chapter 6 on developing functional equations up to the fifth order, based on arguments of the form $\pm z^m(1-z)^r(1+z)^s$, supported the hope that this structure might be extended to the sixth order, since there exists an unlimited range of integers $m$, $r$, $s$. However, at the fifth order, the development of the known two families of equations came to a halt due to an absence of any available further usable elements. The problem became one of how to locate more permissible sets of integers that could be utilized in this way. Since the substitutions $z = \pm\rho$, $\pm 1/\rho$, $\pm\omega$, $\pm 1/\omega$ always give a power of $\rho$ or $\omega$, it is clear that the permissible values of $m$, $r$, and $s$ must be constrained so as to give only the indices and related factors of the $\rho$ and $\omega$ ladders. This means that there is an upper limit to these integers since the highest indices known are $N = 24$ and 42 respectively. On the assumption that there are no further supernumary equations, it is possible to locate all other usable integer-trios and exhaustively investigate the resulting formulas.

4.4.2. *Permissible powers.* From the known ladders, the $\rho$ and $\omega$ indices and powers are restricted, respectively, to the sets $R$ and $\Omega$, where

(4.116)
$$R = \{24, 20, 12, 10, 8, 6, 4, 3, 2, 1, 0\},$$
(4.117)
$$\Omega = \{42, 30, 28, 21, 18, 15, 14, 12, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0\}.$$

When inserted into $\mathrm{Li}_2[+z^m(1-z)^r(1+z)^s]$, some terms may involve a net negative index, in which case the argument is inverted to fit in with (4.116) or (4.117). If the numerical sign is negative, as, for example, it would be with $z = -\rho$ and $m$ odd, the duplication formula is used to give a positive argument with double the index. This leads to the formulation that the following quantities must belong to $R$:

(i) $m + 2r - s$;
(ii) $2(m + 2r - s)$ if $m + s$ is odd;
(iii) $m - r + 2s$ if $m$ or $r$ or both are even; and
(iv) $2(m - r + 2s)$ if $m$ or $r$ or both are odd.

Similarly, from the $\omega$-substitution, the following quantities must belong to $\Omega$:

(i) $m + 5r - 2s$;
(ii) $m - 2r + 5s$ if $m$ is even;
(iii) $2(m - 2r + 5s)$ if $m$ is odd;
(iv) $m - 4r + 3s$ if $r$ is even;
(v) $2(m - 4r + 3s)$ if $r$ is odd;
(vi) $m + 3r - 4s$ if $m + s$ is even;
(vii) $2(m + 3r - 4s)$ if $m + s$ is odd.

A similar formulation for $\mathrm{Li}_2[-z^m(1-z)^r(1+z)^s]$ leads to the requirement that the following belong to $R$:

(i) $m + 2r - s$ if $m + s$ is odd;
(ii) $2(m + 2r - s)$ all $m$, $r$, $s$;
(iii) $m - r + 2s$ if $m$ or $r$ or both are odd;
(iv) $2(m - r + 2s)$ if $m$ or $r$ or both are even.



The following belong to $\Omega$:

(i) $2(m + 5r - 2s)$ all $m$, $r$, $s$;
(ii) $m - 2r + 5s$ if $m$ is odd;
(iii) $2(m - 2r + 5s)$ if $m$ is even;
(iv) $m - 4r + 3s$ if $r$ is odd;
(v) $2(m - 4r + 3s)$ if $r$ is even;
(vi) $m + 3r - 4s$ if $m + s$ is odd;
(vii) $2(m + 3r - 4s)$ if $m + s$ is even.

4.4.3. *Negative search results.* A computer search with $0 \le m \le 50$, $-50 \le r$, $s \le 50$ gave sets of values up to $m = 26$ for the positive argument and $m = 15$ for the negative argument; the maximum sets reached being $(m, r, s) = (26, -2, -2)$ and $(15, -7, -5)$ respectively. No results for larger $m$ were generated, and it is assumed that none occur for $m > 50$, too. A further computer search using the MACSYMA program then looked for factorizations of $1 \mp z^m(1-z)^r(1+z)^s$ with the determined sets of $m$, $r$, $s$. No useful combinations other than those already utilized emerged from this analysis. It is concluded that no extension of these functional equations to $n = 6$ is possible in this way. This demonstration is contingent on the nonexistence of supernumary equations beyond those already known, and the absence of any formulas of the type considered outside the selected range for $m$, $r$, and $s$. Although short of a rigorous proof, this demonstration is sufficient to discourage further efforts in this direction.

H. Gangl (1990) has produced a single-variable functional equation at $n = 6$, with elements based on powers of $z$, $(1-z)$, and $(1-z+z^2)$. This equation is not able to give any of the $\rho$-ladders, but by taking $z$ as $-\omega$ or $\omega^2$ it gives two combinations of $\omega$-ladders.

Note added in proof. During April/May 1991, L. Lewin was supported for a period of four weeks at the Max-Planck-Institut für Mathematik at Bonn. In collaboration with H. Cohen, a number of outstanding matters have been resolved. In particular,

i) The base $\psi$ (where $\psi^4 + \psi^3 = 1$) possesses a further cyclotomic equation of index 60, the three supernumary ladders extending now to give a single valid ladder at $n = 6$. The same holds for the closely related base $\phi$.

ii) Two Salem numbers have been examined in some detail. One possesses 11 cyclotomic equations, and extends to the 8th order. The other, as discussed in Chapter 16, possesses 71, and is conjectured to yield ladder results at the 16th order. Four valid ladders at the 16th order have been generated.

## References

1. H. S. M. Coxeter, *The functions of Schläfli and Lobatschefsky*, Quart. J. Math. Oxford Ser. **6** (1935), 13–29.
2. M. J. Phillips and D. J. Whitehouse, *Two-dimensional discrete properties of random surfaces*, Philos. Trans. Roy. Soc. London Ser. A **305** (1982), 441–468.

3. Ljunggren, Math. Scand. 8 (1960), 65–70.
4. Tverberg, Math. Scand. 8 (1960), 121–126.
5. M. Abouzahra, L. Lewin, and Xiao Hongnian, *Polylogarithms in the field of omega (a root of a given cubic): functional equations and ladders*, Aequationes Math. 33 (1987), 23–45.
6. G. Szekeres, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. 13 (1970), 113–140.
7. Addendum to [5], Aequationes Math. 35 (1988), 304.

CHAPTER 5

# Supernumary Ladders

M. ABOUZAHRA AND L. LEWIN

## 5.1. The concept of supernumary results.

5.1.1. *Cyclotomic equations.* In §3.5 the base family $u^p + u^q = 1$ was introduced, for which there are eight *general* (i.e., for all $p$ and $q$) cyclotomic equations. As far as is known there are no more in the general case, but for *specific* values of $p$ and $q$ there may be additional ones. In some cases these may reduce to nothing more than a combination of the generic equations. In other cases, and particularly if a new index is involved, the formula will be new. We refer to these latter as *supernumary*; but it should be made clear that this does not *necessarily* mean that there is a total of more than eight cyclotomic equations. In some cases the generic indices may be repeated, leading to redundant equations and an impoverished set of generic formulas. Thus, with $p = q = 1$ the number of generic equations degenerates to two. With $p = 2, q = 1$ the generic set is reduced to four. However, in this case there are two further equations, with indices 20 and 24, and these are supernumary in the sense that they are not contained in the generic set; and they lead to a total of six cyclotomic equations for this base.

The quantity $\omega$ of §4.2 of Chapter 4, and which satisfies *two* base equations of the given form, nevertheless possesses a total of only eight generic indices, due to considerable mutual redundancies. However, there are four additional indices associated with four supernumary cyclotomic equations for this base, so it has a total of twelve such formulas allocated to it. Other examples will be given later.

5.1.2. *Supernumary ladders.* A supernumary cyclotomic equation is the result of a simple algebraic rearrangement of the base equation. There are a few guide lines on how to go about constructing them, but almost no basic theory to indicate whether any, and if so, how many, may exist. All those found to date have been consequential on lucky guesses from examining forms like $(1 - u^r)$ for various values of $r$, or from the more systematic methods developed by J. Browkin and by D. Zagier. Once a new cyclotomic equation has been discovered there seems to be no basic reason why it should be associated with a *valid* ladder. Of course, the corresponding supernumary *component*-ladder can be written down by inspection and tested for validity. In the few cases of transparency ($\rho$-ladders of indices 20 and 24 as far as $n = 5$) the method works. In many other cases, in the early studies, a combination with the generic component-ladder of index $p$ (or $q$) also succeeded, much to some initial astonishment; for there seems to be no known reason why a valid ladder should always ensue. Still later work showed the need for a further combination, using also the generic component ladder of index $2(2p - q)$. This seemed to go "against the grain" since this index was often quite unrelated to the factors of $N$, and almost all the early work had placed the presence of these factors into prominence. However, two (and only two) of the eight generic component-ladders are independent in the sense that there are just six generic valid ladder combinations; so combination with an independent pair is to be expected, unusual indices notwithstanding. And in all cases, apart from the exceptions to be discussed, a valid ladder ensued. These are considered in detail in subsequent sections.

5.1.3. *Nonvalid ladders.* To date there is only one known *specific* exception to the rule that supernumary cyclotomic equations always are associated with a valid ladder. This is for the pair $(p, q) = (11, 7)$, where an index-10 component-ladder seemingly has no valid combinations. This case is discussed further in §5.5. (This anomaly has now been resolved [3].)

There is, however, a *semi-generic* form which seems to be *quite barren* in all the cases examined so far. It arises as follows: if, to the base equation $u^p + u^q = 1$, one subtracts from each side a term $u^{(p+q)/2}$, the result takes the form

$$u^p - u^{(p+q)/2} + u^q = 1 - u^{(p+q)/2}. \tag{5.1}$$

The terms on the left can be rearranged as $u^q[u^{p-q} - u^{(p-q)/2} + 1]$ to give, on summation,

$$1 - u^{(p+q)/2} = (1 - u^{3(p-q)})(1 - u^{3(p-q)/2})^{-1}(1 - u^{(p-q)})^{-1}(1 - u^{(p-q)/2}). \tag{5.2}$$

This is a seemingly new generic form of index $(p + q)/2$, but in fact it is only appropriate to use if $(p + q)/2$ is integral. This requires *either* (i) $p$ and $q$ both even, or (ii) $p$ and $q$ both odd. The first case violates the stipulation $(p, q) = 1$. This case seems to give only redundant results, and is discussed further in §5.2.1. But there seems nothing amiss with the second alternative. It turns out that the initial cases $(p, q) = (3, 1), (5, 1), (5, 3)$, and $(7, 5)$ are all redundant versions of some combination of the generic formulas. But the cases $(7, 1)$ and $(7, 3)$, both of which have been numerically examined, yield no valid ladders. Neither does the case $(11, 7)$ of index-9. It is an interim conjecture that when $(p + q)/2$ is integral, and not part of the generic set, the corresponding component-ladder is not capable of combining with the two independent generic components to create a new valid ladder. If correct, the reasons for this are no more understood than is the complementary conjecture that valid combinations always arise from all other supernumary equations—with the specific exception of the index-10 formula for $(p, q) = (11, 7)$, the only currently-known anomaly.



5.1.4. *Other classes of supernumary equations.* Although the family of equations $u^p + u^q = 1$ has been the most thoroughly researched there are other generic equations for which supernumary results can also be found. Examples are discussed in §§5.6 and 5.7. In the latter, particularly, there is only a single generic result known; but *many* valid supernumary ladders have been found in the two cases so far investigated. G. Ray (Chapter 7) has recently succeeded in producing an analytic derivation of them.

One might also draw attention to a different sort of supernumary result, at higher orders. Here, the index is already known from formulas at the lower orders, where the ladders are *accessible.* Starting at the fourth order, or at the fifth, Kummer's equations no longer generate the index. This happens, for example, with the base $\theta$ (see §4.3) where the index 14 comes, at the second and third orders, from working with the combination $(1 - x)/(1 - y)$. At the higher orders the factors $(1 - x)$ and $(1 - y)$ occur separately, so that the properties arising from solely the combination $(1 - x)/(1 - y)$ are no longer available. For the base $\omega$ (see §4.2) the existence of an initial *pair* of base equations ($\omega$ is the *unique* base possessing more than one two-term base equation) gives some flexibility at the fourth order, but leaves the results short of four formulas at the fifth. In all these cases valid results exist nevertheless, as found by numerical computation. Because of their close affinity with the other *inaccessible* ladders for these bases, they may be termed *nonaccessible.*

In the case of $\omega$ the subsequent development of two new families of single-variable functional equations has enabled the nonaccessible $\omega$-results, up to the fifth order, to now be capable of analytic derivation. But difficulties with the family of equations to handle the $\theta$-ladders have so far prevented development in this direction beyond the third order.

Supernumary results for the base $\chi = \tan(\pi/12)$, which satisfies the equation $u^2 - 4u + 1 = 0$, are discussed in §3.7.2, where a cyclotomic equation of index-6 is developed, and gives rise to valid ladders for $n = 2, 3$, and $4$.

**5.2. Supernumary results for $p = 4$.**

5.2.1. *The case* $q = 2$. We shall touch briefly here on the assertion of §5.1.3 that only redundant results ensue if $(p, q) \neq 1$ by a consideration of the equation

$$u^4 + u^2 = 1. \tag{5.3}$$

Clearly, $u = \rho^{1/2}$, so any $\rho$-ladder becomes a corresponding $u$-ladder involving, however, only even powers of $u$. For a novel result to emerge, odd powers would also be required. It has been possible to find only one such result, obtained by adding $u^3$ to both sides of (5.3) to give, after minor manipulation,

$$1 - u^6 = (1 - u^3)^2(1 - u)^{-1}u^2. \tag{5.4}$$

The corresponding ladder is

(5.5) $$L_2(6, u) = \frac{\mathrm{Li}_2(u^6)}{6} - \{\tfrac{2}{3}\mathrm{Li}_2(u^3) - \mathrm{Li}_2(u) - \log^2 u\},$$

where $u = \rho^{1/2}$.

If this ladder is equated to $C\zeta(2)$ then a numerical calculation gives

(5.6) $$C = 0.46928439504570001\cdots,$$

which clearly is not a simple rational. (See Appendix A, §2.)

It is not known if there is a further result that could be combined with (5.5) to produce a rational coefficient. Subtracting rather than adding $u^3$ produces only a trivial change to (5.5). Tentatively one concludes that there are no new ladders of this character in the present case. To extend this conclusion to all cases where $(p, q) \neq 1$ is clearly a big leap. Currently it is just a conjecture with (5.5) and (5.6) as its sole backing.

5.2.2. *The case* $q = 1$. The cases $q = 1$ and $q = 3$ are closely related. Define $\phi$ and $\psi$ to be, respectively, the solutions in $(0, 1)$ of the equations

(5.7) $$u^4 + u = 1,$$

(5.8) $$u^4 + u^3 = 1.$$

Each equation possesses two real roots, the second being negative and numerically greater than unity. The transformation $u \to -1/u$ transforms (5.7) into (5.8) or *vice versa.* As a consequence, any $\psi$-ladder goes over into a related $\phi$-ladder of corresponding structure. The duplication and inversion formulas are needed to put it in standard form, but essentially the one can be written down from the other. Only the coefficients $D_m$ are different, and need to be evaluated separately.

Each base possesses seven generic cyclotomic equations and a corresponding quota of generic ladders up to the fifth order. Each also possesses two supernumary cyclotomic equations, one of which leads to valid ladders up to the fourth order, and the other to the second order only. Thus no trans-Kummer result can be found for these bases based on the present data, despite the presence of the needed pair of additional ladders at $n = 2$.

We shall deal only briefly with (5.7) and give the results for (5.8) in more detail.

A cyclotomic equation of index $N = 21$ can be found as follows:

$$1 + \phi^7 = 1 + \phi^3 \cdot \phi^4 = 1 + \phi^3 - \phi^4 = \phi(1 + \phi^2).$$

Squaring,

$$\begin{aligned}(1 + \phi^7)^2 - \phi^7 &= \phi^2(1 + 2\phi^2 + \phi^4 - \phi^5)\\ &= \phi^2[(1 + \phi^2 + \phi^4) + \phi^2(1 - \phi^3)].\end{aligned}$$



$$\begin{aligned}1 + \phi^7 + \phi^{14} &= \phi^2\left[\frac{1 - \phi^6}{1 - \phi^2} + \phi^2(1 - \phi^3)\right]\\ &= \phi^2\frac{(1 - \phi^3)}{(1 - \phi^2)}[1 + \phi^3 + \phi^2(1 - \phi^2)]\\ &= \phi\frac{(1 - \phi^3)}{(1 - \phi^2)}(1 + \phi^3 - \phi^5) = \frac{1 - \phi^3}{1 - \phi^2}(1 - \phi^6).\end{aligned}$$

Hence,

(5.9) $$1 - \phi^{21} = (1 - \phi^7)(1 - \phi^3)(1 - \phi^6)(1 - \phi^2)^{-1}.$$

This can be combined with the readily verified equations

(5.10) $$1 - \phi^6 = (1 - \phi^3)\phi^{-1},$$

(5.11) $$1 - \phi = \phi^4,$$

to give the preferred combination

(5.12) $$1 - \phi^{21} = (1 - \phi^7)(1 - \phi^6)^2(1 - \phi^2)^{-1}(1 - \phi)^{-1}\phi^5.$$

The corresponding ladder is

(5.13)
$$L_2(21, \phi) = \frac{\mathrm{Li}_2(\phi^{21})}{21} - \left\{\frac{\mathrm{Li}_2(\phi^7)}{7} + \frac{\mathrm{Li}_2(\phi^6)}{3} - \frac{\mathrm{Li}_2(\phi^2)}{2} - \mathrm{Li}_2(\phi) - \frac{5}{2}\log^2\phi\right\}.$$

This can now be combined with the index-one ladder

(5.14) $$L_2(1, \phi) = \mathrm{Li}_2(\phi) + 2\log^2\phi,$$

coming from (5.11).

A numerical search yields

(5.15) $$42L_2(21, \phi) - 37L_2(1, \phi) = 11\zeta(2).$$

This result seems incapable of extension to $n = 3$.

An index-30 cyclotomic equation can be generated as follows:

$$1 - \phi^5 = 1 - \phi(1 - \phi) = 1 - \phi + \phi^2 = \phi^2(1 + \phi^2).$$

Squaring gives

$$1 - 2\phi^5 + \phi^{10} = \phi^4(1 + 2\phi^2 + \phi^4) = \phi^4(2 - \phi + 2\phi^2).$$

Hence

$$1 - \phi^5 + \phi^{10} = 2\phi^4(1 + \phi^2),$$
$$1 + \phi^5 + \phi^{10} = 2\phi^4(1 + \phi + \phi^2).$$

Eliminating the factor 2 between the last two equations and rearranging leads to

(5.16)
$$1 - \phi^{30} = (1 - \phi^{15})^2(1 - \phi^{10})(1 - \phi^5)^{-2}(1 - \phi^3)^{-1}(1 - \phi^2)^{-1}\phi^5.$$

The index-30 ladder corresponding to this is

(5.17)

$$L_2(30,\phi)=\frac{\mathrm{Li}_2(\phi^{30})}{30}-\left\{\frac{2}{15}\mathrm{Li}_2(\phi^{15})+\frac{\mathrm{Li}_2(\phi^{10})}{10}-\tfrac{2}{5}\mathrm{Li}_2(\phi^5)\right.$$
$$\left.-\frac{\mathrm{Li}_2(\phi^3)}{3}-\frac{\mathrm{Li}_2(\phi^2)}{2}-\tfrac{5}{2}\log^2\phi\right\}.$$

This can be combined with (5.14), and a numerical search gives

(5.18) $$30L_2(30,\phi)-2L_2(1,\phi)=13\zeta(2).$$

This equation can be extended numerically as far as the fourth order, but apparently no further. Neither (5.15) nor (5.18) have any current analytical proof. It is not known if there are yet further cyclotomic equations which could be incorporated into these results.

5.2.3. *The case* $q = 3$. If we replace $\phi$ by $-1/\psi$ in (5.12), or alternatively go through a corresponding algebraic development starting with $1-\psi^7=\psi^4(1+\psi^2)$, the index-42 cyclotomic equation so generated leads to

(5.19) $$1-\psi^{42}=(1-\psi^{21})(1-\psi^{14})(1-\psi^7)^{-1}(1-\psi^6)^2(1-\psi^2)^{-2}(1-\psi).$$

The related ladder is

(5.20)

$$L_2(42,\psi)=\frac{\mathrm{Li}_2(\psi^{42})}{42}-\left\{\frac{\mathrm{Li}_2(\psi^{21})}{21}+\frac{\mathrm{Li}_2(\psi^{14})}{14}-\frac{\mathrm{Li}_2(\psi^7)}{7}+\frac{\mathrm{Li}_2(\psi^6)}{3}\right.$$
$$\left.-\mathrm{Li}_2(\psi^2)+\mathrm{Li}_2(\psi)\right\}.$$

(One can note the limitation of the powers $r$ to factors of 42 in this result.)

From the equation $(1-\psi^2)=(1-\psi)\psi^{-3}$ one can define an index-2 ladder

(5.21) $$L_2(2,\psi)=\frac{\mathrm{Li}_2(\psi^2)}{2}-\{\mathrm{Li}_2(\psi)+\tfrac{3}{2}\log^2\psi\}.$$

A numerical search then yields the counterpart of (5.15) in the form

(5.22) $$42L_2(42,\psi)-37L_2(2,\psi)=8\zeta(2).$$

Again, this result has not been extended to the next order. The equivalent of (5.16), found in a similar manner, is

(5.23)

$$1-\psi^{30}=(1-\psi^{15})^2(1-\psi^{10})(1-\psi^5)^{-2}(1-\psi^6)(1-\psi^3)^{-1}(1-\psi)\psi^{-3}.$$

The corresponding ladder, and the generic ladders up to the fifth order, follow. The generic indices are comprised of

(5.24) $$N=\{2,3,4,7,10,18,24\},$$



with the index 3 repeated. Along with the supernumary index 30 the following component-ladders can be defined (these are convenient forms; other legitimate variants exist):

(5.25) $$L_n(2,\psi)=\frac{\mathrm{Li}_n(\psi^2)}{2^{n-1}}-\mathrm{Li}_n(\psi)-\frac{3\log^n\psi}{n!},$$

(5.26) $$L_n(3,\psi)=\frac{\mathrm{Li}_n(\psi^3)}{3^{n-1}}+\frac{4\log^n\psi}{n!},$$

(5.27) $$L_n(4,\psi)=\frac{\mathrm{Li}_n(\psi^4)}{4^{n-1}}+\frac{3\log^n\psi}{n!},$$

(5.28) $$L_n(7,\psi)=\frac{\mathrm{Li}_n(\psi^7)}{7^{n-1}}+\mathrm{Li}_n(\psi)+\frac{10\log^n\psi}{n!},$$

(5.29) $$L_n(10,\psi)=\frac{\mathrm{Li}_n(\psi^{10})}{10^{n-1}}-\frac{\mathrm{Li}_n(\psi^5)}{5^{n-1}}+\mathrm{Li}_n(\psi)+\frac{7\log^n\psi}{n!},$$

(5.30) $$L_n(18,\psi)=\frac{\mathrm{Li}_n(\psi^{18})}{18^{n-1}}-\frac{\mathrm{Li}_n(\psi^9)}{9^{n-1}}-\frac{\mathrm{Li}_n(\psi^6)}{6^{n-1}}+\frac{\mathrm{Li}_n(\psi^2)}{2^{n-1}}+\frac{3\log^n\psi}{n!},$$

(5.31) $$L_n(24,\psi)=\frac{\mathrm{Li}_n(\psi^{24})}{24^{n-1}}-\frac{\mathrm{Li}_n(\psi^{12})}{12^{n-1}}-\frac{\mathrm{Li}_n(\psi^8)}{8^{n-1}}+\frac{\mathrm{Li}_n(\psi^2)}{2^{n-1}}+\frac{4\log^n\psi}{n!},$$

(5.32) $$L_n(30,\psi)=\frac{\mathrm{Li}_n(\psi^{30})}{30^{n-1}}-\frac{2\,\mathrm{Li}_n(\psi^{15})}{15^{n-1}}-\frac{\mathrm{Li}_n(\psi^6)}{6^{n-1}}+\frac{\mathrm{Li}_n(\psi^5)}{5^{n-1}}.$$

These give rise to six accessible ladders, and one of index 30, numerically determined:

(5.33) $$L_n^{(2)}(3,\psi)=3L_n(3,\psi)-L_n(2,\psi)-\zeta(2)\frac{\log^{n-2}\psi}{(n-2)!},$$

(5.34) $$L_n^{(2)}(4,\psi)=4L_n(4,\psi)+L_n(2,\psi),$$

(5.35) $$L_n^{(2)}(7,\psi)=7L_n(7,\psi)+6L_n(2,\psi)-3\zeta(2)\frac{\log^{n-2}\psi}{(n-2)!},$$

(5.36) $$L_n^{(2)}(10,\psi)=10L_n(10,\psi)+7L_n(2,\psi)-4\zeta(2)\frac{\log^{n-2}\psi}{(n-2)!},$$

(5.37) $$L_n^{(2)}(18,\psi)=18L_n(18,\psi)+L_n(2,\psi)-4\zeta(2)\frac{\log^{n-2}\psi}{(n-2)!},$$

(5.38) $$L_n^{(2)}(24,\psi)=24L_n(24,\psi)-7L_n(2,\psi)-10\zeta(2)\frac{\log^{n-2}\psi}{(n-2)!},$$

$$(5.39)\quad L_n^{(2)}(30,\psi)=30L_n(30,\psi)+3L_n(2,\psi)+\zeta(2)\frac{\log^{n-2}\psi}{(n-2)!}.$$

These ladders are all zero at $n=2$. At $n=3$ the following combinations are found:

$$(5.40)\quad L_n^{(3)}(4,\psi)=4L_n^{(2)}(4,\psi)+3L_n^{(2)}(3,\psi),$$

$$(5.41)\quad L_n^{(3)}(10,\psi)=10L_n^{(2)}(10,\psi)-14L_n^{(2)}(7,\psi)-9L_n^{(2)}(3,\psi),$$

$$(5.42)\quad L_n^{(3)}(18,\psi)=6L_n^{(2)}(18,\psi)-14L_n^{(2)}(7,\psi)+17L_n^{(2)}(3,\psi),$$

$$(5.43)\quad L_n^{(3)}(24,\psi)=8L_n^{(2)}(24,\psi)-28L_n^{(2)}(7,\psi)-19L_n^{(2)}(3,\psi),$$

$$(5.44)\quad L_n^{(3)}(30,\psi)=10L_n^{(2)}(30,\psi)+17L_n^{(2)}(3,\psi).$$

At $n=3$ these are equal to $\zeta(3)$ multiplied, respectively, by $1, -1, 1, -5, -2$. Combinations with the $\zeta(3)$ term cancelled are

(5.45)
$$L_n^{(4)}(10,\psi)=10\{5L_n^{(2)}(10,\psi)-7L_n^{(2)}(7,\psi)+2L_n^{(2)}(4,\psi)-3L_n^{(2)}(3,\psi)\},$$

(5.46)
$$L_n^{(4)}(18,\psi)=18\{3L_n^{(2)}(18,\psi)-7L_n^{(2)}(7,\psi)-2L_n^{(2)}(4,\psi)+7L_n^{(2)}(3,\psi)\},$$

(5.47)
$$L_n^{(4)}(24,\psi)=24\{2L_n^{(2)}(24,\psi)-7L_n^{(2)}(7,\psi)+5L_n^{(2)}(4,\psi)-L_n^{(2)}(3,\psi)\},$$

(5.48)
$$L_n^{(4)}(30,\psi)=30\{10L_n^{(2)}(30,\psi)+8L_n^{(2)}(4,\psi)+23L_n^{(2)}(3,\psi)\}.$$

Combinations of these, required either by Kummer's equations at $n=4$, or by numerical computation for the index-30 formula, are

$$(5.49)\quad L_n^{(5)}(18,\psi)=L_n^{(4)}(18,\psi)+L_n^{(4)}(10,\psi)-\frac{61}{3}\zeta(4)\frac{\log^{n-4}\psi}{(n-4)!},$$

$$(5.50)\quad L_n^{(5)}(24,\psi)=L_n^{(4)}(24,\psi)-L_n^{(4)}(10,\psi)+\frac{7}{2}\zeta(4)\frac{\log^{n-4}\psi}{(n-4)!},$$

$$(5.51)\quad L_n^{(5)}(30,\psi)=L_n^{(4)}(30,\psi)+14L_n^{(4)}(10,\psi)-\frac{41}{3}\zeta(4)\frac{\log^{n-4}\psi}{(n-4)!}.$$

These are all zero at $n=4$. At $n=5$ we have

$$(5.52)\quad 9L_5^{(5)}(18,\psi)+24L_5^{(5)}(24,\psi)=28\zeta(5),$$

obtained by Kummer's equation at $n=5$. Subsequent research has combined all these formulas with a new index 60 to reach an equation at $n=6$.

Results exist for the base $\phi$ of §5.2.2 corresponding to these equations.

**5.3. Supernumary results for $p=5$.**

5.3.1. *The case* $q=3$. For $q=1$ the base is $\omega$, as discussed in §4.2. For $q=3$ the equation is $u^5+u^3=1$. No novel cyclotomic formula has been found for this equation. The results obtained by adding (or subtracting) $u^4$ reduce to the generic ones. It is not known if there are any further equations, but in view of the tortuous routes that are often needed to generate new cyclotomic equations it is very difficult to know for sure that no further relevant results exist.

5.3.2. *The case* $q=2$. The cases $q=2$ and 4 do give valid ladders but the two sets have no obvious connection. A cyclotomic equation for the case $q=2$ comes by squaring the equation $u^2=1-u^5$ to give $u^4=1-2u^5+u^{10}$. Multiply by $u$, replace $u^5$ by $1-u^2$ and add $u^{10}$ to both sides to get, on rearrangement, $1-u+u^{10}-u^{11}=u^2-2u^6+u^{10}$ or $(1+u^{10})(1-u)=u^2(1-u^4)^2$.

Hence,

$$(5.53)\quad 1-u^{20}=(1-u^{10})(1-u^4)^2(1-u)^{-1}u^2.$$

The related ladder is

$$(5.54)\quad L_2(20,u)=\frac{\mathrm{Li}_2(u^{20})}{20}-\left\{\frac{\mathrm{Li}_2(u^{10})}{10}+\frac{\mathrm{Li}_2(u^4)}{2}-\mathrm{Li}_2(u)-\log^2 u\right\}.$$

This can be combined with the generic ladder of index $p=5$,

$$(5.55)\quad L_2(5,u)=\frac{\mathrm{Li}_2(u^5)}{5}+\log^2 u,$$

and the generic ladder of index $2(2p-q)=16$,

$$(5.56)\quad L_2(16,u)=\frac{\mathrm{Li}_2(u^{16})}{16}-\left\{\frac{\mathrm{Li}_2(u^8)}{8}+\frac{\mathrm{Li}_2(u^2)}{2}-\mathrm{Li}_2(u)-\log^2 u\right\},$$

to give, via numerical computation,

$$(5.57)\quad 20L_2(20,u)-32L_2(16,u)+55L_2(5,u)=\zeta(2).$$

A formula involving this result is proved in §7.7.1, together with a further result of index 24.

5.3.3. *The case* $q=4$. For $q=4$ we proceed by adding $u^7$ to both sides of $u^4+u^5=1$ and multiplying by $1-u$ to get

$$\begin{aligned}(1+u^7)(1-u)&=u^4(1-u^2+u^3-u^4)=u^4(-u^2+u^3+u^5)\\&=u^6(-1+u+u^3)=u^6(u+u^3-u^4-u^5)\\&=u^7(1+u^2-u^3-u^4)=u^7(u^2-u^3+u^5)\\&=u^8(u-u^2+u^4)=u^8(1+u-u^2-u^5)\\&=u^8[1+u(1-u-u^4)]=u^8[1-u(u-u^5)]\\&=u^8(1-u^7).\end{aligned}$$

Hence

$$(5.58)\quad 1-u^{14}=(1-u^7)^2(1-u)^{-1}u^8.$$

The related ladder is

$$L_2(14,u)=\frac{\mathrm{Li}_2(u^{14})}{14}-\left\{\frac{2\,\mathrm{Li}_2(u^7)}{7}-\mathrm{Li}_2(u)-4\log^2 u\right\}. \tag{5.59}$$

This can be combined with generic ladders of indices $q=4$ and $2(2p-q)=12$:

$$L_2(4,u)=\frac{\mathrm{Li}_2(u^4)}{4}+\frac{5}{2}\log^2 u \tag{5.60}$$

$$L_2(12,u)=\frac{\mathrm{Li}_2(u^{12})}{12}-\left\{\frac{\mathrm{Li}_2(u^6)}{6}-\frac{\mathrm{Li}_2(u^3)}{3}-\frac{1}{2}\log^2 u\right\}. \tag{5.61}$$

A numerical search produced

$$14L_2(14,u)+12L_2(12,u)+44L_2(4,u)=17\zeta(2). \tag{5.62}$$

It is not known if this result extends to $n=3$, or whether there are further supernumary equations. An analytic proof is given in §7.7.2.

**5.4. Supernumary results for $p=6$.**

5.4.1. *General comments.* The only cases for which $(p,q)=1$ are $q=1$ and 5. As with $\phi$ and $\psi$ of §(5.2) these are related, the base equations transforming into each other on replacing $u$ by $-1/u$.

It is not known whether the cases $q=2$, 3 and 4 give other than redundant results, but it is conjectured that this is always so when $(p,q)\neq 1$.

In all the analyses up to now a supernumary cyclotomic equation has always been found to be associated with a *valid* ladder result at $n=2$, and it was conjectured that this would be so in all cases where a novel cyclotomic equation could be generated. The results at $p=6$ confirm this.

5.4.3. *The case* $q=1$. The base equation for $(p,q)=(6,1)$ is

$$u^6+u=1. \tag{5.63}$$

Starting with $1+u^{10}$ and factorizing gives

$$\begin{aligned}1+u^{10}&=(1+u^2)(1-u^2+u^4-u^6+u^8)=(1+u^2)(u-u^3+u^4)\\&=u(1+u^2)(1+u^3-u^2)\\&=u(1+u^2)(1-u^3)^{-1}[1-u^6-u^2(1-u^3)]\\&=u(1+u^2)(1-u^3)^{-1}(u-u^2+u^5)\\&=u^2(1+u^2)(1-u^3)^{-1}(1-u+u^4)\\&=u^2(1+u^2)(1-u^3)^{-1}(u^4+u^6)=u^6(1+u^2)^2(1-u^3)^{-1}.\end{aligned}$$

Hence

$$1-u^{20}=(1-u^{10})(1-u^4)^2(1-u^3)^{-1}(1-u^2)^{-2}u^6, \tag{5.64}$$



with a corresponding ladder

$$L_2(20,u)=\frac{\mathrm{Li}_2(u^{20})}{20}-\left\{\frac{\mathrm{Li}_2(u^{10})}{10}+\frac{\mathrm{Li}_2(u^4)}{2}-\frac{\mathrm{Li}_2(u^3)}{3}-\mathrm{Li}_2(u^2)-3\log^2 u\right\}. \tag{5.65}$$

According to the recipe adopted successfully so far, this must be combined with two independent generic results, up till now taken to be of index $p$ or $q$ and $2(2p-q)$. In the present case the choice $q=1$ gives

$$L_2(1,u)=\mathrm{Li}_2(u)+3\log^2 u. \tag{5.66}$$

The second component-ladder used comes from the index $2(2p-q)=22$, and is

$$L_2(22,u)=\frac{\mathrm{Li}_2(u^{22})}{22}-\left\{\frac{\mathrm{Li}_2(u^{11})}{11}+\frac{\mathrm{Li}_2(u^8)}{8}-\frac{\mathrm{Li}_2(u^4)}{4}-\frac{1}{2}\log^2 u\right\}. \tag{5.67}$$

A numerical determination then produces

$$3\cdot 20L_2(20,u)+16\cdot 22L_2(22,u)-7L_2(1,u)=51\zeta(2). \tag{5.68}$$

A further supernumary result of index 8 is proved analytically in §7.7.3.

5.4.3. *The case* $q=5$. The transformation $u\to -1/u$ changes the base equation into

$$u^6+u^5=1. \tag{5.69}$$

The corresponding change in the ladders requires the use of the inversion formula, and leads to defining

$$\begin{aligned}L_2(20,u)=\frac{\mathrm{Li}_2(u^{20})}{20}-\Bigg\{&\frac{\mathrm{Li}_2(u^{10})}{10}-\frac{\mathrm{Li}_2(u^6)}{6}+\frac{\mathrm{Li}_2(u^4)}{2}\\&+\frac{\mathrm{Li}_2(u^3)}{3}-\mathrm{Li}_2(u^2)-\frac{3}{2}\log^2 u\Bigg\},\end{aligned} \tag{5.70}$$

$$L_2(11,u)=\frac{\mathrm{Li}_2(u^{11})}{11}-\left\{\frac{\mathrm{Li}_2(u^8)}{8}-\frac{\mathrm{Li}_2(u^4)}{4}-3\log^2 u\right\}, \tag{5.71}$$

$$L_2(2,u)=\frac{\mathrm{Li}_2(u^2)}{2}-\{\mathrm{Li}_2(u)+\tfrac{5}{2}\log^2 u\}. \tag{5.72}$$

The *structure* of the ladder is determined by (5.68); only the new coefficient of $\zeta(2)$ has to be found. The resulting formula is

$$3\cdot 20L_2(20,u)+32\cdot 11L_2(11,u)-7L_2(2,u)=64\zeta(2). \tag{5.73}$$

**5.5. Supernumary results for the equation-family $u^{6m+1} + u^{6r-1} = 1$.**

5.5.1. *Construction of the equation.* The difficulty in developing any systematic analysis stems from the lack, until recently, of a synthesis procedure for the cyclotomic equations. In §4.2 the equation $\omega^3 + \omega^2 - 1 = 0$ is converted into $\omega^5 + \omega - 1 = 0$ by multiplying by $\omega^2 - \omega + 1$. Since the roots of the latter are at $\exp(\pm i\pi/3)$ it follows that if either of the above terms $\omega^5$ or $\omega$ are multiplied by any power of $\omega^6$ the resulting equation will still be divisible by $\omega^2 - \omega + 1$. The polynomial quotient so developed can then be examined for the production of cyclotomic equations by algebraic rearrangement. This leads to a consideration of the equation

$$u^{6m+1} + u^{6r-1} = 1, \tag{5.74}$$

for integers $r > 0$, $m \geq 0$. It is fairly easy to reproduce many of the generic formulas this way. In some cases new results ensue. Apparently the pairs $(p, q) = (11, 1), (13, 5)$, and $(13, 11)$ are unproductive, but $(7, 5)$ and $(11, 7)$ give new results by this method. It may be noted that one can always produce a novel cyclotomic equation of index $N = (p+q)/2$ from (5.74) by subtracting $u^{(p+q)/2} = u^{3(m+r)}$ from each side; in fact this method works whenever both $p$ and $q$ are odd. For small values of $p$ and $q$ the resulting equations tend to be contained within the generic results, but otherwise this process doesn't seem to be very useful in producing valid ladders, as discussed in §5.1.3.

5.5.2. *Generic three-term equations.* For the second and third orders, Kummer's equations consist of polylogarithmic arguments of products of powers of $x$, $y$ and $(1-x)/(1-y)$ only. Hence one can generate powers of $u$ by taking

$$x = \pm u^m; \qquad y = -u^{r-s}; \qquad (1-x)/(1-y) = u^s. \tag{5.75}$$

Eliminating $x$ and $y$ gives the three-term equations (2.10) for the base $u$:

$$u^r + u^s \pm u^m = 1. \tag{5.76}$$

There are three generic forms that the two-term equation $u^p + u^q = 1$ takes when rearranged in three-term form. They are

$$u^p + u^{q+p} + u^{2q} = 1, \tag{5.77}$$

$$u^q + u^{q+p} + u^{2p} = 1, \tag{5.78}$$

$$u^q + u^{p-q} - u^{2p-q} = 1. \tag{5.79}$$

The corresponding dilogarithmic equations are (2.11) to (2.16) of Chapter 2. If it is possible to generate a further, nongeneric, three-term equation in a particular case, then the use of these equations will generate an accessible supernumary ladder. The method will be illustrated for the case $p = 7$, $q = 5$.



5.5.3. *The case* $(p, q) = (7, 5)$. From the identity $x^7 + x^5 - 1 = (x^2 - x + 1)(x^5 + x^4 + x^3 - x - 1)$ we get, from the quotient polynomial, with $x = u : 1 + u = u^3(1 + u + u^2)$, or $1 - u^2 = u^3(1 - u^3)$, whence

$$u^2 + u^3 - u^6 = 1. \tag{5.80}$$

This is of the form (5.76) and is nongeneric for $p = 7$, $q = 5$. Before taking $m = 6$, $r = 3$, $s = 2$ in the equations, (5.80) must first be put in cyclotomic form to generate the necessary ladders. The equation

$$1 - u^2 = u^3(1 - u^3) \tag{5.81}$$

suggests indices 2 and 3, neither of which is generic. Now the index 6 comes from both the generic $3(p-q)$ and the nongeneric $(p+q)/2$. These both happen to give the same result which degenerates, in this case, to an index-3 form. Therefore, the index 2 is the new one. To isolate it from (5.81), square the equation to get

$$\begin{aligned}(1-u^2)^2 &= u^6(1-u^3)^2 = u^6(1-u)(1+u+u^2)(1-u^3)\\ &= u^6(1-u)(1+u+u^2-u^3-u^4-u^5) = u^8(1-u),\end{aligned} \tag{5.82}$$

since $u^5 + u^4 + u^3 - u - 1 = 0$. Hence the supernumary component-ladder, on taking the square root of (5.82), is

$$L_2(2, u) = \frac{\mathrm{Li}_2(u^2)}{2} - \left\{\frac{\mathrm{Li}_2(u)}{2} - 2\log^2 u\right\}. \tag{5.83}$$

This has to be combined with, as it happens, only one generic component-ladder to produce a valid result. The index-4 formula suffices, and is given by

$$L_2(4, u) = \frac{\mathrm{Li}_2(u^4)}{4} - \left\{\frac{\mathrm{Li}_2(u^2)}{2} + \tfrac{5}{2}\log^2 u\right\}. \tag{5.84}$$

The resulting supernumary, accessible result is found to be

$$4L_2(4, u) + 2L_2(2, u) + \zeta(2) = 0. \tag{5.85}$$

There is a corresponding accessible result at $n = 3$, but apparently none at $n = 4$. A second supernumary result, of index 24, is proved in §7.7.4.

5.5.4. *The case* $(p, q) = (11, 7)$. For $p = 11$, $q = 7$ we have the identity

$$x^{11} + x^7 - 1 = (x^2 - x + 1)(x^9 + x^8 - x^6 + x^4 + x^3 - x - 1). \tag{5.86}$$

Hence, on multiplying by $u^2$, we get, from the quotient polynomial:

$$u^{11} + u^{10} - u^8 + u^6 + u^5 - u^3 - u^2 = 0. \tag{5.87}$$

This can be rearranged, on putting $u^{11} = 1 - u^7$, as $u^5(1-u) = (1+u^5) \times (1 - u^2 - u^3 + u^5)$ to give a nongeneric index-10 formula

$$1 - u^{10} = (1-u^5)(1-u^3)^{-1}(1-u^2)^{-1}(1-u)u^5. \tag{5.88}$$

There is also an index-9 result from the power $(p+q)/2$, obtained as in §5.1.3. If this is combined with the generic $3(p-q)$ formula one finds

$$1 - u^9 = (1-u^3)^{-1}(1-u^2)u^4. \tag{5.89}$$

The corresponding component-ladders are readily generated, but despite a lengthy numerical search, no valid ladders have been found for them. This is the first apparent exception encountered to the rule that supernumary cyclotomic equations give rise to corresponding valid ladders at $n = 2$. According to D. Zagier, at least one further supernumary component-ladder is needed to generate a valid result, and with the recent discovery of a new index-20 formula, the anomaly for this base has now been resolved [3].

### 5.6. Supernumary results for an irreducible quintic.

5.6.1. *Cyclotomic equations.* In §2.1.3 a generic three-term equation was discussed. Supernumary cases have been little researched, but one notable example comes from the quintic [1]

$$u^5 - u^3 + u^2 + u = 1. \tag{5.90}$$

By rearrangement (5.90) can be put in the form $1 - u^2 = u(1 - u^2 + u^4) = u(1+u^6)/(1-u^2)$, whence

$$u^7 + u^4 + u = 1. \tag{5.91}$$

This is of the form (2.10), leading to some accessible ladders. But it is quite exceptional, in fact unique in its class, because further rearrangements lead to the generation of two supernumary valid ladders of the inaccessible category. Re-expressing (5.91) in the form $1 = u(1 + u^3 + u^6) = u(1-u^9)/(1-u^3)$ gives the further three-term equation

$$u + u^3 - u^{10} = 1, \tag{5.92}$$

which is also of the form (2.10).

A further rearrangement of (5.92) creates the equation $1 + u^5 + u^{10} = u(1 + u^2 + u^4)$, or, after some manipulation,

$$1 - u^{15} = (1-u^5)(1-u^3)(1-u)^2u^{-8}. \tag{5.93}$$

This is a cyclotomic equation from which a corresponding component-ladder can be constructed, and then evaluated numerically. The net outcome is the generation of four accessible and two inaccessible valid ladders.



5.6.2. *Ladders at the second order.* Define the following component-ladders, each corresponding to an easily constructed cyclotomic equation

$$L_2(2,u) = \frac{\mathrm{Li}_2(u^2)}{2} + \mathrm{Li}_2(u) + \tfrac{5}{2}\log^2 u, \tag{5.94}$$

$$L_2(6,u) = \frac{\mathrm{Li}_2(u^6)}{6} - \frac{\mathrm{Li}_2(u^3)}{3} - \mathrm{Li}_2(u) - 2\log^2 u, \tag{5.95}$$

$$L_2(7,u) = \frac{\mathrm{Li}_2(u^7)}{7} - \mathrm{Li}_2(u) - \tfrac{3}{2}\log^2 u, \tag{5.96}$$

$$L_2(8,u) = \frac{\mathrm{Li}_2(u^8)}{8} - \frac{\mathrm{Li}_2(u^4)}{4} - 2\,\mathrm{Li}_2(u) - \tfrac{7}{2}\log^2 u, \tag{5.97}$$

$$L_2(9,u) = \frac{\mathrm{Li}_2(u^9)}{9} - \frac{\mathrm{Li}_2(u^3)}{3} - \tfrac{1}{2}\log^2 u, \tag{5.98}$$

$$L_2(12,u) = \frac{\mathrm{Li}_2(u^{12})}{12} - \frac{\mathrm{Li}_2(u^4)}{4} - \frac{\mathrm{Li}_2(u^3)}{3} + \frac{\mathrm{Li}_2(u^2)}{2}, \tag{5.99}$$

$$L_2(15,u) = \frac{\mathrm{Li}_2(u^{15})}{15} - \frac{\mathrm{Li}_2(u^5)}{5} - \frac{\mathrm{Li}_2(u^3)}{3} - 2\,\mathrm{Li}_2(u) - 4\log^2 u, \tag{5.100}$$

$$L_2(20,u) = \frac{\mathrm{Li}_2(u^{20})}{20} - \frac{\mathrm{Li}_2(u^{10})}{10} - \frac{\mathrm{Li}_2(u^4)}{4} + \frac{\mathrm{Li}_2(u^2)}{2} + \tfrac{1}{2}\log^2 u. \tag{5.101}$$

Then the following six independent valid ladder combinations can be readily found, the first four via (2.11) to (2.16) and the last two by numerical computation:

$$6L_2(6,u) = 4L_2(8,u) + 9L_2(9,u) + 2\zeta(2), \tag{5.102}$$

$$7L_2(7,u) = 4L_2(8,u) + \zeta(2), \tag{5.103}$$

$$6L_2(12,u) = 3L_2(6,u) + 4L_2(2,u) - \zeta(2), \tag{5.104}$$

$$5L_2(20,u) = 2L_2(8,u) + 4L_2(2,u), \tag{5.105}$$

$$8L_2(8,u) - 6L_2(6,u) + 5L_2(2,u) = -2\zeta(2), \tag{5.106}$$

$$15L_2(15,u) - 24L_2(6,u) + L_2(2,u) = -\tfrac{1}{2}\zeta(2). \tag{5.107}$$

5.6.3. *Ladders at the third order.* The second-order ladders of the previous section can be pseudo-integrated, and arranged in the following combinations

$$L_n^{(2)}(6,u) = 6L_n(6,u) - 14L_n(7,u) - 5L_n(2,u), \tag{5.108}$$

$$L_n^{(2)}(7,u) = 7L_n(7,u) - 4L_n(8,u) - \zeta(2)\frac{\log^{n-2} u}{(n-2)!}, \tag{5.109}$$

$$L_n^{(2)}(9,u) = 9L_n(9,u) - 6L_n(6,u) + 4L_n(8,u) + 2\zeta(2)\frac{\log^{n-2} u}{(n-2)!}, \tag{5.110}$$

(5.111)

$$L_n^{(2)}(12,u) = 12L_n(12,u) - 6L_n(6,u) - 8L_n(2,u) + 2\zeta(2)\frac{\log^{n-2} u}{(n-2)!},$$

(5.112)

$$L_2^{(2)}(20,u) = 15L_n(15,u) - 24L_n(6,u) - L_n(2,u) + \tfrac{1}{2}\zeta(2)\frac{\log^{n-2} u}{(n-2)!}.$$

Then it is found analytically from Kummer's equations that

$$5L_3^{(2)}(20,u) - 14L_3^{(2)}(7,u) - 6L_3^{(2)}(12,u) = -2\zeta(3), \tag{5.113}$$

$$3L_3^{(2)}(9,u) + 7L_3^{(2)}(7,u) = \zeta(3), \tag{5.114}$$

or, by elimination of $\zeta(3)$,

$$5L_3^{(2)}(20,u) + 6L_3^{(2)}(9,u) - 6L_3^{(2)}(12,u) = 0. \tag{5.115}$$

No other valid ladder combinations could be discovered, though a computer run came up with

$$29L_3^{(2)}(20,u) - 36L_3^{(2)}(12,u) + 30L_3^{(2)}(6,u) = C\zeta(3), \tag{5.116}$$

where $C = 118.99999282251\ldots$. This result is presented as an example of the sort of artifact one occasionally encounters in such calculations. Double precision was essential to ensure that the erroneous identification $C = 119$ was avoided. No results at the fourth order could be found.

**5.7. Supernumary ladders from a 15-term functional equation.**

5.7.1. *The base equation.* In §2.2.3 a fifteen-term functional equation was derived via Clausen's functional equation. The five variables involved are subject to a single constraint, which essentially is the base equation for the ladder-type relation (2.36). For five arbitrary integers $q_m$ the base equation is the eleven-term relation

$$u^s + 1 = \sum_{m=1}^{5}(u^{q_m} + u^{s-q_m}); \qquad 2s = \sum_{m=1}^{5} q_m, \tag{5.117}$$

and (2.36) is the single (accessible) ladder associated therewith. However, for specific values of the $q_m$, (5.117) can be rearranged in various cyclotomic forms, to each of which corresponds a component-ladder. It is the properties of these supernumary expressions that will be examined here.

5.7.2. *The base* $\Omega$. The case $q_1 = q_2 = 3$, $q_3 = q_4 = q_5 = 2$ was considered in [2] and gives rise to the equation

$$u^6 - 3u^4 - 4u^3 - 3u^2 + 1 = 0. \tag{5.118}$$



There is one root, $\Omega$, $= (1/2)(1 + 3^{1/2} - 12^{1/4})$ in $(0, 1)$, for which there is a single accessible ladder formula:

(5.119)

$$L_2(4,\Omega) = \frac{\mathrm{Li}_2(\Omega^4)}{4} - \left\{\frac{\mathrm{Li}_2(\Omega^3)}{3} + \frac{\mathrm{Li}_2(\Omega^2)}{2} + \mathrm{Li}_2(\Omega) + \frac{1}{2}\log^2\Omega\right\}$$
$$= -(7/12)\zeta(2).$$

This is not quite of the strict ladder form because of the presence of a term in $\Omega^3$, and gives rise to the suspicion that (5.119) is really a second-degree ladder involving ladders of indices 4 and 3. An attempt was made to find a cyclotomic equation involving $(1 - x^3)$ from which the expected structure could be generated. Although it is not known for sure that no such equation exists, it did not seem possible to generate one. However, in the process, several other cyclotomic equations emerged, and the following sequence, not necessarily complete, can be constructed fairly readily from (5.118).

$$(1 - x^4) = (1 - x^3)(1 - x^2)(1 - x)x^{-1}, \tag{5.120}$$

$$(1 - x^6) = (1 - x^2)^3(1 - x)^{-4}x^2, \tag{5.121}$$

$$(1 - x^8) = (1 - x^4)^3(1 - x^2)^{-2}(1 - x)^2x^{-1}, \tag{5.122}$$

$$(1 - x^{10}) = (1 - x^5)(1 - x^3)(1 - x^2)(1 - x)^{-2}x, \tag{5.123}$$

$$(1 - x^{18}) = (1 - x^9)(1 - x^6)^3(1 - x^3)^{-2}(1 - x^2)^{-2}(1 - x). \tag{5.124}$$

Component-ladder structures corresponding to these equations can be written down, but do not necessarily produce rational multiples of $\pi^2$ or $\zeta(2)$. Define component-ladders and constants $A^{(N)}$ (which may not be rational) as follows:

(5.125)

$$L_2(4,\Omega) = \frac{\mathrm{Li}_2(\Omega^4)}{4} - \left\{\frac{\mathrm{Li}_2(\Omega^3)}{3} + \frac{\mathrm{Li}_2(\Omega^2)}{2} + \mathrm{Li}_2(\Omega) + \frac{1}{2}\log^2\Omega\right\}$$
$$= \frac{A^{(4)}}{4}\pi^2,$$

(5.126)

$$L_2(6,\Omega) = \frac{\mathrm{Li}_2(\Omega^6)}{6} - \left\{\frac{3}{2}\mathrm{Li}_2(\Omega^2) - 4\,\mathrm{Li}_2(\Omega) - \log^2\Omega\right\} = \frac{A^{(6)}\pi^2}{6}$$

(5.127)

$$L_2(8,\Omega) = \frac{\mathrm{Li}_2(\Omega^8)}{8} - \left\{\frac{3}{4}\mathrm{Li}_2(\Omega^4) - \mathrm{Li}_2(\Omega^2) + 2\,\mathrm{Li}_2(\Omega) + \frac{1}{2}\log^2\Omega\right\}$$
$$= \frac{A^{(8)}\pi^2}{8},$$

(5.128)

$$L_2(10,\Omega)=\frac{\mathrm{Li}_2(\Omega^{10})}{10}-\left\{\frac{\mathrm{Li}_2(\Omega^5)}{5}+\frac{\mathrm{Li}_2(\Omega^3)}{3}+\frac{\mathrm{Li}_2(\Omega^2)}{2}\right.$$
$$\left.-2\,\mathrm{Li}_2(\Omega)-\frac{1}{2}\log^2\Omega\right\}=\frac{A^{(10)}}{10}\pi^2,$$

(5.129)

$$L_2(18,\Omega)=\frac{\mathrm{Li}_2(\Omega^{18})}{18}-\left\{\frac{\mathrm{Li}_2(\Omega^9)}{9}+\frac{1}{2}\mathrm{Li}_2(\Omega^6)\right.$$
$$\left.-\frac{2}{3}\mathrm{Li}_2(\Omega^3)-\mathrm{Li}_2(\Omega^2)+\mathrm{Li}_2(\Omega)\right\}=\frac{A^{(18)}}{18}\pi^2.$$

Equation (5.125) is, of course, (5.119). Thus we expect $A^{(4)}=-7/18$. The remainder of the $A$'s are not known from any analytic formulas, but can be readily calculated numerically. This gives (to nine decimal places) the values in Table 1, and confirms the value of $A^{(4)}$ as $-7/18$. The others do not appear to be individually rational, but simple rational combinations can easily be selected by inspection. In particular we have the following three relations among the last four coefficients:

(5.130) $$A^{(6)}+A^{(18)}=1,$$

(5.131) $$A^{(6)}+A^{(8)}=1/2,$$

(5.132) $$A^{(8)}+A^{(10)}=5/18.$$

The use to be made of these results is as follows: if we knew, from analysis, the second-degree ladder structures, then (5.130) to (5.132) could be simply verified. But lacking such analysis we can reverse the process to create such ladders, to correspond with (5.130) to (5.132), by adding appropriate multiples of the formulas of (5.125) to (5.129). The resulting second-degree ladders are therefore

(5.133) $$L_2^{(2)}(8,\Omega)=L_2(8,\Omega)+\tfrac{3}{4}L_2(6,\Omega)=\pi^2/16,$$

(5.134) $$L_2^{(2)}(10,\Omega)=L_2(10,\Omega)+\tfrac{4}{5}L_2(8,\Omega)=\pi^2/36,$$

(5.135) $$L_2^{(2)}(18,\Omega)=L_2(18,\Omega)+\tfrac{1}{3}L_2(6,\Omega)=\pi^2/18.$$

See also equation (5.200) for a new ladder of index 16. G. Ray has now provided an analytic proof of these results.

| $N$ | 4 | 6 | 8 | 10 | 18 |
|---|---|---|---|---|---|
| $A^{(N)}$ | −3.88888889 | 1.443141138 | −.943141138 | 1.220918916 | −.443141138 |

TABLE 1. Coefficients for Ω-ladders



5.7.3. *The base* Γ. If in (5.127) we take $q_m=2m$ the base equation becomes

(5.136) $$u^{15}+1=u^{13}+u^{11}+u^{10}+u^9+u^8+u^7+u^6+u^5+u^4+u^2.$$

If roots on the unit circle are removed the equation is reduced to

(5.137) $$u^6-u^5-u^3-u+1=0,$$

with one real root Γ in $(0,1)$. The ladder-like relation, the sole analytic result, becomes

(5.138)

$$\mathrm{Li}_2(\Gamma^{18})+\mathrm{Li}_2(\Gamma^{14})=2\,\mathrm{Li}_2(\Gamma^9)+2\,\mathrm{Li}_2(\Gamma^8)+2\,\mathrm{Li}_2(\Gamma^7)+\mathrm{Li}_2(\Gamma^6)$$
$$+4\,\mathrm{Li}_2(\Gamma^5)+2\,\mathrm{Li}_2(\Gamma^4)+2\,\mathrm{Li}_2(\Gamma^3)+\mathrm{Li}_2(\Gamma^2)$$
$$+2\,\mathrm{Li}_2(\Gamma)+15\log^2\Gamma-2\pi^2/3.$$

To put this into ladder form we note the nine cyclotomic sequences, obtainable through manipulation of (5.136)

(5.139) $$1-\Gamma^5=(1-\Gamma)^{-1}\Gamma^3,$$

(5.140) $$1-\Gamma^6=(1-\Gamma^3)^2(1-\Gamma^4)\Gamma^{-2},$$

(5.141) $$1-\Gamma^8=(1-\Gamma^2)(1-\Gamma)^{-2}\Gamma^4,$$

(5.142) $$1-\Gamma^9=(1-\Gamma^3)(1-\Gamma^4)^2(1-\Gamma^2)^{-2}\Gamma,$$

(5.143) $$1-\Gamma^{10}=(1-\Gamma^5)(1-\Gamma^4)(1-\Gamma^2)^2(1-\Gamma)^{-1}\Gamma^{-1},$$

(5.144) $$1-\Gamma^{12}=(1-\Gamma^6)^2(1-\Gamma^2)^{-1}\Gamma,$$

(5.145) $$1-\Gamma^{14}=(1-\Gamma^7)(1-\Gamma^4)(1-\Gamma)^{-1}\Gamma^2,$$

(5.146) $$1-\Gamma^{18}=(1-\Gamma^9)(1-\Gamma^3)(1-\Gamma^2)(1-\Gamma)^{-2}\Gamma^3,$$

(5.147) $$1-\Gamma^{30}=(1-\Gamma^{15})(1-\Gamma^{10})(1-\Gamma^6)(1-\Gamma^3)^{-1}(1-\Gamma^2)^{-1}\Gamma^2.$$

The corresponding component-ladders can be put in the form

(5.148)

$$L_2(5,\Gamma)=\frac{\mathrm{Li}_2(\Gamma^5)}{5}-\left\{-\mathrm{Li}_2(\Gamma)-\frac{3}{2}\log^2\Gamma\right\}=\frac{A^{(5)}}{5}\pi^2,$$

(5.149)

$$L_2(6,\Gamma)=\frac{\mathrm{Li}_2(\Gamma^6)}{6}-\left\{\frac{2}{3}\mathrm{Li}_2(\Gamma^3)+\frac{\mathrm{Li}_2(\Gamma^4)}{4}+\log^2\Gamma\right\}=\frac{A^{(6)}}{6}\pi^2,$$

(5.150)

$$L_2(8,\Gamma)=\frac{\mathrm{Li}_2(\Gamma^8)}{8}-\left\{\frac{\mathrm{Li}_2(\Gamma^2)}{2}-2\,\mathrm{Li}_2(\Gamma)-2\log^2\Gamma\right\}=\frac{A^{(8)}}{8}\pi^2,$$

(5.151)

$$L_2(9,\Gamma)=\frac{\mathrm{Li}_2(\Gamma^9)}{9}-\left\{\frac{\mathrm{Li}_2(\Gamma^3)}{3}+2\frac{\mathrm{Li}_2(\Gamma^4)}{4}-\mathrm{Li}_2(\Gamma^2)-\frac{1}{2}\log^2\Gamma\right\}=\frac{A^{(9)}}{9}\pi^2,$$

(5.152)

$$L_2(10,\Gamma) = \frac{\mathrm{Li}_2(\Gamma^{10})}{10} - \left\{\frac{\mathrm{Li}_2(\Gamma^5)}{5} + \frac{\mathrm{Li}_2(\Gamma^4)}{4} + \mathrm{Li}_2(\Gamma^2) - \mathrm{Li}_2(\Gamma) + \frac{1}{2}\log^2\Gamma\right\}$$
$$= \frac{A^{(10)}}{10}\pi^2,$$

(5.153)

$$L_2(12,\Gamma) = \frac{\mathrm{Li}_2(\Gamma^{12})}{12} - \left\{\frac{\mathrm{Li}_2(\Gamma^6)}{3} - \frac{\mathrm{Li}_2(\Gamma^2)}{2} + \frac{1}{2}\log^2\Gamma\right\} = \frac{A^{(12)}}{12}\pi^2,$$

(5.154)

$$L_2(14,\Gamma) = \frac{\mathrm{Li}_2(\Gamma^{14})}{14} - \left\{\frac{\mathrm{Li}_2(\Gamma^7)}{7} + \frac{\mathrm{Li}_2(\Gamma^4)}{4} - \mathrm{Li}_2(\Gamma) - \log^2\Gamma\right\} = \frac{A^{(14)}}{14}\pi^2,$$

(5.155)

$$L_2(18,\Gamma) = \frac{\mathrm{Li}_2(\Gamma^{18})}{18} - \left\{\frac{\mathrm{Li}_2(\Gamma^9)}{9} + \frac{\mathrm{Li}_2(\Gamma^3)}{3}1 + \frac{\mathrm{Li}_2(\Gamma^2)}{2} - 2\,\mathrm{Li}_2(\Gamma) - \frac{3}{2}\log^2\Gamma\right\}$$
$$= \frac{A^{(18)}}{18}\pi^2,$$

(5.156)

$$L_2(30,\Gamma) = \frac{\mathrm{Li}_2(\Gamma^{30})}{30} - \left\{\frac{\mathrm{Li}_2(\Gamma^{15})}{15} + \frac{\mathrm{Li}_2(\Gamma^{10})}{10} + \frac{\mathrm{Li}_2(\Gamma^6)}{6}\right.$$
$$\left. - \frac{\mathrm{Li}_2(\Gamma^3)}{3} - \frac{\mathrm{Li}_2(\Gamma^2)}{2} - \log^2\Gamma\right\} = \frac{A^{(30)}}{30}\pi^2.$$

It is only if these ladders are transparent that the $A^{(N)}$ will be rational. However, rational combinations can be sought, with the following interesting findings:

(5.157) $$A^{(5)} = \frac{4}{3} - \frac{1}{2}A^{(30)},$$

(5.158) $$A^{(6)} = \frac{5}{6} - \frac{1}{3}A^{(10)} - \frac{2}{3}A^{(30)},$$

(5.159) $$A^{(8)} = \frac{7}{3} + \frac{2}{3}A^{(10)} - \frac{2}{3}A^{(30)},$$

(5.160) $$A^{(9)} = \frac{17}{6} - A^{(10)} - \frac{3}{2}A^{(30)},$$

(5.161) $$A^{(12)} = 1 - \frac{2}{3}A^{(10)} - \frac{1}{3}A^{(30)},$$

(5.162) $$A^{(14)} = \frac{10}{3} + \frac{1}{3}A^{(10)} - \frac{4}{3}A^{(30)},$$

(5.163) $$A^{(18)} = \frac{41}{6} + \frac{2}{3}A^{(10)} - \frac{8}{3}A^{(30)},$$

(5.164) $$A^{(18)} = 2A^{(14)} + 1/6.$$



The last formula is not independent of the others, but is included because, apart from (5.157), it is the only other two-component relation uncovered.

Equation (5.138) does not give rise to any new relations between the $A^{(N)}$. However, it is the only formula with a current analytic derivation, and it can be expressed in the second-degree ladder form

(5.165)

$$\overline{L}_2^{(2)}(18,\Gamma) = 9L_2(18,\Gamma) + 7L_2(14,\Gamma) - 8L_2(8,\Gamma) - 3L_2(6,\Gamma) - 10L_2(5,\Gamma)$$
$$= -\pi^2/3.$$

One first-degree ladder is obtainable from (5.147):

(5.166)

$$\overline{L}_2(30,\Gamma) = \frac{\mathrm{Li}_2(\Gamma^{30})}{30} - \left\{\frac{\mathrm{Li}_2(\Gamma^{15})}{15} + \frac{\mathrm{Li}_2(\Gamma^{10})}{10} + \frac{\mathrm{Li}_2(\Gamma^6)}{6} - \frac{\mathrm{Li}_2(\Gamma^5)}{15}\right.$$
$$\left. - \frac{\mathrm{Li}_2(\Gamma^2)}{2} - \frac{\mathrm{Li}_2(\Gamma)}{3} - \tfrac{1}{2}\log^2\Gamma\right\}$$
$$= 4\pi^2/45.$$

Of the seven independent formulas implicit in (5.157) to (5.164), six remain after extraction of (5.166) and can be expressed as second-degree ladders. It is necessary first to rearrange the connections between the different $A^{(N)}$ and there is no unique way to do this. The following is one acceptable way:

(5.167) $$4A^{(30)} + 3A^{(14)} - A^{(10)} = 10,$$

(5.168) $$A^{(18)} - 2A^{(14)} = \tfrac{1}{6},$$

(5.169) $$A^{(12)} - 2A^{(6)} + 2A^{(5)} = 2,$$

(5.170) $$A^{(10)} - A^{(8)} + A^{(6)} = -\tfrac{3}{2},$$

(5.171) $$A^{(9)} - 3A^{(6)} + A^{(5)} = \tfrac{5}{3},$$

(5.172) $$A^{(8)} + 2A^{(6)} - 4A^{(5)} = -\tfrac{4}{3}.$$

The corresponding valid second-degree ladders are:

(5.173)

$$L_2^{(2)}(30,\Gamma) = L_2(30,\Gamma) + \tfrac{7}{20}L_2(14,\Gamma) - \tfrac{1}{12}L_2(10,\Gamma) = \pi^2/12,$$

(5.174) $$L_2^{(2)}(18,\Gamma) = L_2(18,\Gamma) - \tfrac{14}{9}L_2(14,\Gamma) = \pi^2/108,$$

(5.175) $$L_2^{(2)}(12,\Gamma) = L_2(12,\Gamma) - L_2(6,\Gamma) + \tfrac{5}{6}L_2(5,\Gamma) = \pi^2/6,$$

(5.176)

$$L_2^{(2)}(10,\Gamma) = L_2(10,\Gamma) - \tfrac{4}{5}L_2(8,\Gamma) + \tfrac{3}{5}L_2(6,\Gamma) = -3\pi^2/2,$$

(5.177) $$L_2^{(2)}(9,\Gamma) = L_2(9,\Gamma) - 2L_2(6,\Gamma) + \tfrac{5}{9}L_2(5,\Gamma) = 5\pi^2/27,$$

(5.178) $$L_2^{(2)}(8,\Gamma) = L_2(8,\Gamma) + \tfrac{3}{2}L_2(6,\Gamma) - \tfrac{5}{2}L_2(5,\Gamma) = -\pi^2/6.$$

G. Ray (Chapter 7) has recently provided an analytic derivation of these results.

5.7.4. *Concluding remarks.* From the foregoing the key role of the cyclotomic equation in generating component ladders for evaluation is very clear. What is missing is some basic understanding as to why the existence of a cyclotomic equation, along with its component-ladder, should be associated with a *valid* ladder result. The existence of the semi-generic barren equations of §5.1.3, and the apparent exception at $p = 11$, $q = 7$, shows that this association is not an absolute one; more research, needed to uncover the real nature of the connection, is discussed by Zagier in Appendix A.

**5.8. Supernumary ladders on the unit circle.**

5.8.1. *The functional-equation conjecture and its consequences.* The supernumary results of §5.7 appear not to be accessible from the 15-term functional equation, and it is a conjecture that the supernumary ladders may be obtainable from *some* functional equation. If this is so, the equation will not distinguish between the different roots of the base equation (5.117). Since the latter will also have roots on the unit circle, the hypothesis predicts corresponding supernumary results for the Clausen function. The existence of such results, if confirmed, would lend support to the conjecture, while their absence would presumably refute it.

Although the Clausen combination corresponding to the *component*-ladders cannot be examined on this basis, the combinations corresponding to valid dilogarithmic ladders, *with the same combining constants*, should, in some sense, be valid. Since all known Clausen-function equations are linear, *and with no constant or logarithmic terms*, the conjecture, if valid, would assert that the sum of Clausen functions corresponding to the dilogarithmic supernumary ladder structure should equal zero. This contention is examined in the next section for the equation leading to the base $\Omega$ of §5.7.2.

5.8.2. *The base equation.* On the unit circle take $u = e^{i\theta}$ so that (5.118) becomes

(5.179) $$\cos 3\theta - 3\cos\theta - 2 = 0.$$

Putting $\cos\theta = C$ this reduces to

(5.180) $$2C^3 - 3C - 1 = 0.$$

Removing the factor $(C+1)$, the ensuing quadratic is $(2C^2 - 2C - 1) = 0$, with the solution for $C$ in $(-1, 1)$ of

(5.181) $$C = -(\sqrt{3} - 1)/2.$$

This has the solution $\theta \approx 111.47^\circ$. In the remainder of this section $\theta$ is understood to take this value.

5.8.3. *Clausen component-ladders.* These correspond to (5.125) to (5.129). $\Omega$ is replaced by $e^{i\theta}$, and, on taking the imaginary part, $\mathrm{Li}_2(e^{i\theta})$ is



replaced by $\mathrm{Cl}_2(\theta)$. The logarithmic and constant terms are omitted. Hence the Clausen component ladders are

(5.182)
$$C(4, \theta) = \frac{\mathrm{Cl}_2(4\theta)}{4} - \left\{\frac{\mathrm{Cl}_2(3\theta)}{3} + \frac{\mathrm{Cl}_2(2\theta)}{2} + \mathrm{Cl}_2(\theta)\right\},$$

(5.183)
$$C(6, \theta) = \frac{\mathrm{Cl}_2(6\theta)}{6} - \left\{\frac{3\,\mathrm{Cl}_2(2\theta)}{2} - 4\,\mathrm{Cl}_2(\theta)\right\},$$

(5.184)
$$C(8, \theta) = \frac{\mathrm{Cl}_2(8\theta)}{8} - \left\{\frac{3\,\mathrm{Cl}_2(4\theta)}{4} - \mathrm{Cl}_2(2\theta) + 2\,\mathrm{Cl}_2(\theta)\right\},$$

(5.185)
$$C(10, \theta) = \frac{\mathrm{Cl}_2(10\theta)}{10} - \left\{\frac{\mathrm{Cl}_2(5\theta)}{5} + \frac{\mathrm{Cl}_2(3\theta)}{3} + \frac{\mathrm{Cl}_2(2\theta)}{2} - 2\,\mathrm{Cl}_2(\theta)\right\},$$

(5.186)
$$C(18, \theta) = \frac{\mathrm{Cl}_2(18\theta)}{18} - \left\{\frac{\mathrm{Cl}_2(9\theta)}{9} + \frac{\mathrm{Cl}_2(6\theta)}{2} - \frac{2\,\mathrm{Cl}_2(3\theta)}{3} - \mathrm{Cl}_2(2\theta) + \mathrm{Cl}_2(\theta)\right\}.$$

Numerical evaluation gives, to six significant figures,

(5.187) $$C(4, \theta) = 0,$$
(5.188) $$C(6, \theta) = 3.6052604,$$
(5.189) $$C(8, \theta) = -2.7039490,$$
(5.190) $$C(10, \theta) = 2.1631585,$$
(5.191) $$C(18, \theta) = -1.2017620.$$

Following the structures of (5.133) to (5.135) it is now readily verified that

(5.192) $$C(8, \theta) + (3/4)C(6, \theta) = 0,$$
(5.193) $$C(10, \theta) + (4/5)C(8, \theta) = 0,$$
(5.194) $$C(18, \theta) + (1/3)C(6, \theta) = 0.$$

These formulas support the contentions of §5.8.1, so far as they go. Further support comes from an independent analysis (Browkin, Chapter 11), in which (5.187) and (5.192) to (5.194) were all found numerically; and a still further result, overlooked in the earlier derivation, comes from the additional cyclotomic equation

(5.195) $$1 - u^{16} = (1 - u^4)^6 (1 - u^2)^{-5} u.$$

This gives rise to the Clausen component-ladder

(5.196) $$16C(16, \theta) = \mathrm{Cl}_2(16\theta) - 24\,\mathrm{Cl}_2(4\theta) + 40\,\mathrm{Cl}_2(2\theta) = -2D_F.$$

Here, $D_F$ is a constant generated by the analysis, and equal numerically to $21.63158\ldots$, and to which all the other Clausen component-ladders are similarly related. Thus $6C(6, \theta) = D_F$ and elimination of $D_F$ gives the additional result of Browkin's

$$C(16, \theta) + (3/4)C(6, \theta) = 0. \tag{5.197}$$

The new cyclotomic equation (5.195) gives rise to a corresponding dilogarithmic component-ladder

$$L_2(16, \Omega) = \frac{\mathrm{Li}_2(\Omega^{16})}{16} - \{\tfrac{3}{2}\,\mathrm{Li}_2(\Omega^4) - \tfrac{5}{2}\,\mathrm{Li}_2(\Omega^2) - \tfrac{1}{2}\log^2\Omega\}. \tag{5.198}$$

Reversing the process predicts that $L_2(16, \Omega) + (3/4)L_2(6, \Omega)$ should be a simple rational multiple of $\zeta(2)$, and numerical computation gives directly

$$L_2(16, \Omega) + (3/4)L_2(6, \Omega) = (25/16)\zeta(2), \tag{5.199}$$

or, from (5.133),

$$L_2(16, \Omega) - L_2(8, \Omega) = (3/2)\zeta(2). \tag{5.200}$$

5.8.4. *The constant term.* In generating (5.133) to (5.135) both the coefficients of $\zeta(2)$ *and* the combining rationals had to be determined *together* by a numerical search. This is in contrast to (5.192) and (5.194) in which only the combinations of the component-ladders are needed. If the studies had been done in reverse order the much easier task of first finding the right combinations would have preceded the consequential unambiguous determination of the $\zeta(2)$ coefficients. This was the process used in generating (5.199), in which only the coefficient $25/16$ needed to be determined directly. Thus, in some sense, the Clausen-function ladders seem to have a sort of precedence over the dilogarithmic results.

This conclusion is supported by Browkin's discoveries related in Chapter 11. Although, there, certain combinations of component-ladders can, in some cases, be deduced from the functional equations, the *component ladders themselves* also possess "valid" results; but the meaning of "valid" has to be seen in relation to the more intricate form of the constant term $D_F$ consequential on the application of Lichtenbaum's conjecture. This raises, but does not answer, the question of whether *all* component-ladders, whether Clausen or dilogarithmic, may possess a valid result, with a suitable development of the corresponding form of the constant term. This aspect has been pursued with great success by D. Zagier.

5.8.5. *Dilogarithmic ladders from Clausen ladders.* Browkin (see Chapter 11) has examined Clausen ladders of angle $\theta = \arccos(1 - \sqrt{2})/2$. With $u = e^{i\theta}$ he finds seven cyclotomic equations for $u$, giving rise to seven corresponding Clausen component-ladders. Each of these is equal to a simple multiple of $D_F = 4.116964\ldots$ appropriate to this field. The elimination of $D_F$ between the seven component-ladders then gives six independent valid Clausen ladders.



The base equation, when written in rational form, also possesses a real root $\Delta$ in $(0, 1)$ and dilogarithmic component-ladders corresponding directly to the Clausen component-ladders can be constructed. They can then be combined in exactly the same proportions as in the six independent valid Clausen ladders above, and equated numerically to a multiple of $\zeta(2)$. If rational multiples are produced by this process then not only are the dilogarithmic ladders thus produced valid, but the conjectured close linkage between the two sets of ladders is confirmed.

With $\cos\theta = \frac{1}{2}(1 - \sqrt{2})$ the equation for $u = e^{i\theta}$ can be put in the form

$$\sqrt{2} = 1 - (u + 1/u). \tag{5.201}$$

The rational form of the base equation comes from squaring this, so the additional roots come from (5.201) by reversing the sign of $\sqrt{2}$. The root in $(0, 1)$ is thus found to be

$$\Delta = \frac{\sqrt{2} + 1 - \sqrt{2\sqrt{2} - 1}}{2}. \tag{5.202}$$

The seven cyclotomic equations and the corresponding dilogarithmic component-ladders are:

$$\begin{gathered} 1 - u^4 = (1 - u^2)(1 - u)^{-2}u^2, \\ 4L_2(4, \Delta) = \mathrm{Li}_2(\Delta^4) - [2\,\mathrm{Li}_2(\Delta^2) - 8\,\mathrm{Li}_2(\Delta) - 4\log^2\Delta]; \end{gathered} \tag{5.203}$$

$$\begin{gathered} 1 - u^6 = (1 - u^3)^2(1 - u^2)u^{-1}, \\ 6L_2(6, \Delta) = \mathrm{Li}_2(\Delta^6) - [4\,\mathrm{Li}_2(\Delta^3) + 3\,\mathrm{Li}_2(\Delta^2) + 3\log^2\Delta]; \end{gathered} \tag{5.204}$$

$$\begin{gathered} 1 - u^7 = (1 - u^4)^3(1 - u^2)^{-3}(1 - u), \\ 7L_2(7, \Delta) = \mathrm{Li}_2(\Delta^7) - \tfrac{1}{4}[21\,\mathrm{Li}_2(\Delta^4) - 42\,\mathrm{Li}_2(\Delta^2) + 28\,\mathrm{Li}_2(\Delta)]; \end{gathered} \tag{5.205}$$

$$\begin{gathered} 1 - u^{10} = (1 - u^5)(1 - u^2)(1 - u)^{-3}u^3, \\ 10L_2(10, \Delta) = \mathrm{Li}_2(\Delta^{10}) - [2\,\mathrm{Li}_2(\Delta^5) + 5\,\mathrm{Li}_2(\Delta^2) - 30\,\mathrm{Li}_2(\Delta) - 15\log^2\Delta]; \end{gathered} \tag{5.206}$$

$$\begin{gathered} 1 - u^{12} = (1 - u^4)(1 - u^3)^5(1 - u)^3u^{-5}, \\ 12L_2(12, \Delta) = \mathrm{Li}_2(\Delta^{12}) - [3\,\mathrm{Li}_2(\Delta^4) + 20\,\mathrm{Li}_2(\Delta^3) + 36\,\mathrm{Li}_2(\Delta) + 30\log^2\Delta]; \end{gathered} \tag{5.207}$$

$$\begin{gathered} 1 - u^{14} = (1 - u^7)(1 - u^2)^3(1 - u)^{-3}u^2, \\ 14L_2(14, \Delta) = \mathrm{Li}_2(\Delta^{14}) - [2\,\mathrm{Li}_2(\Delta^7) + 21\,\mathrm{Li}_2(\Delta^2) - 42\,\mathrm{Li}_2(\Delta) - 14\log^2\Delta]; \end{gathered} \tag{5.208}$$

$$\begin{gathered} 1 - u^{15} = (1 - u^5)^2(1 - u^3)(1 - u)^{-2}u^2, \\ 15L_2(15, \Delta) = \mathrm{Li}_2(\Delta^{15}) - [6\,\mathrm{Li}_2(\Delta^5) + 5\,\mathrm{Li}_2(\Delta^3) - 30\,\mathrm{Li}_2(\Delta) - 15\log^2\Delta]. \end{gathered} \tag{5.209}$$

The corresponding Clausen component-ladders can be written down by replacing $\mathrm{Li}_2(\Delta^n)$ by $\mathrm{Cl}_2(n\theta)$ and omitting the logarithmic term. They are found to be equal, respectively, to $D_F$ multiplied by $2, 1, -7/2, 6, -3, 10$ and $7$. The coefficients to eliminate $D_F$ and create valid Clausen ladders are therefore displayed directly, and if these *same* coefficients are used to create dilogarithmic ladders they can be checked numerically for validity by equating them to a rational multiple of $\zeta(2)$. In this way the following valid dilogarithmic ladders were obtained:

(5.210) $$L_2(4, \Delta) - 3L_2(6, \Delta) = \tfrac{7}{4}\zeta(2),$$

(5.211) $$L_2(7, \Delta) + 3L_2(6, \Delta) = -\tfrac{27}{28}\zeta(2),$$

(5.212) $$5L_2(10, \Delta) - 18L_2(6, \Delta) = 12\zeta(2),$$

(5.213) $$2L_2(12, \Delta) + 3L_2(6, \Delta) = -\tfrac{14}{3}\zeta(2),$$

(5.214) $$7L_2(14, \Delta) - 30L_2(6, \Delta) = 16\zeta(2),$$

(5.215) $$5L_2(15, \Delta) - 14L_2(6, \Delta) = \tfrac{26}{3}\zeta(2).$$

Ray (§7.7.5) has now proven these results analytically, together with two further formulas involving indices 18 and 24.

5.8.6. *Clausen ladders from dilogarithmic ladders.* The quantity $\Gamma$ of §5.7.3 satisfies (5.137), but if $u$ is replaced by $e^{i\phi}$ then it is found that $C = \cos\phi$ satisfies

(5.216) $$8C^3 - 4C^2 - 6C + 1 = 0.$$

There are two values of $C$ in $(-1, 1)$ giving real angles $\phi, \phi'$ where

(5.127) $$\phi \approx 81.05^\circ; \qquad \phi' \approx 137.78^\circ.$$

Corresponding to the results of §5.7.3, and particularly equations (5.157) to (5.163), we can define Clausen component-ladders for these two angles, and corresponding constants. It turns out that *two*, apparently distinct, constants $D_F$ and $\overline{D}_F$ are needed for the $\phi$-ladders, and two further constants $D'_F$ and $\overline{D}'_F$ for $\phi'$. The structure of the Clausen component-ladders follows directly from equations (5.148) to (5.156) and one finds, by numerical computation

(5.218) $$\mathrm{Cl}_2(5\phi) + 5\,\mathrm{Cl}_2(\phi) = D_F = 5.80132\ldots,$$

(5.219) $$\mathrm{Cl}_2(6\phi) - \tfrac{3}{2}\mathrm{Cl}_2(4\phi) - 4\,\mathrm{Cl}_2(3\phi) = \overline{D}_F = 4.82312\ldots,$$

(5.220) $$\mathrm{Cl}_2(8\phi) - 4\,\mathrm{Cl}_2(2\phi) + 16\,\mathrm{Cl}_2(\phi) = 4D_F - 2\overline{D}_F,$$

(5.221) $$\mathrm{Cl}_2(9\phi) - \tfrac{9}{2}\mathrm{Cl}_2(4\phi) - 3\,\mathrm{Cl}_2(3\phi) + 9\,\mathrm{Cl}_2(2\phi) = 3\overline{D}_F - D_F,$$

(5.222)
$$\mathrm{Cl}_2(10\phi) - 2\,\mathrm{Cl}_2(5\phi) - \tfrac{5}{2}\mathrm{Cl}_2(4\phi) - 10\,\mathrm{Cl}_2(2\phi) + 10\,\mathrm{Cl}_2(\phi) = 4D_F - 3\overline{D}_F,$$



(5.223) $$\mathrm{Cl}_2(12\phi) - 4\,\mathrm{Cl}_2(6\phi) + 6\,\mathrm{Cl}_2(2\phi) = 2\overline{D}_F - 2D_F,$$

(5.224) $$\mathrm{Cl}_2(14\phi) - 2\,\mathrm{Cl}_2(7\phi) - \tfrac{7}{2}\mathrm{Cl}_2(4\phi) + 14\,\mathrm{Cl}_2(\phi) = 4D_F - \overline{D}_F,$$

(5.225)
$$\mathrm{Cl}_2(18\phi) - 2\,\mathrm{Cl}_2(9\phi) - 6\,\mathrm{Cl}_2(3\phi) - 9\,\mathrm{Cl}_2(2\phi) + 36\,\mathrm{Cl}_2(\phi) = 8D_F - 2\overline{D}_F,$$

(5.226)
$$\mathrm{Cl}_2(30\phi) - 2\,\mathrm{Cl}_2(15\phi) - 3\,\mathrm{Cl}_2(10\phi) - 5\,\mathrm{Cl}_2(6\phi) + 10\,\mathrm{Cl}_2(3\phi) + 15\,\mathrm{Cl}_2(2\phi) = -2D_F.$$

The structures for the angle $\phi'$ are identical to the above, but $D_F$ and $\overline{D}_F$ are replaced by

(5.227) $$D'_F = 1.59265\ldots; \qquad \overline{D}'_F = 0.79934\ldots.$$

No rational relations between these four constants has currently come to light.

5.8.7. *A Clausen ladder in the field of $\omega$.* Although not a supernumary result, the following analysis has a close affinity with the preceding results, and is therefore reported here. The quantity $\omega$ of §4.2 is the solution in $(0, 1)$ of $z^3 + z^2 - 1 = 0$. This cubic factorizes into $(z - \omega)(z^2 + z/\omega^2 + 1/\omega)$ so the complex roots are

(5.228) $$z = [-1 \pm (1 - 4\omega^3)^{1/2}]/2\omega^2 = re^{\pm i\theta},$$

with $r = \omega^{-1/2}$, $\cos\theta = -1/2\omega^{3/2}$. From this one obtains

(5.229) $$\psi = 2\theta = \cos^{-1}(\omega^4/2).$$

According to (2.34), $\mathrm{Im}\,\mathrm{Li}_2(z^m)$ involves both $\arg(z^m) = m\theta$, and also $\arg(1 - z^m)$. The latter can be put directly in terms of $\theta$, from the relationships satisfied by the base, when $m = 1, 2, 3, 5$, and 7. Also, the form $\arg(1 + z^m)$ can likewise be so expressed when $m = 1, 4$, and 7. Hence the valid ladder structures with indices $N = 2, 3, 5, 8$, and 14 of (4.23) to (4.28) can all be expressed directly in terms of $\theta$ when $\omega$ is replaced by $z$ of (5.228) and the imaginary parts taken.

Rather surprisingly, *they all yield a null result* except for the ladder of index 2. Defining

(5.230) $$C(3, \psi) = \mathrm{Cl}_2(3\psi) - \mathrm{Cl}_2(2\psi) - \mathrm{Cl}_2(\psi),$$

(5.231) $$C(5, \psi) = \mathrm{Cl}_2(5\psi) - \mathrm{Cl}_2(4\psi) - \mathrm{Cl}_2(\psi),$$

the resulting relation, via (4.23), is

(5.232) $$2C(5, \psi) + C(3, \psi) = 0.$$

The value for the component-ladder $C(3, \psi)$ alone is $-1.88541100\ldots$, but it has no known analytic construction at this time.

### REFERENCES

1. M. Abouzahra and L. Lewin, *The polylogarithm in the field of two irreducible quintics*, Aequationes Math. **31** (1986), 315–321.
2. L. Lewin, *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), 345–373.
3. L. Lewin, *Further results on supernumary polylogarithmic ladders*, Max-Planck-Institut Report, 1991.

CHAPTER 6

# Functional Equations and Ladders

L. LEWIN

**6.1. New categories of functional equations.**

6.1.1. *The ladder structure and functional equations.* The ladders examined in the previous chapters fall broadly into two groups—those directly accessible from Kummer's equations, and those not. For $n > 5$ this categorization is not in doubt, since no equations of Kummer's type exist in this region. For $n \leq 5$, and particularly for $n = 2$, the grouping must be considered an interim one, since there is always a possibility that a procedure may be found for deduction of a ladder from the functional equation. For those ladders that are indeed obtainable in this way, the ladder structure can be seen as a diluted version of the functional equation, from which some, but not all, of the structure has been lost. One can ask whether this process can be reversed: to produce a functional equation to explain a particular ladder. The structure need not correspond to a recognizable variant of Kummer's form, and if it does not, the separate question of its accessibility therefrom is raised. This seems relevant to the single-variable forms deduced later in this chapter. One can also ask whether these single-variable forms are special cases of two-variable equations, a matter that has not been answered at this time; though working with them evokes the possibility that underlying them there may exist a two-variable equation more intricate in form than Kummer's.

6.1.2. *The inaccessible $\rho$- and $\omega$-ladders.* Although, as discussed in §4.1.5, Coxeter's index-20 ladder has now been deduced from Kummer's functional equation at $n = 2$, the earlier studies had considered this ladder inaccessible, as also the index-24 $\rho$-ladder. Moreover, the extension of the $\rho$-ladder complex into the trans-Kummer region, $n > 5$, required index-24 and -20 ladders at least as far as $n = 5$. If there was to be any hope of finding functional equations at $n = 6$, then, at the very least, these two ladders would have to be produced analytically to $n = 5$.

The situation with the base $\omega$ is a little more involved. Although the indices $N = 14$, 20, and 28 are not part of the generic set, the corresponding ladders at $n = 2$ *are* accessible from Kummer's equation, and various combinations of them, together with the generic set, are obtainable to $n = 5$. In fact, at $n = 5$, *three* ladder combinations are accessible, and between them pseudo-integrate to give one valid ladder at $n = 6$. There is, however, a presumed-inaccessible ladder of index 42, and numerical computations have produced four ladder-combinations at $n = 5$ that do not seem to be accessible from the functional equation. All of these need to be produced analytically from functional equations of some sort before the existence of such equations in the trans-Kummer region can be addressed.

The fact that both the $\rho$- and $\omega$-ladders can be pseudo-integrated to $n = 9$ suggests some common features. The most obvious come from the forms taken by the base equations, which can both be written

$$(6.1)\qquad 1 - u = u^s, \qquad 1 + u = u^{-t},$$

with $(s, t) = (2, 1)$ for $\rho$, and $(5, 2)$ for $\omega$. This suggests arguments of the form $\pm z^p(1 - z)^q(1 + z)^r$, which give simple powers of $\rho$ or $\omega$ when $z$ is given values $\pm\rho$, $\pm 1/\rho$, or $\pm\omega$, $\pm 1/\omega$; and this structure is pursued in the ensuing sections. No other useful forms of any sort have so far been located, though Gangl's 1990 single-variable equation at $n = 6$ gives results for $\omega$ only.

6.1.3. *Arguments to give an index-20 $\rho$-ladder.* An argument $z^{20}$ can be trivially obtained from an argument $-z^{10}$ via the duplication formula. However, when the dilogarithm is differentiated it leads to $\log(1 + z^{10})$, and the factorization of this form does not lead to any useful formulas. What is needed is a structure of the form $\pm z^p(1 - z)^q(1 + z)^r$ which can give both an $N = 20$ $\rho$-ladder, and also a permissible $\omega$-ladder, with index taken from (4.20). This is quite restrictive, and the only form so far discovered leads to a consideration of the structure, called the "head" term,

$$(6.2)\qquad y = \mathrm{Li}_2\left[-z^7\left(\frac{1-z}{1+z}\right)\right].$$

On differentiation one gets

$$(6.3)\qquad \frac{dy}{dz} = \left[\frac{7}{z} - \frac{1}{1-z} - \frac{1}{1+z}\right][\log(1 + z) - \log(1 + z + z^7 - z^8)].$$

The entire success depends now on the possibility of usefully factorizing $1 + z + z^7 - z^8$. "Useful" here implies the existence of factors which, when later integrated, give ladder-type arguments. Now it is readily verified that

$$(6.4)\qquad 1 + z + z^7 - z^8 = (1 + z^2)(1 + z - z^3)(1 - z^2 + z^3),$$

and this result is the key to all the subsequent equations. We can write $1 + z - z^3$ as either $(1 + z)[1 - z^3/(1 + z)]$ or as $1 + z(1 - z)(1 + z)$ while $1 - z^2 + z^3$ can be similarly expressed as $(1 - z)(1 + z)[1 + z^3/(1 - z)(1 + z)]$ or as $1 - z^2(1 - z)$. If dilogarithms of corresponding arguments are differentiated and inserted into (6.3) modified by (6.4), then it is found that all the various terms can be collected together so as to be integrable. There are, in fact, more terms to be treated than there are free coefficients to handle them—*this feature occurs throughout*—and the process works only because the terms occur in just the right combinations. Why this is so is unknown, but this fortunate feature enables the integrations to proceed, leading to

(6.5)
$$\begin{aligned}\mathrm{Li}_2&\left[-z^7\frac{1-z}{1+z}\right]\\ &= 2\mathrm{Li}_2[z^2(1-z)] + \mathrm{Li}_2\left[\frac{-z^3}{1-z^2}\right] + 2\mathrm{Li}_2\left[\frac{z^3}{1+z}\right] + \mathrm{Li}_2[-z(1-z^2)]\\ &\quad + \frac{7}{4}\mathrm{Li}_2(z^4) - \frac{9}{4}\mathrm{Li}_2(z^2) + \frac{1}{2}\mathrm{Li}_2\left[z\frac{1-z}{1+z}\right] - \frac{1}{2}\mathrm{Li}_2\left[-z\frac{1+z}{1-z}\right]\\ &\quad + \frac{5}{4}\log^2(1+z) + \frac{1}{4}\log^2(1-z) + \frac{3}{2}\log(1+z)\log(1-z).\end{aligned}$$

As will be shown in §6.2 this equation gives rise to what can be called "the $\rho$-family" of equations, integrable to the fifth order, but no further. The formula (6.5) with $z = \rho$ leads to the $N = 20$ ladder first discovered by Coxeter. For $z = \omega$ it gives the already-known $\omega$-ladder of index 28; but otherwise it leads to nothing new.

6.1.4. *Arguments to give an index-42 $\omega$-ladder.* High powers in a dilogarithmic argument give rise, on differentiation, to high-degree polynomials, whose factorizing, in general, does not lead to useful formulas. Most of the accessible ladders of high index arise from terms of half the index and negative sign, via the duplication formula. Thus, as with the construction of (6.2), one can begin by halving the index. Moreover, many of the accessible ladders involve cubes, a clue to be followed where possible. Thus the index 42 for $\omega$ would be reduced to a consideration of $\omega^7$. In a like manner, the index-24 $\rho$-ladder reduces to a consideration of $\rho^4$. To reduce these powers further, note that $1 - \omega = \omega^5$ and $1 - \rho = \rho^2$, so the required powers can be presented as $\omega^2(1 - \omega)$ and $\rho^2(1 - \rho)$. A structure common to both is therefore $z^2(1 - z)$, suggesting the form

$$(6.6)\qquad y = \mathrm{Li}_2[-z^6(1 - z)^3],$$

as an argument capable of generating both the $N = 42$ $\omega$-ladder and the $N = 24$ $\rho$-ladder.

The above form is pursued in §6.3, where it gives rise to what can be called "the $\omega$-family" of equations, integrating as far as the fifth order, but no further.

6.1.5. *Arguments to give an index-24 $\theta$-ladder.* Using the squaring and cubing processes of the previous section reduces the $\theta$-form to $\theta^4$, and since $\theta^3 = 1 - \theta$, this can be written $\theta(1 - \theta)$. This suggests the form

$$(6.7)\qquad y = \mathrm{Li}_2[-z^3(1 - z)^3],$$

which is pursued in §6.4, and give rise to "the $\theta$-family" of equations. These have been integrated as far as the third order but, so far, no further.

**6.2. The $\rho$-family of equations.**

6.2.1. *Integration to the third order.* If $\mathrm{Li}_3[-z^7(\frac{1-z}{1+z})]$ is differentiated, it gives $[\frac{7}{z} - \frac{1}{1-z} - \frac{1}{1+z}]\mathrm{Li}_2[-z^7(\frac{1-z}{1+z})]$. The $\mathrm{Li}_2$ term is eliminated via (6.5) and the groups of terms headed by $1/z$, $1/(1-z)$, and $1/(1+z)$ are segregated; they clearly are not individually integrable. The technique, which subsequently became identifiable as an algorithm for use with the computer program MACSYMA, is to concentrate on the group headed by $1/z$, to find or create $\mathrm{Li}_3$ functional equations carrying exactly the same arguments, to differentiate them, and to use those new terms, headed by $1/z$, to attempt to cancel out all arguments other than those that are a simple power of $z$ (these are directly integrable). If and when this has been done the remnant of terms headed by $1/(1-z)$ and $1/(1+z)$ are collected, equations in $\mathrm{Li}_2$ are utilized to attempt to remove all terms with arguments involving powers of $z$ other than zero, and if this can be done, the remaining terms are integrated. As before, there are more terms to be cancelled than there are free coefficients, but they occur in just the right combinations to permit the process to work.

The $\mathrm{Li}_2$ and $\mathrm{Li}_3$ equations utilized in the above process all come from Kummer's formulas by taking $x = \pm z$ and $y = z$ or $z^2$, or such other combinations of these as correspond to the terms in the harmonic group (2.2). There are very many such formulas, but only a specific few of them carry the requisite arguments, so the transition from $n = 2$ to $n = 3$ was not too difficult to achieve, once the correct *modus operandi* had been found. The resulting equation was found to be

(6.8)

$$\begin{aligned}
&\mathrm{Li}_3\left[-z^7\left(\frac{1-z}{1+z}\right)\right] \\
&\quad - 7\left\{\mathrm{Li}_3[z^2(1-z)] + \frac{1}{3}\mathrm{Li}_3\left[\frac{-z^3}{1-z^2}\right] + \frac{2}{3}\mathrm{Li}_3\left[\frac{z^3}{1+z}\right]\right. \\
&\qquad + \mathrm{Li}_3[-z(1-z^2)] + \frac{7}{16}\mathrm{Li}_3(z^4) - \frac{9}{8}\mathrm{Li}_3(z^2) \\
&\qquad\left. + \frac{1}{2}\mathrm{Li}_3\left[z\left(\frac{1-z}{1+z}\right)\right] - \frac{1}{2}\mathrm{Li}_3\left[-z\left(\frac{1+z}{1-z}\right)\right]\right\} \\
&\quad + 2\mathrm{Li}_3\left[\frac{-z^2}{(1+z)^2(1-z)}\right] + 2\mathrm{Li}_3\left[\frac{-z}{(1-z)^2(1+z)}\right] \\
&\quad + 6\mathrm{Li}_3(z) - 6\mathrm{Li}_3\left(\frac{1}{1+z}\right) \\
&\quad + 4\mathrm{Li}_3[-z(1-z^2)] - \frac{2}{3}\mathrm{Li}_3\left[\frac{-z^3}{1-z^2}\right] - \frac{4}{3}\mathrm{Li}_3\left[\frac{z^3}{1+z}\right] + 6\mathrm{Li}_3(1-z)
\end{aligned}$$



$$\begin{aligned}
&+ \mathrm{Li}_3[z^2(1-z)] - \frac{1}{16}\mathrm{Li}_3\left[\frac{(1-z)^4}{(1+z)^4}\right] + \frac{3}{8}\mathrm{Li}_3\left[\frac{(1-z)^2}{(1+z)^2}\right] \\
&= \frac{5}{16}\zeta(3) + \frac{1}{2}\zeta(2)[11\log(1+z) + 13\log(1-z)] \\
&\quad + \frac{1}{4}\log(z)[11\log^2(1+z) + 2\log(1+z)\log(1-z) - 13\log^2(1-z)] \\
&\quad + \frac{1}{12}[37\log^3(1-z) + 33\log^2(1-z)\log(1+z) \\
&\qquad\qquad + 75\log(1-z)\log^2(1+z) - \log^3(1+z)].
\end{aligned}$$

This is about as far as the method can be readily taken without a more systematic approach. The difficulty now becomes the *very* large number of terms to be handled, and it was at this stage that resort was made to the MACSYMA program [1] for symbolic computation. This required, among other things, for the *ad hoc* integration process to be made into a precise algorithm appropriate for the computer.

6.2.2. *Preparation for MACSYMA.* By straightforward differentiation, it is found that

(6.9)

$$\frac{d}{dz}\mathrm{Li}_n[\pm z^p(1-z)^q(1+z)^r] = \left[\frac{p}{z} - \frac{q}{1-z} + \frac{r}{1+z}\right]\mathrm{Li}_{n-1}[\pm z^p(1-z)^q(1+z)^r].$$

Although it is possible to program a direct translation of (6.9), it is not very convenient since MACSYMA would perform an unwanted "simplification" of the first factor on the right. Moreover, once the differentiation has been achieved, the role of $z$ as a variable is reduced to that of being a mere carrier of the powers $p$, $q$, and $r$. Since these powers are the dominant feature of the calculation, a change of notation is warranted which largely "algebraices" the problem.

6.2.2.1. *The $(pqrs)$ notation.* In (6.9), the order $n$ plays no particular role and it was soon discovered that in any given procedure it was not necessary to specify it. It could always be recovered from the context when needed.

Because of the duplication formulas, which connect variables $v$, $-v$, and $v^2$, the minus signs in the arguments of (6.9) can be converted into plus signs. However, it is not convenient to do this in every case, and some arguments with negative signs are retained. In order to specify the argument unambiguously, we introduce a further parameter $s$ which takes on the values $+1$ or $-1$ according to the sign of the argument. The ensuing simplified notation is then given by the definition

$$(6.10) \qquad M(p, q, r, s) \equiv \mathrm{Li}_n[\pm z^p(1-z)^q(1+z)^r]; \qquad s = \pm 1.$$

The duplication formulas can then be used to eliminate undesired linear relations between $M$'s. For the $n$th order this formula can be put into the form

$$M(p,q,r,-1)=2^{1-n}M(2p,2q,2r,1)-M(p,q,r,1). \tag{6.11}$$

A differentiation operator $D$ is defined so that (6.9) can be written as

$$DM(p,q,r,s)=[p,-q,r]M(p,q,r,s). \tag{6.12}$$

In this context $[p,-q,r]$ implies a set of three segregated terms in which the coefficients of a term are multiplied by $p$, $-q$, and $r$ respectively. (Factors $1/z$, $1/(1-z)$, and $1/(1+z)$ are implied but their presence is not important in the ensuing calculations.)

6.2.2.2. *Rogers' function* $L_n$. One of the very substantial chores in the procedure is keeping track of and integrating the subsidiary logarithmic terms. The use of Rogers' function would eliminate this, but unfortunately it does not satisfy a recurrence formula like (1.28) on which (6.9) and (6.12) are based. The bookkeeping chore increases rapidly in magnitude as the order increases, and since the main concern is to establish the nature of the transcendental terms in any presumptive equation, the procedure adopted was to work with $\mathrm{Li}_n$, to ignore all subsidiary logarithms and constants, and then to interpret a resulting functional equation as applying to $L_n$. Only an integration constant, easily found by putting $z=0$, is then needed. The use of (3.19) will then, in principle, produce the needed equation of $\mathrm{Li}_n$, including all the subsidiary logarithms. The subsequent steps to be discussed are thereby significantly simplified.

6.2.2.3. *Integration.* There is no corresponding formula to (6.12) for integration. An integration by parts, operating on the $z^p$ part of the argument, is

$$\begin{aligned}\int_0^z &\mathrm{Li}_{n-1}[\pm z^p(1-z)^q(1+z)^r]\frac{dz}{z}\\ &=\frac{1}{p}\mathrm{Li}_n[\pm z^p(1-z)^q(1+z)^r]\\ &\quad-\frac{1}{p}\int_0^z\left[\frac{-q}{1-z}+\frac{r}{1+z}\right]\mathrm{Li}_{n-1}[\pm z^p(1-z)^q(1+z)^r]\,dz.\end{aligned} \tag{6.13}$$

This can be expressed in the form

$$IM(p,q,r,s)=\frac{1}{p}M(p,q,r,s)-I[0,-q/p,r/p]M(p,q,r,s), \tag{6.14}$$

where $I$ is an integration operator. In order for the integration in (6.13) to be possible, $p$ must be positive, so any arguments with negative $p$ are first replaced by ones with positive $p$ by using the inversion formula.

The procedure outlined above also requires that $p$ should not be zero. Such terms *must* first be cancelled out by utilizing any suitable pre-existing functional equations in appropriate linear combinations. To date this has always been found possible to achieve, *even at the transition to the sixth order*, by utilizing various forms of Kummer's equations. As elsewhere, there have been more terms to handle than free coefficients, so that the necessary equations are overdetermined. But relations among the coefficients of these terms have been such that the equations in fact do have a valid solution.



The effect of the process is to produce a starting sequence of terms all of whose arguments vanish with $z=0$; there is therefore no constant of integration to be determined since the expression vanishes with $z$. Now, one of the requirements encountered when extrapolating a ladder to higher order is that ladders at odd order $(2m+1)$ should first be combined to eliminate any $\zeta(2m+1)$ term. This is paralleled by the preceding procedure which does so, however, both for odd and for even orders. (The ladder extrapolation process makes no such requirements when the order is even.) The prescribed process clearly obviates any need to take explicit account of integration constants.

The production of the term $(1/p)M(p,q,r,s)$, by integration by parts, as on the right side of (6.14), is an operation of which we make frequent use—we shall refer to this procedure as "partial-integration."

6.2.3. *An algorithm for the transition from order* $n$ *to* $n+1$.

6.2.3.1. List all possible single-variable equations of order $n$ obtainable from Kummer's equations using the variables of §6.2.1. These seem to be the only possible ones giving the $(pqrs)$ form except for $x=y=z^2$, but this latter form introduces several arguments not heretofore encountered and thus does not seem to be a useful variant.

6.2.3.2. Remove all redundant equations from the list of §6.2.3.1, expressing the final result in terms of linearly independent equations. In practice there are many redundant equations and some can be excluded easily. At the fourth order, for example, Kummer's equation is symmetrical in $x$ and $y$, so interchange can give no new results. However, there are also linear combinations of several equations which are difficult to detect "by hand." In general we must consider possible linear relationships

$$\sum_i C_i\sum_j a_{ij}M(p_j,q_j,r_j,s_j)=0, \tag{6.15}$$

with constants $C_i$ where $i$ denotes one of the Kummer equations and the $j$-summation gives the generation of terms in that equation. A program using MACSYMA was written to analyze (6.15) and set up the appropriate linear system in (many) equations with (somewhat fewer) unknowns. This task required picking out and taking apart the terms in the several equations and used the built-in LINSOLVE facility to provide the possible nontrivial solutions $C_i$. The residual set of independent equations is called the *basic-set* at order $n$.

6.2.3.3. From the above-deduced basic-set, determine all independent linear combinations which are free of $p=0$ arguments. These will be known as the *residual-set*. This procedure is computerized by a natural extension of the one used in §6.2.3.2.

6.2.3.4. Starting with the equation with the head term, known as the *head-equation*, e.g., (6.8) for $n = 3$, cancel out all terms with arguments having $p = 0$, using members of the basic-set. If this cannot be done the process stops. If it can be done, and if the residual-set of §6.2.3.3 is nonempty, then the cancellation is nonunique and some parameters are available and may be chosen to achieve some simplicity in the form of what we now call a *modified head-equation*.

6.2.3.5. Partial-integrate the modified head-equation and all members of the residual-set, and combine them with as yet undetermined coefficients, called the *primary-set*. This combination of terms, when the coefficients in the primary-set have been determined, will constitute the bulk of the new equation at order $(n+1)$, and will be called the *bulk-expression*. It is possible to create it in this way because all terms with $p = 0$ have been removed, and terms with negative $p$ have been expressed with arguments with positive $p$ by the inversion relation.

6.2.3.6. Differentiate the bulk-expression and segregate the groups of terms headed by $1/z$, $-1(1-z)$, and $1/(1+z)$, to be known as the $P$-, $Q$- and $R$-terms, respectively. By construction, the $P$-term is identically zero.

6.2.3.7. To the $Q$-term add linear combinations of the basic-set and the head-equation with as yet undetermined coefficients, called the *secondary-set*. Do the same for the $R$-term, with a different secondary set of coefficients. These now constitute the *modified $Q$- and $R$-terms.*

6.2.3.8. Extract from the modified $Q$- and $R$-terms the complete set of net coefficients of all terms for which the arguments have $p \neq 0$ and equate all such coefficients to zero so as to generate a set of equations for the determination of the primary and two secondary sets of coefficients. Up to the fifth order it was found that this set of equations, though very highly overdetermined, had a valid solution including some available undetermined parameters.

6.2.3.9. Solve the equations of §6.2.3.8, put the values for the primary set into the bulk-expression, and put the values for all sets into the modified $Q$- and $R$-terms. These are now reduced to a set of terms all of whose arguments have $p = 0$, and are called the *remnant $Q$- and $R$-terms.*

6.2.3.10. Multiply the remnant $Q$- and $R$-terms by $-1/(1-z)$ and $1/(1+z)$ respectively and integrate, determining any outstanding coefficients to make this possible. For example, $[a/(1-z)+b/(1+z)]\mathrm{Li}_n(1-z^2)$ requires $a=-b$, whereupon it integrates to $b\mathrm{Li}_{n+1}(1-z^2)$; while $[c/(1-z)]\mathrm{Li}_n(1+z)$ is not integrable in the present context, and requires $c = 0$.

6.2.3.11. Subtract the integrated terms of §6.2.3.10 from the bulk-expression, and determine any needed integration constants by taking $z = 0$. The resulting expression, by construction, differentiates to zero, and therefore becomes the sought head-equation at the $(n+1)$th order. The method



of construction makes it possible for it to contain some arbitrary multiples of Kummer's equations at order $(n+1)$, and these can be dealt with in any convenient manner.

Although the preceding algorithm was developed for the $\rho$-family, it applies with very little change to the $\omega$- and $\theta$-families. The main difference is due to the existence of, not one head-term, but six coupled head-terms in each family. However, the details of this coupling have already been determined, and if this information is used, the algorithm should proceed as scheduled.

A minor change for the $\theta$-family is the presence of terms in $1+z^2$ rather than $1+z$. The $(pqrs)$ notation is retained unaltered; however, the selection for Kummer's equation is more limited, and comprises $x = z$, $y = z$, or $y = -z^2$, plus the harmonic group variations. The only other change is the relatively trivial one for the logarithmic differential, which is now $2z/(1+z^2)$, and with a corresponding impact on the integration of the remnant $Q$ and $R$ terms.

6.2.4. *Transition from the third to four order.*

6.2.4.1. *Modification of* (6.8). Some simplification is possible by combining a few terms, but the main alteration needed is to remove terms not vanishing with $z$, i.e., the terms with arguments in $[(1-z)/(1+z)]^4$, $[(1-z)/(1+z)]^2$, $(1-z)$, and $1/(1+z)$. The first two are handled via equations (6.110) of reference [2], the last two by taking $x = z$ in (6.10) and $x = -z$ in (6.11) of the same reference. The resulting equation is somewhat simpler, with 14 transcendental terms:

(6.16)

$$\begin{aligned}
&\mathrm{Li}_3\left[-z^7\frac{1-z}{1+z}\right] - 6\mathrm{Li}_3[z^2(1-z)] - 3\mathrm{Li}_3\left[\frac{-z^2}{1-z^2}\right] - 6\mathrm{Li}_3\left[\frac{z^3}{1+z}\right]\\
&\quad - 3\mathrm{Li}_3[-z(1-z^2)] - 3\mathrm{Li}_3(z^4) + 9\mathrm{Li}_3(z^2) - 4\mathrm{Li}_3\left[z\frac{1-z}{1+z}\right]\\
&\quad + 3\mathrm{Li}_3\left[-z\frac{1+z}{1-z}\right] + 6\mathrm{Li}_3\left[\frac{z}{1+z}\right] - 6\mathrm{Li}_3\left[\frac{-z}{1-z}\right] - 6\mathrm{Li}_3(z)\\
&\quad + 2\mathrm{Li}_3\left[\frac{-z^2}{(1+z)^2(1-z)}\right] + 2\mathrm{Li}_3\left[\frac{-z}{(1-z)^2(1+z)}\right]\\
&= 2[\log^3(1+z) + \log^3(1-z)]\\
&\quad + 3\log(1-z)\log(1+z)[\log(1+z) + 2\log(1-z)].
\end{aligned}$$

The $(pqrs)$ form of (6.16) is

$$\begin{aligned}
&M(7,1,-1,-1) - 6M(2,1,0,1) - 3M(3,-1,-1,-1)\\
&\quad - 6M(3,0,-1,1) - 3M(1,1,1,-1) - 3M(4,0,0,1)\\
&\quad + 9M(2,0,0,1) - 4M(1,1,-1,1) + 3M(1,-1,1,-1) \qquad (6.17)\\
&\quad + 6M(1,0,-1,1) - 6M(1,-1,0,-1) - 6M(1,0,0,1)\\
&\quad + 2M(2,-1,-2,-1) + 2M(1,-2,-1,-1) = 0,
\end{aligned}$$

where the $M(\ )$ may be taken to stand for Rogers' function $L_3$.

Equation (6.17) is the modified head-equation discussed in §6.2.3.4.

6.2.4.2. *The basic set of equations at* $n = 3$. These are obtainable from (6.107) of reference [2]. Three forms of this equation come by taking $y$, $1-y$, and $-(1-y)/y$ for $y$. In these, $y$ takes on the set $\{z, -z, z^2\}$ and $x$ takes on the set $\{z, 1-z, -z/(1-z), -z, 1+z, z/(1+z)\}$ giving a total of 54 functional equations, each with nine transcendental terms. The substitution and subsequent simplification were done entirely by the computer thereby avoiding transcription and algebraic errors. Of the 54 equations, 30 were found to form the basic set, the remainder being redundant. From these 30, 18 combinations were found that were free of $p = 0$ terms and which constitute the residual set.

6.2.4.3. *Fourth-order equation.* The head-equation (6.17) together with the residual set (in practice only a few of them are essential) are partial-integrated and the procedures of §§6.2.3.5 to 6.2.3.11 are followed. Apart from handling the large number of terms (for which the computer is very convenient), no special problems occur, and the resulting expression can be written in terms of Rogers' function $L_4$ with 28 transcendental terms:

(6.18)
$$\begin{aligned} L_4\left[-z^7\frac{1-z}{1+z}\right] &= L_4\left[\frac{-z^5}{1+z}\right] + 3L_4[-z^4(1-z)] - \frac{9}{8}L_4(z^6) + \frac{21}{4}L_4(z^4) \\ &- 3L_4\left[\frac{-z^3}{1-z^2}\right] + 6L_4\left[\frac{z^3}{1+z}\right] + 6L_4[z^2(1-z)] - 3L_4[-z(1-z^2)] \\ &+ 3L_4\left[\frac{-z^2}{1-z^2}\right] - L_4\left[\frac{-z^2}{(1+z)^2(1-z)}\right] + L_4\left[\frac{-z}{(1-z)^2(1+z)}\right] \\ &- \frac{99}{8}L_4(z^2) + 9L_4\left[\frac{-z^2}{(1-z)}\right] + 28L_4\left[\frac{z(1-z)}{(1+z)}\right] - 21L_4\left[\frac{-z(1+z)}{(1-z)}\right] \\ &- 3L_4\left[\frac{-z}{(1-z)^2}\right] + 9L_4[-z(1+z)] - 18L_4\left[\frac{z}{(1+z)}\right] \\ &+ 12L_4\left[\frac{-z}{(1-z)}\right] + 3L_4\left[\frac{z}{(1+z)^2}\right] - 18L_4\left[\frac{-z^2}{1+z}\right] \\ &- 18L_4[z(1-z)] - \frac{3}{4}L_4\left[\frac{(1-z)^4}{(1+z)^4}\right] + 9L_4\left[\frac{(1-z)^2}{(1+z)^2}\right] \\ &+ 3L_4(1-z^2) - 18L_4(1-z) - 12L_4\left[\frac{1}{(1+z)}\right] + \frac{75}{4}L_4(1). \end{aligned}$$

With $z = \pm\rho$, (6.18) gives combinations of ladders of index 12, already known from Kummer's equations, and the sought-for ladder of index 20. It may be noted that the last six terms in (6.18) have $p = 0$.



6.2.5. *Transition from the fourth to the fifth order.*

6.2.5.1. *The basic set at* $n = 4$. Kummer's equation is given by (7.90) of [2], and an extremely large number of equations, many of them redundant, or exhibiting trivial variations, can be created by the variable changes of §6.2.4.2. As mentioned there, it was necessary to use the computer to search for a linearly independent set; this set was found to involve 15 equations, of which 7 combinations free of $p = 0$ terms could be constructed. Combinations to isolate each of the $p = 0$ terms of (6.18) [except for $L_4(1)$] were also constructed and used to clear (6.18) of all $p = 0$ terms. The resulting relation is the modified head-equation, ready for partial-integration.

6.2.5.2. The processes of §§6.2.3.5 to 6.2.3.11 can be carried out without difficulty, and in the form in which the equation at $n = 5$ was first generated it contained five $p = 0$ terms in the following combination

$$\text{(6.19)}\quad \begin{aligned} &-72L_5[(1-z)^2(1+z)] + 153L_5[(1-z)(1+z)^2] - 162L_5(1-z^2) \\ &\quad + 297L_5(1-z) - 378L_5(1+z). \end{aligned}$$

Because Kummer's equation at the fifth order involves all combinations of the harmonic group, only five independent equations can be constructed. These come from $x = y = z$; $x = y = -z$; $x = -y = z$; $x = z^2$, $y = \pm z$. (The variant $x = y = z^2$ is ineffective.) Each of these five equations involves a different $p = 0$ term, so no combinations involving only terms with arguments with $p \neq 0$ can be constructed. The two equations from $x = z^2$, $y = \pm z$ are the only ones to give the arguments $(1-z)^2(1+z)$ and $(1+z)^2(1-z)$ of (6.19). As encountered throughout this study, the equations to eliminate terms, like the five above, are very much over-determined. It was therefore a pleasant surprise to discover that the pair of coefficients needed to cancel the two specified arguments also eliminated the remaining three. The resulting modified head-equation, with no $p = 0$ terms, contains 55 transcendental terms:

(6.20)
$$\begin{aligned} 9L_5\left[-z^7\frac{(1-z)}{(1+z)}\right] &= -\frac{27}{2}L_5(z^6) + 64L_5\left[\frac{-z^5}{1+z}\right] - 16L_5\left[\frac{z^5}{1-z}\right] \\ &+ 11L_5\left[\frac{-z^5}{(1-z)(1+z)^3}\right] - 8L_5\left[\frac{z^5}{(1-z)^3(1+z)}\right] + 56L_5[-z^4(1-z)] \\ &- 8L_5[-z^4(1+z)] + \frac{135}{2}L_5(z^4) - 16L_5\left[\frac{z^4}{(1-z)^2(1+z)^3}\right] \\ &+ 28L_5\left[\frac{z^4}{(1-z)^3(1+z)^2}\right] + 90L_5\left[\frac{-z^3(1-z)}{(1+z)}\right] + 36L_5\left[\frac{z^3}{1+z}\right] \\ &+ 72L_5\left[\frac{-z^3}{(1+z)^3}\right] - 36L_5\left[\frac{z^3(1+z)}{(1-z)}\right] + 72L_5\left[\frac{-z^3}{1-z}\right] + 144L_5\left[\frac{z^3}{1-z^2}\right] \end{aligned}$$

(continues)

$$
\begin{aligned}
&- 306L_5\left[\frac{-z^3}{1-z^2}\right] + 72L_5\left[\frac{-z^3}{(1-z)(1+z)^2}\right] - 72L_5\left[\frac{z^3}{(1-z)^2(1+z)}\right] \\
&- 72L_5\left[\frac{z^3}{(1-z)^3}\right] - 36L_5[z^2(1-z)] + 144L_5[z^2(1+z)] + \frac{9}{2}L_5(z^2) \\
&- 944L_5\left[\frac{z^2}{1+z}\right] - 1656L_5\left[\frac{-z^2}{1+z}\right] - 324L_5\left[\frac{z^2}{(1+z)^2}\right] + 452L_5\left[\frac{z^2}{1-z}\right] \\
&+ 1224L_5\left[\frac{-z^2}{1-z}\right] - 144L_5\left[\frac{-z^2}{(1-z)(1+z)^2}\right] + 324L_5\left[\frac{z^2}{(1-z)^2}\right] \\
&+ 72L_5\left[\frac{-z^2}{(1-z)^2(1+z)}\right] + 19L_5[-z(1-z)^3(1+z)] - 16L_5[z(1-z)(1+z)^3] \\
&+ 72L_5[z(1-z^2)] - 234L_5[-z(1-z^2)] - 936L_5[z(1-z)] \\
&- 1024L_5[-z(1-z)] - 180L_5\left[\frac{z(1-z)}{(1+z)}\right] + 532L_5[z(1+z)] \\
&+ 504L_5[-z(1+z)] - 1872L_5(z) + 1080L_5\left[\frac{z}{1+z}\right] - 648L_5\left[\frac{-z}{1+z}\right] \\
&- 296L_5\left[\frac{z}{(1+z)^2}\right] + 135L_5\left[\frac{-z(1+z)}{(1-z)}\right] + 648L_5\left[\frac{z}{1-z}\right] - 1080L_5\left[\frac{-z}{1-z}\right] \\
&- 1188L_5\left[\frac{z}{1-z^2}\right] + 216L_5\left[\frac{-z}{1-z^2}\right] + 144L_5\left[\frac{z}{(1-z)(1+z)^2}\right] \\
&+ 344L_5\left[\frac{-z}{(1-z)^2}\right] - 216L_5\left[\frac{-z}{(1-z)^2(1+z)}\right] + 20L_5\left[\frac{z}{(1-z)^2(1+z)^3}\right] \\
&- 8L_5\left[\frac{-z}{(1-z)^3(1+z)^2}\right].
\end{aligned}
$$

The ordering of the terms in this equation is in a $(pqrs)$ hierarchy which occurs automatically in the computer algebra performed. It has the disadvantage, however, of separating structurally-related terms, and of inter-mixing the head and subsidiary terms arising from the use of Kummer's results.

At $z = \pm\rho$, (6.20) gives the needed ladder of index 20, and the already-known ladder of index 12. With $z = \pm\omega$, (6.20) yields two of the missing $\omega$-ladders discussed earlier.

6.2.6. Nonextension to the sixth order. Although the first step, elimination of $p = 0$ terms, was readily accomplished, the next step is halted by the impossibility of creating any residual equation (no $p = 0$ terms) at the fifth order from Kummer's formulas. The only other available equation, that obtainable from (6.20) by replacing $z$ by $-z$, and called the *adjoint equation,* is insufficient. Any further progress requires additional equations involving the same arguments as those in (6.20), as generated by Kummer's equation. Whether such equations exist without introducing yet further arguments is not currently known. One of the difficulties is the large number of variants that can be constructed, making it very time-consuming to determine whether a supposedly new result is really novel. At this time no such formulas are known.



**6.3. The $\omega$-family of equations.**

6.3.1. *Head-term arguments, and equations for* $n = 2$. In §6.1.4 it was shown that the argument $[-z^6(1-z)^3]$ could be used to generate results of index 42 for $\omega$ and 24 for $\rho$. Since the differential of $\mathrm{Li}_2(u)$ involves $\log(1-u)$, the factors of $1-u$ are crucial for the purpose of developing such formulas. In the present instance we have

$$1 + z^6(1-z)^3 = (1 + z^2 - z^3)(1 - z + z^2)A(z), \tag{6.21}$$

where

$$A(z) = 1 + z - z^2 - z^3 + z^4. \tag{6.22}$$

This quartic does not possess any simple roots and the key to the whole development is the existence of (at least) five further formulas, of the same character as (6.21) from which $A(z)$ can be mutually eliminated.

The following arguments occur repeatedly, and it is convenient to introduce the temporary shorthand notation:

$$
\begin{aligned}
\alpha &= -z(1-z)^2(1+z), \\
\beta &= -z^4/(1-z)(1+z)^2, \\
\gamma &= -z^6(1-z)^3, \\
\delta &= z^3/(1-z^2)^3, \\
\varepsilon &= -z^9/(1+z)^3, \\
\lambda &= z^5(1-z)/(1+z).
\end{aligned}
\tag{6.23}
$$

Functions of these arguments constitute the *head terms* in the equations.

The following five identities, in addition to (6.21), can be readily verified:

$$1 - \alpha = A(z), \tag{6.24}$$

$$1 - \beta = A(z)/(1-z)(1+z)^2, \tag{6.25}$$

$$1 - \delta = A(z)(1 - z - z^2)/(1 - z^2)^3, \tag{6.26}$$

$$1 - \varepsilon = A(z)(1 + z + z^2)(1 + z + z^3)/(1+z)^3, \tag{6.27}$$

$$1 - \lambda = A(z)(1 + z^2)/(1+z). \tag{6.28}$$

By differentiating dilogarithms of $\gamma$, $\delta$, $\varepsilon$, and $\lambda$, using results from $\alpha$ and $\beta$ to eliminate terms in $A(z)$, and then reintegrating, the following four functional equations at the second order can be constructed. They are all readily confirmed by differentiation.

$$
\begin{aligned}
&2\mathrm{Li}_2(\alpha) + \mathrm{Li}_2(\beta) - \mathrm{Li}_2(\gamma) + 3\mathrm{Li}_2[-z^2(1-z)] + 4\mathrm{Li}_2[z(1-z)] \\
&\quad + \mathrm{Li}_2\left[\frac{-z^2}{1-z}\right] + 4\mathrm{Li}_2(z^2) - 2\mathrm{Li}_2(z) + \log^2(1-z) \\
&\quad + 2\log(1-z)\log(1+z) + 2\log^2(1+z) = 0,
\end{aligned}
\tag{6.29}
$$

$$\begin{aligned}(6.30)\quad & \mathrm{Li}_2(\alpha) + 2\mathrm{Li}_2(\beta) - \mathrm{Li}_2(\varepsilon) + 3\mathrm{Li}_2\left[\frac{-z^3}{1+z}\right] + 4\mathrm{Li}_2\left[\frac{-z^2}{1+z}\right] \\ & + \mathrm{Li}_2[-z(1+z)] + 3\mathrm{Li}_2(z^2) + 2\mathrm{Li}_2(z) + \log^2(1-z) \\ & + 4\log(1-z)\log(1+z) + 3\log^2(1+z) = 0,\end{aligned}$$

$$\begin{aligned}(6.31)\quad & \mathrm{Li}_2(\alpha) - \mathrm{Li}_2(\beta) + \mathrm{Li}_2(\delta) - 3\mathrm{Li}_2\left[\frac{z}{(1-z^2)}\right] - \mathrm{Li}_2(z^2) + 4\mathrm{Li}_2(z) \\ & + \frac{5}{2}\log^2(1-z) + 4\log(1-z)\log(1+z) + \log^2(1+z) = 0,\end{aligned}$$

(6.32)

$$\begin{aligned}& \mathrm{Li}_2(\alpha) + \mathrm{Li}_2(\beta) - \mathrm{Li}_2(\lambda) + \frac{1}{2}\mathrm{Li}_2\left[z\frac{1-z}{1+z}\right] - \frac{1}{2}\mathrm{Li}_2\left[-z\frac{1+z}{1-z}\right] + \frac{5}{4}\mathrm{Li}_2(z^4) \\ & - \frac{3}{4}\mathrm{Li}_2(z^2) + \frac{1}{4}\log^2(1-z) + \frac{5}{2}\log(1-z)\log(1+z) \\ & + \frac{7}{4}\log^2(1+z) = 0.\end{aligned}$$

The following comments seem relevant: (1) the use of Rogers' function $L_2$ removes the logarithmic terms from the equations; (2) the differentiation introduces factors $z^{-1}$, $(1-z)^{-1}$, and $(1+z)^{-1}$. Terms in each of these need to be cancelled out and would normally call for three sets of equations for elimination. The fact that only two are needed comes from an internal structure in the equations giving rise to an "economy" that pervades all orders; (3) except for (6.31) and (6.32) all these equations seem inaccessible from Kummer's formulas. Equations (6.31) and (6.32) are exceptional because, at the second (and third) order the arguments of Kummer's equations involve $x$ and $y$, together with $(1-x)$ and $(1-y)$ only in the ratio $(1-x)/(1-y)$. This makes it possible, for example, to take $x = \alpha(1-\beta)/\beta(1-\alpha) = (1-z)^2(1+z)/z^3$, $y = (1-\alpha)/(1-\beta) = (1-z)(1+z)^2$ giving $(1-y)/(1-x) = -\beta(1-\alpha)/(1-\beta) = z^4$; whereas individually, $1-y = (\alpha-\beta)/(1-\beta)$, $1-x = -(\alpha-\beta)/\beta(1-\alpha)$, which do not give rise to usable arguments. Thus (6.31) and (6.32) become accessible from Kummer's equations. This fortunate happenstance does not hold for the other equations, nor at orders higher than the third; (4) *all* the remaining arguments are generated in Kummer's equations by taking $x$ or $y$ as $\pm z$ or $z^2$, or harmonic-group variants of these. The further integration process can therefore proceed by the analysis explained in §6.2.2; the main difference stems from the appearance of six, coupled head-terms, rather than the single one treated in §6.2.

6.3.2. *Functional equations at* $n = 3$. If, say, $\mathrm{Li}_3(\alpha)$ is differentiated, one gets $\mathrm{Li}_2(\alpha)$ multiplied by $[1/z - 2/(1-z) + 1/(1+z)]$. Each of these three factors must be cancelled out, using (6.29) to (6.32), for any functional equation for $\mathrm{Li}_3(\alpha)$. This requires the other leading terms to appear in a



certain ratio, and were it not for the internal structural economy referred to earlier, three $\mathrm{Li}_2$ equations would be needed to handle one $\mathrm{Li}_3$ formula. In actual fact, only two are needed, so three $\mathrm{Li}_3$ equations can be generated from the four $\mathrm{Li}_2$ results. They can be verified by differentiation. All the logarithmic terms can be cancelled out by utilizing the generalized Rogers' function $L_3$.

The three functional equations at the third order developed in this way are:

(6.33)

$$\begin{aligned}& \mathrm{Li}_3(\alpha) - \mathrm{Li}_3(\beta) - \frac{1}{3}\mathrm{Li}_3(\gamma) + \frac{1}{3}\mathrm{Li}_3(\varepsilon) - 3\mathrm{Li}_3\left[\frac{-z^3}{1+z}\right] + 3\mathrm{Li}_3[-z^2(1-z)] \\ & + \mathrm{Li}_3[-z(1+z)] - 4\mathrm{Li}_3\left[\frac{-z^2}{1+z}\right] - \mathrm{Li}_3\left[\frac{-z^2}{1-z}\right] + 4\mathrm{Li}_3[z(1-z)] \\ & - \frac{8}{3}\mathrm{Li}_3(z^3) + \frac{1}{3}\mathrm{Li}_3(z^6) - \frac{3}{2}\mathrm{Li}_3(z^2) - 2\mathrm{Li}_3(z) - 4\mathrm{Li}_3(1-z^2) \\ & + 2\mathrm{Li}_3(1-z) + 2\mathrm{Li}_3\left[\frac{1}{1+z}\right] + 2\zeta(2)\log[(1-z)(1+z)^3] \\ & + \frac{2}{3}\log^3(1+z) + \frac{1}{3}\log^3(1-z) - 3\log z\log^2(1-z) \\ & + \log(1+z)\log^2(1-z) + 2\log^2(1+z)\log(1-z) - 3\log z\log^2(1+z) \\ & - 8\log z\log(1-z)\log(1+z) = 0,\end{aligned}$$

$$\begin{aligned}(6.34)\quad & 2\mathrm{Li}_3(\alpha) + 2\mathrm{Li}_3(\beta) - \mathrm{Li}_3(\delta) - \mathrm{Li}_3(\lambda) + \frac{3}{2}\mathrm{Li}_3(z^4) + 9\mathrm{Li}_3\left[\frac{z}{1-z^2}\right] \\ & + 3\mathrm{Li}_3\left[z\frac{1-z}{1+z}\right] - 2\mathrm{Li}_3\left[-z\frac{1+z}{1-z}\right] - 3\mathrm{Li}_3(z^2) + 12\mathrm{Li}_3(z) \\ & + 12\mathrm{Li}_3\left[\frac{-z}{1-z}\right] - 12\mathrm{Li}_3\left[\frac{z}{1+z}\right] + \frac{5}{4}\log^3\left[\frac{1-z}{1+z}\right] \\ & + 12\log^2(1-z)\log(1+z) + 4\log^3(1+z) - \log^3(1-z) = 0.\end{aligned}$$

(This equation, or at least a variant of it with the same leading terms, can be deduced from Kummer's equation for $\mathrm{Li}_3$.)

(6.35)

$$
\begin{aligned}
&2\mathrm{Li}_3(\alpha) + \mathrm{Li}_3(\beta) - \frac{1}{3}\mathrm{Li}_3(\gamma) - \frac{2}{3}\mathrm{Li}_3(\delta) + 6\mathrm{Li}_3\left[\frac{z}{1-z^2}\right] + 3\mathrm{Li}_3[-z^2(1-z)] \\
&\quad + 4\mathrm{Li}_3[z(1-z)] - \mathrm{Li}_3\left[\frac{-z^2}{1-z}\right] + \frac{1}{3}\mathrm{Li}_3(z^6) - \frac{4}{3}\mathrm{Li}_3(z^3) + 2\mathrm{Li}_3(z^2) \\
&\quad - 8\mathrm{Li}_3(z) + 8\mathrm{Li}_3\left[\frac{1}{1+z}\right] - 6\mathrm{Li}_3(1-z) - 2\mathrm{Li}_3(1-z^2) \\
&\quad + 8\zeta(2)\log(1-z) + 10\zeta(2)\log(1+z) + 2\log^3(1-z) - \frac{2}{3}\log^3(1+z) \\
&\quad + 2\log z\log^2(1+z) + 5\log^2(1-z)\log\left[\frac{1+z}{z}\right] \\
&\quad + 4\log(1+z)\log(1-z)\log\left[\frac{1+z}{z}\right] = 0.
\end{aligned}
$$

6.3.3. *Functional equations at* $n = 4$. The algebra from here on gets extremely heavy, and the MACSYMA program for symbolic manipulation is used in the manner described in §6.2. Rogers' function $L_4$ is utilized, so no logarithmic terms are in evidence in the formulas.

Since (6.33) to (6.35) do not have a $\zeta(3)$ term the three $\mathrm{Li}_3$ equations can be used directly, and give rise to just two results for $\mathrm{Li}_4$:

(6.36)

$$
\begin{aligned}
&9L_4(\alpha) - 9L_4(\beta) - L_4(\gamma) + 2L_4(\delta) + L_4(\varepsilon) + L_4(z^6) - 20L_4(z^3) \\
&\quad + 18L_4(z^2) - 108L_4(z) - 27L_4\left[\frac{-z^3}{1+z}\right] + 27L_4[-z^2(1-z)] \\
&\quad - 54L_4\left[\frac{-z^2}{1+z}\right] + 45L_4\left[\frac{-z^2}{1-z}\right] - 9L_4\left[\frac{-z^2}{1-z}\right] + 72L_4[z(1-z)] \\
&\quad + 27L_4[-z(1+z)] - 108L_4\left[\frac{-z}{1-z}\right] + 18L_4\left[\frac{z}{1+z}\right] \\
&\quad - 54L_4\left[\frac{z}{1-z^2}\right] - 2L_4[-(1-z)^3] + \frac{27}{4}L_4[(1-z)^2] - 36L_4(1-z) \\
&\quad - 2L_4[-(1+z)^3] + \frac{27}{4}L_4[(1+z)^2] + 90L_4(1+z) - 45L_4(1-z^2) \\
&\quad + 36L_4\left[\frac{1-z}{1+z}\right] - 36L_4\left[-\frac{1-z}{1+z}\right] = \frac{187}{2}L_4(1),
\end{aligned}
$$



(6.37)

$$
\begin{aligned}
&3L_4(\alpha) + 3L_4(\beta) - L_4(\gamma) - L_4(\varepsilon) + 6L_4(\lambda) + L_4(z^6) - \frac{45}{4}L_4(z^4) \\
&\quad + 4L_4(z^3) + 63L_4(z^2) - 36L_4(z) + 54L_4\left[\frac{-z^2}{1+z}\right] - 9L_4\left[\frac{-z^2}{1-z}\right] \\
&\quad - 9L_4\left[\frac{-z^2}{1-z^2}\right] + 27L_4\left[\frac{-z^3}{1+z}\right] + 27L_4[-z^2(1-z)] + 72L_4[z(1-z)] \\
&\quad - 27L_4[-z(1+z)] + 60L_4\left[-z\frac{1+z}{1-z}\right] - 90L_4\left[z\frac{1-z}{1+z}\right] \\
&\quad + 54L_4\left[\frac{z}{1+z}\right] - 36L_4\left[\frac{-z}{1-z}\right] - 2L_4[-(1-z)^3] + \frac{27}{4}L_4[(1-z)^2] \\
&\quad + 36L_4(1-z) + 2L_4[-(1+z)^3] - \frac{27}{4}L_4[(1+z)^2] - 18L_4(1+z) \\
&\quad + 18L_4\left[-\left(\frac{1-z}{1+z}\right)^2\right] - 9L_4\left[\left(\frac{1-z}{1+z}\right)^2\right] - 9L_4(1-z^2) = -\frac{63}{4}L_4(1).
\end{aligned}
$$

6.3.4. *Functional equation at* $n = 5$. The two equations (6.36) and (6.37) can be combined to yield a single relation at $n = 5$, using the method given in §6.2. The head terms group into a unique form which yields combinations of (6.36) and (6.37) on differentiation. The equation can take several forms depending on what multiples of Kummer's equations, of which there are five variants, are mixed in with it. It will be recalled that the next preliminary step in trying to go to $n = 6$ is to remove all terms not vanishing with $z$. There are eleven such displayed at $n = 4$, and with only five Kummer's formulas at $n = 5$ the equations for the needed multiples are very much over-determined. Nevertheless the structure of the results permits the elimination to proceed, just as in §6.2.5.2, and the functional equation in this reduced form is as follows:

(6.38)

$$
\begin{aligned}
&L_5(\gamma) + L_5(\delta) + L_5(\varepsilon) - 9L_5(\alpha) - 9L_5(\beta) - 9L_5(\lambda) + 7L_5\left[\frac{-z^5}{1+z}\right] \\
&\quad - 4L_5\left[\frac{z^5}{1-z}\right] + 5L_5\left[\frac{-z^5}{(1-z)(1+z)^3}\right] - 2L_5\left[\frac{z^5}{(1-z)^3(1+z)}\right] \\
&\quad + 5L_5[-z^4(1-z)] - 2L_5[-z^4(1+z)] - 4L_5\left[\frac{z^4}{(1-z)^2(1+z)^3}\right] \\
&\quad + 7L_5\left[\frac{z^4}{(1-z)^3(1+z)^2}\right] + 36L_5\left[-z^3\frac{1-z}{1+z}\right] - 38L_5(z^3) - 39L_5(-z^3) \\
&\quad - 45L_5\left[\frac{z^3}{1+z}\right] - 81L_5\left[\frac{-z^3}{1+z}\right] - 9L_5\left[z^3\frac{1+z}{1-z}\right] - 63L_5\left[\frac{-z^3}{1-z^2}\right] \\
&\quad + 18L_5\left[\frac{-z^3}{(1-z)(1+z)^2}\right] - 18L_5\left[\frac{z^3}{(1-z)^2(1+z)}\right] + 18L_5\left[\frac{-z^3}{1-z}\right]
\end{aligned}
$$

(continues)

$$
\begin{aligned}
&+19L_5\left[\frac{-z^3}{(1+z)^3}\right]-19L_5\left[\frac{z^3}{(1-z)^3}\right]+36L_5\left[\frac{z^3}{1-z^2}\right]\\
&-63L_5[z^2(1-z)]-81L_5[-z^2(1-z)]-380L_5\left[\frac{z^2}{1+z}\right]\\
&-558L_5\left[\frac{-z^2}{1+z}\right]+113L_5\left[\frac{z^2}{1-z}\right]+306L_5\left[\frac{-z^2}{1-z}\right]\\
&-45L_5\left[\frac{-z^2}{(1-z)(1+z)^2}\right]+216L_5(-z^2)+18L_5\left[\frac{-z^2}{(1-z)^2(1+z)}\right]\\
&+36L_5[z^2(1+z)]+7L_5[-z(1-z)^3(1+z)]+18L_5[z(1-z^2)]\\
&-45L_5[-z(1-z^2)]-369L_5[z(1-z)]-400L_5[-z(1-z)]\\
&-54L_5\left[z\frac{1-z}{1+z}\right]+133L_5[z(1+z)]+117L_5[-z(1+z)]+2628L_5(z)\\
&+3123L_5(-z)-1080L_5\left[\frac{z}{1+z}\right]-74L_5\left[\frac{z}{(1+z)^2}\right]+36L_5\left[-z\frac{1+z}{1-z}\right]\\
&-288L_5\left[\frac{-z}{1-z}\right]-378L_5\left[\frac{z}{1-z^2}\right]+95L_5\left[\frac{-z}{(1-z)^2}\right]\\
&-63L_5\left[\frac{-z}{(1-z)^2(1+z)}\right]+5L_5\left[\frac{z}{(1-z)^2(1+z)^3}\right]+171L_5\left[\frac{z}{1-z}\right]\\
&-2L_5\left[\frac{-z}{(1-z)^3(1+z)^2}\right]-1539L_5\left[\frac{-z}{1+z}\right]-4L_5[z(1-z)(1+z)^3]\\
&+36L_5\left[\frac{z}{(1-z)(1+z)^2}\right]=0.
\end{aligned}
$$

This equation gives the remaining two "missing" $\omega$-ladders at $n = 5$ by taking $z = \pm\omega$, and the index-24 $\rho$-ladder by putting $z = \rho$. The equation contains 60 transcendental terms.

6.3.5. *Nonextension to the sixth order.* Because a unique grouping of head terms occurs in (6.38) the equation, as it stands, cannot be extended to $n = 6$; neither can it be combined with the corresponding formula of §6.2 to extend the latter. If any extension at all is possible it must utilize some further combination of expressions yielding the same head terms. It may be noted that a completely new equation, called the *adjoint* equation, comes from replacing $z$ by $-z$, but the two each have their own group requirments and cannot be mutually combined to extend results to the next order.

The replacement of $z$ by $1/z$ also converts an equation into its adjoint, though not in a direct manner, since Kummer's equations are again needed to remove terms not vanishing with $z$. Thus the group $z, -z, 1/z, -1/z$ yields a total of only two basic equations, and there do not seem to be any other usable variants possible.

The very large coefficients that are beginning to emerge in the ladders at $n = 6$ suggest that if indeed functional equations in the trans-Kummer region exist they may contain extremely large numbers of terms, with some enormous coefficients involved.



**6.4. The $\theta$-family of equations.**

6.4.1. *Head-term arguments, and equations for* $n = 2$. In §6.1.5 it was shown that a dilogarithmic functional equation based on the argument $-z^3(1-z)^3$ could reproduce the supernumary ladder of index 24 by taking $z = \theta$. The development has some similarities with the formulas produced in §6.3 for $\omega$, but there are some important differences. In particular, an unresolved feature seems to block the development beyond $n = 3$, even though the leading terms do not have a natural closure until $n = 5$.

6.4.1.1. Since the differential of $\mathrm{Li}_2(w)$ involves $\log(1-w)$ for any variable $w$, the form taken by $1-w$ becomes crucial for the development. For the argument $-z^3(1-z)^3$ we have

$$(6.39)\quad 1+z^3(1-z)^3 = (1+z-z^2)A(z), \quad \text{where } A(z) = 1-z+2z^2-2z^3+z^4.$$

This quartic is analogous to the one appearing in §6.3 and does not factorize into usable expressions. The analysis hinges on a total of six forms, all of which produce $A(z)$ in a similar manner, enabling its mutual cancellation to be effected. There are six head terms with the following arguments and properties; as a convenient shorthand they are designated $\alpha$, $\beta$, $\gamma$, $\delta$, $\varepsilon$, and $\lambda$, expressions that correspond to the similarly-labeled quantities of §6.3:

$$(6.40)\qquad \alpha = \frac{-z}{(1-z)^2(1+z^2)}, \qquad 1-\alpha = \frac{A(z)}{(1-z)^2(1+z^2)};$$

$$(6.41)\qquad \beta = \frac{-z^5}{(1-z)(1+z^2)^2}, \qquad 1-\beta - \frac{A(z)}{(1-z)(1+z^2)^2};$$

$$(6.42)\qquad \gamma = -z^3(1-z)^3, \qquad 1-\gamma = (1+z-z^2)A(z);$$

$$(6.43)\quad \delta = \frac{z^6}{(1-z)^3(1+z^2)^3}, \qquad 1-\delta = \frac{(1-z+z^2)(1-z-z^3)A(z)}{(1-z)^3(1+z^2)^3};$$

$$(6.44)\qquad \varepsilon = \frac{-z^9}{(1+z^2)^3}, \qquad 1-\varepsilon = \frac{(1+z+z^2)(1+z^2+z^3)A(z)}{(1+z^2)^3};$$

$$(6.45)\qquad \lambda = \frac{z^4(1-z)}{1+z^2}, \qquad 1-\lambda = \frac{(1+z)A(z)}{(1+z^2)}.$$

Since $1-\theta = \theta^3$ and $1+\theta^2 = \theta^{-1}$, all of the above expressions on the left become powers of $\theta$ when $z = \theta$; in fact the forms $(1-z)$ and $(1+z^2)$ were selected for this reason.

In addition to these expressions there is an *adjoint* set, obtained by writing $1/z$ for $z$. They are labeled through the use of an asterisk, and two of them are inverted to ensure positive powers of $z$ in the arguments. The transformation $z \to 1/z$ is the only one which retains the form of the arguments as powers of $z$, $1-z$, and $(1+z^2)$, in contrast to $z \to \pm 1/z$ of §6.3; though the extra dimensions of the latter lead to no new equations in the final formulas.

The quartic $A(z)$ becomes replaced by

$$A^*(z) = 1 - 2z + 2z^2 - z^3 + z^4, \tag{6.46}$$

and the adjoint variables are

$$\begin{aligned} \alpha^* &= \frac{-z^3}{(1-z)^2(1+z^2)}, \quad \beta^* = (1-z)(1+z^2)^2; \quad \gamma^* = \frac{z^6}{(1-z)^3}; \\ \delta^* &= \frac{-z^3}{(1-z)^3(1+z^2)^3}, \quad \varepsilon^* = -z^3(1+z^2)^3; \quad \lambda^* = \frac{-z^3(1+z^2)}{(1-z)}. \end{aligned} \tag{6.47}$$

6.4.1.2. There are four dilogarithmic functional equations for each set, obtainable by differentiating $\mathrm{Li}_2$ of the arguments, eliminating $A(z)$ or $A^*(z)$, rearranging, and reintegrating by using special cases of Kummer's equation in which the variables $x$ and $y$ are taken as $z$ or $-z^2$, or harmonic group variants of these. There are a large number of these Kummer-type arguments, and it is useful to keep them conceptually separated from the head terms, since they mutually combine in different ways. Because Hill's and Schaeffer's equations contain $1-x$ and $1-y$ only in the combination $(1-x)/(1-y)$, and we have also the additional relations

$$\alpha\beta = \delta, \qquad \beta/\alpha = \lambda, \qquad (1-\beta)/(1-\alpha) = (1-z)/(1+z^2); \tag{6.48}$$

(where $(1-z)/(1+z^2)$ is one of the Kummer arguments), it is possible to obtain two of the four relations for each set by taking $x = \alpha$, $\alpha^*$ and $y = \beta$, $\beta^*$ in them. Thus two of the four equations in each set are accessible from Kummer's equation but the other pairs are not, so far as is known. At $n = 3$ only Kummer's equation itself operates in this way, leading to a single equation in each set being accessible.

6.4.1.3. The details of the derivations are straightforward but rather lengthy, and only the final equations are shown here:

$$\mathrm{Li}_2(\beta) + \tfrac{1}{3}\mathrm{Li}_2(\gamma) - \tfrac{2}{3}\mathrm{Li}_2(\varepsilon) = F(z) + f(z), \tag{6.49}$$

$$\mathrm{Li}_2(\alpha) + \tfrac{2}{3}\mathrm{Li}_2(\gamma) - \tfrac{1}{3}\mathrm{Li}_2(\varepsilon) = G(z) + g(z), \tag{6.50}$$

$$\mathrm{Li}_2(\lambda) - \tfrac{1}{3}\mathrm{Li}_2(\gamma) - \tfrac{1}{3}\mathrm{Li}_2(\varepsilon) = H(z) + h(z), \tag{6.51}$$

$$\mathrm{Li}_2(\delta) + \mathrm{Li}_2(\gamma) - \mathrm{Li}_2(\varepsilon) = K(z) + k(z). \tag{6.52}$$

The adjoint equations are

$$\mathrm{Li}_2(\beta^*) + \tfrac{1}{3}\mathrm{Li}_2(\gamma^*) - \tfrac{2}{3}\mathrm{Li}_2(\varepsilon^*) = F^*(z) + f^*(z), \tag{6.53}$$

$$\mathrm{Li}_2(\alpha^*) - \tfrac{2}{3}\mathrm{Li}_2(\gamma^*) + \tfrac{1}{3}\mathrm{Li}_2(\varepsilon^*) = G^*(z) + g^*(z), \tag{6.54}$$

$$\mathrm{Li}_2(\lambda^*) - \tfrac{1}{3}\mathrm{Li}_2(\gamma^*) - \tfrac{1}{3}\mathrm{Li}_2(\varepsilon^*) = H^*(z) + h^*(z), \tag{6.55}$$

$$\mathrm{Li}_2(\delta^*) - \mathrm{Li}_2(\gamma^*) + \mathrm{Li}_2(\varepsilon^*) = K^*(z) + k^*(z). \tag{6.56}$$



In the above, the upper-case functions on the right involve Kummer-type arguments and the lower-case functions are combinations of logarithms, (which could have been avoided by working with Rogers' function instead of $\mathrm{Li}_2$). These expressions are given by

$$\begin{aligned} F(z) = {}& \mathrm{Li}_2[-z(1-z)] - 2\mathrm{Li}_2\left(\frac{-z^3}{1+z^2}\right) - 2\mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) \\ & - \mathrm{Li}_2(z) - \frac{4}{3}\mathrm{Li}_2(z^3), \end{aligned} \tag{6.57}$$

$$f(z) = -\log^2(1+z^2) - 2\log(1+z^2)\log(1-z) - \frac{1}{2}\log^2(1-z), \tag{6.58}$$

$$\begin{aligned} G(z) = {}& 2\mathrm{Li}_2[-z(1-z)] - \mathrm{Li}_2\left(\frac{-z^3}{1+z^2}\right) - \mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) \\ & + 2\mathrm{Li}_2(-z^2) - \frac{2}{3}\mathrm{Li}_2(z^3), \end{aligned} \tag{6.59}$$

$$g(z) = -2\log(1+z^2)\log(1-z) - 2\log^2(1-z), \tag{6.60}$$

$$\begin{aligned} H(z) = {}& -\mathrm{Li}_2[-z(1-z)] - \mathrm{Li}_2\left(\frac{-z^3}{1+z^2}\right) - \mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) \\ & + \mathrm{Li}_2\left(\frac{1-z}{1+z^2}\right) + 2\mathrm{Li}_2(z^2) - \mathrm{Li}_2(z) - \frac{2}{3}\mathrm{Li}_2(z^3), \end{aligned} \tag{6.61}$$

$$h(z) = \log z \log(1-z) - \log z \log(1+z^2) + \frac{1}{2}\log^2(1+z^2) - \zeta(2), \tag{6.62}$$

(6.63)

$$\begin{aligned} K(z) = {}& 3\mathrm{Li}_2[-z(1-z)] - 3\mathrm{Li}_2\left(\frac{-z^3}{1+z^2}\right) - 3\mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) \\ & + 3\mathrm{Li}_2\left(\frac{z}{1+z^2}\right) + 3\mathrm{Li}_2\left(\frac{-z^2}{1-z}\right) + 3\mathrm{Li}_2\left(\frac{z^2}{(1-z)(1+z^2)}\right) \\ & + \frac{3}{2}\mathrm{Li}_2(z^4) - \frac{3}{2}\mathrm{Li}_2(z^2) - 3\mathrm{Li}_2(z) - \mathrm{Li}_2(z^3) - \frac{1}{2}\mathrm{Li}_2(z^6), \end{aligned}$$

$$k(z) = -6\log(1-z)\log(1+z^2) - \frac{3}{2}\log^2(1-z). \tag{6.64}$$

The adjoint functions are

$$\begin{aligned} F^*(z) = {}& \mathrm{Li}_2\left(\frac{z^2}{1-z}\right) - 2\mathrm{Li}_2[-z(1+z^2)] + 2\mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) \\ & - \mathrm{Li}_2(z) - \frac{4}{3}\mathrm{Li}_2(z^3), \end{aligned} \tag{6.65}$$

$$f^*(z) = -\log(z)\log(1-z) - 2\log z\log(1+z^2) - \log^2(1-z) + \log^2(1+z^2) + \zeta(2), \tag{6.66}$$

$$G^*(z) = -2\mathrm{Li}_2\left(\frac{z^2}{1-z}\right) + \mathrm{Li}_2[-z(1+z^2)] - \mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) - 2\mathrm{Li}_2(-z^2) + \frac{2}{3}\mathrm{Li}_2(z^3), \tag{6.67}$$

$$g^*(z) = -2\log(1-z)\log(1+z^2) - \log^2(1+z^2), \tag{6.68}$$

$$H^*(z) = -\mathrm{Li}_2\left(\frac{z^2}{1-z}\right) - \mathrm{Li}_2[-z(1+z^2)] + \mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) - \mathrm{Li}_2\left(\frac{-z(1-z)}{1+z^2}\right) + 2\mathrm{Li}_2(z^2) - \mathrm{Li}_2(z) - \frac{2}{3}\mathrm{Li}_2(z^3), \tag{6.69}$$

$$h^*(z) = \frac{1}{2}\log^2(1-z) + \log(1-z)\log(1+z^2), \tag{6.70}$$

$$\begin{aligned} K^*(z) = &-3\mathrm{Li}_2\left(\frac{z^2}{1-z}\right) + 3\mathrm{Li}_2[-z(1+z^2)] - 3\mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) \\ &- 3\mathrm{Li}_2[z(1-z)] + 3\mathrm{Li}_2\left(\frac{z}{1+z^2}\right) + 3\mathrm{Li}_2\left(\frac{-z}{(1-z)(1+z^2)}\right) \\ &- \frac{3}{2}\mathrm{Li}_2(z^4) + \frac{3}{2}\mathrm{Li}_2(z^2) + 3\mathrm{Li}_2(z) + \mathrm{Li}_2(z^3) + \frac{1}{2}\mathrm{Li}_2(z^6), \end{aligned} \tag{6.71}$$

$$k^*(z) = -6\log(1-z)\log(1+z^2) - 3\log^2(1+z). \tag{6.72}$$

The combinations [(6.49)+(6.50)−(6.52)] and [(6.50)+(6.51)−(6.49)] and the corresponding adjoint relations, are accessible from Kummer's equation.

6.4.2. *Extension to* $n = 3$. To get an equation at $n = 3$ one starts with certain combinations of the leading terms, differentiates, and segregates the factors of $1/z$, $1/(1-z)$, and $2z/(1+z^2)$. These each will involve dilogarithmic forms and logarithms. The leading terms are eliminated via (6.49) to (6.52), and this determines the permitted combinations. As far as the factor of $1/z$ is concerned, the remaining dilogarithms from $F(z)$ to $K(z)$ are eliminated via various combinations of Kummer-type arguments until the surviving ones have arguments that are only simple powers of $z$; these terms are directly integrable. The resulting coefficients of $1/(1-z)$ and $2z/(1+z^2)$ are then collected together and Kummer's equations are used to try to eliminate all dilogarithmic arguments containing, as factors, any powers of $z$, i.e., leaving only those arguments with powers of $(1-z)$ and $(1+z^2)$. If the remaining expressions are then directly integrable—in practice the logarithmic terms generated follow suit—then a functional



equation at the third order is produced. In §6.3 this technique was used successfully to generate formulas both at $n = 3$ and at $n = 4$ and 5, at which point all permissible combinations of the leading terms are used up, so that the method cannot generate formulas at $n = 6$. However, in the present instance we run into a problem. To see how this happens we look at one instance in which the differentiation of the $\mathrm{Li}_3$ combination of terms gives

$$\begin{aligned} \frac{d}{dz}\Bigg\{&\left[\mathrm{Li}_3(\alpha) - \mathrm{Li}_3(\beta) + \frac{1}{3}\mathrm{Li}_3(\varepsilon) - \frac{1}{3}\mathrm{Li}_3(\gamma)\right] - 3\mathrm{Li}_3\left(\frac{-z^3}{1+z^2}\right) \\ &-9\mathrm{Li}_3\left(\frac{-z}{1+z^2}\right) + 3\mathrm{Li}_3[-z(1-z)] - 5\mathrm{Li}_3(z) - 2\mathrm{Li}_3(z^3)\Bigg\} \\ &= \frac{0}{z} + \frac{B}{1-z} + \frac{C\,2z}{1+z^2}. \end{aligned} \tag{6.73}$$

The coefficient of $1/z$ has been constrained to be zero by the choice of the terms with the Kummer-type arguments, and the resulting coefficients $B$ and $C$ are

$$B = 4\mathrm{Li}_2(z^2) + \mathrm{Li}_2(z) + \frac{1}{2}\log^2[(1-z)(1+z^2)^2] - \log^2[(1-z)^2(1+z^2)], \tag{6.74}$$

$$\begin{aligned} C = &\,6\mathrm{Li}_2\left(\frac{-z}{1+z^2}\right) - 2\mathrm{Li}_2(z) - 2\mathrm{Li}_2(z^3) - 2\mathrm{Li}_2(-z^2) \\ &+ \frac{3}{2}\log^2(1+z^2) + \frac{1}{2}\log^2[(1-z)^2(1+z^2)] - \log^2[(1-z)(1+z^2)^2]. \end{aligned} \tag{6.75}$$

The combinations occurring on the right of (6.73) enable all the logarithmic terms to be mutually integrated, but not the remaining dilogarithmic ones. Using the connection between $\mathrm{Li}_2(z)$ and $\mathrm{Li}_2(1-z)$ and $\mathrm{Li}_2(-z/(1-z))$ permits the $\mathrm{Li}_2(z)$ term in $B$ to be handled, but no suitable relation has been found to handle the others, apart, that is, from the restrictive process outlined in the next section.

6.4.3. *Combination with the adjoint set.* Since $\mathrm{Li}_2(-z) + \mathrm{Li}_2(-1/z) = -\zeta(2) - \frac{1}{2}\log^2(z)$, it is possible to approach the integration of the expressions in (6.74) and (6.75) by combining them with corresponding terms with $z$ replaced by $1/z$, i.e., by introducing the adjoint set. Some additional terms turn out to be required, but at least the process enables the integration to proceed. The resulting two equations so generated are:

$$
\begin{aligned}
&\left[\mathrm{Li}_3(\alpha) - \mathrm{Li}_3(\beta) + \tfrac{1}{3}\mathrm{Li}_3(\varepsilon) - \tfrac{1}{3}\mathrm{Li}_3(\gamma)\right] \\
&\qquad + \left[\mathrm{Li}_3(\alpha^*) - \mathrm{Li}_3(\beta^*) + \tfrac{1}{3}\mathrm{Li}_3(\varepsilon^*) - \tfrac{1}{3}\mathrm{Li}_3(\gamma^*)\right] \\
&\qquad = 3\mathrm{Li}_3\left(\frac{-z^3}{1+z^2}\right) - 3\mathrm{Li}_3[-z(1-z)] \\
&\qquad + 6\mathrm{Li}_3\left(\frac{-z}{1+z^2}\right) - 3\mathrm{Li}_3\left(\frac{z^2}{1-z}\right) + 3\mathrm{Li}_3[-z(1+z^2)] \\
&\qquad + \frac{8}{3}\mathrm{Li}_3(z^3) + 5\mathrm{Li}_3(z) - 2\mathrm{Li}_3(-z^2) + \frac{1}{2}\log z \log^2(1-z) \\
&\qquad + 2\log z \log(1-z)\log(1+z^2) + 2\log z\log^2(1+z^2) \\
&\qquad + \frac{3}{2}\log^3(1-z) - \log^3(1+z^2) + 3\log^2(1-z)\log(1+z^2) \\
&\qquad - \zeta(2)\log[(1-z)(1+z^2)^2] - \zeta(3),
\end{aligned}
\tag{6.76}
$$

and

(6.77)

$$
\begin{aligned}
&\left[\mathrm{Li}_3(\beta) + 2\mathrm{Li}_3(\lambda) - \tfrac{2}{3}\mathrm{Li}_3(\varepsilon) - \tfrac{1}{3}\mathrm{Li}_3(\gamma)\right] \\
&\qquad + \left[\mathrm{Li}_3(\beta^*) + 2\mathrm{Li}_3(\lambda^*) - \tfrac{2}{3}\mathrm{Li}_3(\varepsilon^*) - \tfrac{1}{3}\mathrm{Li}_3(\gamma^*)\right] \\
&= -6\mathrm{Li}_3\left(\frac{-z^3}{1+z^2}\right) \\
&\quad - 3\mathrm{Li}_3[-z(1-z)] - 12\mathrm{Li}_3\left(\frac{-z}{1+z^2}\right) + 2\mathrm{Li}_3\left[\frac{-z(1-z)}{1+z^2}\right] \\
&\quad - 3\mathrm{Li}_3\left(\frac{z^2}{1-z}\right) - 6\mathrm{Li}_3[-z(1+z^2)] + 2\mathrm{Li}_3\left[\frac{1-z}{1+z^2}\right] \\
&\quad - \frac{16}{3}\mathrm{Li}_3(z^3) - 19\mathrm{Li}_3(z) + 12\mathrm{Li}_3(z^2) + 4\mathrm{Li}_3(-z^2) \\
&\quad + \frac{1}{2}\log z\log^2(1-z) - 4\log z\log(1-z)\log(1+z^2) \\
&\quad + \log z\log^2(1+z^2) + \log^3(1+z^2) - \frac{1}{2}\log^3(1-z) \\
&\quad + 3\log(1-z)\log^2(1+z^2) + \zeta(2)\log[(1-z)(1+z^2)^4] - \zeta(3).
\end{aligned}
$$

In addition to this "mixed" pair, there are two separate equations, one for each set, obtainable from Kummer's equation by taking $x = \beta$, $y = \alpha$,



$(1-x)/(1-y) = (1-z)/(1+z^2)$; and the corresponding adjoint substitutions:

(6.78)

$$
\begin{aligned}
&[\mathrm{Li}_3(\varepsilon) + \mathrm{Li}_3(\lambda) - 2\mathrm{Li}_3(\alpha) - 2\mathrm{Li}_3(\beta)] \\
&\quad = 2\mathrm{Li}_3\left[\frac{z^5}{(1-z)^2(1+z^2)}\right] \\
&\quad + 2\mathrm{Li}_3\left[\frac{z}{(1-z)(1+z^2)^2}\right] - \mathrm{Li}_3\left[\frac{z^4(1+z^2)}{(1-z)}\right] + 2\mathrm{Li}_3\left[\frac{1-z}{1+z^2}\right] \\
&\quad + 2\mathrm{Li}_3(z^4) + \log^2\left[\frac{1+z^2}{1-z}\right]\log\left[\frac{z}{(1-z)(1+z^2)^2}\right] \\
&\quad + \frac{1}{3}\log^3\left[\frac{1+z^2}{1-z}\right] + 2\zeta(2)\log\left[\frac{1+z^2}{1-z}\right] - 2\zeta(3),
\end{aligned}
$$

and a closely similar adjoint formula.

Unfortunately, this is as far as one can go. The coupling of the two sets in (6.76) and (6.77) ensures that the differentiation of any $\mathrm{Li}_4$ formula that uses these two equations to eliminate the $\mathrm{Li}_3$ leading terms can only end up with the two sets present in equal amounts, whereas it turns out that unequal amounts of each are actually required. If the equations were uncoupled this would not be a problem but the rigid linkage shown by (6.76) and (6.77) prevents any further progress. The real question posed, and not answered at the present time, is whether (6.73) can be integrated by any means other than by introducing the adjoint set.

Acknowledgements. The material on the use of the MACSYMA program of §6.2 first appeared in reference [3] in a paper coauthored with Professor Ernest Rost of the University of Colorado, who was instrumental in handling the computer-related aspects of the research. The corresponding developments of §6.3 using the MACSYMA program were performed by Dr. Xiao Hongnian of the Wuhan Institute of Technology, PRC, working under the guidance of Professor Rost.

## References

1. MACSYMA, A large symbolic manipulation program developed at the MIT Laboratory for Computer Science. MACSYMA is a trademark of Symbolics, Inc.
2. L. Lewin, *Polylogarithms and associated functions*, Elsevier, North-Holland, New York and Amsterdam, 1981.
3. L. Lewin and E. Rost, *Polylogarithmic functional equations: A new category of results developed with the help of computer algebra (MACSYMA)*, Aequationes Math. 31 (1986), 223–242.

CHAPTER 7

# Multivariable Polylogarithm Identities

G. A. RAY

**7.0. Introduction.** Many conjectured relations involving polylogarithms and related functions have been discovered and verified numerically in recent years. In this chapter we develop some general identities which can be used to prove many of these relations. One surprising observation is that the (seemingly) simplest relations have often proved to be the most difficult. Consider for example, the linear dilogarithm relations of Lewin, (see [**8**, **9** and **1**] with M. Abouzahra). These relations involve arguments that are powers of a single algebraic number, called the base. A generalization of Rogers' multivariable identity for the dilogarithm based on a polynomial $r(x, t)$ which is linear in $t$ can be used to directly prove many of Lewin's relations for a large number of bases. However, some conjectured dilogarithm relations have eluded proof. Such difficulties provided the impetus to look for multivariable identities resulting from more general polynomials $r(x, t)$. These more general identities are derived and applied in the following sections.

The first three sections derive some general identities for the dilogarithm, the Bloch-Wigner function $D_2(z)$, the trilogarithm and $D_3(z)$. The form of the identity satisfied by Bloch's function clearly shows that its identities are naturally simpler than those for the dilogarithm. Section 7.4 applies these identities to prove linear power relations involving the dilogarithm. Many of the relations first given in [**9**] are verified in this section. Since all bases for polylogarithm relations discovered so far satisfy one or more cyclotomic equations, §7.5 proves a number of purely algebraic results about those bases. One such result gives an effective method for finding all cyclotomic equations satisfied by a given algebraic number base. The next section describes how work related to Mahler's measure and Salem/Pisot numbers has led to a number of new dilogarithm relations, as well as explaining why certain bases tend to produce prolific numbers of dilogarithm relations. Finally, §7.7 investigates some of the recent relations involving supernumary ladders.

**7.1. A general identity for the dilogarithm.** In this section, we prove a general identity for the dilogarithm, which will be used to verify a large number of relations in later sections. Recall that the dilogarithm function is defined by:

$$\mathrm{Li}_2(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^2}, \qquad (|z| < 1).$$

By using the integral representation

$$\mathrm{Li}_2(z) = -\int_0^z \log(1-\alpha)\,\frac{d\alpha}{\alpha},$$

$\mathrm{Li}_2(z)$ can be analytically continued as a multivalued complex function to the rest of the complex plane. Its principal branch can be defined by taking the principal branch of the logarithm function ($-\pi < \mathrm{Im}(\log(z)) = \arg(z) \le \pi$) in the previous integral. With the path of the above integral being a straight line from 0 to $z$, the principal branch of $\mathrm{Li}_2$ will be analytic except on a branch cut along the real axis from 1 to $\infty$.

Constructing $n$-variable identities can be thought of (roughly) as forming a mapping between certain complex algebraic curves (i.e. polynomials in two variables) of degree $n$ and certain linear combinations of dilogarithm functions which give zero when evaluated at points on the curve. For example, Rogers' formula is a map from monic linear curves $t - p(x)$ of degree $n$ in $x$ with $p(0) = 0$ to

$$\mathrm{Li}_2(t) - \sum_{\substack{\{x|p(x)=t\} \\ \{a|p(a)=1\}}} \mathrm{Li}_2(x/a) + C(p),$$

where $C(p)$ is a constant depending on $p$ and the correct branches of $\mathrm{Li}_2$ must be chosen. The generalization of this map to arbitrary linear curves $q(x)t - p(x)$ with no restrictions on $p$ or $q$ is straightforward. This identity turns out to be useful in proving many of Lewin's ladder relations via a direct evaluation. In this section, we further generalize to curves which are defined by differences of two polynomials $p(x,t) - q(x,t)$, each of which can be rationally transformed so that it factors completely over $\mathbf{C}(t)$. This is true, for example, if both $p$ and $q$ are of total degree 2 and so define genus 0 curves. From this we show that identities derived from linear curves may split into several distinct identities via certain birational transformations.

The main identity to be proven is given in Theorem 7.1.5. Since the statement of the main theorem is rather involved, we present a corollary which is the special case involving linear curves.

COROLLARY 7.1.1. *Define the polynomial*

$$r(x,t) = (1-t)Ax^{n_0} \prod_{k=1}^{n_p} (1-\beta_k x) - \prod_{j=1}^{n_q} (1-\alpha_j x) \tag{7.1}$$

*of degree $n_r$ in $x$ for nonzero complex numbers $\{A, \alpha_j, \beta_k\}$. Let $\{r_l(t)\}$,* *$l = 1, \ldots, n_r$ be the algebraic functions satisfying $r(r_l(t), t) = 0$ and set*

$$F_l(t) = \sum_{k=1}^{n_p} \mathrm{Li}_2(\beta_k r_l(t)) - \sum_{j=1}^{n_q} \mathrm{Li}_2(\alpha_j r_l(t)) - \frac{n_0}{2}\log^2(r_l(t)).$$

*Then for some choice of branches of the $\mathrm{Li}_2$ and log functions, we have*

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b = G(t)\big|_a^b,$$

*where*

$$G(t) = \begin{cases} 0, & n_0 > 0,\ n_0 + n_p < n_r, \\ -\mathrm{Li}_2\left(\frac{t-\xi_1}{1-\xi_1}\right), & n_0 = 0,\ n_p < n_r, \\ -\log(A(1-\xi_2))\log(t-\xi_2) + \mathrm{Li}_2\left(\frac{t-\xi_2}{1-\xi_2}\right), & n_0 > 0,\ n_q < n_r, \\ -\log(A(1-\xi_2))\log(t-\xi_2) + \mathrm{Li}_2\left(\frac{t-\xi_2}{1-\xi_2}\right) & \\ \quad -\mathrm{Li}_2\left(\frac{t-\xi_1}{1-\xi_1}\right), & n_0 = 0,\ n_p = n_q = n_r. \end{cases}$$

*Here $\xi_1 = 1 - 1/A$,*

$$\xi_2 = 1 - \frac{\prod_{j=1}^{n_q}(-\alpha_j)}{A\prod_{k=1}^{n_p}(-\beta_k)},$$

*and $a$ and $b$ can be any complex numbers as long as all the algebraic functions and log terms are defined and finite.*

In particular, it is easy to see that if $a, b \neq \xi_1, \xi_2$, then each $r_l(t)$ for $t = a, b$ is finite and nonzero and therefore each term in the above identity is well defined.

As a further corollary, if we restrict each $\mathrm{Li}_2$ and logarithm term to its principal value, the previous identity must be modified by adding integral multiples of terms coming from the difference in value between consecutive branches of the multivalued functions involved. More specifically, we have

COROLLARY 7.1.2. *For the principal branches of $\mathrm{Li}_2$ and logarithm in $F$ and $G$, we have*

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b - G(t)\big|_a^b = 2\pi i \log(C(t))\big|_a^b \tag{7.2}$$

*for some piece-wise algebraic function $C(t)$ defined on a path $\Gamma$ between $a$ and $b$ and some branch of the logarithm on the right hand side. Here $C(t)$ is real if $\{t, A, \alpha_j, \beta_k, r_l(t)\}$ are all real. Furthermore, if $\{t, A, \alpha_j, \beta_k\}$ are algebraic numbers, so is $C(t)$.*

APPLICATION TO THE BASE $\rho = (-1+\sqrt{5})/2$. Theorem 7.1.1 can be applied directly to prove many of the interesting relations which have been discovered recently for the dilogarithm. One interesting example which illustrates how to apply Corollary 7.1.2 is the relation discovered by Lewin ((34) of [9])

$$\mathrm{Li}_2(\rho^{24}) = 6\mathrm{Li}_2(\rho^8) + 8\mathrm{Li}_2(\rho^6) - 6\mathrm{Li}_2(\rho^4) + \pi^2/30. \tag{7.3}$$

One can form a quadratic polynomial $r(x, t)$ so that all the $\alpha_j$'s and $\beta_k$'s are powers of $\rho$ and both roots of $r(x, a)$ and $r(x, b)$ are powers of $\rho$ as follows.

$$r(x, t) = (1-t)(1+\rho^4 x)(1-\rho^8 x) - (1+\rho^6 x)(1+\rho^{12} x), \tag{7.4}$$

with $a = 1$, $b = \rho^6$. Then it is easy to calculate that $r_1(1) = -\rho^{-6}$, $r_2(1) = -\rho^{-12}$, $r_1(\rho^6) = 1$, and $r_2(\rho^6) = \rho^{-6}$. Since $n_0 = 0$, $A = 1$, $\xi_1 = 0$, and $\xi_2 = 1 + \rho^6$, Corollary 7.1.2 shows that $\rho$ satisfies

(7.5)

$$\begin{aligned}\mathrm{Li}(\rho^6) &- \mathrm{Li}_2(\rho^{-6}) + \mathrm{Li}_2(-\rho^4) + \mathrm{Li}_2(-\rho^{-2}) + \mathrm{Li}_2(\rho^8) + \mathrm{Li}_2(\rho^2) \\ &+ 2\mathrm{Li}_2(\rho^0) + \mathrm{Li}_2(\rho^{-6}) + \mathrm{Li}_2(\rho^6) - \mathrm{Li}_2(\rho^{-2}) - \mathrm{Li}_2(\rho^{-8}) - \mathrm{Li}_2(-\rho^2) \\ &- \mathrm{Li}_2(-\rho^{-4}) - \mathrm{Li}_2(-\rho^6) - \mathrm{Li}_2(-\rho^0) - \mathrm{Li}_2(-\rho^{12}) - \mathrm{Li}_2(-\rho^6) \\ &- \log(\rho^6)\log(-\rho^6) = 0,\end{aligned}$$

for some choice of branches of $\mathrm{Li}_2$ and logarithm. Using the classical identities:

$$\begin{aligned}\mathrm{Li}_2(z) + \mathrm{Li}_2(-z) &= \mathrm{Li}_2(z^2)/2, \\ \mathrm{Li}_2(-z) + \mathrm{Li}_2(-1/z) &= -\mathrm{Li}_2(1) - \log^2(z)/2, \qquad z > 0, \\ \mathrm{Li}_2(z) + \mathrm{Li}_2(1/z) &= 2\mathrm{Li}_2(1) - \log^2(z) + 2\pi i \log(z), \qquad z > 1,\end{aligned}$$

together with the relations $\mathrm{Li}_2(1) = \pi^2/6$ and $\mathrm{Li}_2(-1) = -\pi^2/12$, (7.5) can be simplified to yield

$$\mathrm{Li}_2(\rho^{24}) - 6\mathrm{Li}_2(\rho^8) - 8\mathrm{Li}_2(\rho^6) + 6\mathrm{Li}_2(\rho^4) - 8\mathrm{Li}_2(\rho^2) = -\pi^2/2 + 8\log^2(\rho). \tag{7.6}$$

If the principal branches are chosen in the above identity, Corollary 7.1.2 shows that the residual term can be written as $2\pi i \log(R)$ where $R$ is a *real* number. Since every term in (7.6) is real, it follows that the residual term could only be an integral multiple of $\pi^2$. An evaluation then proves that this multiple must be zero and (7.6) is an identity for the principal branches. Now, using the classical relation $\mathrm{Li}_2(\rho^2) = \pi^2/15 - \log^2(\rho)$, (7.3) can be obtained. Section 7.4 gives many other applications of Theorem 7.1.2 to relations of this type.

7.1.1. *Definitions and notation.* To prove our main theorem, Theorem 7.1.5, we begin by picking two polynomials $p$ and $q$ in $\mathbf{C}[x, t]$ of degree $n_p + n_0$ and $n_q$ in the variable $x$ such that $n_0$ is the highest power of $x$ dividing $p$. Let $n_r$ be the degree of $r(x, t) = p(x, t) - q(x, t)$. Then



$r(x, t)$ can be factored as

$$r_0(t)\prod_{l=1}^{n_r}(x - r_l(t))$$

for $r_0(t) \in \mathbf{C}[t]$ and certain algebraic functions $\{r_l(t)\}_{l=1,2,\dots,n_r}$. Similarly, one can factor

$$p(x, t) = p_0(t)x^{n_0}\prod_{j=1}^{n_p}(1 - x/p_j(t)),$$

and

$$q(x, t) = q_0(t)\prod_{k=1}^{n_q}(1 - x/q_k(t)),$$

for $p_0(t)$, $q_0(t) \in \mathbf{C}[t]$ and certain algebraic functions $\{p_j(t)\}_{j=1,2,\dots,n_p}$ and $\{q_k(t)\}_{k=1,2,\dots,n_q}$ which are not identically zero. The indices $j$, $k$, and $l$ will run from $1, 2, \dots, n_p$, $1, 2, \dots, n_q$ and $1, 2, \dots, n_r$ respectively in what follows unless otherwise noted.

The rational function $R(x, t) = p(x, t)/q(x, t)$ then becomes

$$R(x, t) = \frac{p(x, t)}{q(x, t)} = x^{n_0}\frac{p_0(t)\prod_{j=1}^{n_p}(1 - x/p_j(t))}{q_0(t)\prod_{k=1}^{n_q}(1 - x/q_k(t))}. \tag{7.7}$$

Finally, let $\Gamma$ denote a path (a piecewise continuously differentiable mapping) in the complex plane from $a$ to $b$ which avoids all the critical points of $p(x, t)$, $q(x, t)$ and $r(x, t)$ as well as the roots of the lowest order coefficients (as polynomials in $x$) of $p(x, t)$, $q(x, t)$ and $r(x, t)$. This ensures the previous algebraic functions are finite, differentiable and nonzero. From now on, $t$ will take values along $\Gamma$ unless otherwise specified.

From $r(r_l(t), t) = 0$, it follows that for each $l$, $R(r_l(t), t) = 1$. Thus we obtain the following $n_r$ equations:

$$r_l(t)^{n_0}\frac{p_0(t)}{q_0(t)}\frac{\prod_{j=1}^{n_p}(1 - r_l(t)/p_j(t))}{\prod_{k=1}^{n_q}(1 - r_l(t)/q_k(t))} = 1, \tag{7.8}$$

for $l = 1, 2, \dots, n_r$.

So far no restrictions have been placed on the original polynomials $p$ and $q$. However, in order to explicitly evaluate certain integrals, we must suppose that $p(x, t)$ and $q(x, t)$ can be rationally transformed so that they factor completely over $\mathbf{C}(t)$. In other words, we assume there exists rational functions $s_p(t)$, $\{s_{p_j}(t)\}$, $s_q(t)$, $\{s_{q_k}(t)\} \in \mathbf{C}(t)$ such that

$$p(x, s_p(t))/p_0(s_p(t)) = x^{n_0}\prod_j(1 - x/s_{p_j}(t)),$$

and

$$q(x, x_q(t))/q_0(s_q(t)) = \prod_k(1 - x/s_{q_k}(t)).$$

Here we have numbered the rational functions so that $p_j(s_p(t)) = s_{p_j}(t)$ and $q_k(s_q(t)) = s_{q_k}(t)$.

One interesting nontrivial case is when $p(x, t) = 0$ and $q(x, t) = 0$ define curves of degree two. Then they both have genus 0 and so there exists birational transformations $(x, t) \to (z_p(t), y_p(t))$ and $(x, t) \to (z_q(t), y_q(t))$ such that $p(z_p, y_p) = q(z_q, y_q) = 0$ (see §7.5 of [6] for example). Because one of the roots of each of $p(x, y_p(t)) = q(x, y_q(t)) = 0$ is rational, both are rational and therefore both factor completely over the rationals.

Finally, suppose that well-defined branches of the inverse functions $\{s_p^{-1}(t)_j\}$, $j = 1, \dots, n_p$ and $\{s_q^{-1}(t)_k\}$, $k = 1, \dots, n_q$ have been chosen so that if $x = s_p^{-1}(t)_j$ and $y = s_q^{-1}(t)_k$, then $s_{p_j}(x) = p_j(t)$ and $s_{q_k}(y) = q_k(t)$ on $\Gamma$. Then, the plan of the proof of Theorem 7.1.5 is as follows. First, a bimultiplicative to biadditive integration operator $T_2$ will be defined. Then $T_2$ will be applied to each factor in each equation (7.8) for $l = 1, 2, \dots, n_r$. Finally, these terms will be summed over $l$.

DEFINITION 7.1.3. If $f$ and $g$ are any piecewise continuously differentiable functions from $\Gamma$ to $\mathbf{C}$ for which the following integral is finite, define the operator

$$T_2(f, g)(t) = T_2(f, g)|_a^t{}_\Gamma = \int_{a\,\Gamma}^{t} \int_1^{f(y)} \frac{dx}{x} \frac{dg(y)}{g(y)}.$$

Here the path of the inner integral will be a straight line from 1 to $f(y)$. Note that $T_2(1 - x, y)(t) = -\mathrm{Li}_2(t)$ where the path $\Gamma$ is a straight line from 0 to $t$.

$T_2$ is a bimultiplicative to biadditive operator (except for certain trivial cases). To see this for the second parameter, note that

$$\frac{dg_1}{g_1} + \frac{dg_2}{g_2} - \frac{d(g_1 g_2)}{g_1 g_2} = 0. \tag{7.9}$$

For the first parameter, note that the derivative of

$$\int_1^{f_1} \frac{dt}{t} + \int_1^{f_2} \frac{dt}{t} - \int_1^{f_1 f_2} \frac{dt}{t} \tag{7.10}$$

is zero by (7.9). Therefore (7.10) is constant and in fact zero if $f_1$, $f_2$ pass through 1 by a continuity argument. Continuity of $f$ and $g$ also allows the removal of (apparent) singularities of $T_2(f, g)$, such as when $f(y) < 0$ and the path of the inner integral goes through zero.

Therefore, $T_2(f_1 f_2, g) = T_2(f_1, g) + T_2(f_2, g)$ and $T_2(f, g_1 g_2) = T_2(f, g_1) + T_2(f, g_2)$. Also, $T_2(f, c) = 0$ if $c$ is constant. Because of these properties, to evaluate $T_2$ on pairs of rational functions, we need only compute it for pairs of linear factors as follows.

LEMMA 7.1.4. *For $e \neq gf$ and $a, b \neq f$,*

$$T_2(gx - e, y - f)|_a^b = -\mathrm{Li}_2\left(\frac{gt - gf}{e - gf}\right)\Big|_a^b + \log(gf - e)\log(t - f)|_a^b, \tag{7.11}$$



*and for $g \neq 0$,*

$$T_2(gx - gf, y - f)|_a^b = \log^2(gt - gf)/2|_a^b, \tag{7.12}$$

*for some choice of branches for the logarithm and dilogarithm.*

PROOF. Suppose $e \neq gf$. Then the derivative of

$$-\mathrm{Li}_2\left(\frac{gt - gf}{e - gf}\right) - C\log(t - f)$$

times $t - f$ for some constant $C$ gives

$$\log\left(\frac{gt - e}{gf - e}\right) - C, \tag{7.13}$$

while the derivative of $T_2(gx - e, y - f)(t)$ times $t - f$ gives

$$\int_1^{gt - e} \frac{dx}{x}.$$

Thus the previous integral is simply $\log(gt - e)$ for some choice of branch determined as $t$ varies on $\Gamma$. Hence (7.11) is correct if $C$ is chosen to be $\log(gf - e)$ and the correct branch of log is chosen.

The proof of (7.12) is a straightforward change of variables. □

Using bimultiplicativity, the previous lemma shows that $T_2$ for arbitrary pairs of rational functions can be written explicitly in terms of the dilogarithm and logarithm.

7.1.2. *Proof of the main theorem.* We can now state and prove the main theorem.

THEOREM 7.1.5. *Suppose that the polynomials $x^{-n_0}p(x, t) = 0$, $q(x, t) = 0$, and $r(x, t) = p - q = \hat{r}_{n_r}(t)x^{n_r} + \cdots + \hat{r}_0(t) = 0$ define $n_p$, $n_q$ and $n_r$ algebraic functions $\{p_j(t)\}$, $\{q_k(t)\}$ and $\{r_l(t)\}$, resp. where $x$ does not divide $x^{-n_0}p$ or $q$. Let $p_0(t) = x^{-n_0}p(x, t)|_{x=0}$ and $q_0(t) = q(0, t)$. Suppose $s_p(u)$, $s_{p_j}(u)$, $s_q(u)$ and $s_{q_j}(u)$ are rational functions such that $p_j(s_p(u)) = s_{p_j}(u)$ and $q_k(s_q(u)) = s_{q_k}(u)$. Set the rational functions*

$$\prod_{l=1}^{n_r}(1 - r_l(s_p(u))/s_{p_j}(u)) \stackrel{\mathrm{def}}{=} w_{p_j}(u)$$

*and*

$$\prod_{l=1}^{n_r}(1 - r_l(s_q(u))/s_{q_k}(u)) \stackrel{\mathrm{def}}{=} w_{q_k}(u)$$

*and suppose the inverse functions $x = s_p^{-1}(t)_j$ and $y = s_q^{-1}(t)_k$ are chosen so that $s_{p_j}(x) = p_j(t)$ and $s_{q_k}(y) = q_k(t)$. Then if*

$$F_l(t) = \sum_{j=1}^{n_p} \mathrm{Li}_2(r_l(t)/p_j(t)) - \sum_{k=1}^{n_q} \mathrm{Li}_2(r_l(t)/q_k(t)) - \frac{n_0}{2}\log^2(r_l(t)),$$

*and*

$$G(t)=\sum_{j=1}^{n_p} T_2(w_{p_j}, s_{p_j})(s_p^{-1}(t)_j)-\sum_{k=1}^{n_q} T_2(w_{q_k}, s_{q_k})(s_q^{-1}(t)_k)+T_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t),$$

*the following identity holds:*

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b = G(t)|_a^b, \tag{7.14}$$

*for some choice of branches of the dilogarithm and logarithm functions in each term.*

PROOF. By applying $T_2(\cdot, r_l(t))$ to each equation $R(r_l(t), t)=1$, $l=1, \ldots, n_r$, it follows using the bimultiplicative to biadditive property of $T_2$ that

(7.15)

$$T_2(p_0/q_0, r_l)+n_0T_2(r_l, r_l)+\sum_j T_2(1-r_l/p_j, r_l)-\sum_k T_2(1-r_l/q_k, r_l)=0.$$

Since $\prod_l r_l(t)=\hat{r}_0/\hat{r}_{n_r}$,

$$\sum_l T_2(p_0/q_0, r_l)=T_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right),$$

and a simple change of variables shows that

$$T_2(r_l, r_l)|_a^b=\log^2(r_l)/2|_a^b,$$

from Lemma 7.1.4 for some branch of logarithm. To handle the remaining terms, note that via a change of variables and Lemma 7.1.4,

$$-\mathrm{Li}_2(r_l/p_j)|_a^b=T_2((1-r_l/p_j), r_l/p_j)|_a^b,$$

and

$$-\mathrm{Li}_2(r_l/q_k)|_a^b=T_2((1-r_l/q_k), r_l/q_k)(t)|_a^b,$$

for some branch of $\mathrm{Li}_2$. Therefore again using the bimultiplicative to biadditive property of $T_2$,

$$T_2(1-r_l/p_j, r_l)(b)=T_2(1-r_l/p_j, p_j)(b)-\mathrm{Li}_2(r_l/p_j)|_a^b, \tag{7.16}$$

and similarly

$$T_2(1-r_l/q_k, r_l)(b)=T_2(1-r_l/q_k, q_k)(b)-\mathrm{Li}_2(r_l/q_k)|_a^b. \tag{7.17}$$

If we sum the unknown terms $T_2(1-r_l/p_j, p_j)$ and $T_2(1-r_l/q_k, q_k)$ over $l$, it follows that

$$T_2\left(\prod_{l=1}^{n_r}(1-r_l/p_j), p_j\right)(t)=\sum_l T_2(1-r_l/p_j, p_j)(t), \tag{7.18}$$



and

$$T_2\left(\prod_{l=1}^{n_r}(1-r_l/q_k), q_k\right)(t)=\sum_l T_2(1-r_l/q_k, q_k)(t). \tag{7.19}$$

Therefore under the change of variables $t\to s_p(u)$ and $t\to s_q(u)$, the integrals on the left hand side of (7.18) and (7.19) become

$$T_2(w_{p_j}, s_{p_j})|_{(s_p^{-1})_j(\Gamma)}=T_2(w_{p_j}, s_{p_j})(s_p^{-1}(t)_j)|_a^b,$$

and

$$T_2(w_{q_k}, s_{q_k})|_{(s_q^{-1})_k(\Gamma)}=T_2(w_{q_k}, s_{q_k})(s_q^{-1}(t)_k)|_a^b.$$

Therefore if we sum (7.15) over $l=1, 2, \ldots, n_r$ and evaluate at $b$, it follows that

$$\begin{aligned} &T_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t)\Big|_a^b+n_0\sum_l \log^2(r_l(t))/2\Big|_a^b \\ &-\sum_{j,l}\mathrm{Li}_2(r_l(t)/p_j(t))\Big|_a^b+\sum_j T_2(w_{p_j}, s_{p_j})(s_p^{-1}(t)_j)\Big|_a^b \\ &+\sum_{k,l}\mathrm{Li}_2(r_l(t)/q_k(t))\Big|_a^b-\sum_k T_2(w_{q_k}, s_{q_k})(s_q^{-1}(t)_k)\Big|_a^b=0. \quad \square \end{aligned} \tag{7.20}$$

COROLLARY 7.1.6. *With the assumptions in Theorem 7.1.5, we have the following identity for the principal branches of* $\mathrm{Li}_2$ *and log in* $F_l(t)$ *and* $G(t)$:

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b - G(t)|_a^b = 2\pi i\log(C(t))|_a^b \tag{7.21}$$

*for some piecewise algebraic function* $C(t)$ *defined on a path* $\Gamma$ *between* $a$ *and* $b$ *and some branch of the logarithm on the right hand side. Furthermore, if* $p$ *and* $q$ *have algebraic coefficients and* $a$ *and* $b$ *are algebraic numbers,*

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b - G(t)|_a^b = 2\pi i\log(C), \tag{7.22}$$

*for some algebraic number* $C$.

PROOF. If $\Delta\mathrm{Li}_2(t)$ denotes the difference in value between two successive branches of $\mathrm{Li}_2(t)$ in their common region of definition, then since $t\mathrm{Li}_2'(t)=\mathrm{Li}_1(t)=-\log(1-t)$, $\Delta\mathrm{Li}_1(t)=2\pi i$, and $\Delta\mathrm{Li}_2(t)=2\pi i\log(t)$. Thus if $c(t)$ is some algebraic function,

$$\mathrm{Li}_2(c(t))=\mathrm{P.B.}\{\mathrm{Li}_2(c(t))\}+2\pi i n(t)\log(c(t)),$$

for some branch of log where P.B.$\{\mathrm{Li}_2\}$ denotes the principal branch of the dilogarithm. Here $n(t)$ is some integer-valued function determined by how $c(\Gamma)$ winds around the branch point of $\mathrm{Li}_2$. Furthermore, $n(t)$ can only change value at finitely many points where $c(t)$ crosses the branch cut of $\mathrm{Li}_2$ (if $c(t)$ lies on the branch cut over an interval, $n(t)$ can only change values at the end points). Thus we can find a piecewise algebraic function $C(t)$ such that

$$\mathrm{Li}_2(c(t)) = \mathrm{P.B.}\{\mathrm{Li}_2(c(t))\} + 2\pi i \log(C(t)),$$

for each $\mathrm{Li}_2$ term in (7.14) and some branch of log. Similarly, it is easy to show that if $d(t)$, $e(t)$ are algebraic functions,

$$\log(d(t))\log(e(t)) = \mathrm{P.B.}\{\log(d(t))\} \times \mathrm{P.B.}\{\log(e(t))\} + 2\pi i \log(D(t)),$$

for some piecewise algebraic function $D(t)$. Note that if $c(t)$, $d(t)$ and $e(t)$ are algebraic functions defined over some algebraic extension of $\mathbf{Q}$, then $C(t)$ and $D(t)$ are algebraic numbers for $t$ algebraic.

Combining these results for all terms in (7.14), it follows that for the principal branches of $\mathrm{Li}_2$ and log in $F$ and $G$ and $t$ on $\Gamma$,

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b - G(t)\Big|_a^b = 2\pi i \log(C(t))\Big|_a^b, \tag{7.23}$$

for some piecewise algebraic function $C(t)$ and some branch of logarithm on the right-hand side. Furthermore, $C(t)$ will be algebraic at algebraic points if $p$ and $q$ have algebraic coefficients. □

7.1.3. *Corollaries and generalizations.* A special case of Theorem 7.1.5 is Corollary 7.1.1, which can be proven as follows.

Let $p(x, t) = (1-t)Ax^{n_0}\prod_{k=1}^{n_p}(1-\beta_k x)$ and $q(x, t) = \prod_{j=1}^{n_q}(1-\alpha_j x)$ for nonzero $\{A, \alpha_j, \beta_k\} \subset \mathbf{C}$. Then since $p_j$ and $q_k$ are constant, so are $s_{p_j}$ and $s_{q_k}$ and therefore $T_2(w_{p_j}, s_{p_j}) = T_2(w_{q_k}, s_{q_k}) = 0$.

Furthermore, it is easy to show that

$$T_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t) = T_2(A(1-t), c(t)),$$

where

$$c(t) = \begin{cases} 0, & n_0 > 0,\ n_0 + n_p < n_r, \\ (t-\xi_1), & n_0 = 0,\ n_p < n_r, \\ 1/(t-\xi_2), & n_0 > 0,\ n_q < n_r, \\ (t-\xi_1)/(t-\xi_2), & n_0 = 0,\ n_p = n_q = n_r. \end{cases} \tag{7.24}$$

Here $\xi_1 = 1 - 1/A$ and $\xi_2$ equals

$$1 - \frac{\prod_{j=1}^{n_q}(-\alpha_j)}{A\prod_{k=1}^{n_p}(-\beta_k)}.$$

This gives the results in Corollary 7.1.1.

The proof of Corollary 7.1.2 follows by noting that if $\{t, A, \alpha_j, \beta_k, r_l(t)\}$ are all real (resp. algebraic), it follows that each argument of each $\mathrm{Li}_2$ and log term is real (resp. algebraic). Hence the proof of Corollary 7.1.6 shows that $C(t)$ is real (resp. algebraic).



Theorem 7.1.5 is not the most general identity possible because (7.14) may "split", i.e., the identity may actually be a sum of separate identities. More precisely, what can happen is that in certain cases, an identity exists involving summing $F_l(t)$ over a certain subset of values of $l$ instead of from 1 to $n_r$. There will generally be some additional terms which must be included however.

THEOREM 7.1.7. *With the assumptions in Theorem 7.1.5, suppose further that for some subset $I \subseteq \{1, 2, \ldots, n_r\}$, there exists rational functions $s_p'(v)$, $s_q'(v)$, and $s_r(v)$ such that*

$$w'_{p_j}(v) \stackrel{\mathrm{def}}{=} \prod_{l\in I}(1 - r_l(s_p\circ s_p')/(s_{p_j}\circ s_p'))(v),$$
$$w'_{q_k}(v) \stackrel{\mathrm{def}}{=} \prod_{l\in I}(1 - r_l(s_q\circ s_q')/(s_{q_k}\circ s_q'))(v),$$

*and*

$$s_f(v) \stackrel{\mathrm{def}}{=} \prod_{l\in I} r_l(s_r(v))$$

*are rational functions. Then we have the following identity:*

$$\begin{aligned} &\sum_{j=1}^{n_p}\sum_{l\in I}\mathrm{Li}_2(r_l(t)/p_j(t))\Big|_a^b - \sum_{k=1}^{n_q}\sum_{l\in I}\mathrm{Li}_2(r_l(t)/q_k(t))\Big|_a^b - n_0\sum_{l\in I}\log^2(r_l(t))/2\Big|_a^b \\ &\quad = \sum_{j=1}^{n_p} T_2(w'_{p_j}, s_{p_j}\circ s_p')((s_p\circ s_p')^{-1}(t)_j)\Big|_a^b \\ &\qquad - \sum_{k=1}^{n_q} T_2(w'_{q_k}, s_{q_k}\circ s_q')((s_q\circ s_q')^{-1}(t)_k)\Big|_a^b + T_2\left(\frac{p_0\circ s_r}{q_0\circ s_r}, s_f\right)(s_r^{-1}(t))\Big|_a^b \end{aligned} \tag{7.25}$$

*for some choices of branches of the dilogarithm and logarithm. Here the inverse functions $x = (s_p\circ s_p')^{-1}(t)_j$, $y = (s_q\circ s_q')^{-1}(t)_k$, and $z = s_r^{-1}(t)$ are assumed to be well defined for $t$ on $\Gamma$ and chosen so that $s_{p_j}(x) = p_j(t)$, $s_{q_k}(y) = q_k(t)$, and $s_f(z) = \prod_{l\in I} r_l(t)$.*

PROOF. The proof is a minor modification to that of Theorem 7.1.5. Consider the integrals corresponding to (7.18) and (7.19) where the sum over $l$ is restricted to the set $I$:

$$T_2\left(\prod_{l\in I}(1 - (r_l\circ s_p)(u)/s_{p_j}(u)), s_{p_j}(u)\right)(s_p^{-1}(t)_j);$$

and

$$T_2\left(\prod_{l\in I}(1 - (r_l\circ s_q)(u)/s_{q_k}(u)), s_{q_k}(u)\right)(s_q^{-1}(t)_j).$$

By assumption, we can change variables in each via $u \to s_p'(v)$ and $u \to s_q'(v)$ respectively to obtain

$$T_2(w'_{p_j}, s_{p_j} \circ s_p')|_{((s_p \circ s_p')^{-1})_j(\Gamma)} = T_2(w'_{p_j}, s_{p_j} \circ s_p')((s_p \circ s_p')^{-1}(t)_j)|_a^b,$$

and

$$T_2(w'_{q_k}, s_{q_k} \circ s_q')|_{((s_q \circ s_q')^{-1})_k(\Gamma)} = T_2(w'_{q_k}, s_{q_k} \circ s_q')((s_q \circ s_q')^{-1}(t)_k)|_a^b,$$

which can be computed explicitly using Lemma 7.1.4 since all arguments are rational functions. Similarly, since the change of variables $t \to s_r(v)$ will make $f(t) = \prod_{l \in I} r_l(t)$ into a rational function, the integral

$$\sum_{l \in I} T_2(p_0/q_0, r_l)\Big|_\Gamma = T_2(p_0/q_0, \prod_{l \in I} r_l)(t)\Big|_\Gamma = T_2\left(\frac{p_0 \circ s_r}{q_0 \circ s_r}, s_f\right)\Big|_{s_r^{-1}(\Gamma)}.$$

Thus every term of (7.15) summed over $l \in I$ can be computed explicitly and we obtain the desired identity. □

Note that if

$$\prod_{l \in I} r_l(t),$$

$$\prod_{l \in I} (1 - (r_l \circ s_p)/s_{p_j})(t),$$

*and*

$$\prod_{l \in I} (1 - (r_l \circ s_q)/s_{q_k})(t),$$

are algebraic functions of genus 0, then there exists a rational parametrization such that these functions become rational and hence satisfy the conditions of Theorem 7.1.7. An important special case of this more general result is when $r(x, t)$ has total degree two (as a polynomial in two variables) and each $p_j$, $q_k$ is constant. In this case, $r_1$ and $r_2$ themselves are algebraic functions of genus 0, so the identity given in Corollary 7.1.2 will split into two separate identities. However, each of these new identities will generally have some additional terms resulting from the birational map. For example, the term

$$T_2\left(\frac{p_0 \circ s_r}{q_0 \circ s_r}, s_f\right)(s_r^{-1}(t))$$

will generally produce more dilogarithm terms than the corresponding term

$$T_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t),$$

if the rational function $s_r$ has degree $> 1$ in numerator or denominator. We will see an application of this splitting in the next section with Bloch's function.



**7.2. A general identity for the Bloch-Wigner function.** The Bloch-Wigner function $D_2(z)$ is closely related to the dilogarithm and satisfies identities very much like those in the previous section. Recall that $D_2(z)$ is defined by:

$$D_2(z) = \arg(1-z)\log|z| + \mathrm{Im}(\mathrm{Li}_2(z)).$$

More specifically, one can write this as follows:

$$D_2(z) = \mathrm{Im}\left(\int_1^{1-z} \frac{dx}{x}\right)\log|z| - \mathrm{Im}\left(\int_0^z \int_1^{1-y} \frac{dx}{x}\frac{dy}{y}\right).$$

Henceforth in this section, $\arg(z)$ will mean the integral $\mathrm{Im}(\int_1^z dx/x)$. It is easy to show that $D_2(z)$ is continuous on $\mathbf{C} \cup \{\infty\}$, zero on $\mathbf{R}$ and real-analytic everywhere on $\mathbf{C}$ except at 0 and 1. Moreover, it is also well defined, independent of the path of integration. $D_2(z)$ has the property that it can be expressed in terms of Clausen's function $\mathrm{Cl}_2(\theta) = \sum_{n=1}^\infty \sin(n\theta)/n^2$ using the identity

$$D_2(z) = \frac{1}{2}\left(D_2\left(\frac{z}{\bar{z}}\right) + D_2\left(\frac{1-1/z}{1-1/\bar{z}}\right) + D_2\left(\frac{1/(1-z)}{1/(1-\bar{z})}\right)\right).$$

$D_2(z)$ also satisfies simpler identities than $\mathrm{Li}_2(z)$. For example,

$$\begin{aligned} D_2(z) &= D_2(1-1/z) = D_2(1/(1-z)) = -D_2(1/z) \\ &= -D_2(1-z) = -D_2(z/(z-1)). \end{aligned} \tag{7.26}$$

Bloch and Wigner introduced this function in [3] as an analog to the logarithm for higher $K$ groups.

As one might expect, since the Bloch-Wigner function is continuous on the whole complex plane as well as at infinity, multi-valued functions with singularities like those of $\log(x)$ should not be present in an identity involving $D_2(x)$. In fact, the multi-variable identities that $D_2(x)$ satisfy turn out to be of the same form as those of the dilogarithm without any additional logarithm terms and no unspecified branches! For example, Rogers' formula for $D_2(x)$ would be a map from monic linear curves $t - p(x)$ of degree $n$ in $x$ with $p(0) = 0$ to

$$D_2(t) - \sum_{\substack{\{p(x)=t\} \\ \{p(a)=1\}}} D_2(x/a) + \sum_{\substack{\{p(x)=1\} \\ \{p(a)=1\}}} D_2(x/a),$$

with no additional terms and no unspecified branches involved. Thus this expression evaluates to zero at any desired point, as one can easily verify on a computer. The generalization of this map to arbitrary linear curves $q(x)t - p(x)$ with no restrictions on $p$ or $q$ is straightforward. Just as with the dilogarithm, this identity can be evaluated to prove many ladder style relations for $D_2(x)$. As before, the linear curve case is a specific case of the main theorem in this section, Theorem 7.2.3, which gives an identity resulting from differences of two polynomials $p(x, t) - q(x, t)$, each of which can be rationally transformed so that it factors completely over $\mathbf{C}(t)$. As with the dilogarithm, identities derived from linear curves may split into several distinct identities via certain birational transformations. The linear curve case for $D_2(z)$ corresponding to Corollary 7.1.1 for $\mathrm{Li}_2(z)$ is presented below.

COROLLARY 7.2.1. *With the notation as in Corollary 7.1.1, set*

$$F_l(t) = \sum_{j=1}^{n_q} D_2(\alpha_j r_l(t)) - \sum_{k=1}^{n_p} D_2(\beta_k r_l(t)).$$

*Then we have the following identity*

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b = G(t)|_a^b,$$

*where*

$$G(t) = \begin{cases} 0, & n_0 > 0,\ n_0 + n_p < n_r, \\ D_2\left(\frac{t-\xi_1}{1-\xi_1}\right), & n_0 = 0,\ n_p < n_r, \\ -D_2\left(\frac{t-\xi_2}{1-\xi_2}\right), & n_0 > 0,\ n_q < n_r, \\ D_2\left(\frac{t-\xi_1}{1-\xi_1}\right) - D_2\left(\frac{t-\xi_2}{1-\xi_2}\right), & n_0 = 0,\ n_p = n_q = n_r. \end{cases}$$

Compare this result to that in Corollary 7.1.1.

7.2.1. *Proof of main theorem.* Assuming the notation as given in §7.1, we could define a bimultiplicative to biadditive operator such as

$$\arg(f(t))\log|g(t)| - \mathrm{Im}(T_2(f, g)(t)),$$

and apply it to the following equations

$$r_l(t)^{n_0}\frac{p_0(t)}{q_0(t)}\frac{\prod_{j=1}^{n_p}(1 - r_l(t)/p_j(t))}{\prod_{k=1}^{n_q}(1 - r_l(t)/q_k(t))} = 1, \tag{7.27}$$

for each $l = 1, 2, \ldots, n_r$. However, it turns out that a much easier proof can be given by using the results of applying $T_2$ in the last section, taking imaginary parts and then modifying each dilogarithm term $-\mathrm{Im}(\mathrm{Li}_2(f(t)))$ by adding $\arg(1 - f(t))\log|f(t)|$. Thus each such dilogarithm term will be converted to $D_2(f(t))$. A simple argument involving the absence of monodromy and singularities will then show that no residual logarithm terms are left.

For notational purposes, we instead define operators on pairs of rational functions which give (essentially) the dilogarithm and logarithm parts of $T_2(f, g)$ separately and the corresponding Bloch function term.



DEFINITION 7.2.2. If $f, g \in \mathbf{Q}(t)$, define three operators $U_2(f, g)$, $V_2(f, g)$, and $W_2(f, g)$ as follows.

$$U_2(gx - e, y - f) = -\mathrm{Li}_2\left(\frac{gt - gf}{e - gf}\right),$$

for $e \neq fg$ and zero otherwise. Set

$$V_2(gx - e, y - f) = \log(gf - e)\log(t - f),$$

for $e \neq fg$ and

$$V_2(gx - gf, y - f) = \log^2(gt - gf)/2.$$

Finally, set

$$W_2(gx - e, y - f) = -D_2\left(\frac{gt - gf}{e - gf}\right),$$

for $e \neq fg$ and zero otherwise. Then extend each operator to pairs of rational functions in a bimultiplicative to biadditive fashion, where if the second parameter is constant, the operator is zero.

Lemma 7.1.4 shows $T_2 = U_2 + V_2$ for some choice of branches of the functions involved. The main $D_2$ identity can now be stated.

THEOREM 7.2.3. *Suppose that the polynomials $x^{-n_0}p(x, t) = 0$, $q(x, t) = 0$, and $r(x, t) = p - q = \hat{r}_{n_r}(t)x^{n_r} + \cdots + \hat{r}_0(t) = 0$ define $n_p$, $n_q$, and $n_r$ algebraic functions $\{p_j(t)\}$, $\{q_k(t)\}$, and $\{r_l(t)\}$, resp. where $x$ does not divide $x^{-n_0}p$ or $q$. Let $p_0(t) = x^{-n_0}p(x, t)|_{x=0}$ and $q_0(t) = q(0, t)$. Suppose $s_p(u)$, $s_{p_j}(u)$, $s_q(u)$, and $s_{q_j}(u)$ are rational functions such that $p_j(s_p(u)) = s_{p_j}(u)$ and $q_k(s_q(u)) = s_{q_k}(u)$. Set the rational functions*

$$\prod_{l=1}^{n_r}(1 - r_l(s_p(u))/s_{p_j}(u) \stackrel{\text{def}}{=} w_{p_j}(u) \quad \textit{and} \quad \prod_{l=1}^{n_r}(1 - r_l(s_q(u))/s_{q_k}(u) \stackrel{\text{def}}{=} w_{q_k}(u)$$

*and suppose the inverse functions $x = s_p^{-1}(t)_j$ and $y = s_q^{-1}(t)_k$ are chosen so that $s_{p_j}(x) = p_j(t)$ and $s_{q_k}(y) = q_k(t)$. Then if*

$$F_l(t) = \sum_{k=1}^{n_q} D_2(r_l(t)/q_k(t)) - \sum_{j=1}^{n_p} D_2(r_l(t)/p_j(t)),$$

*and*

$$G(t) = \sum_{j=1}^{n_p} W_2(w_{p_j}, s_{p_j})(s_p^{-1}(t)_j) - \sum_{k=1}^{n_q} W_2(w_{q_k}, s_{q_k})(s_q^{-1}(t)_k) + W_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t);$$

*the following identity holds*

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b = G(t)|_a^b. \tag{7.28}$$

**Proof.** From (7.20) it follows that

$$\begin{aligned}
&V_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t) + \sum_j V_2(w_{p_j}, s_{p_j})(s_p^{-1}(t)_j) - \sum_k V_2(w_{q_k}, s_{q_k})(s_q^{-1}(t)_k) \\
&\quad = \sum_{j,l} \mathrm{Li}_2(r_l(t)/p_j(t)) - \sum_{k,l} \mathrm{Li}_2(r_l(t)/q_k(t)) - \frac{n_0}{2}\sum_l \log^2(r_l(t)) \\
&\quad + \sum_k U_2(w_{q_k}, s_{q_k})(s_q^{-1}(t)_k) - \sum_j U_2(w_{p_j}, s_{p_j})(s_p^{-1}(t)_j) \\
&\quad - U_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t).
\end{aligned} \tag{7.29}$$

After taking the imaginary part of (7.29), the left-hand side consists of terms of the form $\arg(f(u))\log|g(u)|$ evaluated at $u = h(t)$ where $f$ and $g$ are rational functions and $h(t)$ is an algebraic function. For each term $\mathrm{Im}(\mathrm{Li}_2(H(t)))$ on the right-hand side of (7.29) and $H$ an algebraic function, add the term $\arg(1 - H(t))\log|H(t)|$ to both sides. Also, add $n_0 \arg(r_l(t))\log|r_l(t)|$ for each $l = 1, \dots, n_r$ to both sides. It then follows that the right-hand side of (7.29) becomes

$$\begin{aligned}
&\sum_{j,l} D_2(r_l(t)/p_j(t)) - \sum_{k,l} D_2(r_l(t)/q_k(t)) + \sum_k W_2(w_{q_k}, s_{q_k})(s_q^{-1}(t)_k) \\
&\quad - \sum_j W_2(w_{p_j}, s_{p_j})(s_p^{-1}(t)_j) - W_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t).
\end{aligned} \tag{7.30}$$

Since $D_2$ is bounded on $\mathbf{C} \cup \{\infty\}$, (7.30) has no singularities. Also, since $D_2$ is independent of the integration path $\Gamma$, it follows that at worst (7.30) has finite monodromy at the branch points of the algebraic functions $r_l(s_p^{-1})_j$, $p_j$, $(s_q^{-1})_k$, and $q_k$.

Now consider the left hand side of (7.29) after adding the $\arg(\cdot)\log|\cdot|$ terms. It will still consist of terms of the form $\arg(f(u))\log|g(u)|$ evaluated at $u = h(t)$ for $f, g \in \mathbf{Q}(u)$ and $h(t)$ an algebraic function. To see this, note that those terms $\arg(1 - H(t))\log|H(t)|$ where $H$ is not necessarily rational, i.e.,

$$\sum_j \arg(1 - r_l/p_j)\log|r_l| - \sum_k \arg(1 - r_l/q_k)\log|r_l| + \sum_l n_0 \arg(r_l)\log|r_l|$$

sum up to be

$$-\sum_l \arg\left(r_l^{n_0}\frac{\prod_j(1 - r_l/p_j)}{\prod_k(1 - r_l/q_k)}\right)\log|r_l|,$$



which is simply

$$\arg(p_0/q_0)\log\left|\frac{\hat{r}_0}{\hat{r}_{n_r}}\right|.$$

Furthermore, by factoring and combining all like arg terms, we can assume that the left-hand side can be written as

$$\begin{aligned}
&\sum_{f,g \text{ nonconstant}} a_{f,g} \arg(f(u))\log|g(u)|\big|_{u=h_{f,g}(t)} \\
&\quad + \sum_{g \text{ nonconstant}} b_g \log|g(u)|\big|_{u=h_g(t)} + c,
\end{aligned} \tag{7.31}$$

for $a_{f,g}, b_g, c \in \mathbf{C}$ where each nonconstant argument $f$ in each term in the first sum is a linear factor with a unique zero not shared by any other parameter of arg. Therefore, by letting the path $\Gamma$ wind exclusively around the unique zero of $f$, it follows that any nonzero term in the first summation of (7.31) forces the left-hand side to have infinite monodromy. Thus (7.31) must be of the form

$$\sum_{g \text{ nonconstant}} b_g \log|g(t)| + c, \tag{7.32}$$

where each $g(t)$ is linear with a unique zero. However, any nonzero term of the first sum of (7.32) forces the left-hand side of (7.29) to have singularities unless each $g(t)$ is constant. This implies the left-hand side is in fact constant and therefore zero. □

7.2.2. *Corollaries and generalizations.* A special case of Theorem 7.2.3 is Corollary 7.2.1, which can be proven as follows.

Let $p(x, t) = (1-t)Ax^{n_0}\prod_{k=1}^{n_p}(1 - \beta_k x)$ and $q(x, t) = \prod_{j=1}^{n_q}(1 - \alpha_j x)$ for nonzero $\{A, \alpha_j, \beta_k\} \subset \mathbf{C}$. Then since $p_j$ and $q_k$ are constant, $W_2(w_{p_j}, s_{p_j}) = W_2(w_{q_k}, s_{q_k}) = 0$.

Furthermore, it is easy to show that

$$W_2\left(\frac{p_0}{q_0}, \frac{\hat{r}_0}{\hat{r}_{n_r}}\right)(t) = W_2(A(1-t), c(t)),$$

where $c(t)$ is defined in (7.24) in §7.1. This gives the result in Corollary 7.2.1.

The identity in Theorem 7.2.3 may split as did the main identity in the previous section and for exactly the same reasons.

**Theorem 7.2.4.** *With the assumptions in Theorem 7.2.3, suppose further that for some subset $I \subseteq \{1, 2, \dots, n_r\}$, there exists rational functions $s_p'(v)$, $s_q'(v)$, and $s_r(v)$ such that*

$$w_{p_j}'(v) \stackrel{\text{def}}{=} \prod_{l \in I}(1 - r_l(s_p \circ s_p')/(s_{p_j} \circ s_p')(v),$$

$$w_{q_k}'(v) \stackrel{\text{def}}{=} \prod_{l \in I}(1 - r_l(s_q \circ s_q')/(s_{q_k} \circ s_q')(v),$$

*and*

$$s_f(v) \stackrel{\text{def}}{=} \prod_{l\in I} r_l(s_r(v)),$$

*are rational functions. Then we have the following identity*

$$\begin{aligned}
&\sum_{k=1}^{n_q}\sum_{l\in I} D_2(r_l(t)/q_k(t))|_a^b - \sum_{j=1}^{n_p}\sum_{l\in I} D_2(r_l(t)/p_j(t))|_a^b \\
&= \sum_{j=1}^{n_p} W_2(w'_{p_j}, s_{p_j}\circ s'_p)((s_p\circ s'_p)^{-1}(t)_j)|_a^b \\
&\quad - \sum_{k=1}^{n_q} W_2(w'_{q_k}, s_{q_k}\circ s'_q)((s_q\circ s'_q)^{-1}(t)_k)|_a^b \\
&\quad + W_2\left(\frac{p_0\circ s_r}{q_0\circ s_r}, s_f\right)(s_r^{-1}(t))\Big|_a^b.
\end{aligned} \tag{7.33}$$

*Here the inverse functions* $x = (s_p\circ s'_p)^{-1}(t)_j$, $y = (s_q\circ s'_q)^{-1}(t)_k$, *and* $z = s_r^{-1}(t)$ *are assumed to be well defined so that* $s_{p_j}(x) = p_j(t)$, $s_{q_k}(y) = q_k(t)$, *and* $s_f(z) = \prod_{l\in I} r_l(t)$.

Proof. The proof is a minor modification to that of Theorem 7.2.3. □

As in the previous section, an important special case of this splitting result is when $r(x, t)$ is a quadratic and each $p_j$, $q_k$ is constant. In this case, $r_1$ and $r_2$ are algebraic functions of genus 0, so the identity given in Theorem 7.2.3 will split into two separate identities. We give an application of this below.

*An identity for* $L(2, \chi_{-8})$. Let $\chi_{-8}$ denote the odd quadratic character of conductor 8 associated with the field $\mathbf{Q}(\sqrt{-2})$. Since $\chi$ is odd, very little is known about the value of the Dirichlet $L$-function $L(2, \chi_{-8})$. For example, it is not known if it is even irrational. Let us form the polynomial

$$\begin{aligned}
r(x, t) &= (t-1)(1-\zeta_8 x)(1-\zeta_8^3 x) - (t+1)(1-\zeta_8^5 x)(1-\zeta_8^7 x) \\
&= (t-1)(x^2 - \sqrt{-2}x - 1) - (t+1)(x^2 + \sqrt{-2}x - 1),
\end{aligned}$$

where $\zeta_8 = e^{2\pi i/8}$ with $a = 0$ and let $b \to \infty$. This polynomial is of total degree 2 and each algebraic function $r_l$ can be parameterized by a rational function $s_r$ so that $r_l(s_r)$ is rational. Therefore, choosing the algebraic function $r_1$ where $r_1(0) = 1$ and $r_1(t) \to 0$ as $t \to \infty$, we see that Theorem 7.2.4 implies (after much simplification) that

$$\begin{aligned}
&D_2\left(\frac{-i}{1-\sqrt{2}-i}\right) \\
&\quad - D_2\left(\frac{i}{1-\sqrt{2}+i}\right) + D_2\left(\frac{-i}{-1-\sqrt{2}-i}\right) - D_2\left(\frac{i}{-1-\sqrt{2}+i}\right) \\
&= D_2(\zeta_8) + D_2(\zeta_8^3) - D_2(\zeta_8^5) - D_2(\zeta_8^7) = \sqrt{8}L(2, \chi_{-8}).
\end{aligned}$$



But the identities in (7.26) can be used to show that all four $D_2$ terms above are identical (up to sign), giving

$$4D_2\left(\frac{i}{1-\sqrt{2}+i}\right) = \sqrt{8}L(2, \chi_{-8}).$$

Since $|i/(1-\sqrt{2}+i)| \approx .3827$, the argument of $D_2$ is well inside the unit circle. Therefore, if irrationality/transcendence results for a single term of Bloch's function with complex argument near to the origin can be obtained (similar to those of Bombieri or Chudnovsky for the dilogarithm), then similar irrationality/transcendence results will apply to $\sqrt{8}L(2, \chi_{-8})$. (Note: If this calculation had been done using both roots of $r(x, t)$, only a trivial result would have emerged.)

**7.3. A general identity for the trilogarithm and $D_3(z)$.** In this section, we look at the higher polylogarithms and prove general identities for the trilogarithm as well as Ramakrishnan's generalization of Bloch's function $D_3(x)$.

As with the dilogarithm, $n$-variable identities can be formed by a mapping between certain complex algebraic curves of degree $n$ and certain linear combinations of trilogarithm functions or $D_3(x)$ which give zero when evaluated at points on the curve. Sandham's $n$-variable identity [15] is a map from monic linear curves $t - p(x)$ of degree $n$ in $x$ with $p(0) \neq 0$ to a certain (complicated) linear combination of trilogarithms. The generalization of this map to curves defined by $p(x, t) - q(x, t)$ where $p(x, t) = p_0(t)p_1(x)$ and $q(x, t) = q_0(t)q_1(x)$ is straightforward. However, the generalization to curves which can be rationally transformed so that they factor completely over $\mathbf{C}(t)$ has not been possible due to some difficulties with certain integrals. Because of the complexity of the identities involved in this section, we leave the results in a somewhat less explicit form than for the dilogarithm.

In many ways, the results about $\mathrm{Li}_k$ and $D_k$ for $k > 2$ are less than satisfactory. First, no one has yet discovered a way to construct general multivariable identities for $\mathrm{Li}_k$ and $D_k$ for $k > 3$. Second, the identities satisfied by $D_3$ (and presumably by the higher Bloch-Wigner functions for $k$ odd) have logarithmic correction terms in them, in contrast to $D_2$. Third, $D_3$ cannot apparently be written in terms of Clausen functions as was $D_2$. Finally, even the definitions of the higher Bloch-Wigner functions are not universally agreed upon, owing to attempts to alter the definitions to correct some of the above problems.

7.3.1. *Definitions and Notations.* Recall that the polylogarithm function is defined by:

$$\mathrm{Li}_k(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^k}, \qquad (|z| < 1).$$

for $k = 1, 2, \ldots$. By using the integral representation

$$\mathrm{Li}_k(z) = \int_0^z \frac{\mathrm{Li}_{k-1}(\alpha)}{\alpha}\, d\alpha,$$

$\mathrm{Li}_k(z)$ can be analytically continued to the rest of the complex plane minus a single branch cut as with the dilogarithm.

Ramakrishnan's generalization of Bloch's function (as given explicitly by Zagier in [18]) is defined as follows. Let

$$L_k(z) = \sum_{j=1}^{k} \frac{(-\log|z|)^{k-j}}{(k-j)!} \mathrm{Li}_j(z). \tag{7.34}$$

Then

$$D_k(z) = \begin{cases} \mathrm{Im}(L_k(z)) & (k \text{ even}), \\ \mathrm{Re}(L_k(z)) & (k \text{ odd}). \end{cases}$$

$D_k(z)$ can be real-analytically continued to $\mathbf{C}\backslash\{0, 1\}$ and is continuous on all of $\mathbf{C}$. In addition, if $k$ is even, $D_k(z)$ is continuous on $\mathbf{C} \cup \{\infty\}$ and is zero on $\mathbf{R} \cup \{\infty\}$.

The above definition was modified by Zagier in [17] and [18] for $k$ odd by adding $\log^k|z|/2(k!)$ so that the inversion formula becomes $D_k(z) = (-1)^{k-1} D_k(1/z)$ for all $k$. This however introduced a discontinuity at 0 for $k$ odd and it still did not get rid of the presence of logarithm terms in the other identities satisfied by $D_3$, e.g.,

$$\begin{aligned} &D_3((z-1)/z) + D_3(1-z) + D_3(z) \\ &\quad = \mathrm{Li}_3(1) + \frac{1}{12} \log|(1-z)(z)| \log\left|\frac{z}{(1-z)^2}\right| \log\left|\frac{z^2}{1-z}\right|. \end{aligned}$$

However if $D_3$ is further modified by subtracting the term

$$\log\left|\frac{(1-z)^2}{z}\right| \log\left|\frac{1-z}{z}\right| \log|1-z|/12,$$

then the reader can check that $D_3$ would now satisfy

$$D_3((z-1)/z) + D_3(1-z) + D_3(z) = \mathrm{Li}_3(1),$$

with no logarithm terms. But now $D_3$ would not satisfy the duplication formula $D_3(z^2)/4 = D_3(z) + D_3(-z)$ anymore! We will continue to use Ramakrishnan's definition in what follows.

As in §7.1, we can define an operator $T_k$ for any $k \geq 1$. Then to develop a trilogarithm identity, $T_3$ will be applied to the polynomial equation:

$$r(x, t) = p_0(t) x^{n_0} \prod_{k=1}^{n_p} (1 - \beta_k x) - q_0(t) \prod_{j=1}^{n_q} (1 - \alpha_j x) = 0 \tag{7.35}$$

of degree $n_r$ in $x$ for nonzero complex numbers $\{\alpha_j, \beta_k\}$.



Definition 7.3.1. If $\{f_1, f_2, \ldots, f_k\}$ are any piecewise continuously differentiable functions from $\Gamma$ to $\mathbf{C}$ for which the following integral is finite, define the operator

$$\begin{aligned} T_k(f_1, f_2, \ldots, f_k)(t) &= T_k(f_1, f_2, \ldots, f_k)|_{a\,\Gamma}^{t} \\ &= \int_{a\,\Gamma}^{t} \left[\int_0^{\alpha_{k-1}} \cdots \left[\int_0^{\alpha_2} \left[\int_1^{f_1(\alpha_1)} \frac{d\alpha}{\alpha}\right] \frac{df_2(\alpha_1)}{f_2(\alpha_1)}\right] \cdots \frac{df_{k-1}(\alpha_{k-2})}{f_{k-1}(\alpha_{k-2})}\right] \frac{df_k(\alpha_{k-1})}{f_k(\alpha_{k-1})}, \end{aligned}$$

for $k = 1, 2, \ldots$, where the path of the inner integrals are straight lines.

As before, $T_k$ is $k$-multiplicative to $k$-additive whenever $f_1$ passes through 1. Note that $T_3(1-x, y, z)(t) = -\mathrm{Li}_3(t)$ for a straight line path from 0 to $t$.

As in previous sections, it is necessary to compute $T_k$ on rational functions. However, things are not as straightforward for general values of $k$. It is easy to see that this question is equivalent (by using several applications of integration by parts) to the classical problem of integrating the following types of integrals:

$$\int \prod_{j=1}^{k-1} \log(x - a_j) \frac{dx}{x}.$$

In Chapter 8 of this monograph, Wechsung shows that this general integral cannot be reduced to polylogarithms for $k > 3$. Therefore, we restrict our attention to the case $k = 3$ in the rest of this section.

Since $T_3$ is trimultiplicative to triadditive and zero if any parameter except the first is constant, the following lemma is sufficient for computing $T_3$ on any triple of rational functions.

Lemma 7.3.2. *$T_3(hx-e, y-f, z-g)$ can be evaluated explicitly in terms of the $\mathrm{Li}_3$, $\mathrm{Li}_2$, and log functions.*

Proof. Suppose that $g \neq f$. Then via a change of variables sending $t \to (g-f)t + f$, it follows that $T_3(hx-e, y-f, z-g)|_\Gamma$ becomes

$$\begin{aligned} &T_3((hf-e) - h(f-g)x, y, z-1))|_{(\Gamma-f)/(g-f)} \\ &\quad - T_2((hf-e) - h(f-g)x, y)|_0^{f/(f-g)} T_1(z-1)|_{(\Gamma-f)/(g-f)}. \end{aligned}$$

The second term can clearly be written in terms of dilogarithms and logarithms. Using integration by parts and pulling out a constant from the inner integral, the first term can be reduced to evaluating an integral of the form

$$\int_0^u \log(1-x)\log(1-cx)\frac{dx}{x},$$

where $c = h(f-g)/(hf-e)$. From (8.111) in [7], this integral can be evaluated as

$$\begin{aligned}&\mathrm{Li}_3\left(\frac{1-uc}{1-u}\right)+\mathrm{Li}_3(1/c)+\mathrm{Li}_3(1)-\mathrm{Li}_3(1-cu)\\&-\mathrm{Li}_3(1-u)-\mathrm{Li}_3\left(\frac{1-cu}{c(1-u)}\right)+\log(1-cu)[\mathrm{Li}_2(1/c)-\mathrm{Li}_2(u)]\\&+\log(1-u)[\mathrm{Li}_2(1-cu)-\mathrm{Li}_2(1/c)+\mathrm{Li}_2(1)]+\log(c)\log^2(1-u)/2,\end{aligned}$$

for some choice of branches of $\mathrm{Li}_3$, $\mathrm{Li}_2$, and log. If $g=f$, then sending $t\to t+f$ gives

$$T_3(x-e, y-f, z-f)|_\Gamma=T_3(x-(e-f), y, z)|_{\Gamma-f}.$$

This integral can obviously be evaluated in terms of the trilogarithm, dilogarithm and logarithm since the inner integrals can be evaluated in terms of the dilogarithm and logarithm. □

7.3.2. *Proof of main theorem.* The main identity for the trilogarithm is given below.

THEOREM 7.3.3. *Define the polynomial*

$$\begin{aligned}r(x,t)&=p_0(t)x^{n_0}\prod_{k=1}^{n_p}(1-\beta_k x)-q_0(t)\prod_{j=1}^{n_q}(1-\alpha_j x)\\&=\hat{r}_{n_r}(t)x^{n_r}+\cdots+\hat{r}_0(t),\end{aligned}\tag{7.36}$$

*of degree $n_r$ in $x$ for nonzero complex numbers $\{\alpha_j, \beta_k\}$. Let $\{r_l(t)\}$, $l=1,\dots,n_r$ be the algebraic functions satisfying $r(r_l(t), t)=0$ and set*

$$\begin{aligned}F_l(t)=&\sum_l n_0^2T_3(x,y,z)(r_l(t))+n_0\sum_l\sum_j T_3(x, 1-\alpha_j y, z)(r_l(t))\\&-n_0\sum_l\sum_k T_3(x, 1-\beta_k y, z)(r_l(t))\\&+n_0\sum_l\sum_j T_3(1-\alpha_j x, y, z)(r_l(t))\\&+\sum_l\sum_{j'}\sum_j T_3(1-\alpha_{j'}x, 1-\alpha_j y, z)(r_l(t))\\&-\sum_l\sum_j\sum_k T_3(1-\alpha_j x, 1-\beta_k y, z)(r_l(t))\\&-n_0\sum_l\sum_k T_3(1-\beta_k x, y, z)(r_l(t))\\&-\sum_l\sum_k\sum_j T_3(1-\beta_k x, 1-\alpha_j y, z)(r_l(t))\\&+\sum_l\sum_{k'}\sum_k T_3(1-\beta_{k'}x, 1-\beta_k y, z)(r_l(t)),\end{aligned}$$

*and*

$$G(t)=T_3(q_0/p_0, q_0/p_0, \hat{r}_0/\hat{r}_{n_r})(t).$$



*Here $j, j'=1,\dots,n_q$ and $k, k'=1,\dots,n_p$. Then for some choice of branches of $\mathrm{Li}_3$, $\mathrm{Li}_2$, and log, we have*

$$\sum_{l=1}^{n_r}F_l(t)\Big|_a^b=G(t)|_a^b.\tag{7.37}$$

*$a$ and $b$ can be any complex numbers as long as all the algebraic functions and log terms are defined and finite.*

PROOF. Let

$$R(x,t)=x^{n_0}\frac{\prod_{j=1}^{n_p}(1-\alpha_j x)}{\prod_{k=1}^{n_q}(1-\beta_k x)}.$$

The proof of Theorem 7.3.3 will proceed by applying $T_3$ to the following equations

$$R(r_l(t), t)=q_0(t)/p_0(t),$$

obtaining $T_3(R, R, r_l)=T_3(q_0/p_0, q_0/p_0, r_l)$ and summing over $l=1, 2, \dots, n_r$. Using trimultiplicativity to tri-additivity, we obtain:

(7.38)

$$\begin{aligned}\sum_l T_3(q_0/p_0, q_0/p_0, r_l)\\=n_0^2\sum_l T_3(r_l, r_l, r_l)+n_0\sum_l\sum_j T_3(r_l, 1-\alpha_j r_l, r_l)\\-n_0\sum_l\sum_k T_3(r_l, 1-\beta_k r_l, r_l)+n_0\sum_l\sum_j T_3(1-\alpha_j r_l, r_l, r_l)\\+\sum_l\sum_{j'}\sum_j T_3(1-\alpha_{j'}r_l, 1-\alpha_j r_l, r_l)\\-\sum_l\sum_j\sum_k T_3(1-\alpha_j r_l, 1-\beta_k r_l, r_l)\\-n_0\sum_l\sum_k T_3(1-\beta_k r_l, r_l, r_l)\\-\sum_l\sum_k\sum_j T_3(1-\beta_k r_l, 1-\alpha_j r_l, r_l)\\+\sum_l\sum_{k'}\sum_k T_3(1-\beta_{k'}r_l, 1-\beta_k r_l, r_l),\end{aligned}$$

where $j'$ and $k'$ run from 1 to $n_p$ and 1 to $n_q$ resp.

The sum over $l$ of the term on the left-hand side of (7.38) can be pulled into the third parameter. Then using $\prod_l r_l=\hat{r}_0/\hat{r}_{n_r}$, all the parameters of $T_3$ are rational functions and we obtain the right hand side of (7.37). The change of variables sending $t\to r_l(t)$ will convert all the remaining $T_3$ terms so that all parameters are rational functions as well and can be explicitly computed in terms of $\mathrm{Li}_3$, $\mathrm{Li}_2$ and log by Lemma 7.3.2. □

There is a corresponding split version of the above theorem, whose proof is a trivial modification of the previous one. If $I\subseteq\{1,\dots,n_r\}$ and there exist rational functions $s_r$ and $s_f$ such that $s_f(t) = \prod_{l\in I} r_l(s_r(t))$, then the following identity holds:

$$\sum_{l\in I} F_l(t)\Big|_a^b = \widehat{G}(t)|_a^b,$$

where $\widehat{G}(t) = T_3(q_0/p_0, q_0/p_0, s_f)(s_r^{-1}(t))$.

There is a similar multivariable identity for $D_3$, which can be derived from the previous results for $\mathrm{Li}_3$ as was the identity for $D_2$. There is, however, a fundamental difference between the two multivariable identities. Since $D_3$ is infinite at 0 and $\infty$, there will be logarithm terms in the identity for $D_3$. In fact they are of the form

$$\log|e(z)|\log|f(z)|\log|g(z)||_{z=h(t)}, \tag{7.39}$$

where $e, f, g \in \mathbf{C}(t)$ and $h$ is an algebraic function. We sketch here a method which can be used to derive the multivariable identity for $D_3$ (modulo the exact form of the log terms).

Let $U_3(hx-e, y-f, z-g)$ be the formal sum of non-constant $\mathrm{Li}_3$ terms resulting from computing $T_3(hx-e, y-f, z-g)$ as in Lemma 7.3.2 and let $W_3(hx-e, y-f, z-g)$ be the corresponding sum of $D_3$ terms with the same arguments as $U_3$. Then as with $D_2$, rewrite the identity $\sum_l F_l(t) - G(t) = 0$ in Theorem 7.3.3 by keeping all $\mathrm{Li}_3$ terms on the left hand side and putting all other logarithm and dilogarithm terms on the right hand side. Then take real parts and to each term $\mathrm{Re}(\mathrm{Li}_3(f(z))$ add

$$\log^2|f(z)|\log|1-f(z)|/2 - \log|f(z)|\mathrm{Re}(\mathrm{Li}_2(f(z)))$$

to both sides so that by (7.34), each trilogarithm term on the right-hand side becomes the corresponding $D_3$ term. Now by a monodromy argument (similar to that in Theorem 7.2.3) applied to each term on the left hand side, it is easy to show that no nonconstant terms with factors of the form $\mathrm{Re}(\mathrm{Li}_2(f(z)))$, $\mathrm{Im}(\mathrm{Li}_2(f(z)))$, or $\arg(f(z))$ can be present. Thus the only terms on the left hand side are of the form in (7.39) and the identity for $D_3$ becomes

$$\sum_{l=1}^{n_r} F_l(t)\Big|_a^b - \dot{G}(t)|_a^b = \sum_{e,f,g} \log|e(z)|\log|f(z)|\log|g(z)|\Big|_{z=h_{e,f,g}(a)}^{z=h_{e,f,g}(b)},$$

for some collection $\{e(z), f(z), g(z)\}$ of rational functions and algebraic functions $\{h_{e,f,g}(t)\}$ where now

$$\begin{aligned}
F_l(t) = {} & \sum_l n_0^2 W_3(x, y, z)(r_l(t)) + n_0\sum_l\sum_j W_3(x, 1-\alpha_j y, z)(r_l(t)) \\
& - n_0\sum_l\sum_k W_3(x, 1-\beta_k y, z)(r_l(t)) \\
& + n_0\sum_l\sum_j W_3(1-\alpha_j x, y, z)(r_l(t)) \\
& + \sum_l\sum_{j'}\sum_j W_3(1-\alpha_{j'} x, 1-\alpha_j y, z)(r_l(t)) \\
& - \sum_l\sum_j\sum_k W_3(1-\alpha_j x, 1-\beta_k y, z)(r_l(t)) \\
& - n_0\sum_l\sum_k W_3(1-\beta_k x, y, z)(r_l(t)) \\
& - \sum_l\sum_k\sum_j W_3(1-\beta_k x, 1-\alpha_j y, z)(r_l(t)) \\
& + \sum_l\sum_{k'}\sum_k W_3(1-\beta_{k'} x, 1-\beta_k y, z)(r_l(t)),
\end{aligned}$$

and

$$G(t) = W_3(q_0/p_0, q_0/p_0, \hat{r}_0/\hat{r}_{n_r})(t).$$

For example, the identity resulting when $p(x, t) = (1-t)$ and $q(x, t) = (1-\alpha_1 x)(1-\alpha_2 x)$ can be rewritten in rational form as the explicit two variable formula given in [17] and reproduced below

$$\begin{aligned}
& D_3\left(\frac{x(1-y)^2}{y(1-x)^2}\right) + D_3(xy) + D_3\left(\frac{x}{y}\right) - 2D_3\left(\frac{x(1-y)}{y(1-x)}\right) \\
& - 2D_3\left(\frac{x(1-y)}{x-1}\right) - 2D_3\left(\frac{y(1-x)}{y-1}\right) \\
& - 2D_3\left(\frac{1-y}{1-x}\right) - 2D_3(x) - 2D_3(y) + 2D_3(1) \\
& = -\frac{1}{4}\log|xy|\log\left|\frac{x}{y}\right|\log\left|\frac{x(1-y)^2}{y(1-x)^2}\right|
\end{aligned}$$

If the above methods are used on more general polynomials $p(x, t)$ and $q(x, t)$ such as those that have a rational parameterization which factors completely over $\mathbf{C}(t)$, several terms result which seem to be difficult to compute explicitly.

**7.4. Linear power relations among dilogarithms.** Lewin's ladder relations involve rational linear combinations of polylogarithms of the following form:

$$L_k(N, u) = \frac{\mathrm{Li}_k(u^N)}{N^{k-1}} - \left\{\sum_{r=1}^{N-1} \frac{A_r}{r^{k-1}}\mathrm{Li}_k(u^r) + A_0\log^k(u)/k!\right\}. \tag{7.40}$$

The argument $u$ is an algebraic number called the *base* and $\{A_j\} \subseteq \mathbf{Q}$. Since the multivariable identity for $\mathrm{Li}_3$ is very complicated, we concentrate on dilogarithm ladders in this section. Relations of the type $L_2(N, u) = a_2\zeta(2) = a_2\pi^2/6$, $a_2$ rational, are termed *ladder relations*. In most cases, Lewin observed that the sum over $r$ could be restricted to summing only over $r$ dividing $N$ and $\{A_j\} \subseteq \mathbf{Z}$. He then observed that the base seemed to always satisfy a cyclotomic polynomial equation of the following form:

$$1 - u^N = u^{-A_0} \prod_{r|N} (1 - u^r)^{A_r}. \tag{7.41}$$

In some cases, however, while each individual ladder $L_2(N_k, u)$ was not equal to a rational times $\zeta(2)$ when $r$ is restricted to divide $N$, certain linear combinations did work. In other words, there were rationals $\{B_k\}$ such that $\sum_k B_k L_2(N_k, u) = a_2\zeta(2)$. These were called second degree ladder relations.

To facilitate our investigation, we modify the definition slightly and define a dilogarithm *linear power relation* as a relation of the form

$$\sum_{k=1}^{N} C_k \mathrm{Li}_2(u^{c_k}) = 2\pi i \log(d) + C_0 \log^2(u) \tag{7.42}$$

where $\{C_k, c_k\}$ are integers and $u$ and $d$ are (complex) algebraic numbers. Throughout this section, we take the principal branches of $\mathrm{Li}_2$ and log in each equation unless otherwise noted. Clearly a ladder relation (as well as a second degree ladder relation) is a linear power relation. We also define a cyclotomic equation as one of the following form:

$$x^{D_0} \prod_k (1 - x^{d_k})^{D_k} = \pm 1, \tag{7.43}$$

where $\{D_k, d_k\}$ are integers.

We first show how these relations arise by evaluating the identity in Corollary 7.1.1. We then analytically derive many of the relations in [9], including those which had only been verified numerically. We begin by developing sufficient conditions so that the arguments of all $\mathrm{Li}_2$, log log and $\log^2$ terms in the identity from Corollary 7.1.2 (reproduced below) are integral powers of a single algebraic number $u$. If we suppose that $\{A, a, b, \alpha_j, \beta_k\}$ are algebraic numbers, then by Theorem 7.1.2

(7.44)

$$\sum_{l=1}^{n_r}\sum_{k=1}^{n_p} \mathrm{Li}_2(\beta_k r_l(t))\Big|_a^b - \sum_{l=1}^{n_r}\sum_{j=1}^{n_q} \mathrm{Li}_2(\alpha_j r_l(t))\Big|_a^b - \frac{n_0}{2}\sum_{l=1}^{n_r} \log^2(r_l(t))\Big|_a^b - G(t)\Big|_a^b$$
$$= 2\pi i \log(C) - 4\pi^2 W.$$

Here all the arguments of $\mathrm{Li}_2$ and logarithm are algebraic numbers, all branches are principal and therefore the log term on the right hand side



of (7.2) must be modified by adding the term $4\pi^2 W$ for some integer $W$ because of its unspecified branch. Also,

$$G(t) = \begin{cases} 0, & n_0 > 0,\ n_0 + n_p < n_r, \\ -\mathrm{Li}_2\left(\frac{t-\xi_1}{1-\xi_1}\right), & n_0 = 0,\ n_p < n_r, \\ -\log(A(1-\xi_2))\log(t-\xi_2) & \\ \qquad + \mathrm{Li}_2\left(\frac{t-\xi_2}{1-\xi_2}\right), & n_0 > 0,\ n_q < n_r, \\ -\log(A(1-\xi_2))\log(t-\xi_2) & \\ \qquad + \mathrm{Li}_2\left(\frac{t-\xi_2}{1-\xi_2}\right) - \mathrm{Li}_2\left(\frac{t-\xi_1}{1-\xi_1}\right), & n_0 = 0,\ n_p = n_q = n_r, \end{cases}$$

with $\xi_1 = 1 - 1/A$ and

$$\xi_2 = 1 - \frac{\prod_{j=1}^{n_q}(-\alpha_j)}{A\prod_{k=1}^{n_p}(-\beta_k)}.$$

However, since the duplication identity $\mathrm{Li}_2(-z) = \mathrm{Li}_2(z^2)/2 - \mathrm{Li}_2(z)$ can be used to express terms whose arguments are minus an integral power of $u$ as a sum of two terms whose arguments are integral powers of $u$, we need only develop conditions so that all arguments are plus or minus a power of $u$.

If each of $\{A, \alpha_j, \beta_k, r_l(a), r_l(b)\}$ is (up to sign) a power of $u$, then $1-\xi_1$ and $1-\xi_2$ will be (up to sign) a power of $u$ as well. Therefore each argument of each $\mathrm{Li}_2$, $\log^2$ and log( )log( ) term on the left-hand side in (7.44) will be (up to sign) a power of $u$. By pulling out the exponent of $u$ in each $\log^2$ and log( )log( ) term, each can be written in the form $c\log^2(u)$ for some integer $c$ with possible additional terms of the form $2\pi i\log(u) - 4\pi^2 W'$ for some integer $W'$. These additional terms can be incorporated into the right term $2\pi i \log(D) - 2\pi^2 W$ of (7.44). The above ensures all arguments on the right hand side are powers of $u$; four cases then handle the remaining arguments on the left side of (7.44).

(I) If $n_0 > 0$ and $n_0 + n_p < n_r$, nothing else is needed.

(II) If $n_0 = 0$ and $n_p < n_r$, it is sufficient that $a - \xi_1$ and $b - \xi_1$ be (up to sign) powers of $u$.

(III) If $n_0 > 0$ and $n_0 + n_p = n_r$, it is sufficient that $a - \xi_2$ and $b - \xi_2$ be (up to sign) powers of $u$.

(IV) If $n_0 = 0$ and $n_p = n_r$, it is sufficient that $a - \xi_1$, $b - \xi_1$, $a - \xi_2$, and $b - \xi_2$ be (up to sign) powers of $u$.

An inspection shows that given the above conditions, (7.44) reduces to a linear power relation. We have

THEOREM 7.4.1. *Suppose that $\{\alpha_j, \beta_k, r_l(a), r_l(b), A\}$ are (up to sign) integral powers of a single algebraic number $u$ and the conditions of the relevant case* (I)–(IV) *are met. Then* (7.44) *generates a linear power relation with base $u$.*

The conditions of Theorem 7.4.1 imply that in all except case (I), $u$ must satisfy one or more cyclotomic equations of the form (7.43). To see this, let

$$\begin{aligned} \alpha_j &= \pm u^{a_j}, & 1 &\le j \le n_q, \\ \beta_k &= \pm u^{b_k}, & 1 &\le k \le n_0 + n_p, \\ r_l(a) &= \pm u^{d_l}, & 1 &\le l \le n_r, \\ r_l(b) &= \pm u^{d_l'}, & 1 &\le l \le n_r, \\ 1-\xi_2 &= \pm u^{e}, & A &= \pm u^{f}, \\ a-\xi_1 &= \pm u^{g}, & b-\xi_1 &= \pm u^{g'}, \\ a-\xi_2 &= \pm u^{h}, & b-\xi_2 &= \pm u^{h'}. \end{aligned}$$

for integers $\{a_j, b_k, d_l, d_l', e, f, g, g', h, h'\}$. If the above values are substituted into the equation $r(r_l(a), a) = 0$, we obtain the following equations satisfied by $u$ for $l = 1, 2, \ldots, n_r$:

$$u^{n_0 d_l + f}\left\{\begin{matrix} u^{-f} \pm u^{g} \\ u^{e} \pm u^{h} \end{matrix}\right\}\prod_j (1 \pm u^{b_k - d_l})\prod_k (1 \pm u^{a_j - d_l})^{-1} = \pm 1, \tag{7.45}$$

for some particular choices of signs in each factor. The upper, lower or both values in the braces of (7.45) are chosen depending on the relevant case (II), (III) or (IV), respectively. It is then clear that (7.45) can be transformed into an equation of the form (7.43) by multiplying by appropriate factors of the form $1 - u^n$. However, notice that in case (I), there are no conditions on $a$. Therefore $r(r_l(a), a) = 0$ is not necessarily a cyclotomic equation.

Similarly, if the above values are substituted into the equation $r(r_l(b), b) = 0$, $u$ satisfies

$$u^{n_0 d_l + f}\left\{\begin{matrix} u^{-f} \pm u^{g'} \\ u^{e} \pm u^{h'} \end{matrix}\right\}\prod_j (1 \pm u^{b_k - d_l'})\prod_k (1 \pm u^{a_j - d_l'})^{-1} = \pm 1. \tag{7.46}$$

As before, in case (I) there are no conditions on $b$ and so $r(r_l(b), b) = 0$ is not necessarily a cyclotomic equation.

Thus we see that, except for case (I), $u$ satisfies one or more cyclotomic equations. The problem of finding cyclotomic equations satisfied by a given algebraic number and the problem of finding algebraic numbers which satisfy cyclotomic equations are addressed in the next section.

To illustrate Theorem 7.4.1, we analyze several collections of ladder relations for various bases from [9].

7.4.1. *The base* $\rho$. The base $\rho = (-1 + \sqrt{5})/2$ satisfies the following cyclotomic equations:

$$(1-\rho) = \rho^2, \tag{7.47a}$$

$$(1-\rho^2) = \rho, \tag{7.47b}$$



$$(1-\rho)(1+\rho^3) = (1-\rho^3)(1-\rho^2), \tag{7.47c}$$

$$(1+\rho^6) = (1+\rho^2)(1-\rho^3), \tag{7.47d}$$

$$(1-\rho^2)(1+\rho^{10}) = (1-\rho^4)^3, \tag{7.47e}$$

$$(1+\rho^{12})(1-\rho^3) = (1-\rho^8)(1-\rho^2). \tag{7.47f}$$

These were derived from [9, (33a)-(33f)] by canceling common factors on both sides. The first two equations (7.47a) and (7.47b) lead to the classical relations $\mathrm{Li}_2(\rho) = \pi^2/10 - \log^2(\rho)$ and $\mathrm{Li}_2(\rho^2) = \pi^2/15 - \log^2(\rho)$. The remaining equations lead to more recently discovered relations. We can form $r(x, t)$, obtaining

$$(1-\rho x)(1+\rho^3 x) = (1-\rho^3 x)(1-t), \qquad a = 1, b = \rho^2, \tag{7.48a}$$

$$(1+\rho^6 x) = (1+\rho^2 x)(1-t), \qquad a = 1, b = \rho^3, \tag{7.48b}$$

$$(1-\rho^2 x)(1+\rho^{10} x) = (1-\rho^4 x)^2(1-t), \qquad a = 1, b = \rho^4, \tag{7.48c}$$

$$(1+\rho^{12} x)(1-\rho^3 x) = (1-\rho^8 x)(1-t), \qquad a = 1, b = \rho^2. \tag{7.48d}$$

These four polynomials produce (after some simplification) the four relations (12a)–(12c) and (34) in [9]. The first two relations can be proven by Kummer's two variable functional equation, while the story for the latter two is more complicated. Coxeter had proven (12c) (given below) using a series argument:

$$\mathrm{Li}_2(\rho^{20}) = 2\mathrm{Li}_2(\rho^{10}) + 15\mathrm{Li}_2(\rho^4) - 10\mathrm{Li}_2(\rho^2) + \pi^2/5,$$

while Lewin discovered (34) in [9] (given here in (7.3)) numerically. It had appeared that neither one could be proven by Kummer's identity. It now appears that there are at least three different proofs of (12c) and (34) in [9]. Besides the multivariable identity proof (given in detail in §7.1), there are proofs using a family of single variable functional equations with arguments of the form $\pm z^m(1-z)^r(1+z)^s$ and (despite appearances) a proof by H. Gangl by applying Kummer's identity a number of times.

7.4.2. *The base* $\Omega$. The base $\Omega = (1+\sqrt{3}-\sqrt[4]{12})/2$, a root of $x^6 - 3x^4 - 4x^3 - 3x^2 + 1 = 0$, satisfies the cyclotomic equations (52a)–(52e) from [9]. Lewin's variant of Abel's functional equation verified analytically the ladder relation corresponding to (52a), but three other ladder relations (these were actually second degree ladders), (55a)–(55c) from [9], were verified numerically.

A search among quadratic polynomials $r(x, t)$ found many which satisfy Theorem 7.4.1 for the base $\Omega$ and so give linear power relations. In fact, four independent relations were discovered including a new second degree ladder relation (also found by Browkin) satisfied by $\Omega$. The following four polynomials

$$\begin{aligned}(7.49a)\qquad &(1+\Omega x)(1-\Omega^2 x)(1-t)\\ &\quad = (1+\Omega^3 x)(1+\Omega^4 x), \qquad a=1,\ b=\Omega^4,\end{aligned}$$

$$\begin{aligned}(7.49b)\qquad &(1+\Omega)(1-\Omega^3 x)(1-t)\\ &\quad = (1+\Omega^2 x)(1+\Omega^5 x), \qquad a=1,\ b=\Omega^3,\end{aligned}$$

$$\begin{aligned}(7.49c)\qquad &(1+\Omega^3 x)(1-\Omega^6 x)(1-t)\\ &\quad = (1+\Omega^4 x)(1+\Omega^9 x), \qquad a=1,\ b=\Omega^4,\end{aligned}$$

$$\begin{aligned}(7.49d)\qquad &(1+\Omega^2 x)(1-\Omega^6 x)(1-t)\\ &\quad = (1+\Omega^3 x)(1+\Omega^8 x), \qquad a=1,\ b=\Omega^3,\end{aligned}$$

give four linear power relations:

$$\begin{aligned}(7.50a)\qquad &\mathrm{Li}_2(\Omega^4)/2+\mathrm{Li}_2(\Omega^2)+3\mathrm{Li}_2(\Omega^3)-2\mathrm{Li}_2(\Omega)-\mathrm{Li}_2(\Omega^6)\\ &\quad -\mathrm{Li}_2(\Omega^8)+\mathrm{Li}_2(\Omega^5)-\mathrm{Li}_2(\Omega^{10})/2+\log^2(\Omega)=0,\end{aligned}$$

$$\begin{aligned}(7.50b)\qquad &3\mathrm{Li}_2(\Omega^3)+\mathrm{Li}_2(\Omega^2)/2+2\mathrm{Li}_2(\Omega)+2\mathrm{Li}_2(\Omega^4)+\mathrm{Li}_2(\Omega^5)\\ &\quad -\mathrm{Li}_2(\Omega^{10})/2-\mathrm{Li}_2(\Omega^8)/2+3\log^2(\Omega)/2-\pi^2/4=0,\end{aligned}$$

$$\begin{aligned}(7.50c)\qquad &3\mathrm{Li}_2(\Omega^4)-2\mathrm{Li}_2(\Omega^3)+3\mathrm{Li}_2(\Omega^6)+\mathrm{Li}_2(\Omega)-\mathrm{Li}_2(\Omega^8)\\ &\quad +5\mathrm{Li}_2(\Omega^2)/2+\mathrm{Li}_2(\Omega^9)-\mathrm{Li}_2(\Omega^{18})/2+\log^2(\Omega)-\pi^2/6=0,\end{aligned}$$

$$\begin{aligned}(7.50d)\qquad &6\mathrm{Li}_2(\Omega^3)-3\mathrm{Li}_2(\Omega^2)+3\mathrm{Li}_2(\Omega^4)/2+2\mathrm{Li}_2(\Omega)+\mathrm{Li}_2(\Omega^8)\\ &\quad -\mathrm{Li}_2(\Omega^{16})/2+\log^2(\Omega)-\pi^2/6=0.\end{aligned}$$

Equations (55a)–(55c) from **[9]** can be derived as linear combinations of (7.50a)–(7.50c), while (7.50d) gives the new relation.

7.4.3. *The base* $\Gamma$. The base $\Gamma$, the real root of $x^6-x^5-x^3-x+1=0$ between 0 and 1, satisfies many cyclotomic equations and, as with $\Omega$, one ladder relation was analytically verified by Lewin but six additional independent relations were actually discovered and checked numerically. A similar search was done for appropriate polynomials $r(x,t)$. Many were discovered and in fact all of Lewin's conjectured second degree ladder relations on $\Gamma$ can now be analytically verified in this manner.

7.4.4. *Additional remarks.* The method used of searching among quadratic $r(x,t)$ for appropriate polynomials which satisfy the conditions of Theorem 7.4.1 has in fact verified many of the other dilogarithm relations



reported in **[8, 9 and 1]**. Notable exceptions were the families of quadratic unit bases, namely

$$(7.51)\qquad \{\theta\mid \theta^2-i\theta+1=0,\ 0<\theta<1,\ i=2,3,4,*5,*6,7,*8,*10\},$$

and

$$(7.52)\qquad \{\theta\mid \theta^2+i\theta-1=0,\ 0<\theta<1,\ i=1,2,*3,4\}.$$

Actually, only the starred values of $i$ are the ones for which no quadratic $r(x,r)$ was found with base $\theta_i$ to satisfy Theorem 7.4.1. The remaining ones lead to dilogarithm relations which are accessible from Kummer's equations. It is possible that either a more thorough search of quadratic $r(x,t)$'s or a search using the more general split identity of Theorem 7.1.7 of §7.1 or a search among higher degree $r(x,t)$'s is necessary to prove the resulting linear relations.

There are also obvious directions in which one could generalize Theorem 7.4.1. One possible generalization is the following. Suppose that $u$ is a *real* algebraic number in Theorem 7.4.1 and we assume that $\{\alpha_j, \beta_k, r_l(a), r_l(b)\}$ are (up to sign) either integral powers of $u$ or are roots of unity and the conditions in the relevant case (I)–(IV) are met, where in each case we modify the statement to say that the appropriate quantities are either (up to sign) powers of $u$ or roots of unity. Then (7.44) will still generate a linear power relation with base $u$.

To see this, suppose that (7.44) is written in the form

$$(7.53)\qquad \sum_k C_k\mathrm{Li}_2(u^{c_k})+\sum_k C_k'\mathrm{Li}_2(\zeta^{c_k'}) = 2\pi i\log(d)+C_0\log^2(u)+C_0'\log^2(\zeta),$$

where $\zeta=e^{2\pi i/n}$ is chosen so that all the aforementioned roots of unity are powers of it. Then taking the real part of (7.53) gives

$$\begin{aligned}(7.54)\qquad &\sum_k C_k\mathrm{Li}_2(u^{c_k})+\sum_k C_k'\mathrm{Re}(\mathrm{Li}_2(\zeta^{c_k'}))\\ &\quad = 2\pi i\log(d/|d|)+C_0\log^2(u)+C_0'\log^2(\zeta).\end{aligned}$$

But

$$\mathrm{Re}(\mathrm{Li}_2(\zeta^{c_k'})) = \pi^2((c_k'/n)^2-c_k'/n+1/6),$$

and

$$\log^2(\zeta) = -4\pi^2/n^2.$$

Therefore (7.54) can be written as

$$\sum_k C_k\mathrm{Li}_2(u^{c_k}) = 2\pi i\log(d/|d|)+C_0\log^2(u)+C_0'\mathrm{Li}_2(1);$$

for some rational $C_0'$ and is in fact a linear power relation with real base $u$.

Another direction of generalization is to note that it is possible to use the factorization formula

$$(7.55)\qquad (1/n)\mathrm{Li}_2(u^n) = \sum_{k=0}^{n-1} \mathrm{Li}_2(\zeta^k u)$$

to express certain sums of terms whose arguments are roots of unity times powers of $u$ in terms of sums whose arguments are powers of $u$. Thus the arguments do not necessarily have to be powers of $u$ up to sign.

Lastly, we verify a version of Lewin's conjecture about ladder "clusters". He noted that if $u$ was a base for a ladder relation, then the conjugates of $u$ tended to be bases for ladder relations as well. We prove

THEOREM 7.4.2. *If $r(x, t)$ a and b satisfy the conditions of Theorem 7.4.1 so that (7.44) gives a linear power relation with base $u$, then each conjugate of $u$ will also be a base for some linear power relation.*

PROOF. Suppose that $r(x, t)$, $a$ and $b$ are chosen so that each of $\{\alpha_j, \beta_k, r_i(a), r_i(b), A\}$ is a power of $u$ and the relevant case (I)–(IV) of Theorem 7.4.1 is satisfied. Let $u'$ denote a conjugate of $u$ under some field automorphism of $\mathbf{Q}(u)$ and let $r'(x, t)$ be that conjugation of $r(x, t)$ over $\mathbf{Q}(x, t)$. Then each root $r_i'(a')$ of $r'(x, a') = 0$ and $r_i'(b')$ of $r'(x, b') = 0$ will be a power of $u'$ (up to sign) since each root of $r(x, a)$ and $r(x, b)$ is a power of $u$ (up to sign). Also each $\alpha_j'$ and $\beta_k'$ will be powers of $u'$ (up to sign) as well. Finally, by conjugation any additional quantities in the corresponding case will be powers of $u'$ (up to sign) as well. Therefore the conditions of Theorem 7.4.1 are satisfied and $u'$ must be a base for a linear power relation. □

**7.5. Cyclotomic equations and bases for polylogarithm relations.** We have shown in the previous section that Theorem 7.4.1 implies linear power relations with base $u$ exist if $u$ satisfies cyclotomic equations of a certain form. Two broad goals would be to characterize all algebraic numbers which satisfy (nontrivial) cyclotomic equations and for each such $u$, to find all cyclotomic equations which it satisfies. The first theorem of this section solves both problems if $u$ is rational. Furthermore, if $u$ is a nonzero algebraic number which is not a root of unity, then Theorem 7.5.3 shows that $u$ satisfies only finitely many independent cyclotomic equations and in fact these can be effectively determined. Finally, we develop some methods for handling infinite collections of algebraic integers of a fixed degree $d$. Specifically, we show that all cyclotomic equations satisfied by any one of the quadratic units in the infinite collections

$$\{\theta \mid \theta^2 - m\theta + 1 = 0,\ m = 1, 2, \dots\},$$

and

$$\{\theta \mid \theta^2 + m\theta - 1 = 0,\ m = 1, 2, \dots\}$$

can be determined.



Two cyclotomic equations of the form

$$(7.56)\qquad x^{C_0}\prod_{k=1}^{N}(1 - x^{c_k})^{C_k} = \pm 1$$

can be "multiplied" in a natural way by multiplying the left and right hand sides of each. Thus it makes sense to speak of collections of cyclotomic equations being multiplicatively independent if no product of these equations gives a left hand side of 1. The left-hand side of (7.56) is identically 1 iff the vector $(C_0, C_1, \dots, C_N)$ is zero. Hence cyclotomic equations are multiplicatively independent iff their exponent vectors are additively independent. Note that a base $u$ which satisfies two cyclotomic equations will also satisfy their product. Note also that if $u$ satisfies the cyclotomic equation

$$x^{C_0}\prod_{k=1}^{N}(1 - x^{c_k})^{C_k} = (-1)^n, \qquad n = 0, 1,$$

then $1/u$ satisfies the cyclotomic equation

$$x^{-(C_0 + \sum_k c_k C_k)}\prod_{k=1}^{N}(1 - x^{c_k})^{C_k} = (-1)^{n + \sum_k C_k}.$$

For notational purposes, we restrict the form of a cyclotomic equation to eliminate trivial cases. Such an equation will be written as (7.56) where $C_0 \in \mathbf{Z}$, each $c_k$ is a positive integer with $c_j < c_k$ if $j < k$, and each $C_k$ is a nonzero integer for $k = 1, 2, \dots, N$. If $N = 0$, (7.56) becomes the equation $x^{C_0} = \pm 1$ and the roots are roots of unity. The above definition ensures that (7.56) is a nontrivial rational equation with at most $|C_0| + \sum_k |C_k c_k|$ roots provided $C_0 \neq 0$ or $N \geq 1$.

We first examine the case where a rational satisfies a cyclotomic equation.

THEOREM 7.5.1. *The only rational numbers $x$ which satisfy cyclotomic equations are $0, \pm 1, \pm 1/2, \pm 2, \pm 1/3$, and $\pm 3$.*

(Note: these rationals all lead to linear power relations for polylogarithms. In fact Zagier in **[18]** discovered numerically the following interesting relation for $D_3$:

$$\frac{67}{24}D_3(1) = 6D_3\left(1 - \frac{1}{3}\right) + 3D_3\left(1 - \frac{1}{2^2}\right) - 3D_3\left(\frac{1}{2}\right) - D_3\left(1 - \frac{1}{3^2}\right) - 2D_3\left(\frac{1}{3}\right) + D_3\left(-\frac{1}{3}\right).$$

PROOF. We suppose $N \geq 1$ since $N = 0$ is trivial. We also assume $x \neq 0$. Let $x = w/v$, $(w, v) = 1$, and $v > w \in \mathbf{N}$ (the previous comments cover the case when $v < w$). Then

$$(7.57)\qquad w^{C_0}\prod_k (v^{c_k} - w^{c_k})^{C_k} = \pm v^{C_0 + \sum_k c_k C_k}.$$

Birkhoff and Vandiver [2] showed that if $n > 2$, then any integer of the form $v^n - w^n$ has a *primitive divisor* $p$, with precisely one exception: $v = 2$, $w = 1$, and $n = 6$. To say that $p$ is a primitive divisor means that $p | (v^n - w^n)$ but $p \nmid (v^m - w^m)$ for $1 \le m < n$.

Suppose first that $w/v \neq 1/2$ and $c_N = \max\{c_k\} > 2$. Then there exists a prime primitive divisor $p | v^{c_N} - w^{c_N}$, but not dividing any other factor in (7.57). This is impossible. Now suppose that $w/v \neq 1/2$ but $c_N = 1$ or 2.

If $c_N = 1$, then $N = 1$ and (7.57) becomes

$$w^{C_0}(v - w)^{C_1} = \pm v^{C_0 + C_1}.$$

By the above equation, if $v - w = 1$, $v$ and $w$ must have a prime factor in common, a contradiction. Thus $v - w \neq 1$, so let $p$ be a prime dividing $v - w$. Then by the above equation, $p$ must also divide one of $v$ or $w$ and this implies $p$ divides both $v$ and $w$, a contradiction.

If $c_N = 2$, then (7.57) becomes either

$$w^{C_0}(v^2 - w^2)C_1 = \pm v^{C_0 + 2C_1}, \tag{7.58}$$

or

$$w^{C_0}(v - w)^{C_1}(v^2 - w^2)^{C_2} = \pm v^{C_0 + C_1 + 2C_2}. \tag{7.59}$$

Since $v^2 - w^2 > 1$, if $p$ is a prime dividing $v^2 - w^2$ in (7.58), we obtain a contradiction as before. Therefore, assuming (7.59) is satisfied, suppose $C_0 \neq 0$. If $w > 1$, any prime dividing $w$ cannot divide any other factor, a contradiction. Similarly, assuming $C_0 + C_1 + 2C_2 \neq 0$, any prime dividing $v$ cannot divide any other factor as well. If both are 0, then $C_1 = -2C_2$ and $(v + w)^{C_2} = \pm(v - w)^{C_2}$, or $v + w = \pm(v - w)$, which is impossible. Thus $w$ must be 1 and $C_0 + C_1 + 2C_2$ must be 0, so

$$(v - 1)^{C_1 + C_2}(v + 1)^{C_2} = \pm 1. \tag{7.60}$$

If $C_1 + C_2 = 0$, this implies $v = 0$ or $-2$. Otherwise if a prime $p | v - 1$, then $p | v + 1$, so $p = 2$ and $v - 1 = 2^k$ for some $k$. But $v + 1 = 2(2^{k-1} + 1)$, so $v + 1$ has odd factors for $k > 1$, which is not possible. Hence $v = 3$ and $w = 1$ is the only solution and

$$(1/3)^{C_2}(1 - 1/3)^{-3C_2}(1 - (1/3)^2)^{C_2} = 1, \tag{7.61}$$

for an arbitrary nonzero integer $C_2$ is the most general cyclotomic equation satisfied by 1/3.

The next case is $w/v = 1/2$. If $c_N = 1$, then

$$x^{-C_1}(1 - x)^{C_1} = 1 \tag{7.62}$$

is the most general cyclotomic equation of this type satisfied by 1/2, where $C_1$ is an arbitrary integer. If $c_N \neq 6$, then there is a primitive divisor of $2^{c_N} - 1$, so there are no cyclotomic equations of this type.



By Birkhoff and Vandiver, the final case is the exceptional one with $v = 2$, $w = 1$, and $n = 6$. Then $2^6 - 1$ has no primitive divisors and it easily follows that

$$x^{C_4 - C_1}(1 - x)^{C_1}(1 - x^2)^{-2C_4}(1 - x^3)^{-C_4}(1 - x^6)^{C_4} = 1 \tag{7.63}$$

is the most general cyclotomic equation satisfied by 1/2, and includes (7.62) above as a special case. Here $C_1$ and $C_4$ are arbitrary integers.

Suppose now that $w/v$ is negative. Then any cyclotomic equation satisfied by $-w/v$ corresponds to a cyclotomic equation satisfied by $w/v$ since

$$(1 - (-w/v)^{c_k})^{C_k} = \begin{cases} (1 - (w/v)^{c_k})^{C_k}, & \text{if } c_k \text{ is even}, \\ ((1 - (w/v))^{2c_k}/(1 - (w/v))^{c_k})^{C_k}, & \text{if } c_k \text{ is odd}. \end{cases} \tag{7.64}$$

For example, $-1/2$ satisfies

$$(-1/2)^{C_4 - C_1}(1 - (-1/2))^{-C_1}(1 - (-1/2)^2)^{C_1 - 2C_4}(1 - (-1/2)^3)^{C_4} = (-1)^{C_4 - C_1},$$

and $-1/3$ satisfies

$$(-1/3)^{C_2}(1 - (-1/3)^2)^{-2C_2}(1 - (-1/3))^{3C_2} = (-1)^{C_2}.$$

This concludes all the possibilities in the rational case. □

If $u = \theta$ is now an arbitrary algebraic number of degree $d$ which is not a root of unity, we first show that it can only satisfy finitely many multiplicatively independent cyclotomic equations. The assumption that $\theta$ is not a root of unity is necessary since any root of unity would clearly satisfy infinitely many independent cyclotomic equations. In fact the Kubert identities (see [11]) essentially determine all cyclotomic equations satisfied by a root of unity. These identities are the set of relations

$$m^{1-s} f(mx) = \sum_{k=0}^{m-1} f((x + k)/m),$$

for fixed $s$ and $x \in \mathbf{Q}/\mathbf{Z}$ or $\mathbf{R}/\mathbf{Z}$, satisfied by many classical functions including $f(x) = \mathrm{Li}_s(e^{2\pi i x})$ and in particular $f_1(x) \stackrel{\text{def}}{=} \log|1 - e^{2\pi i x}|$. For $f_1$, the theorem of Bass states:

> Every $\mathbf{Q}$-linear relation between the numbers $f_1(x)$ for rational $x$ is a consequence of the Kubert identities for $s = 1$, together with evenness.

This result is essentially equivalent to the statement that the additive group generated by $f_1(k/m)$, $k = 0, 1, \ldots, m - 1$, has rank $\phi(m)/2 + \pi(m) - 1$, where $\phi$ is Euler's $\phi$ function and $\pi(m)$ is the number of distinct prime factors of $m$.

If one takes the absolute value of (7.56) and then the logarithm, it is clear that any cyclotomic equation satisfied by $\theta = e^{2\pi i k/m}$ implies a $\mathbf{Q}$-linear relation among the numbers $f_1(k/m)$, and this mapping from (7.56) to such a $\mathbf{Q}$-linear relation is one-to-one if we assume that $0 < c_k < m$ for all $k$.

Thus as a corollary to Bass' theorem, we have

COROLLARY 7.5.2. *If $\theta$ is an $m$th root of unity, then $\theta$ satisfies at most $\phi(m)/2 + \pi(m) - 1$ multiplicatively independent cyclotomic equations where (7.56) is assumed to be written so that $0 < c_k < m$ for all $k$.*

If $\theta$ is not a root of unity, we have a much different result.

THEOREM 7.5.3. *Let $\theta$ be a nonzero algebraic number of degree $d$ which is not a root of unity. Then if $\theta$ satisfies the cyclotomic equation (7.56), $c_N$ and $N$ can be bounded by an effectively computable integer depending only on $d$. Therefore $\theta$ satisfies at most finitely many multiplicatively independent cyclotomic equations.*

PROOF. Suppose $\theta$ satisfies the cyclotomic equation (7.56). Then if we bound $c_N$ (and as a consequence $N$) by an effectively determined function depending on $d$, the theorem follows at once. To do this we use Schinzel's results on primitive divisors in algebraic number fields:

THEOREM 7.5.4. (Schinzel (Theorem 1 of [**16**])). *If $A$ and $B$ are relatively prime algebraic integers with $A/B$ of degree $d$ and $A/B$ not a root of unity, then $A^n - B^n$ has a primitive divisor for all $n > N_0(d)$, and $N_0(d)$ is effectively computable.*

Let $\theta = A/B$, with $A$ and $B$ algebraic integers in $\mathbf{Q}(\theta)$ and $(A, B) = 1$. Then (7.56) can be written as

$$A^{C_0}\prod_{k=1}^{N}(B^{c_k} - A^{c_k})^{C_k} = \pm B^{C_0+\sum_k c_k C_k}, \tag{7.65}$$

and from Schinzel's result, if $c_N$ is sufficiently large, $B^{c_N} - A^{c_N}$ has a primitive prime divisor which divides no other factor of (7.65). This is a contradiction. Hence $c_N$ (as well as $N$) must be $\leq N_0(d)$. □

We can get a more precise result for specific families of algebraic numbers. Recall that two quadratic families of bases, (7.51) and (7.52), were discovered which give ladder relations. Lewin found no other values of $i$ which give bases $\theta$ which satisfy any cyclotomic equations. From the previous results we know that a *fixed* $\theta$ can satisfy only finitely many multiplicatively independent cyclotomic equations, but of course this does not allow one to effectively determine all cyclotomic equations satisfied by $\theta$'s in the infinite collections

$$\{\theta_i | \theta_i^2 + i\theta_i - 1 = 0,\ i = 1, 2, \dots\}, \tag{7.66}$$

and

$$\{\theta_i | \theta_i^2 - i\theta_i + 1 = 0,\ i = 1, 2, \dots\}. \tag{7.67}$$

As a corollary of the following theorem, we show in fact that only finitely many $\theta_i$ satisfy any cyclotomic equations.



THEOREM 7.5.5. *Let $\theta_i$ be the collection of algebraic integers in (7.67) or (7.66) where each $\theta_i$ is chosen to lie between 0 and 1. Then there are at most a finite number of $\theta_i$ which satisfy cyclotomic equations and they are effectively computable.*

PROOF. Let $\phi_n(x, y)$ denote the $n$th cyclotomic polynomial in homogeneous form and let $\phi_0(x, y) = y$. Writing a general cyclotomic equation (in ideal form) satisfied by $\theta_i$ in terms of the ideals $(\phi_n(1, \theta_i))$, we have

$$(1) = \prod_{n=0}^{N}(\phi_n(1, \theta_i))^{C_n}, \qquad C_n \in \mathbf{Z}. \tag{7.68}$$

Taking norms of the above equation gives

$$1 = \prod_{n=0}^{N}\mathrm{N}(\phi_n(1, \theta_i))^{C_n}, \qquad C_n \in \mathbf{Z}.$$

First we prove that there exists a constant $I_0$ such that if $C_n \neq 0$ for some $n > 6$, then $i < I_0$. This follows by showing that $\mathrm{N}(\phi_n(1, \theta_i))$ has a primitive divisor for all $i \geq I_0$ if $n > 6$. Estimating $|\mathrm{N}(\phi_n(1, \theta))|$ as in (5) of [**16**],

$$\log|\mathrm{N}(\phi_n(1, \theta_i))| = \sum_\sigma \mu(n/m)\log|1 - (\theta_i^m)^\sigma| = \phi(n)\log(1/\theta_i) + \varepsilon_n(i),$$

where the sum over $\sigma$ is over the conjugates of $\theta_i$, $\mu$ denotes the Moebius function, and $\varepsilon_n(i)$ denotes a function which approaches 0 as $i \to \infty$. Since $1/\theta_i \to \infty$, there exists $I_0$ such that for all $i \geq I_0$,

$$|\mathrm{N}(\phi_n(1, \theta_i))| = (1/\theta_i)^{\phi(n)} > n^2. \tag{7.69}$$

Therefore, by the following lemma, $\phi_n(1, \theta_i)$ has a primitive divisor for all $n > 2(2^2 - 1) = 6$ if $i \geq I_0$, since otherwise (7.69) would be contradicted.

LEMMA 1 (Lemma 4 of [**16**]). *Let $A$ and $B$ be algebraic integers of a number field $K$ with $A/B$ of degree $d$. If $P$ is a prime ideal of $K$, $n > 2(2^d - 1)$, $P|(\phi_n(A, B))$, and $P$ is not a primitive divisor of $A^n - B^n$, then $\mathrm{ord}_P(\phi_n(A, B)) \leq \mathrm{ord}_P(n)$.*

Next we assume $N \leq 6$ and suppose $\theta_i$ satisfies $\theta_i^2 + i\theta_i - 1 = 0$. We need to show that only finitely many such $\theta_i$ satisfy a cyclotomic ideal equation of the form

$$(1) = (\phi_1(1, \theta_i))^{C_1}(\phi_2(1, \theta_i))^{C_2}(\phi_3(1, \theta_i))^{C_3}(\phi_4(1, \theta_i))^{C_4} \times (\phi_5(1, \theta_i))^{C_5}(\phi_6(1, \theta_i))^{C_6}.$$

Taking norms,

$$1 = (i)^{C_1}(-i)^{C_2}(i^2+3)^{C_3}(i^2+4)^{C_4}(i^4+7i^2+13)^{C_5}(i^2+3)^{C_6}.$$

It is straightforward to show that if $i > 2$, each $N(\phi_j)$, $j = 1, 2, \ldots, 6$, has a factor which does not divide any other $N(\phi_{j'})$, $j' < j$, with two exceptions, $N(\phi_2) = N(\phi_1)$ and $N(\phi_6) = N(\phi_3)$. This implies that only $C_n$, $n = 1, 2, 3, 6$, are nonzero and it is easy to show that in fact $C_1 = C_2$ and $C_3 = C_6$. Therefore,

(7.70) $(\phi_1(1, \theta_i))^{C_1}(\phi_3(1, \theta_i))^{C_3} = (\phi_2(1, \theta_i))^{C_2}(\phi_6(1, \theta_i))^{C_6}$, $c_n \in \mathbf{N} \cup \{0\}$.

If 3 does not divide $i$, $\{(\phi_1), (\phi_2)\}$ have no prime ideal factors in common with those of $\{(\phi_3), (\phi_6)\}$, so $(\phi_1) = (\phi_2)$. Thus $u(1-\theta_i) = 1+\theta_i$ for some unit $u$ in $\mathbf{Q}(\theta_i)$. Solving for $u$, we see that since

$$u = -2/i + \sqrt{i^2 + 4}/i \tag{7.71}$$

must be a quadratic integer, $i$ must be 1, 2 or 4.

Now suppose $3|i$. Then without loss of generality, we may suppose $\gcd((\phi_1), (\phi_3)) = P$, a prime ideal of norm 3. Then it is easy to show that $(\phi_2)|(3\phi_1)$, implying $i = 3, 6$ or 12.

Hence we have shown that in either case, only finitely many such $\theta_i$ can satisfy a cyclotomic equation. The argument for $\theta_i$ satisfying $\theta^2 - i\theta + 1 = 0$, (7.67), is similar. □

**7.6. Mahler's measure and Salem/Pisot numbers.** Mahler's measure $M$ is a natural generalization of Jensen's formula to polynomials in several variables. Its definition is

$$M(p) = \exp\left(\left(\frac{1}{2\pi}\right)^n \int_0^{2\pi} \cdots \int_0^{2\pi} \log|p(e^{i\theta_1}, \ldots, e^{i\theta_n}|d\theta_1 \cdots d\theta_n\right).$$

For a polynomial $p(x) = p_n x^n + \cdots + p_0$ of one variable, Jensen's formula shows that

$$M(p) = |p_n| \prod_{p(\sigma)=0} (\max\{|\sigma|, 1\}).$$

where the product runs over the roots of $p$. One important application of Mahler's measure is its connection with Lehmer's classical question:

> Given $\varepsilon > 0$, are there polynomials $p$ with integer coefficients in one variable for which $1 < M(p) < 1 + \varepsilon$?

There are two families of algebraic integers which have been important with regard to this question. A real algebraic integer $\theta > 1$ is a *Salem number* if all of its conjugates except its reciprocal lie on the unit circle. Similarly, $\theta$ is a *Pisot number* if its conjugates lie strictly inside the unit circle. In either case, if $p_\theta$ is the monic irreducible polynomial satisfied by $\theta$, then $M(\theta) \stackrel{\text{def}}{=} M(p_\theta) = \theta$. Boyd in [4, 5] has done extensive calculations and produced lists of Salem numbers with measure close to 1 in order to test Lehmer's conjecture. The noncyclotomic algebraic integer with the smallest known



measure is still Lehmer's example, the Salem number $\alpha \approx 1.1762808183$ satisfying

$$\alpha^{10} + \alpha^9 - \alpha^7 - \alpha^6 - \alpha^5 - \alpha^4 - \alpha^3 + \alpha^1 + 1 = 0.$$

An unexpected connection between Mahler's measure and values of Dirichlet $L$-series was also discovered recently when Smyth showed that the logarithm of Mahler's measure of a certain two variable polynomial gave essentially $L(2, \chi_{-3})$. More specifically,

$$\log(M(1 + x + y)) = \frac{3\sqrt{3}}{4\pi} L(2, \chi_{-3}) = L'(-1, \chi_{-3}),$$

where $\chi_{-3}$ is the odd quadratic character of conductor 3.

In this section, we investigate the connections between dilogarithms, $L$-series, cyclotomic equations and Mahler's measure. It turns out that

1. Using the multivariable identities for the dilogarithm, other polynomials with integer coefficients in two variables can be found where the log of Mahler's measure is a rational times the derivative of an $L$-series value.
2. Using the multivariable identity for the Bloch-Wigner function, a question about twisted $L$-series relations posed in [13] can now be resolved.
3. Salem and Pisot numbers which have small measure tend to satisfy a large number of cyclotomic equations and therefore are good candidates for bases of linear power relations of dilogarithms. Previously, Lewin's most prolific example of an algebraic integer base satisfying many ladder relations was $\omega$, the reciprocal of the smallest Pisot number. Now, the most prolific example is $\alpha$, the smallest known Salem number! Chapter 16 in this monograph examines the many cyclotomic equations and corresponding relations satisfied by $\alpha$.

The first two points are covered in the next section while the third is addressed in §7.6.2.

7.6.1. *Mahler's measure and L-series values.* If $\chi$ is any character of even order and conductor $N$, define the polynomial of degree $n_r = \phi(N)/2$,

$$r_\chi(x, t) = t\widehat{p_\chi}(x) - p_\chi(x),$$

where

$$p_\chi(x) = \prod_{\chi(k)=1} (1 - x\zeta_N^k),$$

and

$$\widehat{p_\chi}(x) = \prod_{\chi(k)=-1} (1 - x\zeta_N^k).$$

Then a direct calculation using the definition of Mahler's measure gives

THEOREM 7.6.1.

$$\log(M(r_\chi)) = \frac{1}{2\pi}\int_0^{2\pi}\left[\frac{1}{2\pi}\int_0^{2\pi}\log|r_\chi(e^{i\theta_1}, e^{i\theta_2}|d\theta_1\right]d\theta_2$$
$$= \frac{1}{2\pi}\mathrm{Re}\left[\sum_{\chi(k)=\pm1} i\chi(k)\sum_j[\mathrm{Li}_2(\mu_{2j}\zeta_N^k) - \mathrm{Li}_2(\mu_{2j-1}\zeta_N^k)]\right].$$

*Here* $\bigcup_j[\mu_{2j-1}, \mu_{2j}]$ *denotes the set of circular intervals on the unit circle where* $|p_\chi(x)| \ge |\widehat{p_\chi}(x)|$.

For details the reader should consult [13].

Now, if $\chi = \chi_{-N}$ is an odd quadratic character of conductor $N$, set $g_\chi = r_\chi\overline{r_\chi}$ where the bar denotes complex conjugation of the coefficients. Then $\log(M(g_\chi)) = 2\log(M(r_\chi))$ will have integer coefficients and a result in [13] states that $\log(M(g_\chi))$ will be a rational times $L'(-1, \chi_{-N})$ for $N = 3, 4, 8, 20, 24$, and 7. The only means to prove this result for $N = 7$ turns out to involve the multivariable identity for the dilogarithm given in Corollary 7.1.1. More specifically, we have

COROLLARY 7.6.2. *When the conductor* $N = 3, 4, 8, 20$ *or* 24,

$$\log(M(g_{g_{\chi_{-N}}})) = \frac{8 - 2\chi_{-N}(2)}{N}L'(-1, \chi_{-N}),$$

and

COROLLARY 7.6.3. *When the conductor* $N = 7$,

$$\log(M(g_{\chi_7})) = \log(M((y-1)^2(x^6+x^5+x+1) + (y^2+5y+1)(x^4+x^2) + (y^2+12y+1)x^3)$$
$$= \frac{8}{7}L'(-1, \chi_{-7}).$$

The above results follow from the following theorem in [13].

THEOREM 3 OF [13]. *Let* $\chi$ *be an odd quadratic character of conductor* $N$ *and let* $\Delta_\chi(t)$ *be the discriminant of* $r_\chi(x, t)$ *considered as a polynomial in* $x$. *Suppose that for some real number* $b < 1$, $\Delta_\chi(t) \ne 0$ *for all* $t$ *between* 0 *and* $b$. *Then for each such* $t$, $r_\chi(x, t) = 0$ *has* $n_r = \phi(N)/2$ *distinct roots* $r = r_1(t), r_2(t), \ldots, r_{n_r}(t)$, *all with absolute value one, and*

$$\sum_{l=1}^{n_r}\mathrm{Re}(L_{r_l(b)}(2, \chi)) = \frac{\mu(N)}{2}L(2, \chi). \tag{7.72}$$

Here $\mu(N)$ denotes the Moebius function defined as $\mu(1) = 1$ and $\mu(N) = (-1)^k$, for $N$ square-free, where $k$ is the number of distinct prime factors of $N$. Otherwise, $\mu(N) = 0$. Also $L_\rho(2, \chi)$ denotes a "twisted" $L$-function



defined as follows:

$$L_\rho(2, \chi) = \frac{1}{\tau(\chi)}\sum_{k=1}^{N}\overline{\chi(k)}\mathrm{Li}_2(\rho\zeta^k), \tag{7.73}$$

where $\zeta = e^{2\pi i/N}$ and $\tau(\chi)$ is the Gaussian sum for the character $\chi$.

It was observed in [13] that the identity in (7.72) seems to be true whenever all the roots $\{r_i(t)\}$ are on the unit circle, regardless of the discriminant condition. It is possible to prove this more general result using Corollary 7.2.1 involving Bloch's function rather than the corresponding result Corollary 7.1.1 for dilogarithms. We have

COROLLARY 7.6.4. *Let* $\chi$ *be an odd quadratic character of conductor* $N$. *Then whenever* $r_\chi(x, b) = 0$ *has* $n_r = \phi(N)/2$ *distinct roots* $x = r_1(b), r_2(b), \ldots, r_{n_r}(b)$ *of absolute value one,*

$$\sum_{l=1}^{n_r}\mathrm{Re}(L_{r_l(b)}(2, \chi)) = \frac{\mu(N)}{2}L(2, \chi). \tag{7.74}$$

PROOF. *Sketch.* From Corollary 7.2.1 applied to $r_\chi(x, 1-t)$, it follows that $A = 1$, $\xi_1 = \xi_2 = 0$, and with $a = 1$, we obtain

$$\sum_{l=1}^{n_r}\sum_{i=1}^{n_r}\chi(i)D_2(\zeta_N^i r_l(t))\Big|_1^b = 0.$$

Note that $\{r_l(1)\} = \{\zeta_N^{-l}|\chi(l) = 1\}$. Therefore the above left-hand side evaluated at 1 gives

$$\sum_{\{l|\chi(l)=1\}}\sum_{i=1}^{n_r}\chi(i)D_2(\zeta_N^{i-l}),$$

which from Lemma 13 and Corollary 15 of [13] equals

$$\frac{\mu(N)}{2}\mathrm{Im}(\tau(\chi))L(2, \chi),$$

since $D_2(x) = \mathrm{Im}(\mathrm{Li}_2(x))$ whenever $|x| = 1$.

Evaluating at $b$ where it is assumed that all $\{r_i(b)\}$ have absolute value one,

$$\sum_{l=1}^{n_r}\sum_{i=1}^{n_r}\chi(i)D_2(\zeta_N^i r_l(b))$$

becomes

$$\mathrm{Im}(\tau(\chi))\sum_{l=1}^{n_r}L_{r_l(b)}(2, \chi),$$

since

$$\sum_{i=1}^{n_r}\chi(i)\mathrm{Li}_2(\zeta_N^i x) = \mathrm{Im}(\tau(\chi))\mathrm{Re}(L_x(2, \chi)),$$

when $|x| = 1$ again by Lemma 13 of [13]. This gives (7.74). □

Now a direct application of the previous theorem to $r_\chi(x, 1-t)$ for $\chi = \chi_{-7}$, $a = 0$ and $b = -1$ shows that $1$, $\mu$ and $\bar{\mu}$ are the roots of $r_\chi(x, 2) = 0$ where $\mu = (-3 + \sqrt{-7})/4$ and therefore

$$\mathrm{Re}(L_\mu(2, \chi)) = -\tfrac{3}{4}L(2, \chi).$$

Then Corollary 7.6.3 follows from Theorem 7.6.1 as follows. The intervals on the unit circle where $|p_\chi(t)| \ge |\tilde{p}_\chi(t)|$ are $[1, \mu]$ and $[-1, \bar{\mu}]$. Therefore

$$\log(M(r_\chi)) = \frac{\sqrt{7}}{2\pi}\mathrm{Re}[L_1(2, \chi) + L_{-1}(2, \chi) - L_\mu(2, \chi) - L_{\bar{\mu}}(2, \chi)] = \frac{\sqrt{7}}{\pi}L(2, \chi).$$

Using the functional equation for Dirichlet $L$ functions, we can write $L(s, \chi)$ for $s = 2$ in terms of its derivative at $s = -1$ as

$$L(2, \chi) = \frac{4\pi}{7\sqrt{7}}L'(-1, \chi).$$

Thus

$$\log(M(g_{\chi_7})) = \frac{8}{7}L'(-1, \chi_{-7}),$$

giving Corollary 7.6.3.

7.6.2. *Salem and Pisot bases for linear power relations.* The reason that Salem and Pisot numbers which have small measure tend to satisfy a large number of cyclotomic equations is as follows. Recall that in the previous section, the estimate used to bound the size of the exponents $(N_0(d))$ of a cyclotomic equation satisfied by an algebraic integer of degree $d$ depended on a result by Schinzel. The crucial equation (5) from [16] states that for an algebraic integer $\theta$ of degree $d$ between 0 and 1,

$$|\mathrm{N}(\phi_n(1, \theta))| = \phi(n)w(1/\theta) + O(d + w(1/\theta))2^{v(n)}\log(n) \tag{7.75}$$

where the constant in $O$ depends only on $d$ and $v(n)$ takes on values in the positive integers. By Schinzel's definition, it turns out that $w(1/\theta) = \log(M(\theta))$, so (7.75) shows that Mahler's measure is related to the magnitude of $\log|\mathrm{N}(\phi_n(1, \theta))|$ which is in turn related to the number $N_0(d)$.

Roughly speaking, if one could hold all other variables constant in (7.75) while increasing $M(\theta)$ (this is of course not possible), then the value of $n$ such that $\log|\mathrm{N}(\phi_n(1, \theta))| > d\log(n)$ would decrease. Thus a large $M(\theta)$ would decrease the number of possible cyclotomic factors in any cyclotomic equation which $\theta$ satisfies. Hence it is worth looking for additional linear power relations among $\theta$ which have small measure.

It turns out that the smallest Salem number $\alpha \approx 1.1762808183$ satisfies many cyclotomic equations and many linear power relations. For example, the following quadratic polynomial

$$r(x, t) = (1-t)(1+\alpha x)(1-\alpha^2 x) - (1-\alpha^5 x)(1-\alpha^9 x)$$



gives the following linear power relation satisfied by $\alpha$:

$$\begin{aligned}\mathrm{Li}_2(\alpha^{18})/2 - \mathrm{Li}_2(\alpha^{17}) + \mathrm{Li}_2(\alpha^{16})/2 - \mathrm{Li}_2(\alpha^{13}) + 2\mathrm{Li}_2(\alpha^{11}) + \mathrm{Li}_2(\alpha^{10})\\ - 2\mathrm{Li}_2(\alpha^9) - \mathrm{Li}_2(\alpha^8)/2 + \mathrm{Li}_2(\alpha^7) - \mathrm{Li}_2(\alpha^5) - \mathrm{Li}_2(\alpha^4) + \mathrm{Li}_2(\alpha^3)\\ + 3\mathrm{Li}_2(\alpha^2)/2 - \mathrm{Li}_2(\alpha) - \log^2(\alpha)/2 = 0.\end{aligned}$$

$\alpha$ also satisfies many other linear power relations of higher degree. These are examined in more detail in Chapter 16.

Two other Salem numbers from Boyd's list of smallest Salem numbers produced linear power relations; their minimal polynomials are

$$x^{10} - x^6 - x^5 - x^4 + 1,$$

and

$$x^8 - x^5 - x^4 - x^3 + 1.$$

Several small Pisot numbers not included in [9] and [1] also produced linear powers relations; their minimal polynomials are:

$$\begin{gathered} x^4 - x^3 - x^2 - x - 1, \\ x^3 - 2x^2 + x - 1, \quad x^4 - x^3 - x^2 - 1, \\ x^5 - x^4 - x^3 - 1, \quad x^6 - x^5 - x^4 - 1. \end{gathered}$$

These were found by a cursory search of a small number of quadratic $r(x, t)$ using Theorem 7.4.1. It is possible that by analyzing more such $r(x, t)$, especially of larger degree, many more such relations could be discovered.

**7.7. Recent results for supernumary ladders.** Supernumary ladder relations for the dilogarithm are defined as those linear power relations not given by the generic formulae (Equations 2.21–2.26 in [10]) for which the base $u$ satisfies a polynomial of the form

$$u^p + u^q - 1 = 0.$$

The generic relations can be proven using Kummer's functional equation.

In this section, we apply Corollary 7.1.1 and Theorem 7.4.1 to find and prove the validity of several supernumary ladder relations. Each example gives a particular quadratic polynomial $r(x, t)$ of the form

$$(1-t)(1 \pm u^{m_1}x)(1 \pm u^{m_2}x) - (1 \pm u^{m_3}x)(1 \pm u^{m_4}x),$$

so that all the $\alpha_j$'s and $\beta_k$'s are powers of $u$. In addition $x_1(b)$ and $x_2(b)$ are given, namely the two roots of $r(x, b) = 0$, which are both powers of $u$ as well. In all cases $a = 1$ and the roots of $r(x, a)$ are obvious.

7.7.1. *The case* $(p, q) = (5, 2)$. The algebraic base $u$ between 0 and 1 satisfying the polynomial

$$u^5 + u^2 - 1 = 0, \tag{7.76}$$

gives a couple of supernumary relations for low powers of $u$ in the factors of $r(x, t)$. The first had been numerically predicted in (6.5) of [10]. Choosing $r(x, t)$ to be

$$(1-t)(1+u^2x)(1-u^7x)-(1-u^5x)(1-u^{12}x), \qquad x(u^3)=\{1, u^{-8}\},$$

gives

$$\begin{aligned}2\mathrm{Li}_2(u^3)+5\mathrm{Li}_2(u^6)-\mathrm{Li}_2(u^4)-3\mathrm{Li}_2(u^{12})+2\mathrm{Li}_2(u^7)-2\mathrm{Li}_2(u^1)\\ +2\mathrm{Li}_2(-u^{10})-4\mathrm{Li}_2(u^2)=-3\log^2(u)+\frac{\pi^2}{3}.\end{aligned} \tag{7.77}$$

The second gives an apparently new supernumary relation. Choosing $r(x, t)$ to be

$$(1-t)(1-u^5x)(1+u^{12}x)-(1+u^4x)(1-u^8x), \qquad x(-u^2)=\{1, u^{-14}\},$$

gives
(7.78)

$$\begin{aligned}&2\mathrm{Li}_2(u^5)-2\mathrm{Li}_2(u^9)+2\mathrm{Li}_2(-u^{12})-2\mathrm{Li}_2(-u^4)-5\mathrm{Li}_2(u^8)\\ &\quad -2\mathrm{Li}_2(-u^1)+2\mathrm{Li}_2(u^3)+2\mathrm{Li}_2(-u^{10})+2\mathrm{Li}_2(u^6)=4\log^2(u)-\frac{\pi^2}{3}.\end{aligned}$$

7.7.2. *The case* $(p, q)=(5, 4)$. The algebraic base $u$ between 0 and 1 satisfying the polynomial

$$u^5+u^4-1=0, \tag{7.79}$$

gives a supernumary relation which had been numerically predicted in (6.10) of [10]. Choosing $r(x, t)$ to be

$$(1-t)(1-u^1x)(1+u^{12}x)-(1-u^6x)(1-u^8x), \qquad x(-u^{-3})=\{1, -u^{-8}\},$$

gives

$$\begin{aligned}&-2\mathrm{Li}_2(-u^3)+2\mathrm{Li}_2(u^9)+2\mathrm{Li}_2(u^1)-\mathrm{Li}_2(u^{14})+\mathrm{Li}_2(u^{24})+4\mathrm{Li}_2(u^4)\\ &\quad +2\mathrm{Li}_2(u^5)+4\mathrm{Li}_2(u^7)-3\mathrm{Li}_2(u^{12})-3\mathrm{Li}_2(u^8)+2\mathrm{Li}_2(-u^2)\\ &\quad =79\log^2(u)-\frac{5}{6}\pi^2.\end{aligned} \tag{7.80}$$

7.7.3. *The case* $(p, q)=(6, 1)$. The algebraic base $u$ between 0 and 1 satisfying the polynomial

$$u^6+u-1=0, \tag{7.81}$$

gives at least one apparently new supernumary relation for low powers of $u$ in the factors of $r(x, t)$. Choosing $r(x, t)$ to be

$$(1-t)(1-u^3x)(1-u^{14}x)-(1-u^6x)(1+u^7x), \qquad x(-u^1)=\{1, u^{-10}\},$$



gives
(7.82)

$$\begin{aligned}&2\mathrm{Li}_2(-u^1)+2\mathrm{Li}_2(u^2)+2\mathrm{Li}_2(u^3)+2\mathrm{Li}_2(u^7)+2\mathrm{Li}_2(u^4)-\mathrm{Li}_2(u^8)-\mathrm{Li}_2(u^6)\\ &\quad =20\log^2(u)-\frac{\pi^2}{3}.\end{aligned}$$

7.7.4. *The case* $(p, q)=(7.5)$. The algebraic base $u$ between 0 and 1 satisfying the polynomial

$$u^7+u^5-1=0, \tag{7.83}$$

gives at least one apparently new supernumary relation for low powers of $u$ in the factors of $r(x, t)$. Choosing $r(x, t)$ to be

$$(1-t)(1+u^4x)(1+u^4x)-(1-u^5x)(1+u^{12}x), \qquad x(u^2)=\{1, -u^{-11}\},$$

gives

$$\begin{aligned}4\mathrm{Li}_2(u^2)-2\mathrm{Li}_2(u^4)+6\mathrm{Li}_2(u^8)-4\mathrm{Li}_2(u^7)-2\mathrm{Li}_2(u^5)+2\mathrm{Li}_2(-u^6)\\ -2\mathrm{Li}_2(-u^{12})-6\mathrm{Li}_2(u^1)=-37\log^2(u)+\frac{2}{3}\pi^2.\end{aligned} \tag{7.84}$$

7.7.5. *Analytic relations for* $\Delta$. The algebraic base $\Delta$ between 0 and 1 satisfying the polynomial

$$\Delta^4-2\Delta^3+\Delta^2-2\Delta+1=0 \tag{7.85}$$

is not strictly a candidate for a supernumary base, but comes from Browkin's (5.201), as discussed in §5.8.5. By examining $r(x, t)$ for small powers of $\Delta$ in the factors, it is possible to discover and prove 8 independent dilogarithm relations. Choosing $r(x, t)$ to be

$$\begin{aligned}&(1-t)(1+\Delta^1x)(1-\Delta^2x)-(1+\Delta^4x)(1+\Delta^7x), && x(\Delta^8)=\{1, \Delta^5\},\\ &(1-t)(1+\Delta^1x)(1-\Delta^2x)-(1+\Delta^5x)(1+\Delta^5x), && x(\Delta^7)=\{1, \Delta^4\},\\ &(1-t)(1+\Delta^1x)(1-\Delta^4x)-(1+\Delta^2x)(1+\Delta^7x), && x(\Delta^4)=\{1, \Delta^{-1}\},\\ &(1-t)(1+\Delta^1x)(1-\Delta^4x)-(1+\Delta^3x)(1+\Delta^5x), && x(\Delta^3)=\{1, \Delta^{-2}\},\\ &(1-t)(1+\Delta^2x)(1-\Delta^3x)-(1+\Delta^5x)(1+\Delta^6x), && x(\Delta^6)=\{1, \Delta^1\},\\ &(1-t)(1+\Delta^4x)(1-\Delta^7x)-(1+\Delta^5x)(1+\Delta^{12}x), && x(\Delta^6)-\{1, \Delta^{-5}\},\\ &(1-t)(1-\Delta^5x)(1+\Delta^9x)-(1-\Delta^4x)(1-\Delta^{15}x), && x(\Delta^5)=\{1, \Delta^{-9}\},\end{aligned}$$

and

$$(1-t)(1+\Delta^9x)(1-\Delta^{10}x)-(1-\Delta^{14}x)(1-\Delta^{15}x), \quad x(\Delta^{10})=\{1, \Delta^{-9}\},$$

gives successively

$$\begin{aligned}3\mathrm{Li}_2(\Delta^8)+2\mathrm{Li}_2(-\Delta^1)+3\mathrm{Li}_2(\Delta^{12})+4\mathrm{Li}_2(\Delta^7)+2\mathrm{Li}_2(\Delta^3)\\ +3\mathrm{Li}_2(\Delta^4)+2\mathrm{Li}_2(-\Delta^5)-2\mathrm{Li}_2(-\Delta^9)-\mathrm{Li}_2(\Delta^{14})\\ -\mathrm{Li}_2(\Delta^{24})=\log^2(\Delta),\end{aligned} \tag{7.86}$$

$$\begin{aligned}4\mathrm{Li}_2(\Delta^7) + 2\mathrm{Li}_2(-\Delta^1) - 2\mathrm{Li}_2(-\Delta^5) + 2\mathrm{Li}_2(\Delta^2) + 2\mathrm{Li}_2(\Delta^6)\\ + 4\mathrm{Li}_2(\Delta^4) + 4\mathrm{Li}_2(-\Delta^3) - 4\mathrm{Li}_2(-\Delta^9) = \log^2(\Delta),\end{aligned} \tag{7.87}$$

$$\begin{aligned}6\mathrm{Li}_2(\Delta^4) + 2\mathrm{Li}_2(\Delta^1) + 5\mathrm{Li}_2(\Delta^6) - 4\mathrm{Li}_2(-\Delta^2) - 2\mathrm{Li}_2(-\Delta^7)\\ - \mathrm{Li}_2(\Delta^{12}) = 5\log^2(\Delta) - \frac{\pi^2}{2},\end{aligned} \tag{7.88}$$

$$\begin{aligned}8\mathrm{Li}_2(\Delta^3) - 2\mathrm{Li}_2(-\Delta^1) + 4\mathrm{Li}_2(\Delta^4) + 4\mathrm{Li}_2(\Delta^2) - 2\mathrm{Li}_2(\Delta^6)\\ - 2\mathrm{Li}_2(-\Delta^5) = 7\log^2(\Delta) - \tfrac{2}{3}\pi^2,\end{aligned} \tag{7.89}$$

$$\begin{aligned}10\mathrm{Li}_2(\Delta^6) + 4\mathrm{Li}_2(-\Delta^2) + 4\mathrm{Li}_2(\Delta^4) - 2\mathrm{Li}_2(-\Delta^5) - 2\mathrm{Li}_2(\Delta^{12})\\ - 2\mathrm{Li}_2(-\Delta^7) = \log^2(\Delta),\end{aligned} \tag{7.90}$$

$$\begin{aligned}4\mathrm{Li}_2(\Delta^6) + 3\mathrm{Li}_2(\Delta^8) + 3\mathrm{Li}_2(\Delta^2) + 4\mathrm{Li}_2(\Delta^7) + 4\mathrm{Li}_2(\Delta^1)\\ - 3\mathrm{Li}_2(\Delta^4) - 2\mathrm{Li}_2(-\Delta^{12}) - \mathrm{Li}_2(\Delta^{14}) = 4\log^2(\Delta) - \frac{\pi^2}{2},\end{aligned} \tag{7.91}$$

$$\begin{aligned}10\mathrm{Li}_2(\Delta^5) - 4\mathrm{Li}_2(\Delta^4) + 2\mathrm{Li}_2(-\Delta^9) - 2\mathrm{Li}_2(\Delta^1) + \mathrm{Li}_2(\Delta^{10})\\ + \mathrm{Li}_2(\Delta^{12}) - 2\mathrm{Li}_2(\Delta^{15}) - 4\mathrm{Li}_2(\Delta^6) = -\log^2(\Delta) + \frac{\pi^2}{6},\end{aligned} \tag{7.92}$$

and

$$\begin{aligned}7\mathrm{Li}_2(\Delta^{10}) + 2\mathrm{Li}_2(-\Delta^9) + 2\mathrm{Li}_2(\Delta^1) + \mathrm{Li}_2(\Delta^{12}) + 2\mathrm{Li}_2(\Delta^4)\\ - 2\mathrm{Li}_2(\Delta^5) - 2\mathrm{Li}_2(\Delta^{14}) - 2\mathrm{Li}_2(\Delta^{15}) - 4\mathrm{Li}_2(\Delta^6)\\ = \log^2(\Delta) - \frac{\pi^2}{6}.\end{aligned} \tag{7.93}$$

These comprise the six relations $(5 \cdot 210)$ to $(5 \cdot 215)$, together with two additional results involving indices 18 and 24.

## REFERENCES

1. M. Abouzahra and L. Lewin, *The polylogarithm in algebraic number fields*, J. Number Theory **21** (1985), 214–244.
2. G. D. Birkhoff and H. S. Vandiver, *On the integral divisors of $a^n - b^n$*, Ann. Math. (2) **5** (1904), 173–180.
3. S. Bloch, *Applications of the dilogorithm function in algebraic K-theory and algebraic geometry*, Internat. Sympos. on Algebraic Geometry, Kyoto, Springer-Verlag, New York, 1977, pp. 103–114.
4. D. Boyd, *Small Salem numbers*, Duke Math. J. **44** (1977), 315–328.
5. ——, *Pisot and Salem numbers in intervals of the real line*, Math. Comp. **32** (1978), 1244–1260.
6. W. Fulton, *Algebraic curves*, W. A. Benjamin, Inc., New York, 1969.
7. L. Lewin, *Dilogarithms and associated functions*, North Holland, Amsterdam, 1981.
8. ——, *The dilogarithm in algebraic fields*, J. Austral. Math. Soc. A **33** (1981), 302–330.



9. ——, *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), pp. 345–373.
10. L. Lewin and M. Abouzahra, *Supernumary polylogarithm ladders and related functional equations*, Aequationes Math. **39** (1990), 210–253.
11. J. Milnor, *On polylogarithms, Hurwitz zeta functions, and the Kubert identities*, Enseign. Math. **39** (1983), 281–322.
12. D. Ramakrishnan, *Analogs of the Bloch-Wigner function for higher polylogarithms*, Applications of algebraic K-theory to algebraic geometry and number theory, Boulder, Co., Springer-Verlag, New York, 1983.
13. G. Ray, *Relations between Mahler's measure and values of L-series*, Canad. J. Math. XXXIX (1987), 694–732.
14. L. J. Rogers, *On function sum theorems connected with the series $\sum_{n=1}^{\infty} x^n/n^2$*, Proc. London Math. Soc. **4** (1907), 169–189.
15. H. F. Sandham, *A logarithmic transcendent*, J. London Math. Soc. **24** (1949), 83–91.
16. A. Schinzel, *Primitive divisors of the expression $A^n - B^n$ in algebraic number fields*, J. Reine Angew. Math. **268/269** (1974), 27–33.
17. D. Zagier, *The remarkable dilogarithm*, J. Math. Phys. Sci. **22** (1988), 131–145.
18. ——, *The Bloch-Wigner-Ramakrishnan polylogarithm function*, Math. Ann. (1989), 613–624.

CHAPTER 8

# Functional Equations of Hyperlogarithms

G. WECHSUNG

**8.1. Hyperlogarithms.** Following E. Kummer [1] we call a function of the form

$$(8.1)\qquad f(z)=\int_{b_n}^{z}dt_n r_n(t_n)\int_{b_{n-1}}^{t_n}dt_{n-1}r_{n-1}(t_{n-1})\cdots\int_{b_1}^{t_2}dt_1 r_1(t_1),$$

with rational functions $r_1(z),\ldots,r_n(z)$ of the complex variable $z$ a *logarithmic integral of order* $n$ provided that the values $b_1,\ldots,b_n$ are chosen in such a way that the integral exists. Clearly, by making use of decomposition into partial fractions and partial integration we can reduce (8.1) to integrals of the form

$$(8.2)\qquad F_n\left(\begin{matrix}a_1,\ldots,a_n\\ b_1,\ldots,b_n\end{matrix}\middle| z\right)=\int_{b_n}^{z}\cdots\int_{b_2}^{t_3}\int_{b_1}^{t_2}\frac{dt_1}{t_1-a_1}\frac{dt_2}{t_2-a_2}\cdots\frac{dt_n}{t_n-a_n},$$

and further functions which can be written using less than $n$ integrations. As pointed out by H. Poincaré [2] the functions (8.2) are of importance for solving differential equations. Following Poincaré's ideas, Lappo-Danilewski was able to find an algorithmic solution of the Riemann-Poincaré problem [3]. In [3] the functions (8.2) have been called *hyperlogarithms* (of order $n$). Notice that

$$\mathrm{Li}_n(z)=-F_n\left(\begin{matrix}1,0,\ldots,0\\ 0,0,\ldots,0\end{matrix}\middle| z\right).$$

Let $\mathrm{LI}_n$ be the set of all logarithmic integrals of order $n$. It is evident that $\mathrm{LI}_n$ is a vector space over the field $C$ of the complex numbers ($n=1,2,\ldots$). A function $f$ is called a *proper logarithmic integral of order* $n$ if $f\in\mathrm{LI}_n$, but $f\notin\mathrm{LI}_{n-1}$. In §8.2 we shall prove that the hierarchy

$$\mathrm{LI}_1\subset\mathrm{LI}_2\subset\cdots\subset$$

is proper. This means that for every $n\geq 1$ there exist proper logarithmic integrals of order $n$.

**8.2. Logarithmic singularities.**

8.2.1. *Singularities of the hyperlogarithms.* A profound understanding of the behaviour of $F_n(\begin{smallmatrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{smallmatrix}|z)$ at its singularities $a_1, \ldots, a_n$ and $\infty$ is required for all questions we shall deal with in the following.

Our first theorem concerns the order of growth of the hyperlogarithms in the neighbourhood of their singularities.

THEOREM 8.1. *If the $a_i$ are pairwise different, then*

$$F_n\left(\begin{matrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{matrix}\middle| z\right) = 0((z-a_i)^{n-i}\log(z-a_i)) \quad \textit{for } z \to a_i,$$

$$F_n\left(\begin{matrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{matrix}\middle| z\right) = 0((\log z)^n) \qquad \textit{for } z \to \infty.$$

*$a_1, \ldots, a_n$ and $\infty$ are the only singularities of $F_n(\begin{smallmatrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{smallmatrix}|z)$.*

The proof follows by induction using the recursive relationship

$$F_n\left(\begin{matrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{matrix}\middle| z\right) = \int_{b_n}^{z} \frac{F_{n-1}(\begin{smallmatrix} a_1,\ldots,a_{n-1} \\ b_1,\ldots,b_{n-1} \end{smallmatrix}|t)}{t-a_n}\,dt. \tag{8.3}$$

The Riemann surface of $F_n(\begin{smallmatrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{smallmatrix}|z)$ is a covering surface $R(a_1, \ldots, a_n, \infty)$ of $C' = C\backslash\{a_1, \ldots, a_n, \infty\}$. In order to define unique branches of $F_n$ over $C'$ we cut the complex plane from each of the points $a_1, \ldots, a_n$ to $\infty$ without crossings. The main branch of $F_n$ arises from the main branch of $F_{n-1}$ by integrating only along paths which do not cross the cuts, and the main branch of $F_1(\begin{smallmatrix} a_1 \\ b_1 \end{smallmatrix}|z)$ is uniquely determined by $F_1(\begin{smallmatrix} a_1 \\ b_1 \end{smallmatrix}|b_1) = 0$.

The branching behaviour of $F_n$ is completely determined by

THEOREM 8.2. *Starting at a point $z$ with the value $F_n(\begin{smallmatrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{smallmatrix}|z)$ of the main branch and going around $a_j$ (and not around $a_k$, $k \neq j$) counterclockwise one reaches the value*

$$F_n\left(\begin{matrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{matrix}\middle| z\right) + \Delta_j F_n$$

*where, for $1 \le j \le n$,*

$$\Delta_j F_n = 2\pi i F_{j-1}\left(\begin{matrix} a_1,\ldots,a_{j-1} \\ b_1,\ldots,b_{j-1} \end{matrix}\middle| a_j\right) F_{n-j}\left(\begin{matrix} a_{j+1},\ldots,a_n \\ a_i,\ldots,a_i \end{matrix}\middle| z\right) \tag{8.4}$$

*and $F_0$ is defined to be equal to 1.*

PROOF (INDUCTION).

$n = 1$. $\Delta_1 F_1 = 2\pi i$ describes the known behaviour of

$$F_1\left(\begin{matrix} a \\ b \end{matrix}\middle| z\right) = \log\left(\frac{z-a}{b-a}\right) \quad \text{at } z = a.$$

Assume (8.4) to be correct for $n$. We prove this formula to be correct for $n+1$.



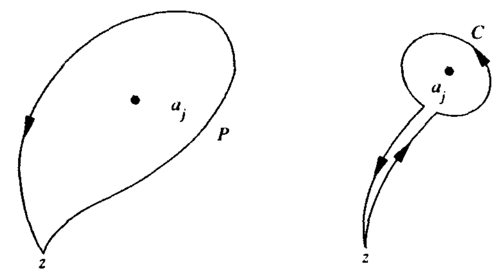



Figure 8.1. Deforming the path $P$.

CASE 1. $1 \le j \le n$.

$$\Delta_j F_{n+1} = \oint_P \frac{F_n(\begin{smallmatrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{smallmatrix}|t)}{t-a_{n+1}}\,dt.$$

The integral has to be taken along a path $P$ leading from $z$ back to $z$ and going around $a_j$ (and not around another $a_k$, $k \neq j$) counterclockwise. This path may be deformed as shown in Figure 8.1.

Here $C$ is a small circle around $a_j$. Because of Theorem 8.1 the integral along $C$ tends to zero if the radius of $C$ tends to zero. There remains

$$\begin{aligned} \Delta_j F_{n+1} &= \int_z^{a_j} \frac{F_n(\begin{smallmatrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{smallmatrix}|t)}{t-a_{n+1}}\,dt + \int_{a_j}^{z} \frac{F_n(\begin{smallmatrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{smallmatrix}|z) + \Delta_j F_n}{t-a_{n+1}}\,dt \\ &= \int_{a_j}^{z} \frac{\Delta_j F_n\,dt}{t-a_{n+1}}, \end{aligned}$$

and by the induction hypothesis we get

$$= 2\pi i F_{j-1}\left(\begin{matrix} a_1,\ldots,a_{j-1} \\ b_1,\ldots,b_{j-1} \end{matrix}\middle| a_j\right) F_{n-j+1}\left(\begin{matrix} a_{j+1},\ldots,a_{n+1} \\ a_j,\ldots,a_j \end{matrix}\middle| z\right).$$

CASE 2. $j = n+1$.

$$\Delta_{n+1} F_{n+1} = 2\pi i F_n\left(\begin{matrix} a_1,\ldots,a_n \\ b_1,\ldots,b_n \end{matrix}\middle| a_{n+1}\right)$$

follows from the residue theorem.

Theorem 8.2 allows a complete understanding of the hyperlogarithms on their Riemann surface. One reaches new branches by analytic continuation along arbitrary paths surrounding the branch points. Theorem 8.2 describes the additive terms which arise in this way. Note that these terms are also hyperlogarithms which themselves also yield additive contributions on further analytic continuation. A detailed study of the Riemann surfaces is necessary only for the polylogarithms (see §8.2.2).

Now we are able to classify the singularities of logarithmic integrals. A branch point $a$ of a logarithmic integral $f(z)$ is called a *logarithmic singularity of order* $k$ if analytic continuation of $f(z)$ along a path surrounding $a$ (and no other branch point) leads to $f(z)+g(z)$ where $g(z)$ is a proper logarithmic integral of order $k-1$. For $k=1$, $g(z)$ must be a constant. (8.4) shows

COROLLARY 8.3. *If $f \in \mathrm{LI}_n$, then $f$ has logarithmic singularities of order no greater than $n$.*

This allows us to prove

THEOREM 8.9. *$F_n(\begin{smallmatrix} a_1, \dots, a_n \\ b_1, \dots, b_n \end{smallmatrix}|z)$ is a proper logarithmic integral of order $n$.*

PROOF (INDUCTION).
The logarithm is known not to be expressible by rational functions and hence is a proper logarithmic integral of order 1.

Assume $F_n$ to be a proper logarithmic integral of order $n$. Then, by Theorem 8.2, $F_{n+1}$ has a logarithmic singularity of order $n+1$. But then, by Corollary 8.3, the function $F_{n+1}$ cannot belong to $\mathrm{LI}_n$. So $F_{n+1} \in \mathrm{LI}_{n+1} \setminus \mathrm{LI}_n$ which shows $F_{n+1}$ to be a proper logarithmic integral of order $n+1$.

8.2.2. *Singularities of the polylogarithms.* For $\mathrm{Li}_n(z) = -F_n(\begin{smallmatrix} 1, 0, \dots, 0 \\ 0, 0, \dots, 0 \end{smallmatrix}|z)$ Theorem 8.1 remains valid for $z \to 0$, and for $z=0$ we get $\mathrm{Li}_n(0)=0$ which follows from $\mathrm{Li}_n(z) = \sum_{m=1}^{\infty} z^m/m^n$. This shows that the main branch of $\mathrm{Li}_n(z)$ is not split at $z=0$. The same argument as in the proof of Theorem 8.2 yields

$$\Delta_1 \mathrm{Li}_n = -\frac{2\pi i}{(n-1)!}(\log z)^{n-1}. \tag{8.5}$$

This allows a complete understanding of the Riemann surface of the polylogarithms. The Riemann surface of $\mathrm{Li}_2(z)$ has been described by N. Nielsen [4]. The main branch of $\mathrm{Li}_n(z)$ is defined by $0 \le \arg(z-1) < 2\pi$ and $\mathrm{Li}_n(0)=0$. This means a cut from 1 to $\infty$ along the real axis, and the upper border of this cut belongs to the main branch. We get the best insight into the branching behaviour of $\mathrm{Li}_n$ by introducing $n$-fold indices for differentiating the branches of $\mathrm{Li}_n(z)$. Let $\mathrm{Li}_n^{(0,\dots,0)}(z)$ denote the main branch as defined above and let

$$\begin{aligned} \mathrm{Li}_n^{(k_0,\dots,k_{n-1})}(z) &= \mathrm{Li}_n^{(0,\dots,0)}(z) \\ &+ \frac{1}{(n-1)!}\sum_{m=0}^{n-1} k_m \binom{n-1}{m}(2\pi i)^{m+1}(\log^{(0)} z)^{n-m-1}. \end{aligned} \tag{8.6}$$

where $\log^{(0)} z$ is the main branch of the logarithm.



For the formulation of the final result concerning the structure of the Riemann surface of $\mathrm{Li}_n$ we need the symbolic notation

$$(1+k)^0 = k_0,$$
$$(1+k)^m = k_0 + \binom{m}{1}k_1 + \cdots + \binom{m}{n}k_m.$$

THEOREM 8.5. *The branches of $\mathrm{Li}_n(z)$ are linked along the cuts $\mathrm{Re}(z) > 1$ and $\mathrm{Re}(z) < 0$ as indicated in Figure 8.2.*

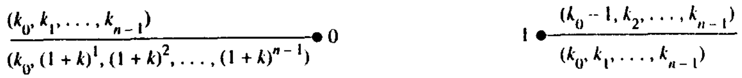



FIGURE 8.2. The Riemann surface of $\mathrm{Li}_n(z)$. The lines $\mathrm{Re}(z) < 0$ and $\mathrm{Re}(z-1) > 0$ are the cuts. At the cuts it is indicated which branches of $\mathrm{Li}_n(z)$ are pasted together.

REMARKS. 1. This figure means for instance: If one surrounds the point 1 counterclockwise in such a way that the point 0 is not surrounded then this path ends in branch $(k_0-1, \dots, k_{n-1})$ if it starts in $(k_0, \dots, k_{n-1})$.

2. This shows that not only the main branch is unramified at $z=0$, but all branches with indices $(0, \dots, 0, k_{n-1})$.

PROOF OF THEOREM 8.5. Formula 8.6 shows that only $\mathrm{Li}_n^{(0,\dots,0)}(z)$ yields a contribution if $z=1$ is surrounded. Formula (8.5) shows that this contribution is the additive term $-(1/(n-1)!)(\log z)^{n-1}$ which shows the correctness of the neighbourhood indicated at the cut $\mathrm{Re}(z) > 0$.

If $z=0$ is surrounded counterclockwise, then formula (8.6) shows that the new branch

$$\mathrm{Li}_n^{(0,\dots,0)}(z) + \frac{1}{(n-1)!}\sum_{m=0}^{n-1} k_m \binom{n-1}{m}(2\pi i)^{m-1}[\log^{(0)} z + 2\pi i]^{n-m-1}$$

is reached. We consider the sum:

$$\begin{aligned} &\sum_{m=0}^{n-1} k_m \binom{n-1}{m}(2\pi i)^{m+1}[\log^{(0)} z + 2\pi i]^{n-m-1} \\ &= \sum_{m=0}^{n-1}\sum_{j=0}^{n-m-1} k_m \binom{n-1}{m}\binom{n-m-1}{j}(2\pi i)^{m+j+1}(\log^{(0)} z)^{n-m-j-1}. \end{aligned}$$

Because of the identity

$$\binom{n-1}{m}\binom{n-m-1}{j} = \binom{n-1}{j+m}\binom{j+m}{m},$$

we get

$$= \sum_{m=0}^{n-1} \sum_{j=0}^{n-m-1} k_m \binom{n-1}{j+m} \binom{j+m}{m} (2\pi i)^{m+j+1} (\log^{(0)} z)^{n-(m+j)-1}$$

$$= \sum_{r=0}^{n-1} \sum_{m=0}^{r} \binom{r}{m} k_m \binom{n-1}{r} (2\pi i)^{r+1} (\log^{(0)} z)^{n-r-1}.$$

This shows that the indices of the new branch are

$$k'_r = \sum_{m=0}^{r} \binom{r}{m} k_m = (1+k)^r \quad \text{for } r = 0, \dots, n-1.$$

**8.3. The linear spaces $\mathrm{LI}_n$ and $\mathrm{PLI}_n$.** When we are dealing with logarithmic integrals of order $n$ we are often interested only in that part which is a proper logarithmic integral of order $n$. This is well described by the quotient space $\mathrm{PLI}_n = \mathrm{LI}_n / \mathrm{LI}_{n-1}$, the space of proper logarithmic integrals of order $n$. More precisely, the elements of $\mathrm{PLI}_n$ are the equivalence classes modulo $\mathrm{LI}_{n-1}$, $f^* = \{f + g : g \in \mathrm{LI}_{n-1}\}$.

If $f$ belongs to $\mathrm{LI}_{n-1}$, then $f^* = 0$. Note that $F_n(\begin{smallmatrix} a_1, \dots, a_n \\ b_1, \dots, b_n \end{smallmatrix} | z)$ and $F_n(\begin{smallmatrix} a_1, \dots, a_n \\ c_1, \dots, c_n \end{smallmatrix} | z)$ differ only by a logarithmic integral from $\mathrm{LI}_{n-1}$. Hence, they belong to the same equivalence class, and the following notation is justified

$$F_n \left( \begin{matrix} a_1, \dots, a_n \\ b_1, \dots, b_n \end{matrix} \middle| z \right)^* = F_n(a_1, \dots, a_n; z).$$

These equivalence classes will also be called hyperlogarithms in the sequel. In §8.2 we have seen that the branches of a logarithmic integral differ from each other by the function which is to be added to the function when a singularity is surrounded. We call them *branching terms.*

Let $f^* \in \mathrm{PLI}_n$ and $f_0 \in f^*$. Then an arbitrary element of $f^*$ can be represented in the form $f_0 + g$ where $g \in \mathrm{LI}_{n-1}$. This and Corollary 8.3 show: If $\varphi_0$ is the branching term of $f_0$ at one of its singularities $a$, then each element of $f^*$ has a branching term in $\varphi_0^* \in \mathrm{PLI}_{n-1}$.

This allows one to describe the structure of the linear space $\mathrm{PLI}_{n-1}$ as follows. First of all, if $X$ is a linear space over some field $F$ and $A$ is an arbitrary set, then $X^A$ is the linear space whose elements are all mappings from $A$ into $X$. Addition and scalar multiplication in $X^A$ are defined as follows: Let $f, g \in X^A$ and $a \in A$ and $\alpha \in F$. Then

$$(f+g)(a) = f(a) + g(a),$$
$$(\alpha f)(a) = \alpha f(a).$$

By $(X^A)_{\mathrm{fin}}$ we understand the set of all elements $f$ from $X^A$ such that $f(a) \neq 0$ is true only for finitely many elements. The elements from $(X^A)_{\mathrm{fin}}$ can be represented in the form

$$\sigma = \begin{pmatrix} a_1, \dots, a_m \\ x_1, \dots, x_m \end{pmatrix},$$



which means the mapping $\sigma : A \to X$ defined by $\sigma(a_i) = x_i$ for $i = 1, \dots, m$.

With this notation we define a mapping

$$\beta : \mathrm{PLI}_n \to (\mathrm{PLI}^C_{n-1})_{\mathrm{fin}}$$

which is defined as follows

$$\beta(f^*) = \begin{pmatrix} a_1, \dots, a_m \\ \varphi_1^*, \dots, \varphi_m^* \end{pmatrix}, \tag{8.7}$$

if and only if $f$ has exactly the logarithmic singularities $a_1, \dots, a_m$ of order $n$ with branching terms $\varphi_1 + \psi_1, \dots, \varphi_m + \psi_m$, respectively, where $\psi_1, \dots, \psi_m$ are in $\mathrm{LI}_{n-2}$. Now we can prove

THEOREM 8.6. *$\beta$ is an isomorphism between $\mathrm{PLI}_n$ and $(\mathrm{PLI}^C_{n-1})_{\mathrm{fin}}$.*

PROOF. 1. Clearly, $\beta$ is a homomorphism.

2. $\beta$ is injective. To show this, let $\beta(f^*) = \beta(g^*)$. This implies $\beta(f^* - g^*) = 0$. This shows $f - g$ does not have logarithmic singularities of order $n$. Hence, $f - g \in \mathrm{LI}_{n-1}$ and thus $f^* = g^*$.

3. Finally, $\beta$ is surjective. Let $\sigma \in (\mathrm{PLI}^C_{n-1})_{\mathrm{fin}}$. Then $\sigma$ is representable in the form $\sigma = \sigma_1 + \dots + \sigma_m$, where $\sigma_i = \binom{a_i}{\varphi_i^*}$. As pointed out in §8.1 we can restrict ourselves to $\varphi_i^* = F_{n-1}(c_1, \dots, c_{n-1}; z)$. By Theorem 8.2, $f_i^* = (1/2\pi i) F_n(a_i, c_1, \dots, c_{n-1}; z)$ is an element in $\mathrm{PLI}_n$ such that $\beta(f_i^*) = \binom{a_i}{\varphi_i^*}$. As $\beta$ is a homomorphism, we get $\beta(f_1^* + \dots + f_m^*) = \sigma$. Thus every element of $(\mathrm{PLI}^C_{n-1})_{\mathrm{fin}}$ is the image of some element of $\mathrm{PLI}_n$ under $\beta$.

From §8.1 it follows that $\mathrm{PLI}_n$ is generated by the equivalence classes $F_n(a_1, \dots, a_n; z)$ of the hyperlogarithms. By induction on $n$ we see immediately that they are linearly independent.

The set of functions $\{L_1(a; z) : a \in C\}$ is linearly independent. If the set of all hyperlogarithms $F_{n-1}(a_1, \dots, a_{n-1}; z)$ is linearly independent (induction hypothesis), then the same is true for the set of the hyperlogarithms of order $n$. This shows

THEOREM 8.7. *The hyperlogarithms $F_n(a_1, \dots, a_n; z)$ form a basis of $\mathrm{PLI}_n$.*

**8.4. Functional equations of hyperlogarithms.**

8.4.1. *Generalizing Euler's formula.* Euler's [5] famous formula

$$\mathrm{Li}_2(z) + \mathrm{Li}_2(1-z) = \frac{\pi^2}{6} - \log(z)\log(1-z) \tag{8.8}$$

(see also §1.1.2) can be understood to be a very special case of a general relationship between hyperlogarithms.

Let $f$ and $g$ be two functions with branching terms $\varphi$ and $\psi$, respectively, at a common singular point $a$. Then, $fg$ has the branching term $f\psi + g\varphi + \varphi\psi$ at $a$.

For the formulation of the next result we need a definition. For $m < n$ let $P_{m,n}$ be the set of all permutations $\pi$ of $\{1, \ldots, n\}$ which have the properties

(1) The numbers $1, \ldots, m$ occur in the natural order in $\pi$.
(2) The numbers $m+1, \ldots, n$ occur in the natural order in $\pi$.

EXAMPLE. $(1, 3, 2, 4)$, $(3, 4, 1, 2) \in P_{2,4}$ and $(2, 1, 3, 4) \notin P_{2,4}$.

The next result shows that products of hyperlogarithms are also hyperlogarithms. It is convenient to define $F_m(a_1, \ldots, a_m; z)F_{n-m}(a_{m+1}, \ldots, a_n; z)$ to be the equivalence class

$$F_m\left(\begin{matrix} a_1, \ldots, a_m \\ b_1, \ldots, b_m \end{matrix}\middle| z\right) F_{n-m}\left(\begin{matrix} a_{m+1}, \ldots, a_n \\ b_{m+1}, \ldots, b_n \end{matrix}\middle| z\right) + \mathrm{LI}_{n-1}.$$

THEOREM 8.8.

$$\begin{aligned} F_m(a_1, \ldots, a_m; z)&F_{n-m}(a_{m+1}, \ldots, a_n; z) \\ &= \sum_{\sigma \in P_{mn}} F_n(a_{\sigma(1)}, \ldots, a_{\sigma(n)}; z). \end{aligned}$$

PROOF (INDUCTION ON $n$).

1. $n = 2$. Using the mapping $\beta$ (see (8.7)) also for $f^*(z) = F_1(a_1; z) \times F_1(a_2; z)$ we get

$$\beta(f^*) = \begin{pmatrix} a_1 & a_2 \\ 2\pi i F_1(a_2; z) & 2\pi i F_1(a_1; z) \end{pmatrix}$$

On the other hand, for $g^*(z) = F_2(a_1, a_2; z) + F_2(a_2, a_1; z)$ we have by (8.7)

$$\beta(g^*) = \begin{pmatrix} a_1 & a_2 \\ 2\pi i F_1(a_2; z) & 2\pi i F_1(a_1; z) \end{pmatrix}.$$

Hence, for any $f \in f^*$ and $g \in g^*$ the difference $f - g$ has only logarithmic singularities of first order. But then, by Theorem 8.6, there exists a function $h \in \mathrm{LI}_1$ such that $f - g - h$ is nowhere ramified. By Theorem 8.1, $f - g - h$ is for all $z$ regular and hence it must be a constant. This shows

$$F_2(a_2, a_1; z) + F_2(a_1, a_2; z) = F_1(a_1; z)F_1(a_2; z). \tag{8.9}$$

2. Let the statement be proved for $n-1$. For

$$f^*(z) = F_m(a_1, \ldots, a_m; z)F_{n-m}(a_{m+1}, \ldots, a_n; z)$$

we use $\beta$ defined in (8.7) and get

$$\beta(f^*) = \begin{pmatrix} a_1 & a_{m+1} \\ 2\pi i F_{m-1}(a_2, \ldots, a_m; z) & 2\pi i F_m(a_1, \ldots, a_m; z) \\ \cdot F_{n-m}(a_{m+1}, \ldots, a_n; z) & \cdot F_{n-m-1}(a_{m+2}, \ldots, a_n; z) \end{pmatrix}.$$

By induction hypothesis, these products may be written as

$$\begin{aligned} F_{m-1}(a_2, \ldots, a_m; z)&F_{n-m}(a_{m+1}, \ldots, a_n; z) \\ &= \sum_{\sigma \in P_1} F_{n-1}(a_{\sigma(1)}, \ldots, a_{\sigma(n-1)}; z), \end{aligned}$$



and

$$\begin{aligned} F_m(a_1, \ldots, a_m; z)&F_{n-m-1}(a_{m+2}, \ldots, a_n; z) \\ &= \sum_{\sigma \in P_2} F_{n-1}(a_{\sigma(1)}, \ldots, a_{\sigma(n-1)}; z), \end{aligned}$$

where $P_1$ and $P_2$ are defined as follows:

$P_1$ is the set of all permutations of $\{2, \ldots, n\}$ such that both $2, \ldots, m$ occur in the natural order and $m+1, \ldots, n$ occur in the natural order. $P_2$ is the set of all permutations of $\{1, \ldots, m, m+2, \ldots, n\}$ such that both $1, \ldots, m$ occur in the natural order and $m+2, \ldots, n$ occur in the natural order.

From this it follows that

$$\begin{aligned} g^*(z) &= \sum_{\sigma \in P_1} F_n(a_1, a_{\sigma(1)}, \ldots, a_{\sigma(n-1)}; z) \\ &\quad + \sum_{\sigma \in P_2} F_n(a_{m+1}, a_{\sigma(1)}, \ldots, a_{\sigma(n-1)}; z) \\ &= \sum_{\sigma \in P_{m,n}} F_n(a_{\sigma(1)}, \ldots, a_{\sigma(n)}; z) \end{aligned}$$

has the property $\beta(f^*) = \beta(g^*)$. As in the induction beginning we conclude from this $f^* = g^*$.

Theorem 8.8 allows two different interpretations. The first is that a product of two hyperlogarithms of orders $n$ and $k$ is a logarithmic integral of order $n+k$. The second is that the hyperlogarithms satisfy certain linear functional equations: suitable linear combinations of hyperlogarithms of order $n$ are expressible by means of lower-order hyperlogarithms. The most important example is (8.8) whose existence follows from (8.9) for $a_1 = 0$, $a_2 = 1$, and $n = 2$. Notice that the argument $1 - z$ in (8.8) effects a permutation of the singularities 0 and 1:

$$\begin{aligned} \mathrm{Li}_2^*(z) &= -F_2(1, 0; z), \\ \mathrm{Li}_2^*(1-z) &= -F_2(0, 1; z). \end{aligned}$$

Thus (8.9) yields

$$\mathrm{Li}_2(z) + \mathrm{Li}_2(1-z) = -\log(z)\log(1-z) + \varepsilon(z), \tag{8.10}$$

where $\varepsilon(z) \in \mathrm{LI}_1$. Of course, (8.10) must be identical with (8.8).

An evident generalization of Theorem 8.8 shows that arbitrary products of hyperlogarithms are logarithmic integrals. Products of the form $F_{n_1}F_{n_2}\cdots F_{n_k}$ will be called functions of type $\langle n_1, \ldots, n_k\rangle$. Functions having at one singularity a branching term of type $\langle n_1, \ldots, n_k\rangle$ will be called of type $\{n_1, \ldots, n_k\}$. If $n_1 + \cdots + n_k = n$, then by Theorem 8.6 its equivalence class in $\mathrm{LI}_n$ modulo $\mathrm{LI}_{n-1}$ is uniquely determined. In particular, we will make use of classes $G_n(a_1, \dots, a_{n-1}; a_0; z) \in \mathrm{PLI}_n$ of those functions having only one finite singularity $a_0$ with the branching term $2\pi i \log(z-a_1) \cdot \log(z-a_2) \cdots \log(z-a_{n-1})$. Notice that

$$(n-1)!\,\mathrm{Li}_n(z) = -G_n(0, 0, \dots, 0; 1; z)$$

(cf. (8.5)). Another notable relationship is

$$G_n(a_1, \dots, a_{n-1}; a_0; z) = \sum_{\sigma \in P_{n-1}} F_n(a_0, a_{\sigma(1)}, \dots, a_{\sigma(n-1)}; z), \tag{8.11}$$

where $P_{n-1}$ is the set of all permutations of $\{1, \dots, n-1\}$. (8.11) follows from

$$F_1(a_1; z) \cdots F_1(a_{n-1}; z) = \sum_{\sigma \in P_{n-1}} F_{n-1}(a_{\sigma(1)}, \dots, a_{\sigma(n-1)}; z)$$

which, in its turn, is proved by induction making repeated use of Theorem 8.8.

8.4.2. *Generalizing the factorization theorem.* Here we ask for functional equations of the hyperlogarithms which are allowed to have rational functions of one variable as arguments. A typical example is the factorization theorem for $\mathrm{Li}_2$:

$$\frac{1}{r}\mathrm{Li}_2(z^r) = \mathrm{Li}_2(z) + \mathrm{Li}_2(\omega z) + \cdots + \mathrm{Li}_2(\omega^{r-1} z),$$

where $\omega = e^{2\pi i/r}$.

The most general functional equation of the given type is

$$c_1 F_n(a_{11}, \dots, a_{1n}; f_1(z)) + \cdots + c_m F_n(a_{m1}, \dots, a_{mn}; f_m(z)) = 0,$$

with rational functions $f_1(z), \dots, f_m(z)$.

We prove the existence of such functional equations. We consider the branching behaviour of $F_n(a_1, \dots, a_n; f(z))$. Its singularities of order $j$ are the zeros and poles of $f(z) - a_{n-j+1}$. Let $b_{j1}, \dots, b_{jp_j}$ be the zeros and poles of $f(z) - a_j$ with the multiplicities $\alpha_{j1}, \dots, \alpha_{jp_j}$, respectively. Note that the multiplicity of a pole is negative. The proof of the following theorem rests on the observation:

(8.12) If $z$ surrounds $b_{jk}$ (and no further $b_{jl}$) counterclockwise, then $f(z)$ surrounds $a_j$ exactly $\alpha_{jk}$ times counterclockwise (note: "$-2$ times counterclockwise" means "2 times clockwise").

THEOREM 8.9.

$$F_n(a_1, \dots, a_n; f(z)) = \sum_{j_1=1}^{p_1} \cdots \sum_{j_n=1}^{p_n} \alpha_{1j_1} \cdots \alpha_{nj_n} F_n(b_{1j_1}, \dots, b_{nj_n}; z).$$

*($f(z)$, the $\alpha$'s, and $b$'s have the meaning just explained.)*



PROOF (INDUCTION ON $n$). 1. $n = 1$.

$$F_1(a_1; f(z)) = \log(f(z) - a_1) = \log \prod_{j=1}^{p_1}(z - b_{1j}) = \sum_{j=1}^{p_1} \log(z - b_{1j})$$
$$= \sum F_1(b_{1j}; z).$$

2. Let

$$F_{n-1}(a_2, \dots, a_n; f(z)) = \sum_{j_2=1}^{p_2} \cdots \sum_{j_n=1}^{p_n} F_n(b_{2j_2}, \dots, b_{nj_n}; z)$$

be proved.

3. To prove the equation for $n$, we simply have to compare the branching terms at the singularities of order $n$. $F_n(a_1, \dots, a_n; f(z))$ has the singularities $b_{11}, \dots, b_{1p_1}$. For surrounding $z = b_{1j}$ counterclockwise we get the branching term $2\pi i \alpha_{1j} F_{n-1}(a_2, \dots, a_n; f(z))$, and the right-hand side of the claim yields

$$2\pi i \alpha_{1j} \sum_{j_2=1}^{p_2} \cdots \sum_{j_n=1}^{p_n} \alpha_{2j_2} \cdots \alpha_{nj_n} F_{n-1}(b_{2j_2}, \dots, b_{nj_n}; z).$$

By the induction hypothesis both expressions are equal.

**8.5. A reduction problem.** A. Block and G. Guillaumin [6] and P. Müller [7] conjecture that the volume of orthoschemes in spaces of constant curvature can be expressed by polylogarithms. J. Böhm [8] observed that in the 7-dimensional space the function $F_4\left({}^{1,0,0,z_0}_{0,0,0,0}\middle| z\right)$ with $z_0 \neq 0$ occurs. If this hyperlogarithm is not expressible by polylogarithms then the above conjecture is disproved.

We start by briefly discussing the cases $n = 2$ and $n = 3$. $F_2(a_1, a_2; z)$ is essentially $\mathrm{Li}_2$: a linear transformation of the argument can map $a_1$ and $a_2$ into 1 and 0, respectively. $F_3$ can be expressed by lower-order functions and $G_3$ as follows: Theorem 8.8 yields.

$$F_1(a_1; z) F_2(a_2, a_3; z)$$
$$= F_3(a_1, a_2, a_3; z) + F_3(a_2, a_1, a_3; z) + F_3(a_2, a_3, a_1; z).$$

By (8.11) we get

$$G_3(a_1, a_3; a_2; z) = F_3(a_2, a_1, a_3; z) + F_3(a_2, a_3, a_1; z).$$

This gives

$$F_3(a_1, a_2, a_3; z) = F_1(a_1; z) F_2(a_2, a_3; z) - G_3(a_1, a_3; a_2; z).$$

**Theorem** 8.10. *It is impossible to represent the general hyperlogarithm of order 4 as a linear combination of functions of types* $\langle 1, 3\rangle$, $\langle 2, 2\rangle$, $\langle 2, 1, 1\rangle$, $\langle 1, 1, 1, 1\rangle$, *and* $G_4$. *This remains true even if the arguments of these functions are allowed to be rational functions.*

**Remark.** As $G_4$ is a generalization of $\mathrm{Li}_4$ it follows from this theorem that $F_4$ is not expressible by lower-order functions and $\mathrm{Li}_4$.

**Proof of Theorem** 8.10. Given $F_4(a_1, a_2, a_3, a_4; z)$, we assume the most general case, namely that the $a_i$ are pairwise different. Our question concerns the 24-dimensional subspace $R$ of $\mathrm{PLI}_4$ which consists of the equivalence classes of those logarithmic integrals having singularities at the points $a_1$, $a_2$, $a_3$, $a_4$, and $\infty$.

The subspace $R$ has the basis $\{e_1, \dots, e_{24}\}$, where

$$e_1 = F_4(a_1, a_2, a_3, a_4; z),$$
$$e_2 = F_4(a_1, a_2, a_4, a_3; z), \dots, e_{24} = F_4(a_4, a_3, a_2, a_1; z).$$

Here, we think of the lexicographic order of the permutations of $\{a_1, a_2, a_3, a_4\}$, and the $i$th permutation corresponds to $e_i$.

For the sake of brevity we use the following abbreviations:

$$(ijkl) = F_4(a_i, a_j, a_k, a_l; z),$$
$$(i)(jkl) = F_1(a_i; z)F_3(a_j, a_k, a_l; z),$$

etc., and

$$[ijkl] = G_4(a_i, a_j, a_k; a_l; z).$$

Because of

$$(i)(j)(k)(l) = [i, j, k, l] + [j, k, l, i] + [k, l, i, j] + [l, i, j, k],$$

we can omit all functions of type $\langle 1, 1, 1, 1\rangle$. By Theorem 8.8 we get

$$(ij) + (ji) = (i)(j), \tag{8.13}$$
$$(ijk) + (jik) + (jki) = (i)(jk). \tag{8.14}$$

Multiplying (8.12) with $(kl)$ we get $(ij)(kl) + (ji)(kl) = (i)(j)(kl)$.

This shows that it is sufficient to consider only the following three $\langle 2, 2\rangle$-functions: $(12)(34)$, $(13)(24)$, $(14)(23)$. From (8.14) we get

$$(l)(ijk) + (l)(jik) + (l)(jki) = (l)(i)(jk).$$

This shows that all functions of type $\langle 2, 1, 1\rangle$ are completely redundant.

By now we have 31 elements $v_1, \dots, v_{31}$ in $R$: 24 of type $\langle 3, 1\rangle$, three of type $\langle 2, 2\rangle$ and four of type $G_4$. As $R$ has dimension 24, further linear dependencies must exist. Instead of searching for them we immediately show that no relationship of the form

$$e_1 = c_1 v_1 + \cdots + c_{31} v_{31} \tag{8.15}$$



can exist. To this end we represent the $v_i$ as linear combinations of $e_1, \dots, e_{24}$. Then we put these representations in (8.15) and get a linear relationship

$$b_1 e_1 + \cdots + b_{24} e_{24} = 0, \tag{8.16}$$

which requires all coefficients $b_1, \dots, b_{24}$ to vanish. This gives a system of 24 linear equations in the unknowns $c_1, \dots, c_{31}$, and we shall show that this system is unsolvable.

$$\begin{aligned}
v_1 &= (1)(234) = e_1 + e_7 + e_9 + e_{10},\\
v_2 &= (1)(243) = e_2 + e_8 + e_{11} + e_{12},\\
v_3 &= (1)(324) = e_3 + e_{13} + e_{15} + e_{16},\\
v_4 &= (1)(342) = e_4 + e_{14} + e_{17} + e_{18},\\
v_5 &= (1)(423) = e_5 + e_{19} + e_{21} + e_{22},\\
v_6 &= (1)(432) = e_6 + e_{20} + e_{23} + e_{24},\\
v_7 &= (2)(134) = e_1 + e_3 + e_4 + e_7,\\
v_8 &= (2)(143) = e_2 + e_5 + e_6 + e_8,\\
v_9 &= (2)(314) = e_9 + e_{13} + e_{14} + e_{15},\\
v_{10} &= (2)(341) = e_{10} + e_{16} + e_{17} + e_{18},\\
v_{11} &= (2)(413) = e_{11} + e_{19} + e_{20} + e_{21},\\
v_{12} &= (2)(431) = e_{12} + e_{21} + e_{23} + e_{24},\\
v_{13} &= (3)(124) = e_1 + e_2 + e_3 + e_{13},\\
v_{14} &= (3)(142) = e_4 + e_5 + e_6 + e_{14},\\
v_{15} &= (3)(214) = e_7 + e_8 + e_9 + e_{15},\\
v_{16} &= (3)(241) = e_{10} + e_{12} + e_{13} + e_{16},\\
v_{17} &= (3)(412) = e_{17} + e_{19} + e_{20} + e_{23},\\
v_{18} &= (3)(421) = e_{18} + e_{21} + e_{22} + e_{24},\\
v_{19} &= (4)(123) = e_1 + e_2 + e_5 + e_{19},\\
v_{20} &= (4)(132) = e_3 + e_4 + e_6 + e_{20},\\
v_{21} &= (4)(213) = e_7 + e_8 + e_{11} + e_{21},\\
v_{22} &= (4)(231) = e_9 + e_{10} + e_{12} + e_{22},\\
v_{23} &= (4)(312) = e_{13} + e_{14} + e_{17} + e_{23},\\
v_{24} &= (4)(321) = e_{15} + e_{16} + e_{18} + e_{24},\\
v_{25} &= (12)(34) = e_1 + e_3 + e_4 + e_{13} + e_{14} + e_{17},\\
v_{26} &= (13)(24) = e_1 + e_2 + e_3 + e_7 + e_8 + e_{11},\\
v_{27} &= (14)(23) = e_1 + e_2 + e_5 + e_7 + e_8 + e_9,\\
v_{28} &= [2341] = e_1 + e_2 + e_3 + e_4 + e_5 + e_6,
\end{aligned}$$

$$v_{29} = [1342] = e_7 + e_8 + e_9 + e_{10} + e_{11} + e_{12},$$
$$v_{30} = [1243] = e_{13} + e_{14} + e_{15} + e_{16} + e_{17} + e_{18},$$
$$v_{31} = [1234] = e_{19} + e_{20} + e_{21} + e_{22} + e_{23} + e_{24}.$$

If these representations are put into (8.15) we get (8.16). As $e_1, \dots, e_{24}$ are linearly independent, we have $b_1 = \dots = b_{24} = 0$. We consider in particular $b_1, b_2, b_7, b_8, b_{17}, b_{18}, b_{23}$, and $b_{24}$ and get

$$b_1 = c_1 + c_7 + c_{13} + c_{19} + c_{25} + c_{26} + c_{27} + c_{28} - 1 = 0,$$
$$b_2 = c_2 + c_8 + c_{13} + c_{19} + c_{26} + c_{27} + c_{28} = 0,$$
$$b_7 = c_1 + c_7 + c_{15} + c_{21} + c_{26} + c_{27} + c_{29} = 0,$$
$$b_8 = c_2 + c_8 + c_{15} + c_{21} + c_{26} + c_{27} + c_{29} = 0,$$
$$b_{17} = c_4 + c_{10} + c_{17} + c_{23} + c_{25} + c_{30} = 0,$$
$$b_{18} = c_4 + c_{10} + c_{18} + c_{24} + c_{30} = 0,$$
$$b_{23} = c_6 + c_{12} + c_{17} + c_{23} + c_{31} = 0,$$
$$b_{24} = c_6 + c_{12} + c_{18} + c_{24} + c_{31} = 0.$$

From this system we get

$$b_1 - b_2 - b_7 + b_8 = c_{25} - 1 = 0$$

and

$$b_{17} - b_{18} - b_{23} + b_{24} = c_{25} = 0,$$

a contradiction.

It remains to show that an equation (8.15) cannot be achieved by admitting rational arguments on the right-hand side. This follows from the fact that $F_n(a_1, \dots, a_n; r(z))$ belongs to $\mathrm{PLI}_n$ for rational functions $r(z)$. Hence it must be representable by a linear combination of $F_n$-functions with argument $z$. The same argument is true for the $G_4$-functions. Therefore, an equation (8.15) with rational arguments would enhance such an equation in its pure form which has been shown to be impossible.

## References


1. E. E. Kummer, *Über die Transzendenten, welche aus wiederholten Integrationen rationaler Funktionen entstehen*, J. Reine Angew. Math. **21** (1840), 74–90.
2. H. Poincaré, *Oevres*, vol. 2, Paris, 1916.
3. J. A. Lappo-Danilewski, *Resolution algorithmique des problèmes reguliers de Poincaré et de Riemann*, J. Soc. Physico-Mathematique de Leningrade II (1) (1928), 94–154.
4. N. Nielsen, *Über den Eulerschen Dilogarithms*, Nova Acta Leopoldina **90** (1909), 154–155.
5. L. Euler, *Memoires de l'Academie de St.-Peterbourg*, vol. 3, 1911, p. 38.
6. A. Bloch and G. Guillaumin, *Sur le volume des polyèdres non euclidiens*, C. R. Acad. Sci. Paris Sér. I Math. **224** (1947), 1690–1692.
7. P. Müller, *Über Simplexinhalte in nichteuklidischen Räumen*, Ph.D. thesis, Bonn, 1954.
8. J. Böhm, *Zu Coxeters Integrationsmethode in gekrümmten Räumen*, Math. Nachr. **27** (1964).


CHAPTER 9

# Kummer-Type Functional Equations of Polylogarithms

G. WECHSUNG

**9.1. Automorphic functions.** For the polylogarithms $\mathrm{Li}_2$, $\mathrm{Li}_3$, $\mathrm{Li}_4$ and $\mathrm{Li}_5$ N. H. Abel [1] (for $\mathrm{Li}_2$) and E. E. Kummer [2] have found functional equations in two variables of the form

$$\sum_{i=1}^{N} a_i \,\mathrm{Li}_n(f_i(z)g_i(w)) = \mathscr{E}(z, w) \tag{9.1}$$

where the $f_i$ and $g_i$ are rational functions and $\mathscr{E}$ is built up from logarithmic integrals of order at most $n$ and depending on only one variable (see also [3, 4]). All these equations share a common structural property: their arguments are essentially automorphic functions with respect to subgroups of a finite group of linear automorphisms of the complex plane.

From the time of F. Klein all finite groups of linear automorphisms of the complex plane have been known (see [5, 6]). They are exactly the cyclic groups $C_k$, the dihedral groups $D_{2k}$, (i.e., the automorphism groups of a regular $k$-gon) and the automorphism groups of the tetrahedron, octahedron and icosahedron.

Let $G$ be some group of linear automorphisms of the complex plane. If $G = \{t_1, \dots, t_k\}$, then each $t_i$ is of the form

$$t_i(z) = \frac{a_i z + b_i}{c_i z + d_i} \quad \text{with} \quad \begin{vmatrix} a_i & b_i \\ c_i & d_i \end{vmatrix} \neq 0.$$

The group composition is the usual substitution

$$t_i t_j(z) = t_i(t_j(z))$$

which is known to be associative. That $G$ is a group with respect to this operation means

1. for any $t_i, t_j \in G$ the product $t_i t_j$ belongs to $G$,
2. for any $t_i \in G$ the inverse $t_i^{-1}$ belongs to $G$, and
3. $G$ contains the identity mapping $\mathrm{id}(z) = z$.

For arbitrary complex $a_0$ the set $A = \{t(a_0): t \in G\}$ is called a *complete system of equivalent points with respect to* $G$. A fixed point of a transformation $t$ of order $m$ will be written down $m$ times.

EXAMPLE. The group $Z = \{z, z/(z-1)\}$ has for instance the following complete systems of equivalent points

$$A_1 = \{0, 0\}$$
$$A_2 = \{1, \infty\}$$
$$A_3 = \{1/2, -1\}.$$

A function $f$ is called *automorphic with respect to the group* $G$ if for each $t \in G$ we have $f(z) \equiv f(t(z))$. The following theorem states our knowledge on functions automorphic with respect to finite groups of linear automorphisms of the plane (see [5, 6]).

THEOREM. (1) *If* $A = \{a_1, \dots, a_k\}$ *and* $B = \{b_1, \dots, b_k\}$ *are two disjoint complete sets of equivalent points with respect to* $G$, *then*

$$f(z) = \frac{(z-a_1)\cdots(z-a_k)}{(z-b_1)\cdots(z-b_k)}$$

*is automorphic with respect to* $G$.

(2) *Any automorphic function with respect to* $G$ *can be given up to a constant factor in this form.*

EXAMPLE. $f(z) = z^2/(z-1)$ is automorphic with respect to $Z = \{z, z/(z-1)\}$. This follows from the theorem for $A = \{0, 0\}$ and $B = \{1, \infty\}$, but it can also be simply verified:

$$f\left(\frac{z}{z-1}\right) = \frac{(\frac{z}{z-1})^2}{\frac{z}{z-1}-1} = \frac{z^2}{z-1} = f(z).$$

**9.2. Kummer-type functional equations.** The theory of Kummer-type functional equations for the polylogarithms to be presented in §§9.2–9.3.4 has been developed in [7].

A first observation in analyzing the conditions under which an equation (9.1) may exist is the following: Consider $w$ as a (fixed) parameter. Assume $f_i(z)g_i(w) = 1$ for $z = e(w)$. As $\mathrm{Li}_n(x)$ has a singularity for $x = 1$ with a branching term $-(2\pi i/(n-1)!)\log^{n-1} x$ there must exist several arguments, say $f_i(z)g_i(w), \dots, f_m(z)g_m(w)$, such that $f_j(z)g_j(w) = 1$ for $z = e(w)$. If $z$ goes around $e(w)$ counterclockwise exactly once the left-hand side of (9.1) yields the branch term

$$-\frac{2\pi i}{(n-1)!}[a_1 \log^{n-1} f_1(z)g_1(w) + \cdots + a_m \log^{n-1} f_m(z)g_m(w)],$$

which must vanish identically in $z$ and $w$ because the right-hand side of (9.1) cannot have the branch point $z = e(w)$. Remember that the functions occurring in $\mathscr{E}(z, w)$ depend on either $z$ or on $w$.



All known functional equations (Kummer's $\mathrm{Li}_4$-equation is slightly exceptional) have the following properties.

(K1) $f_1(z), \dots, f_m(z)$ are rational functions. $f_i$ is automorphic with respect to a group $G_i$ of linear transformations of order degree $f_i$. The groups $G_i$ need not be pairwise different.

(K2) $G_1, \dots, G_m$ are subgroups of some finite group $G$.

(K3) With a second independent variable $w$ an equation

$$I(z, w) = \sum_{i=1}^{s} c_i \log^{n-1} \frac{f_i(z)}{f_i(w)} = 0$$

is satisfied identically in $z$ and $w$. $I(z, w) = 0$ is called an *identity belonging to* $G$.

(K4) With

$$K(z, w) =_{df} \sum_{i=1}^{s} \frac{c_i}{\operatorname{degree} f_i} \mathrm{Li}_n\left(\frac{f_i(z)}{f_i(w)}\right),$$

one can write (9.1) in the form

$$\sum_{t \in G} K(z, t(w)) = \mathscr{E}(z, w). \tag{9.2}$$

A functional equation (9.1) satisfying the properties (K1)–(K4) will be called a *Kummer-type* equation.

Before considering an example we try to illustrate the role of the group $G$. For fixed $w$ the singularities of the left-hand side coming from $f_i(z)/f_i(w) = 1$ for at least one $i$ form a complete set $A$ of equivalent points with respect to $G$. Without loss of generality we can assume $w \in A$. Because of (K3) $z = t(w)$ is not a singularity of the linear combination

$$C(z, t(w)) = \sum_{i=1}^{s} c_i \mathrm{Li}_n\left(\frac{f_i(z)}{f_i(t(w))}\right).$$

The facts that the $f_i$ are automorphic functions and that $G$ is finite ensure that the set of arguments occurring in all these combinations (where $t$ is running through $G$) is finite.

EXAMPLES. We consider Kummer's equation for $\mathrm{Li}_5$ (see [2, 4]). It belongs to the symmetric group $S_3$ containing the six elements $t_0(z) = z$, $t_1(z) = 1/z$, $t_2(z) = (z-1)/z$, $t_3(z) = z/(z-1)$, $t_4(z) = 1/(1-z)$, $t_5(z) = 1-z$. $S_3$ has three cyclic subgroups of order 2, namely $Z_1 = \{t_0, t_1\}$, $Z_2 = \{t_0, t_3\}$, $Z_3 = \{t_0, t_5\}$. These groups have the automorphic functions $f_1(z) = z/(1-z)^2$, $f_2(z) = z^2/(1-z)$, $f_3(z) = z(1-z)$, respectively. One easily verifies the identity

$$\begin{aligned} &\log^4 \frac{z}{(1-z)^2}\cdot\frac{(1-w)^2}{w} + \log^4 \frac{z^2}{1-z}\cdot\frac{1-w}{w^2} + \log^4 \frac{z(1-z)}{w(1-w)} \\ &\quad - 9\log^4 \frac{z}{1-z}\cdot\frac{1-w}{w} - 9\log^4 \frac{z}{w} - 9\log^4 \frac{1-z}{1-w} = 0. \end{aligned} \tag{9.3}$$

Defining

$$K(z, w) = \frac{1}{2}\mathrm{Li}_5\left(\frac{z(1-w)^2}{w(1-z)^2}\right) + \frac{1}{2}\mathrm{Li}_5\left(\frac{z^2(1-w)}{w^2(1-z)}\right) + \frac{1}{2}\mathrm{Li}_5\left(\frac{z(1-z)}{w(1-w)}\right)$$
$$- 9\,\mathrm{Li}_5\left(\frac{z(1-w)}{w(1-z)}\right) - 9\,\mathrm{Li}_5\left(\frac{z}{w}\right) - 9\,\mathrm{Li}_5\left(\frac{1-z}{1-w}\right),$$

we get the main part of Kummer's equation simply in the form

$$\sum_{t \in S_3} K(z, t(w)). \tag{9.4}$$

We write this sum in such a way that three rows and three columns occur. The terms within each row have one singularity in common, but because of the cooperation of these terms according to (9.3) this singularity of the single terms is no longer a singularity of (9.4). The same is true for the columns. The rows correspond to the singularities $z = w$, $z = 1/(1-w)$, $z = (w-1)/w$, and the columns (from left to right) correspond to the singularities $z = 1/w$, $z = w/(1-w)$, $z = 1-w$. We get the detailed form for (9.4) as shown on page 189.

A graph representation of how the group controls the cooperation of the arguments is given as follows. To each singularity $z = t(w)$ occurring in at least one of the left-hand side terms of (9.1) by reaching $f_i(z) = f_i(t(w))$, we associate in a one-one manner a vertex of the graph. To each of these singularities there corresponds a combination $C(z, t(w))$ (see (9.3)) which guarantees that $z = t(w)$ is not a singularity of the left-hand side of (9.1). Therefore, we label this vertex by $C(z, t(w))$. Two vertices are linked by an edge if $C(z, t_1, (w))$ and $C(z, t_2(w))$ have at least one $\mathrm{Li}_n$-term in common (i.e., there is an argument $f_i(z)/g_i(w)$ such that $f_i(z)/g_i(t_1(w)) = f_i(z)/g_i(t_2(w)))$.

The edge from the vertex labelled $z = t_1(w)$ to the vertex labelled $z = t_2(w)$ is labelled by $(t_0, t_1^{-1}t_2)$.

From the detailed form for (9.4) we deduce the graph in Figure 9.1. (See page 190.)

The way of labelling the edges by pairs of transformations indicates that there is also the possibility of applying the transformation to $z$ or to both variables. Kummer's equation for $\mathrm{Li}_4$ corresponds to this possibility. It is based on the identity

$$\log^3 \frac{z^2(1-w)}{w^2(1-z)} + \log^3 \frac{z(1-w)^2}{w(1-z)^2} - 6\log^3 \frac{z(1-w)}{w(1-z)} - 3\log^3 \frac{z}{w} - 3\log^3 \frac{1-w}{1-z} = 0.$$

The first two arguments of this identity are automorphic functions with respect to $Z_2 = \{t_0, t_3\}$ and $Z_1 = \{t_0, t_1\}$, respectively. These two groups generate the group $S_3$, and hence, according to our experience with the $\mathrm{Li}_5$-equation, we would expect a $\mathrm{Li}_4$-equation whose main part has the form



$$\begin{array}{llllll}
\mathrm{Li}_5\left(\frac{z(1-w)^2}{w(1-z)^2}\right) & + \mathrm{Li}_5\left(\frac{z^2(1-w)}{w^2(1-z)}\right) & + \mathrm{Li}_5\left(\frac{z(1-z)}{w(1-w)}\right) & - 9\,\mathrm{Li}_5\left(\frac{z(1-w)}{w(1-z)}\right) & - 9\,\mathrm{Li}_5\left(\frac{z}{w}\right) & - 9\,\mathrm{Li}_5\left(\frac{1-z}{1-w}\right) \\
+ \mathrm{Li}_5\left(\frac{z^2w(1-w)}{1-z}\right) & + \mathrm{Li}_5\left(-\frac{z(1-z)(1-w)^2}{w}\right) & + \mathrm{Li}_5\left(\frac{zw^2}{(1-w)(1-z)^2}\right) & - 9\,\mathrm{Li}_5\left(-\frac{zw}{1-z}\right) & - 9\,\mathrm{Li}_5(z(1-w)) & - 9\,\mathrm{Li}_5\left(-\frac{(1-z)(1-w)}{w}\right) \\
+ \mathrm{Li}_5\left(-\frac{w^2z(1-z)}{1-w}\right) & + \mathrm{Li}_5\left(\frac{-z}{w(1-w)(1-z)^2}\right) & + \mathrm{Li}_5\left(\frac{z^2w}{(1-w)^2(1-z)}\right) & - 9\,\mathrm{Li}_5\left(-\frac{z}{(1-z)(1-w)}\right) & - 9\,\mathrm{Li}_5\left(-\frac{zw}{1-w}\right) & - 9\,\mathrm{Li}_5(w(1-z)) \\
- 9\,\mathrm{Li}_5\left(-\frac{z(1-w)}{1-z}\right) & - 9\,\mathrm{Li}_5\left(-\frac{zw}{1-z}\right) & - 9\,\mathrm{Li}_5\left(\frac{zw}{(1-w)(1-z)}\right) & & & \\
- 9\,\mathrm{Li}_5(zw) & - 9\,\mathrm{Li}_5\left(\frac{z(w-1)}{w}\right) & - 9\,\mathrm{Li}_5\left(\frac{z}{1-w}\right) & & & \\
- 9\,\mathrm{Li}_5\left(-\frac{w(1-z)}{1-w}\right) & - 9\,\mathrm{Li}_5((1-z)(1-w)) & - 9\,\mathrm{Li}_5\left(\frac{1-z}{w}\right) & & &
\end{array}$$

Detailed form for (9.4)

188

Defini

$K(z,$

we get

(9.4)

We wr
The te
of the
single
columı
$z = ($
singula
form f

A gı
argume
least o
we ass
singula
guaran
Theref
an edg
commc
$f_i(z)/g$
The
$t_2(w)$ i
Fror
page 1
The
there is
variabl
based c

$\log^3 \frac{z^2}{w}$

The fir
spect t
generat
equatio



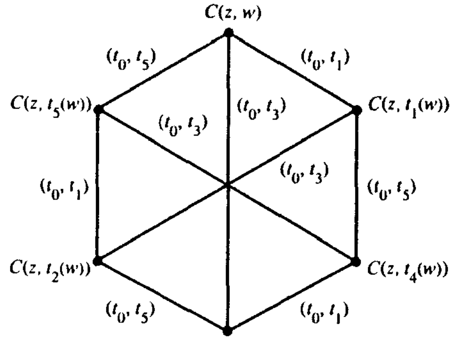



FIGURE 9.1

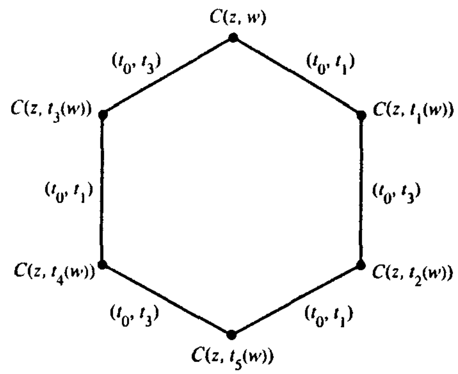



FIGURE 9.2

$\sum_{t\in S_3} K_4(z, t(w))$ where

$$K_4(z, w) = \frac{1}{2}\,\mathrm{Li}_4\left(\frac{z^2(1-w)}{w^2(1-z)}\right) + \frac{1}{2}\,\mathrm{Li}_4\left(\frac{z(1-w)^2}{w(1-z)^2}\right) - 6\,\mathrm{Li}_4\left(\frac{z(1-w)}{w(1-z)}\right)$$
$$- 3\,\mathrm{Li}_4\left(\frac{z}{w}\right) - 3\,\mathrm{Li}_4\left(\frac{1-w}{1-z}\right).$$

This equation has 24 $\mathrm{Li}_4$-terms with arguments depending on both variables. Its graph is shown in Figure 9.2.

However, Kummer found an easier equation corresponding not to $S_3$ but to $Z_2 \times Z_1$. Its main part is

$$\sum_{s\in Z_2,\, t\in Z_1} K_4(s(z), t(w)),$$



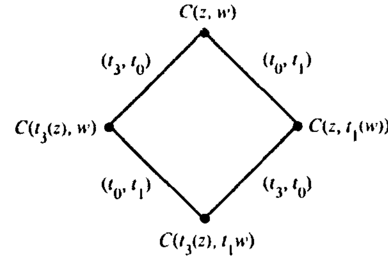



FIGURE 9.3

containing 16 $\mathrm{Li}_4$-terms with two variable arguments and having the graph given in Figure 9.3.

**9.3. A method to construct functional equations.** In the preceding section the notion of Kummer-type functional equations has been defined as a result of analyzing some equations given by Kummer. In this section we prove that the properties (K1)–(K3) are sufficient for the existence of Kummer-type equations.

THEOREM 9.1. *If the properties* (K1)–(K3) *are satisfied, then there exists an equation for* $\mathrm{Li}_n$ *as described in* (K4).

PROOF. Let (K1)–(K3) be satisfied. This means: We have a finite group $G$ of linear transformations, and for the subgroups $G_1, \dots, G_m$ of $G$ we have automorphic functions $f_1, \dots, f_m$ with degree $(f_i) = \mathrm{order}(G_i)$, $i = 1, \dots, s$, such that

$$I(z, w) = \sum_{i=1}^{s} c_i \log^{n-1} \frac{f_i(z)}{f_i(w)} = 0,$$

identically in $z$ and $w$. The claim is that with

$$K(z, w) =_{df} \sum_{i=1}^{s} \frac{c_i}{\deg(f_i)}\,\mathrm{Li}_n\left(\frac{f_i(z)}{f_i(w)}\right) \quad \text{and} \quad S(z, w) =_{df} \sum_{t\in G} K(z, t(w)),$$

there exists a function $\mathscr{E}(z, w)$ built up from functions depending on either $z$ or $w$ and that $S(z, w) = \mathscr{E}(z, w)$.

To prove this claim we initially fix $w$ to be an arbitrary complex value different from the zeros of the functions $f_1(w), \dots, f_m(w)$. Furthermore, we choose $z_0$ in such a way that none of the arguments of $S$ is equal to $0$, $1$ or $\infty$. $\mathscr{E}$ is determined under the assumption that $z$ initially varies in a sufficiently small environment $K(z_0)$ of $z_0$, and the $\mathrm{Li}_n$ in $S$ are understood to be the main branches. If $\mathscr{E}$ is determined under this assumption, the validity of the equation $S(z, w) = \mathscr{E}(z, w)$ may then be extended by analytic continuation with respect to $z$ and $w$.

From Theorem 8.5 we know that the main branch of $\mathrm{Li}_n(u)$ has branch points only for $u = 1$ and $u = \infty$.

We prove first that $S(z, w)$ is constructed in such a way that the points $z = e(w)$ for which at least one argument $f_i(z)/f_i(t(w))$ is equal to 1 is not a branch point for $S(z, w)$. By (K1) we know that $f_i(e(w))/f_i(t(w)) = 1$ is true if and only if $e \in G_i t$. $G_i t$ contains $\operatorname{ord} G_i$ elements. This shows that the term $(c_i/\deg f_i)\,\mathrm{Li}_n(f_i(z)/f_i(t(w)))$ occurs exactly $\operatorname{ord} G_i$ times. Because of $\operatorname{ord} G_i = \deg f_i$ the coefficient of $\mathrm{Li}_n(f_i(z)/f_i(t(w)))$ in $S(z, w)$ is just $c_i$.

For $z = t(w)$ exactly the $\mathrm{Li}_n$ terms occurring in $K(z, t(w))$ are singular. Therefore, the branch term for $S(z, w)$ at $z = t(w)$ is

$$c_1 \log^{n-1} \frac{f_1(z)}{f_1(t(w))} + \cdots + c_m \log^{n-1} \frac{f_m(z)}{f_m(t(w))},$$

which vanishes by (K3).

Next consider the poles of the arguments $f_i(z)/f_i(t(w))$. Let $z_1$ be a pole of $f_i(z)/f_i(t(w))$. By Theorem 8.5 the branch term of $\mathrm{Li}_n(f_i(z)/f_i(t(w)))$ for $z = z_1$ is up to a multiplicative constant $\log^{n-1}(f_i(z)/f_i(t(w)))$ which can be reduced to branch terms of the form $\log(z - a_1)\cdots\log(z - a_k) \times \log(w - b_1)\cdots\log(w - b_j)$. As the $w$-part is understood to be a constant, say $h(w)$, we can make use of the function introduced in (8.11) to get a function

$$h(w)G_{k+1}(a_1, \ldots, a_k; z_1; z)$$

having this branch term for $z = z_1$. In this way we get a function $\mathscr{E}_1(z, w)$ such that $S(z, w) - \mathscr{E}_1(z, w)$ has no branch points and is growing everywhere only logarithmically. Hence, by Liouville's theorem, it must be a constant $c(w)$ depending still on $w$. This proves

$$S(z, w) = \mathscr{E}_1(z, w) + c(w) = \mathscr{E}(z, w).$$

REMARKS. 1. For difficult groups it is a tedious work to determine the exact form of $\mathscr{E}$. Therefore, in most cases we shall be satisfied to know that $\mathscr{E}$ exists.

2. $S(z,w)$ is the real main part of the equation. Nevertheless, in $\mathscr{E}(z,w)$ terms may occur like $\mathrm{Li}_n(g(z))$ or $\mathrm{Li}_n(g(w))$, the latter as parts of $c(w)$. Notice, however, that the arguments of these $\mathrm{Li}_n$-terms depend either on $z$ or on $w$, not both. As a consequence we state the following

METHOD FOR CONSTRUCTING FUNCTIONAL EQUATIONS FOR $\mathrm{Li}_n$.

1. Choose a finite group $G$ of linear automorphisms of the complex plane.
2. Determine all of its subgroups $G_1, \ldots, G_r$.
3. Choose automorphic functions $f_1, \ldots, f_m$ with respect to these groups ($r \le s$, and more than one automorphic function can belong to the same group).
4. Check whether an identity as mentioned in (K3) exists.



If such an identity exists, (K4) tells us how to get the left-hand side of an equation. The right-hand side of this equation is determined as explained in the proof of Theorem 9.1.

9.3.1. *Examples for* $\mathrm{Li}_2$ *and* $\mathrm{Li}_3$. Let $G$ be a finite group of order $k$ and let $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$ be two disjoint complete systems of equivalent points of $G$. Then

$$f(z) = \frac{(z - a_1)\cdots(z - a_k)}{(z - b_1)\cdots(z - b_k)}$$

is an automorphic function with respect to $G$, and one easily verifies the identities

$$\log\frac{f(z)}{f(w)} - \sum_{i=1}^{k}\log\frac{z - a_i}{w - a_i} + \sum_{i=1}^{k}\log\frac{z - b_i}{w - b_i} = 0, \tag{9.5}$$

and

(9.6)

$$\log^2\frac{f(z)}{f(w)} - \sum_{i,j=1}^{k}\log^2\frac{(z - a_i)(w - b_j)}{(w - a_i)(z - b_j)} + \sum_{1\le i<j\le k}\log^2\frac{(z - a_i)(w - a_j)}{(w - a_i)(z - a_j)}$$
$$+ \sum_{1\le i<j\le k}\log^2\frac{(z - b_i)(w - b_j)}{(w - b_i)(z - b_j)} = 0.$$

Each of the identities (9.5) gives a functional equation for $\mathrm{Li}_2$, and each of the identities (9.6) gives a functional equation for $\mathrm{Li}_3$. Abel's well-known equation for $\mathrm{Li}_2$ occurs for the unity group and $f(z) = \frac{z-1}{z}$, and Kummer's equation for $\mathrm{Li}_3$ occurs for the cyclic group of order 2 having the transformations $z' = z$, $z' = 1/z$, and $f(z) = \frac{(z-1)^2}{z}$.

We illustrate the method described in the preceding section by Abel's equation. We begin with the identity

$$I(z, w) = \log\frac{(1 - z)w}{(1 - w)z} - \log\frac{1 - z}{1 - w} - \log\frac{w}{z},$$

and get the main part

$$S(z, w) = \mathrm{Li}_2\left(\frac{(1 - z)w}{(1 - w)z}\right) - \mathrm{Li}_2\left(\frac{1 - z}{1 - w}\right) - \mathrm{Li}_2\left(\frac{w}{z}\right).$$

For $z = 0$ the first and third argument occurring in $S$ become infinite. If $z$ goes counterclockwise around 0 along a small circle, then these two arguments go clockwise around 0 and 1 along a large circle. This yields the branch term $2\pi i \log(1 - z) - 2\pi i \log(1 - w)$ which requires

$$\mathscr{E}(z, w) = \log z \log(1 - z) - 2\pi i \log z \log(1 - w) + \mathscr{E}_1(z, w).$$

For $z = 1$ the term $\log z \log(1 - z)$ has a branch point with the branch term $2\pi i \log z$. By (8.5), the function

$$\mathscr{E}(z, w) = \log z \log(1 - z) + \mathrm{Li}_2(z) - \log z \log(1 - w) + \mathscr{E}_2(z, w)$$

has no more branch points. Since it grows sufficiently slowly $\mathscr{E}_2(z, w)$ must be a constant by Liouville's theorem. Putting $z = w$ in the equation $S(z, w) = \mathscr{E}(z, w)$ we get $-\mathrm{Li}_2(1) = +\mathrm{Li}_2(w) + \mathscr{E}_2(w, w)$. This yields finally

$$\mathrm{Li}_2\left(\frac{(1-z)w}{(1-w)z}\right) - \mathrm{Li}_2\left(\frac{1-z}{1-w}\right) - \mathrm{Li}_2\left(\frac{w}{z}\right)$$
$$= \mathrm{Li}_2(z) - \mathrm{Li}_2(w) + \log z \log\frac{1-z}{1-w} - \frac{\pi^2}{6},$$

which is, up to elementary manipulations, identical with Abel's equation.

9.3.2. *Auxiliary algebraic tools.* According to §9.3.1 the main requirement for constructing equations for polylogarithms is to find identities of the form (K3). Our search for such identities is simplified by making use of an isomorphism

$$\log\frac{z-a_i}{w-b_i} \to x_i \qquad \text{for } i = 1, \dots, k.$$

So, for instance, the relation (9.3) becomes

$$(x_1 - 2x_2)^4 + (2x_1 - x_2)^4 + (x_1 + x_2)^4 - 9(x_1 - x_2)^4 - 9x_1^4 - 9x_2^4 = 0,$$

where $x_1 = \log z/w$ and $x_2 = \log((1-z)/(1-w))$. In general, we have to look for a linear dependency of $n$th powers of linear forms in $k$ variables $x_1, \dots, x_k$. The group $G$ induces a permutation group $H$ of $\{x_1, \dots, x_k\}$, and the linear forms corresponding to automorphic functions with respect to a subgroup $G_1$ are invariant with respect to the corresponding subgroup $H_1$ of $H$. Each element $(a_1x_1 + \dots + a_kx_k)^n$ belongs to the module $P_{n,k}$ of all homogeneous polynomials of degree $n$ in the variables $x_1, \dots, x_k$ with integer coefficients. Evidently, a basis in $P_{n,k}$ is formed by all $x_1^{n_1}x_2^{n_2}\cdots x_k^{n_k}$ with $n_1 + \dots + n_k = n$.

The task of finding a relation

$$b_1h_1^n + \dots + b_mh_m^n = 0, \tag{9.7}$$

with linear forms $h_1, \dots, h_m$ where $h_i$ is invariant with respect to the subgroup $H_i$ of $H$ can be simplified considerably by making use of the symmetry given by the group $H$.

To this end we denote by $th$ the linear form which arises from $h$ if $t$ is applied to $h$. Notice that $th$ is invariant with respect to $tH_it^{-1}$ if $h$ is invariant with respect to $H_i$. Let $H_0$ be some subgroup of $H$. From (9.7) we form

$$b_1\sum_{t\in H_0}(th_1)^n + \dots + b_m\sum_{t\in H_0}(th_m)^n = 0. \tag{9.8}$$

Notice that $th_i = t'h_i$ if $tH_i = t'H_i$. We use this as follows. If $N_i$ is a complete system of representatives of the left equivalence classes $t(H_i \cap H_0)$



within $H_0$, we can write (9.8) in the form

$$c_1\sum_{t\in N_1}(th_1)^n + \dots + c_m\sum_{t\in N_1}(th_m)^n = 0, \tag{9.9}$$

where now

$$f_i = \sum_{t\in N_i}(th_i)^n$$

is invariant with respect to $H_0$. So, with a relation (9.7) we always have a relation

$$c_1f_1 + \dots + c_mf_m = 0, \tag{9.10}$$

where the $f_i$ are invariant with respect to $H_0$. (9.10) is a relation in the module $P_{n,k}(H_0)$ of all homogeneous polynomials in $x$ variables of degree $n$ which are invariant with respect to $H_0$. A basis of $P_{n,k}(H_0)$ is described as follows: If $g$ is a basis element of $P_{n,k}$, then $\sum_{t\in H_0} tg$ is a basis element of $P_{n,k}(H_0)$.

The advantage of (9.10) over (9.7) is that $P_{n,k}(H_0)$ has fewer basis elements than $P_{n,k}$. This can considerably simplify the decision of whether a relation (9.7) exists. Examples are given in the next sections.

9.3.3. *Examples for* $\mathrm{Li}_4$. Assume, that a stereographic projection maps the vertices of a regular tetrahedron onto the points $a_1, a_2, a_3$ and $\infty$ of the complex plane. The group of automorphisms of the tetrahedron corresponds to a group $G$ of linear transformation of the complex plane for which $A = \{a_1, a_2, a_3, \infty\}$ is a complete system of equivalent points. The four cyclic subgroups of order 3 of $G$ have the automorphic functions

$$\frac{(z-a_1)^3}{(z-a_3)(z-a_2)}, \qquad \frac{(z-a_2)^3}{(z-a_1)(z-a_3)},$$
$$\frac{(z-a_3)^3}{(z-a_1)(z-a_2)}, \qquad (z-a_1)(z-a_2)(z-a_3).$$

The three cyclic subgroups of order 2 have the automorphic functions

$$\frac{(z-a_1)(z-a_2)}{z-a_3}, \qquad \frac{(z-a_2)(z-a_3)}{z-a_1}, \qquad \frac{(z-a_3)(z-a_1)}{z-a_2}.$$

The mapping $\log(z-a_i) \to x_i$ transforms the logarithms of these functions into the linear forms

$$h_1 = 3x_1 - x_2 - x_3, \qquad h_2 = 3x_2 - x_3 - x_1, \qquad h_3 = 3x_3 - x_1 - x_2,$$
$$h_4 = x_1 + x_2 + x_3, \qquad h_5 = x_1 + x_2 - x_3,$$
$$h_6 = x_2 + x_3 - x_1, \qquad h_7 = x_3 + x_1 - x_2.$$

We choose $H_0$ to be the symmetric group $S_3$ of all permutations of $\{x_1, x_2, x_3\}$. A basis of $P_{3,3}(S_3)$ is given by

$$u_1 = x_1^3 + x_2^3 + x_3^3,$$
$$u_2 = x_1^2x_2 + x_1^2x_3 + x_2^2x_1 + x_2^2x_3 + x_3^2x_1 + x_3^2x_2,$$
$$u_3 = x_1x_2x_3.$$

According to the preceding section we define the forms of $P_{3,3}(S_3)$

$$\begin{aligned} f_1 &= h_1^3 + h_2^3 + h_3^3 = 25u_1 - 21u_2 + 54u_3, \\ f_2 &= h_4^3 = u_1 + 3u_2 + 6u_3, \\ f_3 &= h_4^3 + h_5^3 + h_6^3 = u_1 + 3u_2 - 18u_3, \\ f_4 &= x_1^3 + x_2^3 + x_3^3 = u_1. \end{aligned}$$

One easily verifies the identity

$$f_1 + 3f_2 + 4f_3 - 32f_4 = 0,$$

which corresponds to an identity of the form (K3) with $n = 3$ having 10 terms. By the method explained at the end of 9.3 we get

THEOREM 9.2. *The tetrahedron group allows the construction of a Kummer-type* $\mathrm{Li}_4$*-equation whose main part contains* 70 $\mathrm{Li}_4$*-terms.*

PROOF. The main part of the equation is of the form $\sum_{t\in G} K(z, t(w))$, where $K$ is formed according to (K4). $K$ contains 10 $\mathrm{Li}_4$-terms, four of them being invariant under a cyclic group of order 3, three of them being invariant under a cyclic group of order 2, and three of them being invariant under the unity group. As $G$ has altogether 12 elements, we get in the main part of the equation $4\cdot 12/3 + 3\cdot 12/2 + 3\cdot 12 = 70$ different terms.

If one uses another system of equivalent points for $G$, namely the projections of the middle points of the six edges of the tetrahedron one gets

THEOREM 9.3. *The tetrahedron group allows the construction of a Kummer-type functional equation for* $\mathrm{Li}_4$ *whose main part consists of* 136 $\mathrm{Li}_4$*-terms.*
*Similarly, one gets*

THEOREM 9.4. *The octahedron group allows the construction of an equation whose main part consists of* 324 $\mathrm{Li}_4$*-terms having arguments of degree* 1, 2, 3, *and* 4.

One can expect similar results for the icosahedron group.

9.3.4. *Examples for* $\mathrm{Li}_5$. With the notations of §9.3.3 we consider the basis

$$\begin{aligned} u_1 &= x_1^4 + x_2^4 + x_3^4, \\ u_2 &= x_1^3x_2 + x_1^3x_3 + x_2^3x_1 + x_2^3x_3 + x_3^3x_1 + x_3^3x_2, \\ u_3 &= x_1^2x_2^2 + x_1^2x_3^2 + x_2^2x_3^2, \\ u_4 &= x_1^2x_2x_3 + x_2^2x_3x_1 + x_3^2x_1x_2, \end{aligned}$$



of $P_{4,3}(S_3)$. We define the forms of $P_{4,3}(S_3)$

$$\begin{aligned} f_1 &= h_1^4 + h_2^4 + h_3^4 = 83u_1 - 116u_2 + 114u_3 + 36u_4, \\ f_2 &= h_4^4 = u_1 + 4u_2 + 6u_3 + 12u_4, \\ f_3 &= h_5^4 + h_6^4 + h_7^4 = 3u_1 - 4u_2 + 18u_3 - 12u_4, \\ f_4 &= x_1^4 + x_2^4 + x_3^4 = u_1, \\ f_5 &= (x_1 - x_2)^4 + (x_2 - x_3)^4 + (x_3 - x_1)^4 = 2u_1 - 4u_2 + 6u_3. \end{aligned}$$

From this we get

$$f_1 + f_2 + 4f_3 - 32f_4 - 32f_5 = 0,$$

which corresponds to an identity (K3) for $n = 4$ with 13 terms. This gives

THEOREM 9.5. *The tetrahedron group allows the construction of a Kummer-type equation for* $\mathrm{Li}_5$ *whose main part consists of* 106 $\mathrm{Li}_5$*-terms with arguments of degree up to* 3.

By choosing different systems of equivalent points for the tetrahedron and octahedron group one finds

THEOREM 9.6. 1. *The tetrahedron group allows the construction of a Kummer-type equation for* $\mathrm{Li}_5$ *whose main part consists of* 286 $\mathrm{Li}_5$*-terms with arguments of degree up to* 3.

2. *The octahedron group allows the construction of three Kummer-type equations for* $\mathrm{Li}_5$ *with* 424, 648 *and* 1326 $\mathrm{Li}_5$*-terms with arguments of degree up to four.*

One can expect that the icosahedron group yields $\mathrm{Li}_5$-equations with arguments of degree 5 since this group has subgroups of order 5.

**9.4. The nonexistence of a Kummer-type functional equation for** $\mathrm{Li}_6$. Following [8] we prove on the basis of the theory developed in §§9.2—9.3.4

THEOREM 9.7. $\mathrm{Li}_6$ *does not satisfy a Kummer-type functional equation.*

PROOF. We have to prove that for no finite group $G$ of linear transformations of the complex plane can one find a relation

$$c_1h_1^5 + \cdots + c_mh_m^5 = 0, \tag{9.11}$$

where the $h_1, \dots, h_m$ are linear forms corresponding to automorphic functions with respect to subgroups of $G$.

I. THE TETRAHEDRON GROUP. Let $G$ be the tetrahedron group. The automorphic functions with respect to subgroups of $G$ may be built using one or more complete systems of points equivalent with respect to $G$, say $A = \{t(a);\, t \in G\}$, $B = \{t(b);\, t \in G\}$ etc. A possible relation (K3) based on $A, B, \dots$ would yield another one, if $a, b, \dots$ are identified with fixed points of transformations $t \in G$. Therefore, if a relation (K3) for $G$ exists, then such a relation exists for automorphic functions based on the system of fixed points of the elements from $G$.

The fixed points of the elements of $G$ are the stereographic projections $a_1, \dots, a_4$ of the vertices, $b_1, \dots, b_4$ of the middle points of the faces and $c_1, \dots, c_6$ of the middle points of the edges of the tetrahedron. We assume that $A_1, \dots, A_4, B_1, \dots, B_4, C_1, \dots, C_6$ are the original tetrahedron points, and we assume that $A_i$ and $B_i$ are diametral, $C_{i+2}$ is between $A_i$ and $A_{i+1}$ ($i$ mod 3) and $C_j$ and $C_{7-j}$ are diametral ($j = 1, 2, 3$).

Making use of the mapping

$$\log(z - a_i) = x_i, \qquad \log(z - b_i) = y_i, \qquad \log(z - c_i) = z_i,$$

and of the theory developed in §9.3.2 we show that a relation (9.11) for linear forms $h_i$ in the variables $x_1, \dots, x_4, y_1, \dots, y_4, z_1, \dots, z_6$ cannot exist. Such a relation would be a dependence relation in the linear space $H_{5,14}$ of the homogeneous forms of degree 5 in the 14 variables $x_1, \dots, z_6$. A basis of $H_{5,14}$ is $B = \{x_1^{n_1} \dots z_6^{n_{14}} : n_1 + \dots + n_{14} = 5,\ n_i \geq 0\}$.

The following types of linear forms must be considered

| | | | |
|---|---|---|---|
| $3x_1 - x_2 - x_3 - x_4,$ | (1) | $3y_1 - y_2 - y_3 - y_4,$ | (4) |
| $x_1 + x_2 - x_3 - x_4,$ | (2) | $y_1 + y_2 - y_3 - y_4,$ | (5) |
| $x_1 - x_2,$ | (3) | $y_1 - y_2,$ | (6) |
| $z_1 + z_2 + z_3 - z_4 - z_5 - z_6,$ | (7) | | |
| $z_1 + z_6 - z_2 - z_5,$ | (8) | | |
| $z_1 + z_6 - 2z_3,$ | (9) | | |
| $z_1 + z_6 - 2z_2,$ | (10) | | |
| $z_1 - z_6,$ | (11) | | |
| $z_1 - z_2,$ | (12) | | |
| $3x_1 - y_2 - y_3 - y_4,$ | (13) | | |
| $x_1 + x_2 + x_3 - y_1 - y_2 - y_3,$ | (14) | | |
| $3y_1 - x_2 - x_3 - x_4,$ | (15) | | |
| $x_1 + x_2 - y_1 - y_2,$ | (16) | | |
| $x_1 + x_2 - y_3 - y_4,$ | (17) | | |
| $x_1 + x_2 + x_3 + x_4 - y_1 - y_2 - y_3 - y_4,$ | (18) | | |
| $x_1 - y_1,$ | (19) | | |
| $x_1 - y_2,$ | (20) | | |
| $3x_1 - z_2 - z_3 - z_6,$ | (21) | $3y_1 - z_2 - z_3 - z_6,$ | (32) |
| $3x_1 - z_1 - z_4 - z_5,$ | (22) | $3y_1 - z_1 - z_4 - z_5,$ | (33) |
| $x_1 + x_2 + x_3 - z_1 - z_2 - z_3,$ | (23) | $y_1 + y_2 + y_3 - z_1 - z_2 - z_3,$ | (34) |
| $x_1 + x_2 + x_3 - z_4 - z_5 - z_6,$ | (24) | $y_1 + y_2 + y_3 - z_4 - z_5 - z_6,$ | (35) |
| $x_1 + x_2 - z_1 - z_6,$ | (25) | $y_1 + y_2 - z_1 - z_6,$ | (36) |
| $x_1 + x_2 - z_2 - z_5,$ | (26) | $y_1 + y_2 - z_2 - z_5,$ | (37) |
| $x_1 + x_2 - 2z_3,$ | (27) | $y_1 + y_2 - 2z_3,$ | (38) |
| $x_1 + x_2 - 2z_4,$ | (28) | $y_1 + y_2 - 2z_4,$ | (39) |
| $x_1 + x_2 + x_3 + x_4 - 2z_1 - 2z_6,$ | (29) | $y_1 + y_2 + y_3 + y_4 - 2z_1 - 2z_6,$ | (40) |
| $x_1 - z_1,$ | (30) | $y_1 - z_1,$ | (41) |
| $x_1 - z_2,$ | (31) | $y_1 - z_2,$ | (42) |

Notice that these forms only represent types of forms. We get the whole set of all forms of interest by applying the transformations of $G$ (understood



as permutations of $\{x_1, \dots, z_6\}$) to the forms (1)–(42). For example, $3x_1 - x_2 - x_3 - x_4$, $3x_2 - x_1 - x_3 - x_4$, $3x_3 - x_1 - x_2 - x_4$, and $3x_4 - x_1 - x_2 - x_3$ are all forms of type (1).

We now assume that there exists a relation (9.1) for forms of the types (1)–(42). We show the following:

1. This relation cannot contain forms of the types (13)–(42).
2. If this relation has the form $I = 0$, then $I = I_1 + I_2 + I_3$ such that $I_1 = I_1(x_1, \dots, x_4)$, $I_2 = I_2(y_1, \dots, y_4)$, $I_3 = I_3(z_1, \dots, z_6)$ and $I_1 = 0$, and $I_2 = 0$, and $I_3 = 0$.
3. Nontrivial relations $I_1 = 0$, $I_2 = 0$, and $I_3 = 0$ are impossible.

AD 1. We first consider (18). $(x_1 + x_2 + x_3 + x_4 - y_1 - y_2 - y_3 - y_4)^5$ needs for instance the basis element $x_1 y_1 y_2 y_3 y_4$ which does not occur in any other of the forms (1)–(42). Hence, (18) cannot occur in $I$. Therefore, for similar reasons, (13), (14), and (15) cannot occur in $I$. From this it follows that (16) and (17) can be excluded and this has the consequence that (19) and (20) cannot occur in $I$.

Next, we convince ourselves that (21)–(31) do not occur in $I$. Here we start with (29) which is the only form requiring a basis element like $z_1 x_1 x_2 x_3 x_4$. Then we see that (21)–(24) cannot participate in $I$ and from this it follows that $I$ cannot contain (25) and (26), and hence not (27) and (28), and hence also not (30)–(31). Finally, (32)–(42) cannot occur in $I$ because the set of forms is symmetric in the $x_i$'s and the $y_i$'s.

AD 2. If $I = 0$ exists, then in $I$ we can have only the forms of types (1)–(3), (4)–(6), and (7)–(12). Since the forms of types (1)–(3) need only basis elements built up purely by $x_i$'s and the forms of types (4)–(6) need only basis elements built up purely by $y_i$'s, the forms of these two groups are pairwise linearly independent. Hence, if $I = 0$, then we must have $I = I_1(x_1, \dots, x_4) + I_2(y_1, \dots, y_4) + I_3(z_1, \dots, z_6)$, with $I_1 = I_2 = I_3 = 0$.

AD 3. We assume that $I_1 = 0$ has the form

$$\begin{aligned} &c_1(3x_1 - x_2 - x_3 - x_4)^5 + c_2(3x_2 - x_3 - x_4 - x_1)^5 + c_3(3x_3 - x_4 - x_1 - x_2)^5 \\ &+ c_4(3x_4 - x_1 - x_2 - x_3)^5 + c_5(x_1 + x_2 - x_3 - x_4)^5 + c_6(x_1 + x_3 - x_2 - x_4)^5 \\ &+ c_7(x_2 + x_3 - x_1 - x_4)^5 + c_8(x_1 - x_4)^5 + c_9(x_2 - x_4)^5 + c_{10}(x_3 - x_4)^5 \\ &+ c_{11}(x_1 - x_3)^5 + c_{12}(x_2 - x_3)^5 + c_{13}(x_1 - x_2)^5 = 0. \end{aligned}$$

We express the terms $h^5$ by means of the basis elements of $H_{5,14}$ and consider the coefficients of the 7 basis elements $x_1^2 x_2 x_3 x_4$, $x_1 x_2^2 x_3 x_4$, $x_1 x_2 x_3^2 x_4$, $x_1 x_2 x_3 x_4^2$, $x_1^2 x_2^2 x_3$, $x_1^2 x_2 x_3^2$, $x_1 x_2^2 x_3^2$ which must vanish. This gives the following equations:

$$\begin{aligned}
-9c_1 - 3c_2 - 3c_3 - 3c_4 + c_5 + c_6 - c_7 &= 0,\\
-3c_1 - 9c_2 - 3c_3 - 3c_4 + c_5 - c_6 + c_7 &= 0,\\
-3c_1 - 3c_2 - 9c_3 - 3c_4 - c_5 + c_6 + c_7 &= 0,\\
-3c_1 - 3c_2 - 3c_3 - 9c_4 - c_5 - c_6 - c_7 &= 0,\\
-9c_1 - 9c_2 - 3c_3 - c_4 - c_5 + c_6 + c_7 &= 0,\\
-9c_1 - 9c_2 - c_3 + 3c_4 - c_5 - c_6 - c_7 &= 0,\\
-c_1 + 3c_2 - 9c_3 - 9c_4 + c_5 - c_6 + c_7 &= 0.
\end{aligned}$$

This system has only the trivial solution zero. This implies that also $c_8 = c_9 = \cdots = c_{13} = 0$.

Because of the symmetry of the $x_i$'s and the $y_i$'s we conclude that $I_2 = 0$ can exist only trivially.

Finally we consider $I_3 = 0$. Considering the coefficients of the basis elements $z_1z_2z_3z_4z_5$, $z_1z_2z_3z_4z_6$, $z_1z_2z_3z_5z_6$, $z_1z_2z_4z_5z_6$ and $z_1^2z_2z_3z_4$ shows that the forms of type (7) cannot occur in $I_3$. Then, as the forms of type (8) require basis elements occurring within no other form, the forms of type (8) cannot occur in $I_3$. Hence, (9) and (10) can be excluded, and finally (11) and (12) can be excluded.

II. THE OCTRAHEDRON GROUP AND THE ICOSAHEDRON GROUP. These groups are dealt with in the same way. We omit the tedious computations.

III. THE DIHEDRAL GROUPS. Let $D_{2k}$ be the dihedral group with the elements $\{T, T^2, \ldots, T^k, S_1, \ldots, S_k\}$ where $T(z) = \varepsilon^2 z$, $S_i(z) = \frac{\varepsilon^{2i}}{z}$ and $\varepsilon = e^{\pi i/k}$. $D_{2k}$ has the following subgroups

(a) $k$ cyclic subgroups $C_2^{(j)}$ of order 2 which are generated by $S_j$ $(j = 1, \ldots, k)$.

(b) the cyclic group $C_k = \{T, T^2, \ldots, T^k\}$ and its subgroups $C_j$ where $j$ divides $k$.

(c) the dihedral groups $D_{2j}$ where $j$ divides $k$.

As in the first case we consider the fixed point set $A = \{1, \varepsilon, \varepsilon^2, \ldots, \varepsilon^{2k-1}, 0, \infty\}$. Let

$$\begin{aligned}
x_\nu &= \log(z - \varepsilon^{2\nu-1}), && \nu = 1, \ldots, k,\\
y_\nu &= \log(z - \varepsilon^{2\nu}), && \nu = 1, \ldots, k,\\
u &= \log z.
\end{aligned}$$

We look for a relation (9.1) in $H_{5,2k+1}$ in the variables $x_1, \ldots, x_k, y_1, \ldots, y_k, u$.

From now on we consider the case $k \equiv 1(2)$. The other case is similar. We have to consider the following invariant forms

(a) For $C_2^{(j)}$

$$\begin{aligned}
f_{\lambda,\nu} &= 2x_\lambda - x_{\lambda+\nu} - x_{\lambda-\nu},\\
g_{\lambda,\nu} &= 2x_\lambda - y_{\lambda+\nu} - y_{\lambda-\nu-1},
\end{aligned}$$



$$\begin{aligned}
f'_{\lambda,\nu} &= 2y_\lambda - x_{\lambda+\nu} - x_{\lambda-\nu-1},\\
g'_{\lambda,\nu} &= 2y_\lambda - y_{\lambda+\nu} - y_{\lambda-\nu},\\
h_{\lambda,\nu,\mu} &= x_{\lambda+\nu} + x_{\lambda-\nu-1} - y_{\lambda+\mu} - y_{\lambda-\mu},\\
h'_{\lambda,\nu,\mu} &= x_{\lambda+\nu} + x_{\lambda-\nu} - x_{\lambda+\mu} - x_{\lambda-\mu-1},\\
h''_{\lambda,\nu,\mu} &= y_{\lambda+\mu} + y_{\lambda-\mu} - y_{\lambda+\nu} - y_{\lambda-\nu},\\
f_\lambda &= 2x_\lambda - u,\\
f'_\lambda &= 2x_\lambda,\\
g_\lambda &= 2y_\lambda - u,\\
g'_\lambda &= 2y_\lambda,\\
u_{\lambda,\nu} &= x_{\lambda+\nu} + x_{\lambda-\nu-1} - u,\\
v_{\lambda,\nu} &= y_{\lambda+\nu} + y_{\lambda-\nu} - u,
\end{aligned}$$

(b) For the groups $C_j$ $(j' = k/j)$

$$\begin{aligned}
\varphi_{j,\nu} &= x_\nu + x_{\nu+j'} + \cdots + x_{\nu+(j-1)j'},\\
\psi_{j,\nu} &= y_\mu + y_{\mu+j'} + \cdots + y_{\mu+(j-1)j'},\\
\varphi_{j,\nu,\mu} &= \varphi_{j,\nu} - \varphi_{j,\mu}, \quad \text{for } \nu \not\equiv \mu(j'),\\
\psi_{j,\nu,\mu} &= \psi_{j,\nu} - \psi_{j,\mu}, \quad \text{for } \nu \not\equiv \mu(j'),\\
\chi_{j,\nu,\mu} &= \varphi_{j,\nu} - \psi_{j,\mu},\\
\varphi'_{j,\nu} &= \varphi_{j,\nu} - ju,\\
\psi'_{j,\nu} &= \psi_{j,\nu} - ju,\\
\omega_j &= ju.
\end{aligned}$$

(c) For the groups $D_{2k}$

$$\begin{aligned}
\varphi''_{j,\nu} &= 2\varphi_{j,\nu} - ju,\\
\psi''_{j,\nu} &= 2\psi_{j,\nu} - ju.
\end{aligned}$$

Now, we assume an identity

$$c_1h_1^5 + \cdots + c_mh_m^5 = 0, \tag{9.12}$$

for these forms.

As the basis representation of $\varphi^5_{j,\nu,\mu}$ contains the element

$$x_\nu x_{\nu+j'} x_{\nu+2j'} x_\mu x_{\mu+j'},$$

for $\nu \not\equiv \mu(j')$ and no other form needs this basis element, the forms $\varphi_{j,\nu,\mu}$ cannot occur in (9.12). Similarly we proceed with the forms $\psi_{j,\nu,\mu}$ and $\chi_{j,\nu,\mu}$.

The forms $\varphi'^5_{j,\nu}$ and $\varphi''^5_{j,\nu}$ are the only forms containing the basis elements

$$x_\nu x_{\nu+j'} x_{\nu+2j'} x_{\nu+3j'} u \quad\text{and}\quad x_\nu x_{\nu+j'} x_{\nu+2j'} u^2.$$

Their coefficients can vanish only if

$$c_1 + 16c_2 = 0,$$
$$c_1 + 8c_2 = 0,$$

where $c_1$ and $c_2$ are the coefficients of $\varphi'^5_{j,\nu}$ and $\varphi''^5_{j,\nu}$ in (9.12), respectively. This yields $c_1 = c_2 = 0$.

Similarly, we exclude $\psi'_{j,\nu}$ and $\psi''_{j,\nu}$, and hence the $\varphi_{j,\nu}$ and $\psi_{j,\nu}$ do not occur in (9.12).

Next, each of the forms $h_{\lambda,\nu,\mu}$, $h'_{\lambda,\nu,\mu}$ and $h''_{\lambda,\nu,\mu}$ contains basis elements occurring in no other form. Hence, these forms cannot occur in (9.12). Since $\varphi'_{j,\nu}$, $\varphi''_{j,\nu}$, $\psi'_{j,\nu}$ and $\psi''_{j,\nu}$ are already excluded the same argument is true for $u_{j,\nu}$ and $v_{j,\nu}$. This argument is also true for those forms $f_{j,\nu}$, $g_{j,\nu}$, $f'_{j,\nu}$, $g'_{j,\nu}$ which correspond to automorphic functions whose zeros and singularities are not vertices of a regular triangle.

If this latter case occurs, however, then as an example, with $2x_i - x_j - x_m$ we have also to consider $2x_j - x_m - x_i$ and $2x_m - x_i - x_j$. As an easy computation shows, there does not exist an identity for the 5th powers of these forms and the remaining forms $x_i - x_j$, $x_m - x_j$, $x_j - x_i$.

Finally, the $f_\lambda$ cannot occur in (9.12). Their basis representation contains terms depending simultaneously on $x_\lambda$ and $u$. Help could come only from $x_\lambda - u$. The requirement that the coefficients of $x_\lambda^4 u$ and $x_\lambda^3 u^2$ in the basis representation of $c_1(2x_\lambda - u)^5 + c_2(x_\lambda - u)^5$ vanish leads to the system of equations

$$16c_1 + c_2 = 0,$$
$$8c_1 + c_2 = 0,$$

which has only the trivial solution zero.

This shows that an identity (9.12) can only be built up by the forms $mu$, $u$, $2x_\lambda$, $x_\lambda$, $2y_\lambda$, $y_\lambda$. The only possibilities are

$$(mu)^5 - m^5u^5 = 0,$$
$$u^5 + (-u)^5 = 0.$$

The corresponding identities containing logarithms are

$$(\log z^m)^5 - m^5(\log z)^5 = 0,$$

and

$$(\log z)^5 + (\log 1/z)^5 = 0,$$

which lead to the well-known factorization theorem

$$\mathrm{Li}_6(z^m) - m^5\sum_{\mu=1}^{m}\mathrm{Li}_6(e^{2\pi i\mu/m}z) = 0, \tag{9.13}$$

and the inversion theorems

$$\mathrm{Li}_6(z) + \mathrm{Li}_6(1/z) = P_6(\log z), \tag{9.14}$$



respectively. In the latter equation, $P_6$ is a polynomial of degree 6.

It is evident that the arguments used for $n = 6$ are even more valid if $n$ is larger than 6. This proves the theorem.

The proof shows the reason for the nonexistence of higher-order Kummer-type functional equations for polylogarithms: The finite groups of linear transformations of the complex plane do not have enough subgroups and hence there do not exist enough automorphic functions compared with the dimension of the corresponding $P_{n,k}$ for $n \geq 6$.

## References


1. N. H. Abel, *Note sur la fonction* $\psi(x) = x + (x^2/2^2) + (x^3/3^2) + \cdots$, Oeuvres, vol. II, Christiania, 1839, pp. 189–229.
2. E. E. Kummer, *Über die Transzendenten, welche aus wiederholten Integrationen rationaler Funktionen entstehen*, J. Reine Angew. Math. **21** (1840), 74–90.
3. N. Nielsen, *Über den Eulerschen Dilogarithmus*, Nova acta Leopoldina, **90** (1909), 154–155.
4. L. Lewin, *Polylogarithms and associated functions*, North Holland, New York and Oxford, 1981.
5. F. Klein, *Über das Ikosaeder*, Teubner Arch. Math., Leipzig, 1884.
6. L. R. Ford, *Automorphic functions*, McGraw-Hill, New York, 1929.
7. G. Wechsung, *Lineare Funktionalgleichungen von Polylogarithmen*, Wiss. Z. Friedrich-Schiller Univ. Jena Natur. Reihe, **14** (1965), 401–408.
8. ——, *Über die Unmöglichkeit des Vorkommens von Funktionalgleichungen gewisser Struktur für Polylogarithmen*, Aequationes Math. **5** (1970), 54–62.

CHAPTER 10

# The Basic Structure of Polylogarithmic Functional Equations

ZDZISLAW WOJTKOWIAK

*Pamięci mojej Matki Heleny Wojtkowiak, ofiary czarnej nocy 1981–1989*

**10.1. Introduction.** The function $\log z$ satisfies the functional equation

$$\log x + \log y = \log(x \cdot y).$$

The dilogarithm $\mathrm{Li}_2(z) := \int_0^z -(\log(1-z)/z\,dz$ satisfies the following functional equation

$$\mathrm{Li}_2\left(\frac{x}{1-x}\cdot\frac{y}{1-y}\right) = \mathrm{Li}_2\left(\frac{y}{1-x}\right) + \mathrm{Li}_2\left(\frac{x}{1-y}\right) - \mathrm{Li}_2(x) - \mathrm{Li}_2(y) - \log(1-x)\log(1-y),$$

(see [A]). Let us set $\mathrm{Li}_0(z) := -\log z$, $\mathrm{Li}_1(z) := -\log(1-z)$ and $\mathrm{Li}_n(z) := \int_0^z \mathrm{Li}_{n-1}(z)/z\,dz$ for $n > 1$. It was expected that functions $\mathrm{Li}_n(z)$ will satisfy functional equations similar to functional equations of $\log z$ and $\mathrm{Li}_2(z)$. In fact various functional equations of functions $\mathrm{Li}_n(z)$ for small $n$ were found. The basic reference is Lewin's book (see [L]).

Our aim is to find new functional equations satisfied by these functions and to give some general results about structures of these equations.

Before we formulate our results we shall make one observation concerning polylogarithms. The functions $\mathrm{Li}_n(z)$ are special cases of Chen iterated integrals. We recall their definition. Let $\omega_1, \dots, \omega_n$ be one-forms on a smooth manifold $M$ and let $\gamma$ be a smooth path from $x$ to $z$. Then we define by a recursive formula:

$$\int_\gamma \omega_1, \dots, \omega_n := \int_\gamma \left(\int_{\gamma^t} \omega_1\right)\omega_2, \dots, \omega_n,$$

where $\gamma^t$ denotes the restriction $\gamma|[0, t]$.

(Instead of $\int_\gamma$ we shall also write $\int_{x,\gamma}^z$ or $\int_x^z$.)

It is clear that

$$\mathrm{Li}_n(z) = \int_0^z -\frac{dz}{z-1}, \frac{dz}{z}, \dots, \frac{dz}{z}.$$

We shall usually write a rational function $f\colon P^1(\mathbb{C}) \to P^1(\mathbb{C})$ in the form

$$f(z) = \alpha \prod_{i=1}^{n} (z - a_i)^{n_i} / \prod_{j=1}^{m} (z - b_j)^{m_j},$$

where $\alpha \in \mathbb{C}$, $n_i$ and $m_j$ are positive integers and $a_i$, $b_j$ are complex numbers.

DEFINITION 10.1.1. We say that $f(z)$ is in an irreducible form if $a_i \neq b_j$ for $i = 1, \dots, n$ and $j = 1, \dots, m$.

DEFINITION 10.1.2. Let $a \in \mathbb{C}$. If

$$f(z) - a = \alpha^* \prod_{k=1}^{r} (z - c_k)^{r_k} / \prod_{j=1}^{p} (z - b_j)^{p_j}$$

is an irreducible form then we define a divisor $f^{-1}(a)$ by the formula $f^{-1}(a) := \sum_{k=1}^{r} r_k \cdot c_k$.

Now we state our main results.

THEOREM A. *Let* $f(z) = \alpha \prod_{i=1}^{n} (z - a_i)^{n_i} / \prod_{j=1}^{m} (z - b_j)^{m_j}$ *be a map from* $P^1(\mathbb{C})$ *to* $P^1(\mathbb{C})$ *and let* $f^{-1}(1) = \sum_{k=1}^{r} r_k \cdot c_k$.

We have the following formula

(10.1.3)

$$\begin{aligned}
&\mathrm{Li}_2(f(z)) - \mathrm{Li}_2(f(x)) + \log(1 - f(x))(\log(f(z)) - \log(f(x))\\
&\quad = \sum_{i,k} n_i \cdot r_k \left(\mathrm{Li}_2\left(\frac{z - a_i}{c_k - a_i}\right) - \mathrm{Li}_2\left(\frac{x - a_i}{c_k - a_i}\right) + \log\frac{x - c_k}{a_i - c_k}\log\frac{z - a_i}{x - a_i}\right)\\
&\quad - \sum_{j,k} m_j \cdot r_k \left(\mathrm{Li}_2\left(\frac{z - b_j}{c_k - b_j}\right) - \mathrm{Li}_2\left(\frac{x - b_j}{c_k - b_j}\right) + \log\frac{x - c_k}{b_j - c_k}\log\frac{z - b_j}{x - b_j}\right)\\
&\quad - \sum_{i,j} n_i \cdot m_j \left(\mathrm{Li}_2\left(\frac{z - a_i}{b_j - a_i}\right) - \mathrm{Li}_2\left(\frac{x - a_i}{b_j - a_i}\right) + \log\frac{x - b_j}{a_i - b_j}\log\frac{z - a_i}{x - a_i}\right)\\
&\quad - \sum_{j<j'} m_j \cdot m_{j'} \left(\log\frac{z - b_j}{x - b_j}\right)\left(\log\frac{z - b_{j'}}{x - b_{j'}}\right) - \frac{1}{2}\sum_j m_j^2 \left(\log\frac{z - b_j}{x - b_j}\right)^2.
\end{aligned}$$

The following summation convention is used in the formula 10.1.3 and it will be used throughout the whole chapter.

$$\sum_{i,k} = \sum_{i=1}^{n}\sum_{k=1}^{r}, \qquad \sum_{j<j'} = \sum_{j=1}^{m-1}\sum_{j'=j+1}^{m},$$
$$\sum_{j} = \sum_{j=1}^{m}, \qquad \sum_{i<i',k} = \sum_{i<i'}\sum_{k=1}^{r},$$

and so on for three or more indices.



One of the difficulties in dealing with formula 10.1.3 in Theorem A is that the functions $\mathrm{Li}_2(z)$ and $\log z$ are multivalued. For example, for some values of $\mathrm{Li}_2(z)$ the formula can be satisfied while for others not. It is not clear at all which values of $\mathrm{Li}_2(z)$ to choose.

However, Theorem A is derived from the following formula which has no ambiguity at all.

First we formulate assumptions.

10.1.4. *Let* $f(z) = \alpha \prod_{i=1}^{n} (z - a_i)^{n_i} / \prod_{j=1}^{m} (z - b_j)^{m_j}$ *be in an irreducible form and let* $f^{-1}(1) = \sum_{k=1}^{r} r_k \cdot c_k$. *Let*

$$X = P^1(\mathbb{C}) \setminus \{a_1, \dots, a_n, b_1, \dots, b_m, c_1, \dots, c_r, \infty\}$$

*and let* $Y = P^1(\mathbb{C}) \setminus \{0, 1, \infty\}$. *Let* $\gamma$ *be a smooth path in* $X$ *from* $x$ *to* $z$. *We assume that* $z$ *and* $x$ *are different from* $a_1, \dots, a_n, b_1, \dots, b_m, c_1, \dots, c_r, \infty$.

THEOREM A$'$. (integral form of the functional equation). *Let us assume that* 10.1.4 *holds. Then we have*

(10.1.3$'$)

$$\begin{aligned}
&\int_{f(\gamma)} \frac{-dz}{z - 1}, \frac{dz}{z} = \sum_{i,k} n_i \cdot r_k \int_{f_{ik}(\gamma)} \frac{-dz}{z - 1}, \frac{dz}{z} - \sum_{j,k} m_j \cdot r_k \int_{h_{jk}(\gamma)} \frac{-dz}{z - 1}, \frac{dz}{z}\\
&\quad - \sum_{i,j} n_i \cdot m_j \int_{g_{ij}(\gamma)} \frac{-dz}{z - 1}, \frac{dz}{z} - \sum_{j<j'} m_j \cdot m_{j'} \left(\int_\gamma \frac{dz}{z - b_j}\right)\left(\int_\gamma \frac{dz}{z - b_{j'}}\right)\\
&\quad - \frac{1}{2}\sum_j m_j^2 \int_\gamma \frac{dz}{z - b_j}, \frac{dz}{z - b_j},
\end{aligned}$$

*where* $f_{ik}(z) = (z - a_i)/(c_k - a_i)$, $g_{ij}(z) = (z - a_i)/(b_j - a_i)$, *and* $h_{jk}(z) = (z - b_j)/(c_k - b_j)$.

To get the expression from Theorem A we must calculate an integral $\int_\varphi \frac{-dz}{z-1}, \frac{dz}{z}$ where $\varphi$ is a path from $a$ to $b$. We have

$$\begin{aligned}
\int_\varphi \frac{-dz}{z - 1}, \frac{dz}{z} &= \int_\varphi -(\log(1 - z) - \log(1 - a))\frac{dz}{z}\\
&= \mathrm{Li}_2(b) - \mathrm{Li}_2(a) + \log(1 - a)(\log b - \log a).
\end{aligned}$$

Observe that $\log(1 - a)$ we could choose arbitrary, but when $\log(1 - a)$ is fixed then $\log(1 - z)$ is determined uniquely. $\mathrm{Li}_2(a)$ can be chosen arbitrary, but $\mathrm{Li}_2(b)$ is determined uniquely by $\mathrm{Li}_2(a)$ and $\log(1 - z)$.

Now it is clear which values of $\mathrm{Li}_2$ and log we must choose in formula 10.1.1 in order to have an equality.

Suppose that we have chosen such values of $\mathrm{Li}_2$ and log that we have no equality any more. Then we can always add some expression containing logarithms and constants to the right-hand side so that once more we have an equality. This is due to the fact that different branches of $\mathrm{Li}_2(z)$ and $\log z$ are given by $\mathrm{Li}_2(z) + 2\pi i k \log z$ and $\log z + 2\pi i k$, where $k \in Z$.

This suggests a new formulation of Theorem A.

**Theorem A″.** *Let $f(z)$ and $f^{-1}(1)$ be as in Theorem A. Then we have*

$$
\begin{aligned}
\mathrm{Li}_2(f(z)) - \mathrm{Li}_2(f(x)) = &\sum_{i,j} n_i \cdot r_k (\mathrm{Li}_2(f_{ik}(z) - \mathrm{Li}_2(f_{ik}(x))) \\
&- \sum_{j,k} m_j \cdot r_k (\mathrm{Li}_2(h_{jk}(z)) - \mathrm{Li}_2(h_{jk}(x))) \\
&- \sum_{i,j} n_i \cdot m_j (\mathrm{Li}_2(g_{ij}(z)) - \mathrm{Li}_2(g_{ij}(x))) \\
&+ \mathrm{l.d.t.}(2)
\end{aligned}
\tag{10.1.3″}
$$

*where* l. d. t. (2) *is a polynomial in logarithms and constants.*

We shall show that from the formula 10.1.3 one can get all functional equations of the dilogarithm in one variable. Also we shall show that one can get most known functional equations of the dilogarithm from 10.1.3 by suitably choosing the function $f(z)$.

We have a similar formula for $\mathrm{Li}_3(z)$. In the introduction we state only a special case when the function $f(z)$ is a polynomial function.

**Theorem B.** *Let $f(z) = \alpha \prod_{i=1}^n (z - a_i)^{n_i}$ and let $f^{-1}(1) = \sum_{k=1}^r r_k \cdot c_k$. We have the following formula*

$$
\begin{aligned}
&\mathrm{Li}_3\left(\alpha \prod_{i=1}^n (z-a_i)^{n_i}\right) - \mathrm{Li}_3\left(\alpha \prod_{i=1}^n (x-a_i)^{n_i}\right) \\
&\quad + \mathrm{Cor}\left(\alpha \prod_{i=1}^n (z-a_i)^{n_i}, \alpha \prod_{i=1}^n (x-a_i)^{n_i}\right) \\
&= -\sum_{i<i'} \sum_k n_i \cdot n_{i'} \cdot r_k \left(\mathrm{Li}_3\left(\frac{z-a_i}{z-a_{i'}} \cdot \frac{c_k - a_{i'}}{c_k - a_i}\right) - \mathrm{Li}_3\left(\frac{x-a_i}{x-a_{i'}} \cdot \frac{c_k - a_{i'}}{c_k - a_i}\right)\right. \\
&\qquad\qquad \left. + \mathrm{Cor}\left(\frac{z-a_i}{z-a_{i'}} \cdot \frac{c_k - a_{i'}}{c_k - a_i}; \frac{x-a_i}{x-a_{i'}} \cdot \frac{c_k - a_{i'}}{c_k - a_i}\right)\right) \\
&\quad - \sum_{i',i,k} n_i \cdot n_{i'} \cdot r_k \left(\mathrm{Li}_3\left(\frac{z-a_i}{c_k - a_i}\right) - \left(\mathrm{Li}_3 \frac{x-a_i}{c_k - a_i}\right)\right. \\
&\qquad\qquad \left. + \mathrm{Cor}\left(\frac{z-a_i}{c_k - a_i}; \frac{x-a_i}{c_k - a_i}\right)\right) \\
&\quad + \sum_{i<i'} \sum_k n_i \cdot n_{i'} \cdot r_k \left(\mathrm{Li}_3\left(\frac{z-a_i}{z-a_{i'}}\right) - \mathrm{Li}_3\left(\frac{x-a_i}{x-a_{i'}}\right)\right. \\
&\qquad\qquad \left. + \mathrm{Cor}\left(\frac{z-a_i}{z-a_{i'}}; \frac{x-a_i}{x-a_{i'}}\right)\right),
\end{aligned}
$$

*where* $\mathrm{Cor}(a; b) = -\mathrm{Li}_2(b)\log(a/b) - (1/2)\log(1-b)(\log(a/b))^2$.



We leave to the reader the formulation of the integral form of Theorem B. Then one can also fix values of $\mathrm{Li}_3$, $\mathrm{Li}_2$ and log for which one has an equality.

**Definition 10.1.5.** Let $f(z) = \alpha \prod_{i=1}^n (z-a_i)^{n_i} / \prod_{j=1}^m (z-b_j)^{m_j}$ be a rational function in an irreducible form. We set $\deg f := \max(\sum_{i=1}^n n_i, \sum_{j=1}^m m_j)$ and we call this number the degree of $f$.

**Definition 10.1.6.** Let $n$ be a natural number.

$$\mathrm{l.d.t.}(n)(\text{resp. } \overline{\mathrm{l.d.t.}(n)}) := p(c_1, \dots, c_r, \mathrm{Li}_{t_1}(g_1(z)), \dots, \mathrm{Li}_{t_s}(g_s(z))),$$

where $p(x_1, \dots, x_r, y_1, \dots, y_s)$ is a polynomial with rational coefficients, $c_j = 2\pi i$ or $\mathrm{Li}_k(a_j)$ where $a_j \in \mathbb{C}$ and $k < n$ (resp. $k \le n$) for $j = 1, \dots, r$; $g_i(z)$ are rational functions on $P^1(\mathbb{C})$ and $t_i < n$ for $i = 1, \dots, s$.

Observe that in Theorem A we expressed $\mathrm{Li}_2(f(z))$ as a sum of $\mathrm{Li}_2(g(z))$'s where $g(z)$ are rational functions of degree one, of logarithmic terms and constants. The same holds for $\mathrm{Li}_3(f(z))$. This is not a general phenomena as we shall see in the next theorem.

**Theorem C.** *Let $f(z)$ be a rational function of degree $k$ greater than 1. Let us assume that $f(z)$ is not a $k$th power. Let $n$ be a natural number greater than 3. Then there is no functional equation of the form*

$$\mathrm{Li}_n(f(z)) = \sum_{i=1}^N n_i \mathrm{Li}_n(f_i(z)) + \overline{\mathrm{l.d.t.}(n)},$$

*where the $f_i(z)$ are rational functions of degree 1 and $n_i$ $(i = 1, \dots, N)$ are rational numbers.*

While proving Theorems A and B we met the problem of expressing iterated integrals of the form $\int_x^z \frac{dz}{z-a_1}, \dots, \frac{dz}{z-a_n}$ by classical polylogarithms. The next result related to Theorem C shows that this is usually impossible.

**Theorem D.** *Let $a_1, a_2, a_3, a_4$ be four different points in $\mathbb{C}$.*

(a) *The function*

$$N(z) = \int_x^z \frac{dz}{z-a_1}, \frac{dz}{z-a_2}, \frac{dz}{z-a_3}$$

*can be expressed by classical polylogarithms.*

(b) *Let*

$$L(z) = \int_x^z \frac{dz}{z-a_1}, \frac{dz}{z-a_2}, \frac{dz}{z-a_3}, \frac{dz}{z-a_4}.$$

*There is no polynomial $p(s, t_1, \dots, t_r)$ which depends essentially on $s$ such that*

$$p(L(z), \mathrm{Li}_{n_1}(f_1(z)), \dots, \mathrm{Li}_{n_r}(f_r(z))) \equiv 0,$$

*where* $\mathrm{Li}_{n_k}$ *are classical polylogarithms (and logarithms) and* $f_i(z)$ *are rational functions.*

The principal tools in our investigations are two observations.

1. Functions of the type of polylogarithms are horizontal sections of the canonical unipotent connection on $P^1(\mathbb{C})\backslash\{a_1, \dots, a_n\}$.
2. The functional equations of functions of the type of polylogarithms are consequences of relations between maps induced by regular functions from $P^1(\mathbb{C})\backslash$ several points to $P^1(\mathbb{C})\backslash$ several points on Lie algebras of fundamental groups.

We illustrate the second principal with a few examples.

EXAMPLE 1. The maps $f(x) = x$ and $g(x) = 1 - x$ from

$$X = P^1(\mathbb{C})\backslash\{0, 1, \infty\}$$

into itself induce opposite maps on $\Gamma^2\pi_1(X, x)/\Gamma^3\pi_1(X, x)$, therefore we have a functional equation

$$\mathrm{Li}_2(x) + \mathrm{Li}_2(1 - x) = \overline{\mathrm{l.d.t.}}(2).$$

EXAMPLE 2. The maps $f(x) = x^2$, $g(x) = x$, and $h(x) = -x$ from $X = P^1(\mathbb{C})\backslash\{0, 1, -1, \infty\}$ to $P^1(\mathbb{C})\backslash\{0, 1, \infty\}$ satisfy $f_* - 2g_* - 2h_* = 0$ on $\Gamma^2\pi_1(X, x)/\Gamma^3\pi_1(X, x)$, therefore there is a functional equation

$$\mathrm{Li}_2(x^2) - 2\,\mathrm{Li}_2(x) - 2\,\mathrm{Li}_2(-x) = \overline{\mathrm{l.d.t.}}(2).$$

EXAMPLE 3. Let $f_1(x) = x$, $f_2(x) = 1/(1 - x)$, $f_3(x) = x/(x - 1)$, $f_4(x) = 1/x$ be maps from $X = P^1(\mathbb{C})\backslash\{0, 1, \infty\}$ into itself. In

$$\mathrm{Hom}(\Gamma^3\pi_1(X, x)/\Gamma^4\pi_1(X, x); \Gamma^3\pi_1(X, x)/\Gamma^4\pi_1(X, x) + [V[U, V]]),$$

where $U$ is a loop around 0 and $V$ is a loop around 1 we have $f_{1^*} = f_{4^*}$ and $f_{1^*} + f_{2^*} + f_{3^*} = 0$. Hence there are functional equations

$$\mathrm{Li}_3(x) = \mathrm{Li}_3\left(\frac{1}{x}\right) + \overline{\mathrm{l.d.t.}}(3),$$

and

$$\mathrm{Li}_3(x) + \mathrm{Li}_3\left(\frac{1}{1 - x}\right) + \mathrm{Li}_3\left(\frac{x}{x - 1}\right) = \overline{\mathrm{l.d.t.}}(3).$$

EXAMPLE 4. Let $X = P^1(\mathbb{C})\backslash\{0, 1, \infty\}$, $f(x) = x$ and $g(x) = 1/x$. Let $U$ be a loop around 0 and let $V$ be a loop around 1. On the quotient $\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x) + L$, where $L$ is a subgroup of $\Gamma^n\pi_1(X, x)$ generated by all these commutators which contain $V$ at least twice, we have $f_* = (-1)^{n-1}g_*$. Therefore we have a functional equation

$$\mathrm{Li}_n(z) = (-1)^{n-1}\,\mathrm{Li}_n(1/z) + \overline{\mathrm{l.d.t.}}(n).$$

All these examples follow easily from the following theorem.



THEOREM E. *Let*

$$X = P^1(\mathbb{C})\backslash\{a_1, \dots, a_r, \infty\} \quad \text{and} \quad Y = P^1(\mathbb{C})\backslash\{0, 1, \infty\}.$$

*Let* $U$ *(resp.* $V$*) be a loop around* 0 *(resp.* 1*) in* $Y$. *Let* $f_1, \dots, f_N : X \to Y$ *be regular maps from* $X$ *to* $Y$ *and let* $n_1, \dots, n_N$ *be integers. There is a functional equation*

$$n_1\,\mathrm{Li}_n(f_1(z)) + \cdots + n_N\,\mathrm{Li}_n(f_N(z)) + \overline{\mathrm{l.d.t.}}(n) = 0,$$

*if and only if*

$$n_1 f_{1^*} + \cdots + n_N f_{N^*} = 0,$$

*in the* $Z$*-module*

$$\mathrm{Hom}(\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x); \Gamma^n\pi_1(Y, y)/\Gamma^{n+1}\pi_1(Y, y) + L_n)$$

*where* $L_n$ *is a subgroup of* $\Gamma^n\pi_1(Y, y)/\Gamma^{n+1}\pi_1(Y, y)$ *generated by all commutators which contain* $V$ *at least twice and* $f_{i^*}$ *is the map induced by* $f_i$ *on fundamental groups.*

Theorem E has the following generalization.

THEOREM F. *Let* $X$ *be a smooth quasiprojective algebraic variety over* $\mathbb{C}$. *Let* $f_1, \dots, f_N$ *be regular maps from* $X$ *to* $Y = P^1(\mathbb{C})\backslash\{0, 1, \infty\}$ *and let* $n_1, \dots, n_N$ *be integers. There is a functional equation*

$$n_1\,\mathrm{Li}_n(f_1(x)) + \cdots + n_N\,\mathrm{Li}_n(f_N(x)) + \overline{\mathrm{l.d.t.}}(n) = 0,$$

*if and only if*

$$n_1 f_{1^*} + \cdots + n_N f_{N^*} = 0,$$

*in the* $Z$*-module*

$$\mathrm{Hom}(\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x); \Gamma^n\pi_1(Y, y)/\Gamma^{n+1}\pi_1(Y, y) + L_n).$$

Observe that the definition of $\overline{\mathrm{l.d.t.}}(n)$ should be modified in Theorem F. One requires that $g_i(z)$ in Definition 10.1.5 are regular functions from $X$ to $Y$. This theorem gives an interpretation of functional equations in several variables as well as functional equations of polylogarithms whose arguments are arbitrary algebraic functions. We shall not prove Theorem F in this chapter.

Theorem E is our principal result. From this theorem we derived all our results about functional equations of polylogarithms.

We should point out that D. Zagier obtained a very short and elegant proof of the related result for higher Bloch-Wigner functions using a version of generalized Bloch homomorphisms (see [Z3]).

**10.2. Canonical unipotent connection on** $P^1(\mathbb{C})\backslash\{a_1, \dots, a_{n+1}\}$. Let $X = P^1(\mathbb{C})\backslash\{a_1, \dots, a_{n+1}\}$. Let $A^*(X)$ be a differential, graded subalgebra of

$\Omega^*(X)$ generated by linear combinations with complex coefficients of one-forms $dz/(z-a_i)$ $i = 1, \ldots, n+1$. It is a trivial observation that $(A^1(X))^* \approx H_1(X, \mathbb{C})$. The isomorphism is given by the bilinear form

$$\int : A^1(X) \otimes H_1(X, \mathbb{C}) \to \mathbb{C}, \quad \text{given by} \quad (\omega, \gamma) \to \int_\gamma \omega.$$

Let $L(\pi_1(X, x)) := \varprojlim_N (\bigoplus_{n=1}^N (\Gamma^n \pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x)) \otimes \mathbb{C})$ be a Lie algebra associated with the lower central series of $\pi_1(X, x)$. We equip $L(\pi_1(X, x))$ with the group law given by the Baker-Hausdorff formula and the topology given by the inverse limit of finite dimensional complex vector spaces. This topological group we denote by $\pi(X)$. The Lie algebra of $\pi(X)$ is $L(\pi_1(X, x))$.

We shall define a one-form $\omega_X$ on $X$ with values in $L(\pi_1(X, x))$ in the following way. We have natural isomorphisms

$$(10.2.1) \quad A^1(X) \otimes H_1(X, \mathbb{C}) \approx A^1(X) \otimes (A^1(X))^* \approx \operatorname{Hom}(A^1(X), A^1(X)).$$

DEFINITION 10.2.2. $\omega_X \in A^1(X) \otimes H_1(X, \mathbb{C})$ is the one-form which corresponds to $2\pi i \cdot \mathrm{id}$ under the isomorphisms 10.2.1. (see also [D, 12.5.5]).

We consider $\omega_X$ as an element of $A^1(X) \otimes L(\pi_1(X, x))$ because of the identification $H_1(X, \mathbb{C}) \approx (\pi_1(X, x)/\Gamma^2\pi_1(X, x)) \otimes \mathbb{C}$.

Let $A_i$ be a loop around $a_i$ in $X$ and let $X_i$ be the image of $A_i$ in $H_1(X, \mathbb{C})$.

Then $L(\pi_1(X, x))$ is the completion of the free Lie algebra over $\mathbb{C}$ on $n$ generators $X_1, \ldots, X_n$ with respect to the filtration given by the lower central series.

Let us assume that $a_{n+1} = \infty$ then

$$(10.2.2.1) \qquad \omega_X = \sum_{i=1}^n \frac{dz}{z - a_i} \otimes X_i.$$

If $a_i \neq \infty$ for $i = 1, \ldots, n+1$ then

$$(10.2.2.2) \qquad \omega_X = \sum_{i=1}^n \left( \frac{dz}{z - a_i} - \frac{dz}{z - a_{n+1}} \right) \otimes X_i.$$

Let $\mathbb{C}[[H_1(X, \mathbb{C})]]$ be an algebra of noncommutative, formal power series on $H_1(X, \mathbb{C})$. We shall denote it shortly by $\mathbb{C}[[X]]$. Let $I$ be an augmentation ideal of $\mathbb{C}[[X]]$. Then $C[[X]]/I^n$ is a finite-dimensional complex vector space, $\mathbb{C}[[X]] = \varprojlim C[[X]]/I^n$ and we equip $C[[X]]$ with the topology of an inverse limit of finite-dimensional complex vector spaces. Let $C[[X]]^*$ be the group of invertible elements in $C[[X]]$. From the discussion given above it follows that $C[[X]]^*$ is a topological group, an inverse limit of finite-dimensional complex Lie groups. We shall denote the group $C[[X]]^*$ by $P(X)$. The Lie algebra of Lie elements, possibly of infinite length, in $\mathbb{C}[[X]]$ is naturally identified with $L(\pi_1(X, x))$. After this identification the exponential map



$$\exp: \pi(X) \to P(X),$$
$$\exp(w) = e^w = 1 + \frac{w}{1!} + \frac{w^2}{2!} + \cdots$$

is defined. The exponential map is a continuous monomorphism of topological groups, whose image is a closed subgroup of $P(X)$. The inverse of exp is defined on the subgroup $\exp(\pi(X)) \subset P(X)$ and we denote it by log.

Let $\operatorname{Lie} P(X)$ be the Lie algebra of $P(X)$. We identify $T \in H_1(X; \mathbb{C}) \subset L(\pi_1(X, X))$ with the tangent vector to $P(X)$ in 1 given by $t \to 1 + tT$. After this identification we consider the one-form $\omega_X$ as a one-form with values in $\operatorname{Lie} P(X)$. We denote it by $\overline{\omega}_X$. The homomorphism exp maps $\omega_X$ into $\overline{\omega}_X$.

Let us consider a principal $\pi(X)$-bundle $X \times \pi(X) \to X$ equipped with the integrable connection given by a one-form $\omega_X$, and a principal $P(X)$-bundle; $X \times P(X) \to X$ equipped with the integrable connection given by a one-form $\overline{\omega}_X$.

LEMMA 10.2.3. *The morphism* $\mathrm{id} \times \exp: X \times \pi(X) \to X \times P(X)$ *over* $\mathrm{id}_X$ *maps horizontal sections with respect to* $\omega_X$ *into horizontal sections with respect to* $\overline{\omega}_X$.

PROOF. This is clear from the fact that exp maps $\omega_X$ into $\overline{\omega}_X$.

It is clear that there is no need to distinguish between $\omega_X$ and $\overline{\omega}_X$, hence from now on we shall denote both forms by $\omega_X$.

**10.3. Horizontal sections.** Let $X = P^1(\mathbb{C}) \backslash \{x_1, \ldots, x_{n+1}\}$. Let $\gamma$ be a smooth path in $X$ from $x$ to $z$. We shall denote by $(z, l_X(z; x, \gamma))$ (resp. $(z, \lambda_X(z; x, \gamma))$) or shortly by $(z, l_X(z; x))$ (resp. $(z, \lambda_X(z; x))$) the value at $z$ of the horizontal section of the bundle $X \times \pi(X) \to X$ (resp. $X \times P(X) \to X$) equipped with the connection form $\omega_X$ along the path $\gamma$ with the initial condition $l_X(x; x, \gamma) = 0$ (resp. $\lambda_X(x; x, \gamma) = 1$).

Let us set

$$\omega_i := -\left[ \frac{dz}{z - x_i} - \frac{dz}{z - x_{n+1}} \right] \qquad i = 1, \ldots, n,$$

if $x_i \neq \infty$ for $i = 1, \ldots, n+1$. If one $x_i = \infty$, then we assume that $x_{n+1} = \infty$ and we set $\omega_i := -dz/(z - x_i)$ for $i = 1, \ldots, n$.

Let us define

$$\Lambda_X(\varepsilon_1^{n_1}, \ldots, \varepsilon_k^{n_k})(z) := \int_{x,\gamma}^z \omega_{\varepsilon_k}, \ldots, \omega_{\varepsilon_k}, \ldots, \omega_{\varepsilon_1}, \ldots, \omega_{\varepsilon_1},$$

where $\varepsilon_i \in \{1, \ldots, n\}$ and $\omega_{\varepsilon_1}$ repeats $n_1$-times $, \ldots, \omega_{\varepsilon_k}$ repeats $n_k$-times.

LEMMA 10.3.1. *The application*

$$X \ni z \to \left(z, 1 + \sum \Lambda_x(\varepsilon_1^{n_1}, \dots, \varepsilon_k^{n_k})(z) X_{\varepsilon_1}^{n_1} \cdots X_{\varepsilon_k}^{n_k}\right) \in X \times P(X)$$

*is horizontal with respect to the connection* $\omega_X = \sum_{i=1}^n -\omega_i \otimes X_i$ *and hence it coincides with the map* $z \to (z, \lambda_X(z; x))$. *(The summation is over all noncommutative monomials in variables* $X_1, \dots, X_n$ *where* $X_i$ *is the class in* $H_1(X, \mathbb{C})$ *of a loop around* $a_i$.)

PROOF. This is a straightforward calculation of horizontal liftings.

Let $X = P^1(\mathbb{C})\backslash\{x_1, \dots, x_{p+1}\}$ and let $Y = P^1(\mathbb{C})\backslash\{y_1, \dots, y_{q+1}\}$. Let $f(z) = \alpha \prod_{i=1}^n (z-a_i)^{n_i} / \prod_{j=1}^m (z-b_j)^{m_j}$ be a rational function. Let us assume that $f$ restricts to a regular map $f: X \to Y$. The map $f$ induces

$$f^*: A^*(Y) \to A^*(X),$$
$$H_1(f): H_1(X) \to H_1(Y),$$

and

$$f_\#: \pi_1(X, x) \to \pi_1(Y, f(x)).$$

The maps $H_1(f)$ and $f_\#$ induce the following three maps

$$f_*: L(\pi_1(X, x)) \to L(\pi_1(Y, f(x))),$$
$$f_*: \pi(x) \to \pi(Y),$$
$$f_*: P(X) \to P(Y).$$

In the next proposition $G(X)$ is $\pi(X)$ (resp. $P(X)$) and $G(Y)$ is $\pi(Y)$ (resp. $P(Y)$).

PROPOSITION 10.3.2. *The map* $(f, f \times f_*)$ *of principal fibre bundles*

$$\begin{array}{ccc} X \times G(X) & \xrightarrow{f \times f_*} & Y \times G(Y) \\ {\scriptstyle (1)}\downarrow & & {\scriptstyle (2)}\downarrow \\ X & \xrightarrow{f} & Y \end{array}$$

*satisfies*

$$(\mathrm{id} \otimes f_*)\omega_X = (f^* \otimes \mathrm{id})\omega_Y.$$

PROOF. This is a direct verification for which one can use explicit formulas 10.2.2.1 and 10.2.2.2 for $\omega_X$ and $\omega_Y$.

COROLLARY 10.3.3. *The map* $f \times f_*$ *maps horizontal sections of the bundle* (1) *into horizontal sections of the bundle* (2). *This implies that we have the following equalities*

(10.3.3.1) $$f_*(l_X(z; x, \gamma)) = l_Y(f(z); f(x), f(\gamma)),$$

*and*

(10.3.3.2) $$f_*(\lambda_X(z; x, \gamma)) = \lambda_Y(f(z); f(x), f(\gamma)).$$

PROOF. The corollary is an immediate consequence of Proposition 10.3.2.



**10.4. Easy lemmas about monodromy.** Let $X = P^1(\mathbb{C})\backslash\{x_1, \dots, x_{n+1}\}$. Let $\alpha$ be a loop in $X$ based at $x \in X$ and let $\gamma$ be a path from $x$ to $z$. The function $l_X(z; x): P^1(\mathbb{C})\backslash\{x_1, \dots, x_{n+1}\} \to \pi(X)$ is a multivalued function. This means that in general $l_X(z; x, \gamma\circ\alpha)$ is different from $l_X(z; x, \gamma)$. Let us set $l_X^\alpha(z; x, \gamma) := l_X(z; x, \gamma\circ\alpha)$. We denote the action of $\alpha$ on $l_X(z; x)$ in the following way

$$\alpha: l_X(z; x) \to l_X^\alpha(z; x),$$

and we shall call this action of $\alpha$, the monodromy of the function $l_X(z; x)$ along $\alpha$. We recall that $\pi(X) = L(\pi_1(X, x))$ is the completion of the free Lie algebra over $\mathbb{C}$ on $n$ generators $X_1, \dots, X_n$ where $X_k$ is the class in $H_1(X, \mathbb{C})$ of the loop $A_k$ around $a_k$.

LEMMA 10.4.1. *The monodromy of the function* $l_X(z; x)$ *along the loop* $A_k$ *is given by the following formula*

$$l_X^{A_k}(z; x) = l_X(z; x) \cdot (-2\pi i X_k + \textit{terms of degree} \geq 2).$$

PROOF. The function $l_X(z; x)$ is the horizontal section of the principal $\pi(X)$-bundle. Hence its monodromy along any loop $\alpha \in \pi_1(X, x)$ is given by the following formula

$$l_X^\alpha(z; x) = l_X(z; x) \cdot l_X(x; x, \alpha).$$

Observe that

$$l_X(z; x) = \sum_{k=1}^n (-\log(z - a_k) + \log(x - a_k)) X_k + \text{terms of degree} \geq 2,$$

if $x_{n+1} = \infty$ and

$$l_X(z; x) = \sum_{k=1}^n (-\log(z-a_k) + \log(z-a_{n+1}) + \log(x-a_k) - \log(x - a_{n+1})) X_k$$
$$+ \text{terms of degree} \geq 2 \text{ if } x_k \neq \infty \text{ for } k = 1, 2, \dots, n+1.$$

This implies that $l_X(x; x, A_k) = -2\pi i X_k + \text{terms of degree} \geq 2$.

Let $L$ be a free Lie algebra on generators $x_1, \dots, x_n$. Then for any fixed ordering of elements $x_1, \dots, x_n$ there is a base of $L$ consisting of basic Lie elements corresponding to this ordering (see [MCS]).

Let $B_m = \{e_i\}_{i \in I}$ be a base of

$$\Gamma^m \pi(X)/\Gamma^{m+1}\pi(X) = (\Gamma^m \pi_1(X, x)/\Gamma^{m+1}\pi_1(X, x)) \otimes \mathbb{C}$$

given by basic Lie elements corresponding to the ordering $X_1, X_2, \dots, X_n$. Let $e_i^*$ be a linear functional dual to $e_i$ with respect to the base $B_m$. We shall consider $e_i^*$ as a polynomial function on $\pi(X)$. We are interested in the monodromy of $e_i^*(l_X(z; x))$.

COROLLARY 10.4.2. *Let $e_i$ and $e_j$ belong to $B_m$. The monodromy of $e_i^*(l_X(z;x))$ is trivial on $\Gamma^d\pi_1(X,x)$ for $d > m$. The monodromy of $e_i^*(l_X(z;x))$ on $\Gamma^m\pi(X,x)/\Gamma^{m+1}\pi(X,x)$ is given by the following formula*

$$e_j : e_i^*(l_X(z;x)) \to e_i^*(l_X(z;x)) + (-2\pi i)^m\delta_i^j.$$

PROOF. It follows from Lemma 10.4.1 that the monodromy of $l_X(z;x)$ on $e_i$ is given by $e_i : l_X(z;x) \to l_X(z;x) + (-2\pi i)^m e_i$ + terms of degree $\geq m$. This implies the corollary.

COROLLARY 10.4.3. *The image of the homomorphism $\pi_1(X,x) \to \pi(X)/\Gamma^n\pi(X)$ given by $\pi_1(X,x) \ni \alpha \to l_X(x;x,\alpha) \in \pi(X)/\Gamma^n\pi(X)$ is Zariski dense in $\pi(X)/\Gamma^n\pi(X)$ for each $n \geq 2$.*

PROOF. Lemma 10.4.1 implies that the image of the composite homomorphism

$$\pi_1(X,x) \to \pi(X) \to \pi(X)/\Gamma^2\pi(X)$$

is Zariski dense in $\pi(X)/\Gamma^2\pi(X)$. Hence it follows that for each $n$ the image of the composite homomorphism

$$\pi_1(X,x) \to \pi(X) \to \pi(X)/\Gamma^n\pi(X)$$

is Zariski dense in $\pi(X)/\Gamma^n\pi(X)$.

**10.5. Functional equations.** In this section we present general results about functional equations. Let $X$ be a complex projective line minus several points. Let $G(X)$ be $\pi(X)$ or $P(X)$. Observe that $G(X)$ is an affine pro-algebraic group. Let $\mathrm{Alg}(G(X))$ be an algebra of polynomial, complex-valued functions on $G(X)$.

Now we set $X = P^1(\mathbb{C})\setminus\{x_1,\dots,x_{n+1}\}$ and $Y = P^1(\mathbb{C})\setminus\{y_1,\dots,y_{m+1}\}$. Let $f: X \to Y$ be a regular map. Let $x \in X$ and $z \in X$ and let $\gamma$ be a path in $X$ from $x$ to $z$. The equalities

(10.3.3.1) $$f_*l_X(z;x,\gamma) = l_Y(f(z);f(x),f(\gamma)),$$

and

(10.3.3.2) $$f_*\lambda_X(z;x,\gamma) = \lambda_Y(f(z);f(x),f(\gamma))$$

are our principal tool to derive functional equations. In fact these equalities are special cases of functional equations.

THEOREM 10.5.1. *Let $f_1,\dots,f_N: X \to Y$ be regular functions. Let $\mathscr{C}_1,\dots,\mathscr{C}_N$ belong to $\mathrm{Alg}(G(Y))$ and let $p(t_1,\dots,t_N)$ be a polynomial in variables $t_1,\dots,t_N$.*

(i) *Let $G(\ ) = \pi(\ )$. There is a functional equation*

(1) $p(\mathscr{C}_1(l_Y(f_1(z);f_1(x),f_1(\gamma))),\dots,\mathscr{C}_N(l_Y(f_N(z);f_N(x),f_N(\gamma)))) = 0,$

*if and only if*

(2) $$p(\mathscr{C}_1\circ f_{1*},\dots,\mathscr{C}_N\circ f_{N*}) = 0.$$

(ii) *Let $G(\ ) = P(\ )$. If*

$$p(\mathscr{C}_1\circ f_{1*},\dots,\mathscr{C}_N\circ f_{N*}) = 0,$$

*then*

$$p(\mathscr{C}_1(\lambda_Y(f_1(z);f_1(x),f_1(\gamma))),\dots,\mathscr{C}_N(\lambda_Y(f_N(z);f_N(x),f_N(\gamma)))) = 0.$$

PROOF. Let us assume that we have (2). Corollary 10.3.3 implies that

$$\mathscr{C}_i(f_{i*}(l_X(z;x\gamma))) = \mathscr{C}_i(l_Y(f_i(z);f_i(x),f_i(\gamma))).$$

Replacing $\mathscr{C}_i(f_{i*}(l_X(z;x\gamma)))$ by $\mathscr{C}_i(l_Y(f_i(z);f_i(x),f_i(\gamma)))$ in the formula (2) we get the functional equation (1). The same arguments show also the part ii).

Let us assume that we have a functional equation (1). It follows from Lemma 10.4.3 that the set of values $l_X(x;x,\gamma)$ for all closed loops $\gamma$ is Zariski dense in $\pi(X)/\Gamma^n\pi(X)$ for all $n$. Vanishing of a regular function $p(\mathscr{C}_1\circ f_{1*},\dots,\mathscr{C}_n\circ f_{n*})$ on a Zariski dense subset implies that this regular function is the zero function.

Now we shall construct some elements of $\mathrm{Alg}(\pi(Y))$ which will be particularly interesting for us.

Let $\mathrm{Lie}(\pi_1(Y,y)) := \bigoplus_{n=1}^{\infty}((\Gamma^n\pi_1(Y,y)/\Gamma^{n+1}\pi_1(Y,y))\otimes\mathbb{C})$. Then $\mathrm{Lie}(\pi_1(Y,y))$ is a free Lie algebra on generators $Y_1,\dots,Y_m$ where each $Y_i$ is a class in $\pi_1(Y,y)/\Gamma^2\pi_1(Y,y)$ of a loop around $y_i$. Let us choose a base of $\mathrm{Lie}(\pi_1(Y,y))$ given by basic Lie elements corresponding to the ordering $Y_1,\dots,Y_n$. Let $v \in \mathrm{Lie}(\pi_1(Y,y))$ be a basic Lie element and let $v^*$ be a linear functional on $\mathrm{Lie}(\pi_1(Y,y))$ dual to $v$ with respect to the base of basic Lie elements, i.e., $v^* \in \mathrm{Hom}(\mathrm{Lie}(\pi_1(Y,y));Z)$. The linear functional $v^*$ we consider as an element of $\mathrm{Alg}\,\pi(Y)$. We set

$$\mathscr{L}_v(z;x,\gamma) := v^*(l_Y(z;x,\gamma)).$$

We shall also write $\mathscr{L}_v(z;x)$ instead of $\mathscr{L}_v(z;x,\gamma)$.

COROLLARY 10.5.2. *Let $f_1,\dots,f_N: X \to Y$ be regular functions, let $n_1,\dots,n_N$ be integers and let $v_1,\dots,v_N$ in $\mathrm{Lie}(\pi_1(Y,y))$ be basic Lie elements of degree $n$ not necessarily different. There is a functional equation*

$$\sum_{i=1}^{N}\mathscr{L}_{v_i}(f_i(z);f_i(x),f_i(\gamma)) = 0,$$

*if and only if*

$$\sum_{i=1}^{N} n_i(v_i^*\circ(f_i)_*) = 0,$$

*in $\mathrm{Hom}(\Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x);Z)$ where*

$$(f_i)_*: \Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x) \to \Gamma^n\pi_1(Y,y)/\Gamma^{n+1}\pi_1(Y,y)$$

*is induced by $f_i$.*

PROOF. The corollary follows immediately from Theorem 10.5.1 if one observes that the condition $\sum_{i=1}^{N} n_i(v_i^* \circ (f_i)_*) = 0$ in $\operatorname{Alg}\pi(X)$ is equivalent to the condition $\sum_{i=1}^{N} n_i(v_i^* \circ (f_i)_*) = 0$ in $\operatorname{Hom}(\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x); Z)$ because of the identification

$$(\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x)) \otimes \mathbb{C} \approx \Gamma^n\pi(X)/\Gamma^{n+1}\pi(X).$$

COROLLARY 10.5.3. *Let $b(X)$ be a base of $\operatorname{Lie}(\pi_1(X, x))$ given by basic Lie elements. The functions $\{\mathscr{L}_v(z; x_0) | v \in b(X)\}$ are algebraically independent on $X$.*

PROOF. Let $v_1, \dots, v_n$ be different elements of $b(X)$. Let $p(t_1, \dots, t_n)$ be a polynomial with complex coefficients such that

$$p(\mathscr{S}_{v_1}(z; x_0), \dots, \mathscr{L}_{v_n}(z; x_0)) \equiv 0.$$

It follows from Theorem 10.5.1 that $p(v_1^*, \dots, v_n^*) = 0$ in $\operatorname{Alg}(\pi(X))$. The functions $v_1^*, \dots, v_n^*$ are linearly independent generators of the algebra $\operatorname{Alg}(\pi(X))$. Hence the polynomial $p(x_1, \dots, x_n)$ is equal to 0.

COROLLARY 10.5.4. *The functions $\{\mathscr{L}_v(z; x_0) | v \in b(x)\}$ are algebraically independent on any open disc around $x_0$.*

PROOF. Assume that we have an identity $p(\mathscr{L}_{v_1}(z; x_0), \dots, \mathscr{L}_{v_n}(z; x_0)) \equiv 0$ on a small disc around $x_0$. Then by analytic continuation we have such an equality along any path. Hence Corollary 10.5.3 implies that the polynomial $p(x_1, \dots, x_n)$ is identically equal to zero.

**10.6. Functional equations of polylogarithms.** Now we shall restrict our attention to polylogarithms. The following assumptions will be used through the whole section so we extract them at the very beginning.

(10.6.1.) Let $X = P^1(\mathbb{C}) \setminus \{x_1, \dots, x_{n+1}\}$ and let $Y = P^1(\mathbb{C}) \setminus \{0, 1, \infty\}$. Let $f_1, \dots, f_N: X \to Y$ be regular functions and let $n_1, \dots, n_N$ be integers. Let $x$ and $z$ belong to $X$ and let $\gamma$ be a smooth path in $X$ from $x$ to $z$.

(10.6.2.) Let $U$ and $V$ be loops in $Y$ in a clockwise direction around points 0 and 1 respectively. We consider $U$ and $V$ as elements of the Lie algebra $\operatorname{Lie}(\pi_1(Y, y))$. Let us set $e_0 := U$, $e_1 := V$, $e_2 := [V, U]$, $e_n := [e_{n-1}, U]$ for $n \geq 2$. Let $e_n^*$ be a linear functional on $\operatorname{Lie}(\pi_1(Y, y))$ dual to $e_n$ with respect to the base of $\operatorname{Lie}(\pi_1(Y, y))$ given by basic Lie elements corresponding to the ordering $U$, $V$. We consider $e_n^*$ as an element of $\operatorname{Alg}\pi(Y)$.

DEFINITION 10.6.3. Let $\mathscr{C}_{n+1}: P(Y) = C[[U, V]]^* \to \mathbb{C}$ associate to an element of $P(Y)$ its coefficient at $U^n V$. We set

$$\operatorname{Li}_n(z; x, \gamma) := (-1)^{n-1}\mathscr{C}_n(\lambda_Y(z; x, \gamma)).$$



We shall write also $\operatorname{Li}_n(z; x)$ when we do not specify the path $\gamma$. Observe that

$$\operatorname{Li}_n(z; x, \gamma) = (-1)^{n-1} \int_{x,\gamma}^{z} \frac{-dz}{z-1}, \frac{-dz}{z}, \dots, \frac{-dz}{z}$$
$$= \int_{x,\gamma}^{z} \frac{-dz}{z-1}, \frac{dz}{z}, \dots, \frac{dz}{z},$$

where $-dz/z$ appears $n-1$ times.

The following theorem is an immediate consequence of the results from section 10.5.

THEOREM 10.6.4. (functional equation of polylogarithms; integral form and abstract form). *Assume* 10.6.1 *and* 10.6.2. *Then we have:*

(i) *There is a functional equation*

$$\sum_{i=1}^{N} n_i \mathscr{L}_{e_n}(f_i(z); f_i(x), f_i(\gamma)) = 0, \tag{10.1}$$

*if and only if one of the following equivalent conditions is satisfied.*

(10.2)
$$\sum_{i=1}^{N} n_i e_n^* \circ (f_i)_* = 0 \quad \textit{in the group } \operatorname{Hom}(\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x); Z);$$

(10.3)
$$\sum_{i=1}^{N} n_i e_n^* \circ (f_i)_* = 0 \quad \textit{in the group}$$
$$\operatorname{Hom}(\Gamma^n\pi_1(X, x)/(\Gamma^{n+1}\pi_1(X, x) + [\Gamma^2\pi_1(X, x), \Gamma^2\pi_1(X, x)] \cap \Gamma^n\pi_1(X, x)); Z);$$

(10.4)
$$\sum_{i=1}^{N} n_i (f_i)_* = 0 \quad \textit{in the group}$$
$$\operatorname{Hom}(\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x); \Gamma^n\pi_1(Y, y)/\Gamma^{n+1}\pi_1(Y, y) + L_n),$$

*where $L_n$ is a subgroup of $\Gamma^n\pi_1(Y, y)$ generated by all commutators which contain $V$ at least twice.*

(ii) *If*

$$\sum_{i=1}^{N} n_i \mathscr{C}_n \circ (f_i)_* = 0 \quad \textit{in } \operatorname{Alg}(P(X)), \tag{10.5}$$

*then there is a functional equation*

$$\sum_{i=1}^{N} n_i \operatorname{Li}_n(f_i(z); f_i(x), f_i(\gamma)) = 0. \tag{10.6}$$

The formulas (10.1) and (10.6) are integral forms of functional equations whilst the formulas (10.2), (10.3), (10.4) and (10.5) are abstract forms of functional equations.

PROOF. It follows from Corollary 10.5.2 that (10.1) is equivalent to (10.2). Conditions (10.2), (10.3) and (10.4) are evidently equivalent. Theorem 10.5.1 implies that the condition (10.5) implies the condition (10.6).

Now we shall show that the function $\mathscr{L}_{e_n}(z;x)$ can be expressed by classical polylogarithms.

LEMMA 10.6.5. *We have*

(i) $\mathrm{Li}_n(z;x) = \mathrm{Li}_n(z) - \mathrm{Li}_n(x) + \mathrm{l.d.t.}(n)$.

(ii) $\mathscr{L}_{e_n}(z;x) - \mathrm{Li}_n(z;x) = \mathrm{l.d.t.}(n)$.

PROOF. The point (i) is a direct calculation. Hence it remains to show (ii). We recall that a horizontal section of the bundle $Y \times P(Y) \to Y$ is $\lambda_Y(z;x)$ while a horizontal section of the bundle $Y \times \pi(Y) \to Y$ is $l_Y(z;x)$. It follows from Lemma 10.2.3 that $\exp l_Y(z;x) = \lambda_Y(z;x)$. The coefficient of $\exp l_Y(z;x)$ at $U^nV$ is equal to

$$(-1)^n \mathscr{L}_{e_{n+1}}(z;x) + \sum_{k=2}^{n} \frac{(-1)^n}{k!}\left(\int_x^z \frac{dz}{z}\right)^{k-1} \mathscr{L}_{e_{n-k+2}}(z;x) + \frac{(-1)^n}{(n+1)!}\left(\int_x^z \frac{dz}{z}\right)^n \left(\int_x^z \frac{-dz}{z-1}\right).$$

On the other side the coefficient of $\lambda_Y(z;x)$ at $U^nV$ is equal to $(-1)^n \mathrm{Li}_{n+1}(z;x)$. Comparing these two coefficients it follows by induction and the point (i) of Lemma 10.6.5 that $\mathscr{L}_{e_n}(z;x) - \mathrm{Li}_n(z;x) = \mathrm{l.d.t.}(n)$.

Now we can show the following corollary of Theorem 10.6.4.

COROLLARY 10.6.6. *Assume* 10.6.1 *and* 10.6.2. *Then the following conditions are equivalent:*

(i) *there is a functional equation* $\sum_{i=1}^N n_i(\mathrm{Li}_n(f_i(z) - \mathrm{Li}_n(f_i(x))) + \mathrm{l.d.t.}(n) = 0$;

(ii) *there is a functional equation* $\sum_{i=1}^N n_i \mathrm{Li}_n(f_i(z)) + \overline{\mathrm{l.d.t.}}(n) = 0$;

(iii) $\sum_{i=1}^N n_i(f_i)_* = 0$ *in the group*

$$\mathrm{Hom}(\Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x);\, \mathrm{Hom}(\Gamma^n\pi_1(Y,y)/\Gamma^{n+1}\pi_1(Y,y) + L_n).$$

PROOF. It follows from Lemma 10.6.5 that

$$\mathscr{L}_{e_n}(f_i(z); f_i(x), f_i(\gamma)) = \mathrm{Li}_n(f_i(z)) - \mathrm{Li}_n(f_i(x)) + \mathrm{l.d.t.}(n).$$

Substituting these expressions for $\mathscr{L}_{e_n}(f_i(z); f_i(x), f_i(\gamma))$ in the formula (10.1) from Theorem 10.6.4 we get $\sum_{i=1}^N n_i(\mathrm{Li}_n(f_i(z)) - \mathrm{Li}_n(f_i(x))) + \mathrm{l.d.t.}(n) = 0$. Hence (iii) implies (i). Observe that $\mathrm{Li}_n(f_i(x)) + \mathrm{l.d.t.}(n) = \overline{\mathrm{l.d.t.}(n)}$. Hence (i) implies (ii).



Assume that (ii) is satisfied. Then it follows from Lemma 10.6.5 that $\sum_{i=1}^N n_i(\mathscr{L}_{e_n}(f_i(z); f_i(x), f_i(\gamma)) + \overline{\mathrm{l.d.t.}}(n) = 0$ for some choice of $\overline{\mathrm{l.d.t.}}(n)$. Let $\gamma \in \Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x)$. Observe that the monodromy of $\mathrm{l.d.t.}(n)$ on $\Gamma^n\pi_1(Y,y)$ is trivial. This follows immediately from Corollary 10.4.2 and Lemma 10.6.5. Hence the value $\sum_{i=1}^N n_i(\mathscr{L}_{e_n}(f_i(x); f_i(x), f_i(\gamma)) = c$ where $c$ is a constant which does not depend on $\gamma$. Let $\mathscr{E} = \sum_{i=1}^N n_i e_n^* \circ (f_i)_*$. Then $\mathscr{E}(l_X(x;x,\gamma)) - c = 0$ for each $\gamma \in \Gamma^n\pi_1(x,x)$. Hence $\mathscr{E} - c$ vanishes on a Zariski dense subset of $\Gamma^n\pi(X)/\Gamma^{n+1}\pi(X)$. This implies that $\mathscr{E} - c = 0$. Evaluating $\mathscr{E} - c$ on a constant loop at $x$ we get $c = 0$. Hence $\mathscr{E} = 0$.

Observe that we have just proved Theorem E.

Now we shall prove some general results about functional equations of polylogarithms. In functional equations from Theorem 10.6.4 and Corollary 10.6.6 coefficients $n_i$ were integers. One can ask whether they cannot be arbitrary complex numbers. We have the following result in this direction.

COROLLARY 10.6.7. *If there is a functional equation of the form*

$$\sum_{i=1}^N \alpha_i \mathrm{Li}_n(f_i(z)) + \overline{\mathrm{l.d.t.}}(n) = 0, \tag{10.7}$$

*then there are rational numbers* $c_1, \dots, c_N$ *not all equal zero such that*

$$\sum_{i=1}^N c_i \mathrm{Li}_n(f_i(z)) + \overline{\mathrm{l.d.t.}}(n) = 0.$$

PROOF. The equation (10.7) is equivalent to the relation $\sum_{i=1}^N \alpha_i e_n^* \circ (f_i)_* = 0$ in $\mathrm{Hom}(\Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x);\mathbb{C})$. The functionals $e_n^* \circ (f_i)_*$ belong to the $\mathbb{Q}$-vector space $\mathrm{Hom}(\Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x);\mathbb{Q})$. Therefore if there is a nontrivial relation of the form $\sum_{i=1}^N \alpha_i e_n^* \circ (f_i)_* = 0$ with $\alpha_i \in \mathbb{C}$, then there is also a nontrivial relation $\sum_{i=1}^N c_i e_n^* \circ (f_i)_* = 0$ with $c_i \in \mathbb{Q}$. Hence the corollary follows from Theorem 10.6.4.

Compare this result with a result in **[B]**. In our corollary one would like to replace functions $f_1(z), \dots, f_N(z)$ by algebraic numbers $a_1, \dots, a_N$ and to take $\alpha_i$ in $\overline{\mathbb{Q}}$.

One would like to get new functional equations from the old one. This is possible as we see from the next result, though unfortunately from functional equations of $\mathrm{Li}_n(z)$ we only get functional equations of $\mathrm{Li}_{n-1}(z)$. We do not know any method which allows one to pass from $\mathrm{Li}_n(z)$ to $\mathrm{Li}_{n+1}(z)$.

DEFINITION 10.6.8. Let $f(z)$ be a rational function. We denote by $\nu_{z-a}(f(z))$ the valuation of $f(z)$ at $(z-a)$.

Observe that $f(z) = \prod_{a\in\mathbb{C}}(z-a)^{\nu_{z-a}(f(z))}$.

LEMMA 10.6.9. *Let $f_1, \dots, f_N : X \to Y$ be regular functions. Assume that $\sum_{i=1}^N n_i e_n^* \circ (f_i)_* = 0$ in $\mathrm{Hom}(\Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x); Z)$. Let $a_1, \dots, a_k$ be complex numbers and let $n-k \geq 2$. Then*

$$\sum_{i=1}^N n_i \cdot \nu_{z-a_1}(f_i(z)) \cdot \nu_{z-a_2}(f_i(z)) \cdots \nu_{z-a_k}(f_i(z)) \cdot e_{n-k}^* \circ (f_i)_* = 0,$$

*in* $\mathrm{Hom}(\Gamma^{n-k}\pi_1(X,x)/\Gamma^{n-k+1}\pi_1(X,x); Z)$.

PROOF. This is an easy observation if one writes the map $(f_i)_*$ in terms of a base given by basic Lie elements.

Observe that Lemma 10.6.9 allows one to get functional equations of $\mathrm{Li}_k$ $(2 \leq k < n)$ if we have a functional equation of $\mathrm{Li}_n$. This follows from Theorem 10.6.4 or Corollary 10.6.6. Observe that the number of functional equations of $\mathrm{Li}_k$ grows when $k$ becomes smaller.

Now we show that certain functional equations are impossible.

PROOF OF THEOREM C. Let $f(z) = \alpha \prod_{i=1}^n (z-a_i)^{n_i} / \prod_{j=1}^m (z-b_j)^{m_j}$. It follows from Example 4 in §10.1 that we can assume that $a_1 \neq a_2$. Let $c \in \mathbb{C}$ be such that $f(c) = 1$ with multiplicity $r$. We consider $f$ as a regular map $f : X = P^1(\mathbb{C}) \setminus \{f^{-1}(0) \cup f^{-1}(1) \cup f^{-1}(\infty) \cup \infty\} \to Y = P^1(\mathbb{C}) \setminus \{0, 1, \infty\}$. (Warning: here $f^{-1}(*)$ is the inverse image of $*$.) We choose a base of $H_1(X)$ given by loops around missing points except $\infty$. Let $A_i$ be a loop around $a_i$ and let $C$ be a loop around $c$. Let us set $\alpha_2 := [C, A_1]$, $\alpha_n := [\alpha_{n-1}, A_2]$ and $\beta_3 := [[C, A_1]A_1]$, $\beta_n := [\beta_{n-1}, A_2]$. The only maps of degree one which induce something nontrivial on $\alpha_n$ and $\beta_n$ are

$$g(z) = \frac{z-a_2}{z-a_1} \cdot \frac{c-a_1}{c-a_2} \quad \text{and} \quad h(z) = \frac{z-a_1}{z-a_2} \cdot \frac{c-a_2}{c-a_1}.$$

For these maps we have

$$g_*(\alpha_n) = -e_n, \qquad g_*(\beta_n) = e_n,$$

and

$$h_*(\alpha_n) = (-1)^{n-2} e_n, \qquad h_*(\beta_n) = (-1)^{n-3} e_n,$$

in the group

$$\mathrm{Hom}(\Gamma^n\pi_1(X,x)/\Gamma^{n+1}\pi_1(X,x); \Gamma^n\pi_1(Y,y)/\Gamma^{n+1}\pi_1(Y,y) + L_n). \tag{10.8}$$

Observe that

$$f_*(\alpha_n) = r \cdot n_1 \cdot n_2^{n-2} e_n \quad \text{and} \quad f_*(\beta_n) = r \cdot n_1^2 \cdot n_2^{n-3} e_n. \tag{10.9}$$

Hence the relation of the form $f_* = \sum_{i=1}^N q_i (f_i)_*$, where $q_i \in \mathbb{Q}$ and $\deg f_i = 1$, is impossible in the group (10.8). Therefore Corollary 10.6.6 implies the theorem.

Closely related to Theorem C is the following result.



THEOREM 10.6.10. *Let $a_1, a_2, \dots, a_n$ be $n$ different points of $\mathbb{C}$. Let*

$$L(z) := \int_x^z \frac{dz}{z-a_1}, \frac{dz}{z-a_2}, \dots, \frac{dz}{z-a_n}.$$

*If $n > 3$ then there is no polynomial $p(s, t_1, \dots, t_r)$ which depends essentially on $s$ such that $p(L(z), \mathrm{Li}_{n_1}(f_1(z)), \dots, \mathrm{Li}_{n_r}(f_r(z))) \equiv 0$ where $\mathrm{Li}_{n_k}$ are classical polylogarithms and logarithms and $f_i(z)$ are rational functions.*

PROOF. Let $T = \{0, 1, \infty\}$ and let $S = \bigcup_{i=1}^r f_i^{-1}(T) \cup \{a_1, \dots, a_n, \infty\}$. Observe that singularities of the functions

$$p(z) := p(\mathrm{L}(z), \mathrm{Li}_{n_1}(f_1(z)), \dots, \mathrm{Li}_{n_r}(f_r(z))),$$

$\mathrm{L}(z)$, and $\mathrm{Li}_{n_i}(f_i(z))$ are contained in the set $S$. On $X = P^1(\mathbb{C}) \setminus S$ these functions are analytic and multivalued. Let $A_i$ be a loop around $a_i$ in $X$. The monodromy of

$$\int_x^z \frac{dz}{z-a_1}, \frac{dz}{z-a_2}, \frac{dz}{z-a_3}, \frac{dz}{z-a_4},$$

on the commutator $\alpha = [[A_1, A_2], [A_3, A_4]]$ is equal to $(2\pi i)^4$ up to sign. Hence the monodromy of $L(z)$ on $\alpha$ is also nontrivial.

Now we must calculate the monodromy of $\mathrm{Li}_{n_k}(f_k(z))$ on $\alpha$. We consider the group $G$ of power series $e^{aX} + \sum_{n=0}^\infty b_n X^n Y$ with a multiplication given by

$$\left(e^{aX} + \sum_{n=0}^\infty b_n X^n Y\right)\left(e^{a'X} + \sum_{n=0}^\infty b'_n X^n Y\right) = e^{(a+a')X} + \sum_{n=0}^\infty \left(b_n + b'_n + \left(\sum_{k=0}^\infty \frac{a^k}{k!} b'_{n-k}\right)\right) X^n Y.$$

The monodromy of polylogarithms was calculated in **[R]** and it can be described in the following way. Let $\mathrm{Li}(z) = e^{(-\log z)X} + \sum_{n=0}^\infty (-1)^n \mathrm{Li}_{n+1}(z) X^n Y$. The monodromy of $\mathrm{Li}(z)$ along the loop around 0 is given by the multiplication on the right-hand side by $e^{(-2\pi i)X}$ and the monodromy along the loop around 1 is given by the multiplication on the right-hand side by $1 - 2\pi i Y$. Observe that for any four elements $a, b, c, d$ in $G$ we have $[[a, b], [c, d]] = 1$. Hence the monodromy of $\mathrm{Li}_{n_k}(f_k(z))$ on $\alpha$ is trivial. This implies $p(z) \neq 0$.

Observe that in Theorem D point (b) is a particular case of Theorem 10.6.10. We leave to the reader to show point (a) of Theorem D.

**10.7. Functional equations of lower degree polylogarithms.** In this section we shall prove Theorem A. We shall also give several examples of functional equations of lower-degree polylogarithms.

10.7.1. *Functional equations of the dilogarithm.*

PROOF OF THEOREM $A'$. Let $f(z)=\alpha\prod_{i=1}^n(z-a_i)^{n_i}/\prod_{j=1}^m(z-b_j)^{m_j}$ and let $f^{-1}(1)=\sum_{k=1}^r c_k\cdot r_k$. Let

$$X=P^1(\mathbb{C})\setminus\{a_1,\dots,a_n,b_1,\dots,b_m,c_1,\dots,c_r,\infty\}$$

and let $Y=P^1(\mathbb{C})\setminus\{0,1,\infty\}$. Let

$$P(X)=\mathbb{C}[[A_1,\dots,A_n,B_1,\dots,B_m,C_1,\dots,C_r]]^*$$

where $A_i$ (resp. $B_j$, resp. $C_k$) is the class in $H_1(X,\mathbb{C})$ of a loop around $a_i$ (resp. $b_j$, resp. $c_k$). Let $P(Y)=C[[U,V]]^*$. Let $\gamma$ be a smooth path in $X$ from $x$ to $z$. We have

$$f_*(A_i\cdot C_k)=n_ir_kU\cdot V,\qquad f_*(B_j\cdot C_k)=-m_jr_kU\cdot V,$$
$$f_*(A_i\cdot B_j)=-n_im_jU\cdot V,\qquad f_*(B_j\cdot B_{j'})=m_jm_{j'}U\cdot V.$$

(Only coefficients at $U\cdot V$ are important.) We need maps of degree one from $X$ to $Y$ which induce the same maps on these products. Here there are three families of such maps:

$$f_{ik}(z)=\frac{z-a_i}{c_k-a_i},\qquad (f_{ik})_*(A_i\cdot C_k)=U\cdot V;$$
$$g_{ij}(z)=\frac{z-a_i}{b_j-a_i},\qquad (g_{ij})_*(A_i\cdot B_j)=U\cdot V;$$
$$h_{jk}(z)=\frac{z-b_j}{c_k-b_j},\qquad (h_{jk})_*(B_j\cdot C_k)=U\cdot V.$$

Let $\mathscr{C}_2\colon P(Y)\to\mathbb{C}$ be as in Definition 10.6.3. Let $\psi_{jj'}\colon P(X)\to\mathbb{C}$ be a coefficient at $B_j\cdot B_{j'}$. We have the following identity

$$\mathscr{C}_2\circ f_*=\sum_{i,k}n_ir_k\mathscr{C}_2\circ(f_{ik})_*-\sum_{i,k}n_im_j\mathscr{C}_2\circ(g_{ij})_*-\sum_{i,k}m_jr_k\mathscr{C}_2\circ(h_{jk})_*+\sum_{j,j'}\psi_{jj'}.\tag{10.10}$$

We shall calculate the expression $\sum_{j,j'}\psi_{jj'}(\lambda_X(z;x,y))$. It follows from the formula

$$\int_\gamma\frac{dz}{z-a},\frac{dz}{z-b}+\int_\gamma\frac{dz}{z-b},\frac{dz}{z-a}=\left(\int_\gamma\frac{dz}{z-a}\right)\left(\int_\gamma\frac{dz}{z-b}\right),$$

(see [Ch] 1.5.1) that

$$\sum_{j,j'}\psi_{jj'}(\lambda_X(z;x,y))=\frac12\sum_{j,j'}m_j\cdot m_{j'}\left(\int_\gamma\frac{dz}{z-b_j}\right)\left(\int_\gamma\frac{dz}{z-b_{j'}}\right).$$



Evaluating the identity (10.10) on $\lambda_X(z;x,y)$ and applying the equality 10.3.3.1 we get

$$\int_{f(y)}\omega=\sum_{i,k}n_ir_k\int_{f_{ik}(y)}\omega-\sum_{i,j}n_im_i\int_{g_{ij}(y)}\omega-\sum_{j,k}m_jr_k\int_{h_{jk}(y)}\omega+\sum_{j,j'}m_jm_{j'}\left(\int_\gamma\frac{dz}{z-b_j}\right)\left(\int_\gamma\frac{dz}{z-b_{j'}}\right),$$

where $\omega=\frac{dz}{z-1},\frac{dz}{z}$. Theorem A$'$ follows immediately from this equation.

Observe that Theorems A and A$''$ are immediate corollaries of Theorem A$'$. This was already observed in §10.1.

Now we shall give an abstract form of the functional equation 10.1.3. We shall keep the notation from 10.7.1.

THEOREM 10.7.1.1. *We have*

$$f_*=\sum_{i,k}n_ir_k(f_{ik})_*-\sum_{i,j}n_im_j(g_{ij})_*-\sum_{j,k}m_jr_k(h_{jk})_*,\tag{10.7.1.2}$$

*in the group* $\operatorname{Hom}(\Gamma^2\pi_1(X,x)/\Gamma^3\pi_1(X,x);\Gamma^2\pi_1(Y,y)/\Gamma^3\pi_1(Y,y))$.

Now we shall show that from the functional equation 10.1.3, choosing suitably a function $f(z)$ and a point $x$, we can get functional equations known before.

EXAMPLES. Let $f(z)=z^n$ and $x=0$. Then we get

$$(1/n)\operatorname{Li}_2(z^n)=\sum_{k=1}^n\operatorname{Li}_2(\xi^kz),\quad\text{where }\xi=e^{\frac{2\pi i}{n}}.\tag{10.7.1.3}$$

Let $f(z)=yz/(y-1)(z-1)$ and $x=0$. Then we have

(10.7.1.4)
$$\operatorname{Li}_2\left(\frac{yz}{(y-1)(z-1)}\right)=\operatorname{Li}_2\left(\frac{z}{1-y}\right)-\operatorname{Li}_2\left(\frac{1-z}{y}\right)+\operatorname{Li}_2\left(\frac1y\right)-\operatorname{Li}_2(z)-\log\left(\frac{y-1}{y}\right)\log(1-z)-\frac12(\log(1-z))^2.$$

Let $f(z)=(1-y)z/(z-1)$ and $x=0$. Then we have

(10.7.1.5)
$$\operatorname{Li}_2\left(\frac{(1-y)z}{z-1}\right)=\operatorname{Li}_2(yz)-\operatorname{Li}_2\left(\frac{y-1}{y}(1-z)\right)+\operatorname{Li}_2\left(\frac{y-1}{y}\right)-\operatorname{Li}_2(z)-\log(1-y)\log(z-1)-\frac12\log^2(1-z).$$

Observe that the Abel equation from §10.1 follows from 10.7.1.4.

Let $f(z)=\alpha\prod_{i=1}^n(z-a_i)^{n_i}/\prod_{j=1}^m(z-b_j)^{m_j}$. Then we have

$$\log f(z)=\log\alpha+\sum_{i=1}^n n_i\log(z-a_i)-\sum_{j=1}^m m_j\log(z-b_j).\tag{10.11}$$

Observe that any functional equation of log of the form $\sum_{i=1}^{N} n_i \log f_i(z) = 0$ where $f_i(z)$ are rational functions is a linear combination with rational coefficients of equations (10.11). For the dilogarithm we have a similar situation. We shall formulate a theorem only for abstract functional equations.

**Theorem** 10.7.1.6. *Assume* 10.6.1 *and* 10.6.2. *Then any relation of the form* $\sum_{i=1}^{N} n_i(f_i)_* = 0$ *in*

$$\operatorname{Hom}(\Gamma^2\pi_1(X, x)/\Gamma^3\pi_1(X, x); \Gamma^2\pi_1(Y, y)/\Gamma^3\pi_1(Y, x))$$

*is a linear combination of the relations* 10.7.1.2 *for functions* $f_i$.

**Proof.** The theorem is an easy observation in linear algebra.

10.7.2. *Functional equations of the trilogarithm.*

**Theorem** 10.7.2.1. (integral form of a functional equation). *Let* $f(z) = \alpha\prod_{i=1}^{n}(z-a_i)^{n_i}/\prod_{j=1}^{m}(z-b_j)^{m_j}$ *be in an irreducible form. Let us set* $\mathscr{L}_3(\dots) := \mathscr{L}_{e_3}(\dots)$ *and* $\mathscr{L}_3(f(z; x, y)) := \mathscr{L}_3(f(z); f(x), f(y))$. *Then we have*

(10.7.2.2)

$$\begin{aligned}
\mathscr{L}_3(f(z; x, y)) = {} & \sum_{i<i',k} n_i n_{i'} r_k(\mathscr{L}_3(d_i^{i'}(z; x, y)) - \mathscr{L}_3(e_{ik}^{i'}(z; x, y))) \\
& + \sum_{i,i',k} n_i n_{i'} r_k(\mathscr{L}_3(f_{ik}(z; x, y))) \\
& + \sum_{i,j,k} n_i m_j r_k(\mathscr{L}_3(g_{ik}^{j}(z; x, y)) - \mathscr{L}_3(h_i^{j}(z; x, y)) \\
& \qquad - \mathscr{L}_3(l_{ik}(z; x, y)) - \mathscr{L}_3(p_{jk}(z; x, y))) \\
& + \sum_{j<j',k} m_j m_{j'} r_k(\mathscr{L}_3(q_j^{j'}(z; x, y)) - \mathscr{L}_3(s_{jk}^{j'}(z; x, y))) \\
& + \sum_{j,j',k} m_j m_{j'} r_k(\mathscr{L}_3(t_{jk}(z; x, y))) \\
& + \sum_{i<i',j} n_i n_{i'} m_j(-\mathscr{L}_3(u_i^{i'}(z; x, y)) + \mathscr{L}_3(v_{ij}^{i'}(z; x, y))) \\
& + \sum_{i,i',j} n_i n_{i'} m_j(-\mathscr{L}_3(w_{ij}(z; x, y))) \\
& + \sum_{j<j',i} m_j m_{j'} n_i(\mathscr{L}_3(\varphi_{ji}^{j'}(z; x, y)) - \mathscr{L}_3(\psi_j^{j'}(z; x, y))) \\
& - \sum_{j,j',i} m_j m_{j'} n_i(\mathscr{L}_3(\chi_{ji}(z; x, y))),
\end{aligned}$$



*where*

$$\begin{aligned}
&d_i^{i'}(z) = \frac{z-a_i}{z-a_{i'}}, && e_{ik}^{i'}(z) = \frac{z-a_i}{z-a_{i'}}\cdot\frac{c_k-a_{i'}}{c_k-a_i}, && f_{ik}(z) = \frac{z-a_i}{c_k-a_i}, \\
&g_{ik}^{j}(z) = \frac{z-a_i}{z-b_j}\cdot\frac{c_k-b_j}{c_k-a_i}, && h_i^{j}(z) = \frac{z-a_i}{z-b_j}, && l_{ik}(z) = \frac{z-a_i}{c_k-a_i}, \\
&p_{jk}(z) = \frac{z-b_j}{c_k-b_j}, && q_j^{j'}(z) = \frac{z-b_j}{z-b_{j'}}, && s_{jk}^{j'}(z) = \frac{z-b_j}{z-b_{j'}}\cdot\frac{c_k-b_{j'}}{c_k-b_j}, \\
&t_{jk}(z) = \frac{z-b_j}{c_k-b_j}, && u_i^{i'}(z) = \frac{z-a_i}{z-a_{i'}}, && v_{ij}^{i'}(z) = \frac{z-a_i}{z-a_{i'}}\cdot\frac{b_j-a_{i'}}{b_j-a_i}, \\
&w_{ij}(z) = \frac{z-a_i}{b_j-a_i}, && \varphi_{ji}^{j'}(z) = \frac{z-b_j}{z-b_{j'}}\cdot\frac{a_i-b_{j'}}{a_i-b_j}, && \psi_j^{j'}(z) = \frac{z-b_j}{z-b_{j'}},
\end{aligned}$$

*and*

$$\chi_{ji}(z) = \frac{z-b_j}{a_i-b_j}.$$

**Proof.** One checks that in

$$\operatorname{Hom}(\Gamma^3\pi_1(X, x)/\Gamma^4\pi_1(X, x); \Gamma^3\pi_1(Y, y)/\Gamma^4\pi_1(Y, y) + L_3)$$

there is the identity (abstract form of a functional equation)

$$\begin{aligned}
f_* = {} & \sum_{i<i',k} n_i n_{i'} r_k((d_i^{i'})_* - (e_{ik}^{i'})_*) + \sum_{i,i',k} n_i n_{i'} r_k((f_{ik})_*) \\
& + \sum_{i,j,k} n_i m_j r_k((g_{ik}^{j})_* - (h_i^{j})_* - (l_{ik})_* - (p_{jk})_*) \\
& + \sum_{j<j',k} m_j m_{j'} r_k((q_j^{j'})_* - (s_{jk}^{j'})_*) \\
& + \sum_{j,j',k} m_j m_{j'} r_k((t_{jk})_*) + \sum_{i<i',j} n_i n_{i'} m_j(-(u_i^{i'})_* + (v_{ij}^{i'})_*) \qquad (10.7.2.3) \\
& + \sum_{i,i',j} -n_i n_{i'} m_j((w_{ij})_*) \\
& + \sum_{j<j',i} m_j m_{j'} n_i((\varphi_{ji}^{j'})_* - (\psi_j^{j'})_*) - \sum_{j,j',i} m_j m_{j'} n_i(\chi_{ji})_*.
\end{aligned}$$

The theorem follows from Theorem 10.6.1(i).

We shall not prove Theorem B. We indicate only a general scheme for a proof. First one proves an analog $B'$ of Theorem B in the same way as we proved Theorem A$'$. Then one deduces Theorem B from B$'$. Observe that the abstract form of the functional equation from Theorem B is a particular case of 10.7.2.3.

For the trilogarithm we have an analog of Theorem 10.7.1.6.

THEOREM 10.7.2.4. *Assume* 10.6.1 *and* 10.6.2. *Then any relation of the form* $\sum_{i=1}^{N} n_i(f_i)_* = 0$ *in*

$$\mathrm{Hom}(\Gamma^3\pi_1(X, x)/\Gamma^4\pi_1(X, x); \Gamma^3\pi_1(Y, y)/\Gamma^4\pi_1(Y, y) + L_3)$$

*is a linear combination of the relations* 10.7.2.3 *for functions* $f_i$.

PROOF. The theorem is once more an easy observation in linear algebra.

10.7.3. *The fourth-order polylogarithm.* We shall give an example of a functional equation of the fourth-order polylogarithm which seems not to be reported in the literature.

Let $f_1(z) = -(z-a)(z-b)/(b-a)^2$ and $f_2(z) = (z-a)^2/(b-a)(z-b)$. Let $c_1$, $c_2$ be roots of the equation $f_1(z) - 1 = 0$. Observe that $c_1$ and $c_2$ are also roots of the equation $f_2(z) - 1 = 0$. Let us set

$$g_1(z) = \frac{z-a}{z-b}\cdot\frac{c_1-b}{c_1-a}, \qquad g_2(z) = \frac{z-a}{z-b}\cdot\frac{c_2-b}{c_2-a},$$
$$h_1(z) = \frac{z-a}{c_1-a}, \qquad h_2(z) = \frac{z-a}{c_2-a},$$
$$k_1(z) = \frac{z-b}{c_1-b}, \qquad k_2(z) = \frac{z-b}{c_2-b},$$
$$l_1(z) = \frac{z-a}{z-b}, \qquad l_2(z) = \frac{z-a}{b-a}, \qquad l_3(z) = \frac{z-b}{a-b}.$$

Let $X = P^1(\mathbb{C})\setminus\{a, b, c_1, c_2, \infty\}$ and $Y = P^1(\mathbb{C})\setminus\{0, 1, \infty\}$. Each of the rational functions described above determines a regular map from $X$ to $Y$.

THEOREM 10.7.3.1. (abstract form of a functional equation). *We have*

$$(f_1)_* + (f_2)_* = 3(g_1)_* + 3(g_2)_* + 6(h_1)_* + 6(h_2)_* + 3(k_1)_* + 3(k_2)_* - 2(l_1)_* - 4(l_2)_* - 2(l_3)_*$$

*in the group* $\mathrm{Hom}(\Gamma^4\pi_1(X, x)/\Gamma^5\pi_1(X, x); \Gamma^4\pi_1(Y, y)/\Gamma^5\pi_1(Y, y) + L_4)$.

Notice that this functional equation has less quadratic terms than the Kummer functional equation of the fourth-order polylogarithm.

**10.8. Generalized Bloch groups.**

DEFINITION 10.8.1. Let $K$ be a field. We set

$$B(K) := \bigoplus_{f\in K\setminus\{0,1\}} \mathbb{Z}.$$

The group $B(K)$ is by definition a free abelian group on elements of $K\setminus\{0, 1\}$. The generator of $B(K)$ corresponding to $f \in K\setminus\{0, 1\}$, we shall denote by $[f]$.

We recall the Abel functional equation

$$\mathrm{Li}_2\left(\frac{x}{1-x}\cdot\frac{y}{1-y}\right) - \mathrm{Li}_2\left(\frac{y}{1-x}\right) - \mathrm{Li}_2\left(\frac{x}{1-y}\right) + \mathrm{Li}_2(x) + \mathrm{Li}_2(y) = \log(1-x)\log(1-y).$$



S. Bloch observed that the element

$$\left[\frac{x}{1-x}\cdot\frac{y}{1-y}\right] - \left[\frac{y}{1-x}\right] - \left[\frac{x}{1-y}\right] + [x] + [y] \in B(C(x, y)),$$

belongs to the kernel of the homomorphism

$$\lambda: B(C(x, y)) \to C^*(x, y) \wedge C^*(x, y),$$

where $\lambda([f]) = f \wedge (1-f)$ and $C^*(x, y) \wedge C^*(x, y)$ is an exterior product of $C^*(x, y)$ with itself considered as an abelian group (see [DS]).

The aim of this section is to generalize the phenomena observed by S. Bloch and to put it in the picture described in the previous sections.

Let $A$ be an abelian group and let $\mathrm{L}(A)$ be a free Lie algebra on $A$. Let $\mathrm{L}'(A) = [\mathrm{L}(A), \mathrm{L}(A)]$ and let $\mathrm{L}''(A) = [\mathrm{L}'(A), \mathrm{L}'(A)]$. We set

$$\mathrm{Li}(A) := \mathrm{L}(A)/\mathrm{L}''(A).$$

Let $K$ be a function field and let $k$ be its field of constants. Let $K^*$ and $k^*$ be respectively its multiplicative groups. Let $I(K^* : k^*)$ be a Lie ideal in $L(K^*)$ generated by brackets $[\cdots[f_1 \cdots f_i], [f_{i+1}, \cdots]\cdots f_n]\cdot]$ where at least one $f_i$ is in $k^*$. Let us set

$$\mathcal{L}i(K^*) := \mathrm{Li}(K^*)/I(K^* : k^*).$$

For any $n \geq 2$ we define a homomorphism $B_n : B(K) \to \mathcal{L}i(K^*)$ by the formula

$$B_n([f]) = [\cdots[f-1, f]f]\cdots]f]\cdots].$$

The main result of this section is the following theorem.

THEOREM 10.8.2. *Let*

$$X = P^1(\mathbb{C})\setminus\{a_1, \dots, a_m, \infty\} \quad and \quad Y = P^1(\mathbb{C})\setminus\{0, 1, \infty\}.$$

*Let* $f_1, \dots, f_N \in C(z)^*$ *be regular functions from* $X$ *to* $Y$ *and let* $k_1, \dots, k_N$ *be integers. The following conditions are equivalent:*

(i) *the element* $\sum_{i=1}^{N} k_i[f_i] \in B(C(z))$ *belongs to the kernel of the map*

$$B_n : B(C(z)) \to \mathcal{L}i(C(z)^*);$$

(ii) $\sum_{i=1}^{N} k_i(f_i)_* = 0$ *in the group*

$$\mathrm{Hom}(\Gamma^n\pi_1(X, x)/\Gamma^{n+1}\pi_1(X, x); \Gamma^n\pi_1(Y, y)/\Gamma^{n+1}\pi_1(Y, y) + L_n);$$

(iii) *there is a functional equation* $\sum_{i=1}^{N} k_i \mathrm{Li}_n(f_i(z)) + \overline{\mathrm{l.d.t.}}(n) = 0$.

PROOF. Let $f \in \mathbb{C}(z)^*$ and let

$$f^{-1}(0) = \sum_{i=1}^{n} m_i \cdot \alpha_i, \qquad f^{-1}(\infty) = \sum_{j=1}^{m} m_{n+j}\alpha_{n+j},$$

and $f^{-1}(1) = \sum_{k=1}^{r} r_k c_k$. Observe that $f$ defines a regular function from $S = P^1(\mathbb{C})\setminus\{\alpha_1, \dots, \alpha_{n+m}, c_1, \dots, c_r, \infty\}$ to $Y$. Let $A_i$ (resp. $C_k$) be the class in $H_1(S, \mathbb{C})$ of a loop in $S$ around $\alpha_i$ (resp. $c_k$). Then $\operatorname{Lie}\pi_1(S, s)$ is a free Lie algebra on $A_1, \dots, C_r$. We choose a base of $\operatorname{Lie}\pi_1(S, s)$ given by basic Lie elements corresponding to the ordering $A_1, \dots, A_{n+m}, C_1, \dots, C_r$. In the group $\operatorname{Hom}(\Gamma^n\pi_1(S, s)/\Gamma^{n+1}\pi_1(S, s), Z)$ we have

$$e_n^* \circ f_* = \sum_{i_{n-2}\geq\cdots\geq i_1\geq i}\sum_k r_k m_i m_{i_1}\cdots m_{i_{n-2}}(\dots((C_k, A_i)A_{i_1})\dots A_{i_{n-2}})^*$$
$$+ \sum_{\substack{i_{n-2}\geq\cdots\geq i_1\geq i\\ i\leq n}}\sum_j m_{j+n}m_i m_{i_1}\cdots m_{i_{n-2}}(\dots((A_{j+n}, A_i)A_{i_1})\dots(A_{i_{n-2}})^*.$$

$((\dots)^*$ denotes the dual functional.) On the other side after the identification $(z-a_i)$ (resp. $(z-c_k)$) with $A_i$ (resp. $C_k$) in the Lie algebra $\mathscr{L}i(C(z)^*)$ we have

$$B_n([f]) = \sum_{i_{n-2}\geq\cdots\geq i_1\geq i}\sum_k r_k m_i m_{i_1}\cdots m_{i_{n-2}}$$
$$\cdot(\gamma(k, i, i_1, \dots, i_{n-2})(\dots((C_k, A_i)A_{i_1})\dots A_{i_{n-2}})^*$$
$$-\delta(k, i_p, i, i_1, \dots, i_{n-2})((\cdots((A_{i_p}, A_i)A_{i_1})\dots\hat{A}_{i_p}\dots A_{i_{n-2}})C_k)^*)$$
$$+ \sum_{\substack{i_{n-2}\geq\cdots\geq i_1\geq i\\ i\leq n}}\sum_j m_{j+n}m_i m_{i_1}\cdots m_{i_{n-2}}\mathscr{C}(j+n, i, i_1, \dots, i_{n-2})$$
$$\cdot(\dots((A_{j+n}, A_i)A_{i_1})\dots A_{i_{n-2}})^*$$
$$+ \sum_{\substack{i_{n-1}\geq\cdots\geq i_1\\ n\geq i\geq i_1}} m_i m_{i_1}\cdots m_{i_{n-1}}\psi(i, i_1, \dots, i_{n-1})$$
$$\cdot(\dots((A_i, A_{i_1})A_{i_2})\dots A_{i_{n-1}})^*.$$

This follows from the Jacobi identity and the fact that in the Lie algebra $\mathscr{L}i(C(z)^*)$ we have $(\dots((A, B)A_1)\dots A_m) = (\dots((A, B)A_{\sigma(1)})\dots A_{\sigma(m)})$ where $\sigma$ is any permutation of $n$ elements. The coefficients $\gamma(\cdots)$, $\delta(\cdots)$, $\mathscr{C}(\cdots)$ and $\psi(\cdots)$ do not depend on $f$. For example if $i_{n-2} > \cdots > i_1 > i$ then $\gamma(k, i, i_1, \dots, i_{n-2}) = (n-1)!$.

Now from the formulas for $e_n^* \circ f_*$ and $B_n([f])$ and from Theorem 10.6.4 ((10.2) and (10.3) are equivalent) it follows that the conditions (i) and (ii) are equivalent. By Corollary 10.6.6(ii) and (iii) are also equivalent.

REMARK. The groups $\operatorname{Li}(K^*)$ and $\mathscr{L}i(K^*)$ are graded. The component in degree $n$ we denote by $\operatorname{Li}_n(K^*)$ and $\mathscr{L}i_n(K^*)$ respectively. They are generated additively by brackets of length $n$. D. Zagier in [Z3] considered the group $(\operatorname{Sym}^{n-2}(K^*) \otimes (K^* \wedge K^*)) \otimes \mathbb{Q}$. He found a condition to have a functional equation of higher Bloch-Wigner functions in terms of this group. Observe that there is an epimorphism with nontrivial kernel from $(\operatorname{Sym}^{n-2}(K^*) \otimes (K^* \wedge K^*)) \otimes \mathbb{Q}$ onto $(\operatorname{Li}_n(K^*)) \otimes \mathbb{Q}$ and hence also onto $(\mathscr{L}i_n(K^*)) \otimes \mathbb{Q}$. This follows from the fact that in $\operatorname{Li}_n(K^*)$ and $\mathscr{L}i_n(K^*)$



we have $(\cdots(A, B)A_1)\cdots A_{n-2})\cdots) = (\cdots A, B)A_{\sigma(1)})\cdots A_{\sigma(n-2)})\cdots)$ for any $\sigma \in \Sigma_{n-2}$.

Let $\mathrm{L}_n(z)$ be the higher Bloch-Wigner function considered in [W3] and in [Z2, Z3]. We would like to show that $B_n(\sum_{i=1}^N k_i[f_i(z)]) = 0$ if and only if $\sum_{i=1}^N k_i(\mathrm{L}_n(f_i(z)) - \mathrm{L}_n(f_i(x))) = 0$.

ACKNOWLEDGEMENTS. This chapter is based on ideas in our preprint *A note on functional equations of the dilogarithm* (see [W1] and also [W2]). We would like to thank P. Deligne very much for his comments on our manuscript under the same title, where he reinterpreted our results in terms of Lie algebras of fundamental groups. He also showed us the connection from section 1 in the special case of $\mathbb{C}\setminus\{0, 1\}$. We acknowledge the influence of the lecture of D. Zagier (Bonn, April 1989, see also [Z1] and [Z2]). We acknowledge the influence of papers of L. J. Rogers [Ro], H. F. Sandham [S], and R. F. Coleman [C]). We would like to thank very much J. L. Loday and Ch. Soulé who told us about functional equations of polylogarithms. While writing this chapter we were a visitor at the Max-Planck-Institut für Mathematik in Bonn. We would like to thank Professor F. Hirzebruch very much for this opportunity to visit Bonn. We would also like to thank Y. André and H. Gangl for several useful discussions and L. Lewin for correspondence.

## REFERENCES


[A] N. H. Abel, *Note sur la fonction* $\psi(x) = x + (x^2/2^2) + (x^3/3^2) + \cdots$, Oeuvres complètes, tome II. Grøndahl & Søn, Christiania, 1881, pp. 189–193.

[B] A. Baker, *Linear forms in logarithms of algebraic numbers*. I, Mathematika **13** (1966), 204–216.

[C] R. F. Coleman, *Dilogarithms, regulators and p-adic L-functions*, Invent. Math. **69** (1982), 171–208.

[Ch] K. T. Chen, *Algebra of iterated path integrals and fundamental groups*, Trans. Amer. Math. Soc. **156** (1971), 359–379.

[D] P. Deligne, *Le groupe fondamental de la droite projective moins trois points*, Galois Groups over **Q**, Springer-Verlag, New York, 1989.

[DS] J. L. Dupont and Ch. H. Sah, *Scissors congruences* II, J. Pure and Appl. Algebra **25** (1982), 159–195.

[L] L. Lewin, *Polylogarithms and associated functions*, North Holland, New York and Oxford, 1981.

[MCS] W. Magnus, A. Karras, and D. Solitar, *Combinatorial group theory*, Interscience, New York, 1966.

[R] D. Ramakrishnan, *On the monodromy of higher logarithms*, Proc. Amer. Math. Soc. **85** (1982), 596–599.

[Ro] L. J. Rogers, *On function sum theorems connected with the series* $\sum_{n=1}^{\infty}(x^n/n^2)$, Proc. London Math. Soc. (2) **4** (1907), 169–189.

[S] H. F. Sandham, *A logarithmic transcendent*, J. London Math. Soc. **24** (1949), 83–91.

[W1] Z. Wojtkowiak, *A note on the functional equation of the dilogarithm*, Preprint No. 1, CRM (Bellaterra), October 1984.

[W2] ——, *A note on functional equations of polylogarithms*, Preprint No. 22, MPI, 1990.

[Z1] D. Zagier, *The remarkable dilogarithm*, J. Math. Phys. Sci. **22** (1988), 131–145.

[Z2] ——, *The Bloch-Wigner-Ramakrishnan polylogarithm function*, Math. Ann. **286** (1990), 613–624.

[Z3] ——, *Polylogarithms, Dedekind zeta function and the algebraic K-theory of fields*, preprint No. 44, MPI, 1990.

CHAPTER 11

# $K$-Theory, Cyclotomic Equations, and Clausen's Function

J. BROWKIN

In the present chapter we consider some "dilogarithmic functions" such as the function of Clausen $\mathrm{Cl}_2(t)$, the dilogarithm of Rogers $L(z)$, and the dilogarithm of Wigner and Bloch $D(z)$. Our main purpose here is to obtain **Z**-linear relations between the values of these functions at some particular arguments. We shall try to make clear the deep sources of these relations lying in algebra, geometry, $K$-theory, and arithmetic.

Some of the results presented here are classical or well known, but many others are new or even conjectural. At the end of the chapter we formulate several problems and conjectures which seem to be important or interesting.

## 11.1. Algebraic background.

11.1.1. *The basic exact sequence.* For any field $F$ we consider the following exact sequence

$$(11.1)\qquad 0 \to C(F) \to A(F) \xrightarrow{\lambda} F^* \wedge F^* \xrightarrow{\mathscr{X}} K_2F \to 0,$$

where

(i) $K_2$ is the functor of Milnor. By the theorem of Matsumoto (see e.g. Milnor [71,§11]) we have

$$K_2F = (F^* \otimes F^*)/I,$$

where $I$ is the subgroup of $F^* \otimes F^*$ generated by elements $a \otimes (1-a)$, where $a \in F^*$, $a \neq 1$. The element $a \otimes b \pmod I$ of $K_2F$ we denote by $\{a, b\}$, and we call it the Steinberg symbol.

Since, for $a \in F^*$, $a \neq 1$, we have

$$a \otimes (-a) = a \otimes (1-a) + a^{-1} \otimes (1-a^{-1}) \in I,$$

it follows that $\{a, -a\} = 0$ in $K_2F$.

(ii) $F^* \wedge F^*$ is a modified external product. In the place of the usual condition $a \wedge a = 0$ we take $a \wedge (-a) = 0$, for $a \in F^*$. In other words

$$F^* \wedge F^* = (F^* \otimes F^*)/J,$$

where $J$ is the subgroup of $F^* \otimes F^*$ generated by elements $a \otimes (-a)$, for $a \in F^*$.

(iii) The homomorphism $\mathscr{K}$ is defined by

$$\mathscr{K}(a \wedge b) = \{a, b\}, \quad \text{for } a, b \in F^*.$$

(iv) $A(F)$ is the free abelian group with generators $[a]$, where $a \in F^*$, $a \neq 1$.

(v) The homomorphism $\lambda$ is defined by

$$\lambda([a]) = a \wedge (1-a), \quad \text{for } a \in F^*, \ a \neq 1.$$

(vi) Finally $C(F) := \operatorname{Ker} \lambda$.

From the above definitions it follows easily that the sequence (11.1) is exact.

In general groups appearing in (11.1) are not finitely generated, even if $F$ is a finite extension of the field $\mathbb{Q}$ of rational numbers. In this last case, we shall approximate the sequence (11.1) by an analogous sequence containing finitely generated groups only.

11.1.2. *Good S-units and an S-version of the sequence* (11.1). Let $F$ be a finite extension of the field $\mathbb{Q}$ of rational numbers, with the ring of integers $O_F$. Let $r_1$ (resp. $r_2$) be the number of real (resp. complex) infinite primes of $F$. For a finite set $S$ of primes of $F$ containing all infinite primes, let $U_S$ be the group of $S$-units of $F$, i.e., $a \in U_S$ iff $\operatorname{ord}_p(a) = 0$ for all primes $p$ of $F$ not belonging to $S$. It is well known that the group $U_S$ is the direct product of the cyclic group $\mu(F)$ of all roots of unity in $F$, and of a free abelian group of rank $t = r_1 + r_2 + s - 1$, where $s = \# S$.

The elements of the set $W_S := U_S \cap (1 - U_S)$ we call good $S$-units. It is known (see e.g. Evertse [84, Theorem 1]) that the set $W_S$ is finite. Evidently if $a$ is a good $S$-unit, then also $1-a$ and $1/a$ are good $S$-units. In other words, the symmetric group $S_3$ acts on the set $W_S$ of good $S$-units by the formulas

$$\sigma(a) = 1 - a, \qquad \tau(a) = 1/a,$$

where $\sigma, \tau$ are two transpositions belonging to $S_3$. Evidently this action is defined on the whole closed complex plane $\mathbb{C} \cup \{\infty\}$. It is well known that the following set $T$ is a fundamental domain of this action of $S_3$ on $\mathbb{C} \cup \{\infty\}$:

$$T = \{z \in \mathbb{C} : 0 < \operatorname{Re} z \leq \tfrac{1}{2}, |z - 1| \leq 1\},$$

(see Figure 11.1). Thus to determine all elements of the finite set $W_S$ it is sufficient to look for them in the set $T \cap F$, and then apply the action of $S_3$. In particular, if the field $F$ is real, then $T \cap F \subseteq (0, 1/2)$, and thus it is sufficient to determine all good $S$-units belonging to the interval $(0, 1/2)$ to know all of them.

We consider the following modification of the sequence (11.1):

$$(11.2) \qquad 0 \to C(F)_S \to A(W_S) \xrightarrow{\lambda} U_S \wedge U_S \xrightarrow{\mathscr{K}} (K_2F)_S \to 0,$$



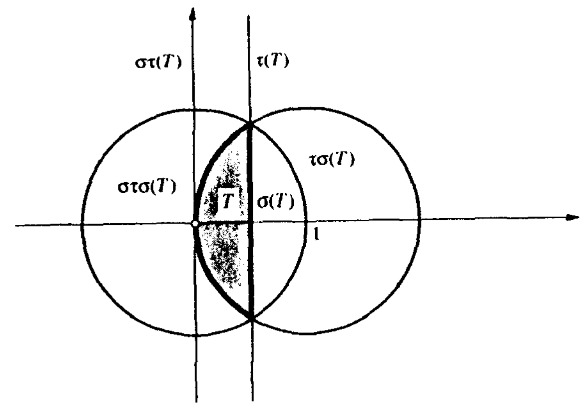



FIGURE 11.1

where

(i) $A(W_S)$ is the subgroup of $A(F)$ generated by elements $[a]$ with $a \in W_S$.

(ii) $C(F)_S := C(F) \cap A(W_S)$.

(iii) $(K_2F)_S$ is the subgroup of $K_2F$ generated by symbols $\{a, b\}$ with $a, b \in U_S$.

Since $K_2F$ is a torsion abelian group, and the group $U_S$ is finitely generated, it follows that the group $(K_2F)_S$ is finite.

From the above definition it follows that the sequence (11.2) is well defined and exact with the possible exception in the term $U_S \wedge U_S$. In any case from $\mathscr{K} \circ \lambda = 0$ it follows that

$$\#(K_2F)_S | (U_S \wedge U_S : \lambda(A(W_S))),$$

provided the last index is finite.

Let $\# \mu(F) = m$, then $\zeta_m := \exp(2\pi i/m)$ is a generator of the group $\mu(F)$. Let $\zeta_m, p_1, \ldots, p_t$ be generators of the group $U_S$. Then the group $U_S \wedge U_S$ is the direct product of the following three groups:

(i) the free abelian group of rank $\binom{t}{2}$ generated by elements

$$p_i \wedge p_j \qquad (1 \leq i < j \leq t);$$

(ii) the product of $t$ cyclic groups of order $m$ generated by $\zeta_m \wedge p_i$ $(1 \leq i \leq t)$;

(iii) the cyclic group generated by $(-1) \wedge \zeta_m$. Its order is equal to 2 if $m \equiv 2 \pmod 4$, and is equal 1 otherwise.

Since $\lambda(A(W_S))$ is a subgroup of $U_S \wedge U_S$, this effective description of generators of $U_S \wedge U_S$ will be useful in the sequel in investigating the group $\lambda(A(W_S))$.

11.1.3. *Elements of* $C(F)$. In the next paragraphs we shall be interested in the effective determination of elements of the group $C(F) = \operatorname{Ker}\lambda$ defined in 11.1. We describe below several methods of constructing such elements.

Some of these elements can be given explicitly.

**Lemma 11.1.** *For $a \in F^*$, $a \neq 1$, the following elements belong to $C(F)$:*
(i) $[a] + [1-a]$,
(ii) $[a] + [1/a]$,
(iii) $[a^k] - k\sum_{j=0}^{k-1}[\zeta_k^j a]$, *provided $\zeta_k \in F$ and $a^k \neq 1$.*
*In particular,*
(iv) $[a^2] - 2[a] - 2[-a]$, *provided $a^2 \neq 1$ (duplication formula).*

**Lemma 11.2.** *For every rational function $f \in F(X)$, $f \neq 0, 1$, and $a \in F$, we have*

$$[f(a)] - [f(\infty)] + \sum_{\substack{b,c\in F\\ b\neq c}} \operatorname{ord}_b(f(X))\cdot \operatorname{ord}_c(1-f(X))\left[\frac{a-b}{a-c}\right] \in C(F),$$

*provided all terms are defined. We assume here that $[0] = [1] = [\infty] = 0$ by definition.*

The proof of both lemmas follows directly from the definition of $\lambda$. □

We call a sequence $(a, b, c, d, e)$ of five elements of a field $F$ a 5-cycle, if

$$a = 1 - cd, \quad b = 1 - de, \quad c = 1 - ea, \quad d = 1 - ab, \quad e = 1 - bc,$$

(see Keune [81] for other equivalent conditions). We say that a 5-cycle is nontrivial if it consists of nonzero elements. Then all its elements are also different from 1.

If $(a, b, c, d, e)$ is a 5-cycle, then it is easy to verify that its cyclic permutation $(b, c, d, e, a)$, and its image under the symmetry $(e, d, c, b, a)$ are also 5-cycles. I.e., the dihedral group $D_5$ acts on the set of all 5-cycles.

We call two 5-cycles equivalent, if they belong to the same orbit under the action of the dihedral group $D_5$.

**Lemma 11.3.** *If $(a, b, c, d, e)$ is a nontrivial 5-cycle, then $[a]+[b]+[c]+[d]+[e] \in C(F)$.*

**Proof.** The lemma follows easily from the definition of a 5-cycle and of $\lambda$. □

All elements of $C(F)$ given in lemmas 11.1–11.3 are trivial in a sense that will be explained later. More important elements of $C(F)$ can be constructed using cyclotomic equations defined as follows. Every equation of the following form we call a (generalized) cyclotomic equation.

$$X^A \prod_{r\in I}(1 - \zeta_m^{k_r} X^r)^{A_r} = 1, \tag{11.3}$$

where $A, k_r, A_r$ are integers, $\zeta_m \in F$, and $I$ is a finite set of natural numbers.



**Theorem 11.1.** *Suppose that $a \in F^*$, $a \neq 1$ satisfies the cyclotomic equation (11.3), and let a natural number $N$ satisfy*
(i) $2 \mid NA$,
(ii) $r \mid NA_r$, $m \mid (NA_r/r)k_r$, *for $r \in I$,*
*then $\sum_{r\in I}(NA_r/r)[\zeta_m^{k_r} a^r] \in C(F)$.*

**Remark.** There always exists $N$ satisfying the assumptions of the theorem: It is sufficient to take $N = \text{l.c.m.}_{r\in I}(2, mr)$. In applications it will be usually important to take $N$ as small as possible.

**Proof.** In view of (11.3) we have

$$\begin{aligned}
0 = a^N \wedge 1 &= a^N \wedge \left(a^A \prod_{r\in I}(1-\zeta_m^{k_r}a^r)^{A_r}\right) \\
&= NA(a\wedge a) + \sum_{r\in I} NA_r(a \wedge (1-\zeta_m^{k_r}a^r)) \\
&= NA((-1)\wedge a) + \sum_{r\in I}\frac{NA_r}{r}(a^r \wedge (1-\zeta_m^{k_r}a^r)) \\
&= \sum_{r\in I}\frac{NA_r}{r}((\zeta_m^{k_r}a^r)\wedge(1-\zeta_m^{k_r}a^r)) \\
&= \lambda\left(\sum_{r\in I}\frac{NA_r}{r}[\zeta_m^{k_r}a^r]\right),
\end{aligned}$$

since $(\zeta_m^{k_r})^{NA_r/r} = 1$ in view of (ii). □

We call (11.3) an ordinary cyclotomic equation if $k_r = 0$ for all $r \in I$.

From the theorem on primitive divisors (see Schinzel [74]) it follows that for every algebraic number $a$, which is not a root of unity, the exponents $r$ in all ordinary cyclotomic equations satisfied by $a$ are uniformly bounded. Therefore there is only a finite number of multiplicatively independent ordinary cyclotomic equations satisfied by $a$. A. Schinzel proved (unpublished) that the same holds also for generalized cyclotomic equations.

If $F$ is a number field, we can determine all elements of $C(F)$ in a more systematic way as follows. Let $S$ be a finite set of primes of $F$ containing all infinite primes. Every $a \in F^*$, $a \neq 1$, is a good $S$-unit if the set $S$ contains all finite primes $p$ of $F$ satisfying $\operatorname{ord}_p(a(1-a)) \neq 0$. Consequently $A(F) = \bigcup_S A(W_S)$ and hence $C(F) = \bigcup_S C(F)_S$, where the sums are extended over all finite sets $S$ of primes of $F$ containing all infinite primes.

Therefore every element of $C(F)$ can be obtained in the following way:
(i) fix a finite set $S$ of primes of $F$ containing all infinite primes;
(ii) determine the (finite!) set $W_S$ of all good $S$-units in $F$;
(iii) let $w_1, w_2, \dots, w_t \in W_S$ be a minimal set with the property that $\lambda([w_1]), \lambda([w_2]), \dots, \lambda([w_t])$ generate $\lambda(A(W_S))$. Then every element

$w \in W_S$ satisfies

$$\lambda([w]) = \sum_{j=1}^{t} c_j \lambda([w_j]),$$

for some $c_j \in Z$. Consequently

$$[w] - \sum_{j=1}^{t} c_j [w_j] \in C(F)_S; \tag{11.4}$$

(iv) every element of $C(F)_S$ evidently is a $\mathbf{Z}$-linear combination of elements (11.4).

**11.2. Analytic background.** Every function $f$ defined on a field $F$ with values in an abelian group can be extended to a homomorphism of the group $A(F)$ by means of the formula

$$f\left(\sum_j k_j[a_j]\right) = \sum_j k_j f(a_j),$$

where $k_j \in Z$, $a_j \in F^*$, $a_j \neq 1$.

We shall be interested in some particular functions called dilogarithms. They have the property (at least conjecturally!) that, for every number field $F$, their values on the subgroup $C(F)$ form a finitely generated $Z$-module.

11.2.1. *Dilogarithms of Euler and of Rogers.* First we fix branches of the multivalued functions $\arg z$ and $\log z$. For $z \in C\backslash(-\infty, 0\rangle$, $z = x + iy$, we define as usual

$$\arg z = \operatorname{sgn} y \cdot \arccos \frac{x}{|z|}. \tag{11.5}$$

Then arg: $C\backslash(-\infty, 0\rangle \to (-\pi, \pi)$ is a real analytic function satisfying:

(i) $\arg(1/z) = \arg \overline{z} = -\arg z$,

(ii) $\arg(-z) = \arg z - \pi \operatorname{sgn}(\operatorname{Im} z)$,

(iii) $\arg z + \arg(1 - 1/z) = \arg(1 - z) + \pi \operatorname{sgn}(\operatorname{Im} z)$,

(iv) $\arg((-1)^{m-1} z^m) = m \sum_{j=0}^{m-1} \arg(\zeta_m^j z)$.

For $z \in \mathbb{C}\backslash(-\infty, 0\rangle$, we define, as usual,

$$\log z = \log|z| + i \arg z. \tag{11.6}$$

From the properties of $\arg z$ given above it follows that

(i) $\log(1/z) = -\log z$,

(ii) $\log \overline{z} = \overline{\log z}$,

(iii) $\log(-z) = \log z - i\pi \operatorname{sgn}(\operatorname{Im} z)$,

(iv) $\log z + \log(1 - 1/z) = \log(1 - z) + i\pi \operatorname{sgn}(\operatorname{Im} z)$,

(v) $\log((-1)^{m-1} z^m) = \sum_{j=0}^{m-1} \log(\zeta_m^j z)$.

For $|z| \le 1$, we define the dilogarithm of Euler $\operatorname{Li}_2(z)$ by the formula

$$\operatorname{Li}_2(z) := \sum_{n=1}^{\infty} \frac{z^n}{n^2}. \tag{11.7}$$



It follows in particular that $\operatorname{Li}_2(0) = 0$ and $\operatorname{Li}_2(1) = \pi^2/6$.

The function $\operatorname{Li}_2(z)$ can be extended to the set $C\backslash(-\infty, 0)$ by the integral

$$\operatorname{Li}_2(z) := \pi^2/6 - \int_1^z \frac{\log(1-t)}{t}\, dt, \tag{11.8}$$

$z \in \mathbb{C}\backslash(-\infty, 0)$.

The dilogarithm of Rogers $L(z)$, which in general has better properties than $\operatorname{Li}_2(z)$, is defined by the formula

$$L(z) := \operatorname{Li}_2(z) + \tfrac{1}{2} \log z \cdot \log(1 - z), \tag{11.9}$$

for $z \in \mathbb{C}\backslash((-\infty, 0) \cup (1, \infty))$. In particular we have $L(0) = 0$, $L(1) = \pi^2/6$.

We state below some useful properties of the dilogarithm of Rogers.

LEMMA 11.4. (i) $L(z) + L(1 - z) = \pi^2/6$,

(ii) $L(z) + L(1/z) = \pi^2/3 + i(\pi/2)\operatorname{sgn}(\operatorname{Im} z) \cdot \log z$,

(iii) $L(z^m) = m \sum_{j=0}^{m-1} L(\zeta_m^j z) + 2\pi \sum_{j=1}^{(m-1)/2} j \arg(1 - \zeta_m^j z)$, *for $m$ odd and $z \in (0, 1)$,*

(iv) *If $(a, b, c, d, e)$ is a 5-cycle, then*

$$L(a) + L(b) + L(c) + L(d) + L(e) = \pi^2/2,$$

(v) $L(\overline{z}) = \overline{L(z)}$,

*provided all terms are defined.*

PROOF. The formulas (i)–(iv) can be proved by differentiation, the formula (v) is obvious. □

From Lemma 11.4(i) with $z = 1/2$ it follows that $L(1/2) = \pi^2/12$.

Thus Lemma 11.4 proves that the values of the dilogarithm of Rogers on elements of $C(F)$ given in Lemmas 11.1 and 11.3, with the exception of the elements $[x] + [1/x]$, are integral multiples of $\pi^2/6$. We shall see below that in general elements of $C(F)$ given in Theorem 11.1 do not have this property.

11.2.2. *Clausen's function.* We define Clausen's function $\operatorname{Cl}_2(t)$, for $t \in R$, as the imaginary part of $\operatorname{Li}_2(e^{it})$,

$$\operatorname{Cl}_2(t) := \operatorname{Im} \operatorname{Li}_2(e^{it}).$$

From $e^{it} = \cos t + i \sin t$ it follows that

$$\operatorname{Li}_2(e^{it}) = \sum_{n=1}^{\infty} \frac{1}{n^2}(\cos nt + i \sin nt),$$

and consequently

$$\operatorname{Cl}_2(t) = \sum_{n=1}^{\infty} \frac{\sin nt}{n^2}, \qquad t \in R. \tag{11.10}$$

Using the integral representation (11.8) of $\mathrm{Li}_2(z)$ we can express $\mathrm{Cl}_2(t)$ in another way:

$$\begin{aligned}\mathrm{Cl}_2(t) &= \mathrm{Im}\,\mathrm{Li}_2(e^{it}) = -\mathrm{Im}\int_1^{e^{it}} \frac{\log(1-u)}{u}\,du \\ &= -\mathrm{Im}\int_0^t \log(1-e^{ir})\cdot i\,dr \\ &= -\mathrm{Re}\int_0^t \log(1-e^{ir})\,dr = -\int_0^t \log|1-e^{ir}|\,dr \\ &= -\int_0^t \log\left|2\sin\frac{r}{2}\right| dr,\end{aligned}$$

where we used the substitution $u = e^{ir}$, and applied the formula $|1-e^{ir}| = |2\sin\frac{r}{2}|$.

Consequently

$$\mathrm{Cl}_2(t) = -\int_0^t \log\left|2\sin\frac{r}{2}\right| dr. \tag{11.11}$$

From (11.10) it is easy to deduce some elementary properties of Clausen's function.

**Lemma** 11.5. (i) $\mathrm{Cl}_2(k\pi) = 0$, *for* $k \in \mathbf{Z}$,
(ii) $\mathrm{Cl}_2(t+2k\pi) = \mathrm{Cl}_2(t)$, *for* $k \in \mathbf{Z}$,
(iii) $\mathrm{Cl}_2(-t) = -\mathrm{Cl}_2(t)$,
(iv) *For* $m \geq 1$, *we have*

$$\frac{1}{m}\mathrm{Cl}_2(mt) = \sum_{j=0}^{m-1} \mathrm{Cl}_2(t+2\pi j/m). \tag{11.12}$$

*In particular*

$$\begin{aligned}\tfrac{1}{2}\mathrm{Cl}_2(2t) &= \mathrm{Cl}_2(t) + \mathrm{Cl}_2(t+\pi) \\ &= \mathrm{Cl}_2(t) - \mathrm{Cl}_2(\pi - t)\end{aligned} \tag{11.13}$$

*(duplication formula).*

**Proof.** Formulas (i)–(iii) follow immediately from (11.10). To prove (iv) we apply the well-known trigonometric formula

$$\sum_{j=0}^{m-1} \sin n(t+2\pi j/m) = \begin{cases} 0, & \text{if } m \nmid n, \\ m\sin(nt), & \text{if } m \mid n. \end{cases}$$



Hence

$$\begin{aligned}\sum_{j=0}^{m-1} \mathrm{Cl}_2(t+2\pi j/m) &= \sum_{n=1}^{\infty}\sum_{j=0}^{m-1} \frac{1}{n^2}\sin n(t+2\pi j/m) \\ &= \sum_{\substack{n=1\\ m\mid n}}^{\infty} \frac{m}{n^2}\sin(nt) \\ &= \sum_{k=1}^{\infty} \frac{m}{(mk)^2}\sin(mkt) = \frac{1}{m}\mathrm{Cl}_2(mt). \quad \square\end{aligned}$$

**Corollary** 11.1. (i) $\mathrm{Cl}_2(2\pi/3) = (2/3)\mathrm{Cl}_2(\pi/3)$,
(ii) $\mathrm{Cl}_2(\pi/6) + \mathrm{Cl}_2(5\pi/6) = (4/3)\mathrm{Cl}_2(\pi/2)$,
(iii) $12\,\mathrm{Cl}_2(\pi/6) = 8\,\mathrm{Cl}_2(\pi/2) + 3\,\mathrm{Cl}_2(\pi/3)$.

**Proof.** (i) Put $t = \pi/3$ in the duplication formula (11.13).
(ii) Put $t = \pi/6$ in the triplication formula: $(1/3)\mathrm{Cl}_2(3t) = \mathrm{Cl}_2(t) + \mathrm{Cl}_2(t+2\pi/3) + \mathrm{Cl}_2(t+4\pi/3)$.
(iii) Put $t = \pi/6$ in the duplication formula (11.13) and apply (ii). $\square$

Let us remark that in the last formula of Corollary 11.1 all arguments belong to the interval $\langle 0, \pi/2\rangle$. One more formula of this kind will be proved below, see (11.21).

**Lemma** 11.6. *The function* $\mathrm{Cl}_2(t)$ *has in the interval* $\langle 0, \pi\rangle$ *its only maximum at* $t = \pi/3$; $\mathrm{Cl}_2(\pi/3) = 1.0149417\ldots$.

**Proof.** We have $\mathrm{Cl}_2(0) = \mathrm{Cl}_2(\pi) = 0$, and in view of (11.11)

$$\mathrm{Cl}_2'(t) = -\log\left|2\sin\frac{t}{2}\right| = 0 \quad \text{iff} \quad t = \pi/3. \quad \square$$

11.2.3. *Partial Clausen's functions.* To prove some relations between the values of Clausen's function in arguments which are rational multiples of $\pi$, we consider the following partial sums of the series (11.10) defining Clausen's function.

For $m \nmid k$ and $t \in R$ let

$$S_{k,m}(t) := \sum_{n\equiv\pm k(2m)} \frac{1}{n^2}\sin(nt). \tag{11.14}$$

We call $S_{k,m}(t)$ the partial Clausen's function.

For $t = j\pi/m$, $j \in Z$, we have evidently

$$S_{k,m}(j\pi/m) = \sin(kj\pi/m)\left(\sum_{n\equiv k(2m)} \frac{1}{n^2} - \sum_{n\equiv -k(2m)} \frac{1}{n^2}\right).$$

For $m$ fixed, we denote $S_{k,m}(j\pi/m)$ simply by $S_k(j)$, and $S_k(1)$ by $S_k$. We put also $T_k(j) := S_k(j) + S_{m-k}(j)$, and $T_k := T_k(1)$.

In the following lemmas we establish some linear relations between the numbers $S_{k,m}(j\pi/m)$ and $\mathrm{Cl}_2(j\pi/m)$, $k, j = 1, 2, \ldots, m-1$.

**Lemma 11.7.** *For $m \nmid 2k$, $j \in Z$, we have*

$$S_{2k,m}(j\pi/m) = \tfrac{1}{2}(\cos(kj\pi/m)S_{k,m}(j\pi/m) + \cos((m-k)j\pi/m)S_{m-k,m}(j\pi/m)),$$

*i.e.,*

$$S_{2k}(j) = \tfrac{1}{2}(\cos(kj\pi/m)S_k(j) + \cos((m-k)j\pi/m)S_{m-k}(j)).$$

**Proof.** Every number $n$ satisfying $n \equiv \pm 2k \pmod{2m}$ is even, $n = 2r$. Moreover $r \equiv \pm k \pmod m$, and this is equivalent to

$$r \equiv \pm k \pmod{2m} \quad \text{or} \quad r \equiv \pm(m-k) \pmod{2m}.$$

Consequently in view of (11.14) we have

$$S_{2k,m}(j\pi/m) = \frac{1}{4}\sum_{r\equiv\pm k(2m)} \frac{1}{r^2}\sin(2rj\pi/m) + \frac{1}{4}\sum_{r\equiv\pm(m-k)(2m)} \frac{1}{r^2}\sin(2rj\pi/m),$$

and from the identity $\sin(2x) = 2\sin x\cos x$ and (11.14) the lemma follows. □

**Lemma 11.8.** *For $l, j, m \in Z$, $m > 0$, we have*

(11.15)
$$\mathrm{Cl}_2((2j+l)\pi/m) - \mathrm{Cl}_2((2j-l)\pi/m) = 2\sum_{k=1}^{m-1}\cos(2kj\pi/m)S_{k,m}(l\pi/m).$$

*In particular*

$$\mathrm{Cl}_2(l\pi/m) = \sum_{k=1}^{m-1} S_{k,m}(l\pi/m).$$

**Proof.** The formula (11.15) follows immediately in view of (11.14) from the trigonometric identity

$$\sin((2j+l)n\pi/m) - \sin((2j-l)n\pi/m) = 2\sin(ln\pi/m)\cos(jn\pi/m).$$

To obtain the second formula it is sufficient to substitute $j=0$ in (11.15). □

**Lemma 11.9.** *For $m \nmid k$ and $j, r \in \mathbb{Z}$, $r > 0$, we have*

$$S_{kr,mr}(j\pi/mr) = \frac{1}{r^2}S_{k,m}(j\pi/m),$$

*and similarly for $T$ in place of $S$.*



**Proof.** We have

$$\begin{aligned} S_{kr,mr}(j\pi/mr) &= \sum_{n\equiv\pm kr(mr)} \frac{1}{n^2}\sin(jn\pi/mr) \\ &\qquad \text{(substitute } n = rt) \\ &= \frac{1}{r^2}\sum_{t\equiv\pm k(m)}\frac{1}{t^2}\sin(jt\pi/m) \\ &= \frac{1}{r^2}S_{k,m}(j\pi/m). \quad \Box \end{aligned}$$

Now we apply the above lemmas to deduce several linear relations for small values of $m$.

$m = 2$. We have

$$\begin{aligned} \mathrm{Cl}_2(\pi/2) &= \sum_{n=1}^{\infty}\frac{1}{n^2}\sin(n\pi/2) = \sum_{n\equiv\pm1(4)}\frac{1}{n^2}\sin(n\pi/2) \\ &= S_{1,2}(\pi/2) = 1 - \frac{1}{3^2} + \frac{1}{5^2} - \cdots =: G \\ &= 0.91596559\ldots \quad \text{(Catalan's constant).} \end{aligned}$$

$m = 3$. From Lemma 11.6 with $m = 3$, $k = j = 1$ we get

$$S_2 = \tfrac{1}{4}(S_1 - S_2), \quad \text{i.e.,} \quad S_1 = 5S_2.$$

Hence Lemma 11.7 gives

(11.16)
$$\mathrm{Cl}_2(\pi/3) = S_1 + S_2 = 6S_2.$$

Thus

$$S_1 = \tfrac{5}{6}\mathrm{Cl}_2(\pi/3), \qquad S_2 = \tfrac{1}{6}\mathrm{Cl}_2(\pi/3).$$

$m = 4$. From Lemma 11.6 with $m = 4$, $k = j = 1$ we get

$$S_2 = \frac{\sqrt{2}}{4}(S_1 - S_3).$$

In view of Lemma 11.8 we have

$$S_2 = S_{2,4}(\pi/4) = \tfrac{1}{4}S_{1,2}(\pi/2) = \tfrac{1}{4}G.$$

Finally, Lemma 11.7 with $m = 4$, $j = 0$, $l = 1$ gives

$$\mathrm{Cl}_2(\pi/4) = S_1 + S_2 + S_4.$$

Hence

(11.17)
$$T_1 = S_1 + S_3 = \mathrm{Cl}_2(\pi/4) - \tfrac{1}{4}G, \qquad S_1 - S_3 = \frac{\sqrt{2}}{2}G,$$

and consequently

$$S_1 = \frac{1}{2}\mathrm{Cl}_2(\pi/4) + \frac{2\sqrt{2}-1}{8}G,$$
$$S_3 = \frac{1}{2}\mathrm{Cl}_2(\pi/4) - \frac{2\sqrt{2}+1}{8}G.$$

$m = 5$. Put $a = 2\cos(\pi/5) = \frac{1}{2}(1+\sqrt{5})$, $b = 2\cos(2\pi/5) = \frac{1}{2}(-1+\sqrt{5})$. We get analogously as above

$$T_1 = S_1 + S_4 = \frac{\sqrt{5}}{5}(\mathrm{Cl}_2(3\pi/5) + b\,\mathrm{Cl}_2(\pi/5)),$$
$$T_2 = S_2 + S_3 = \frac{\sqrt{5}}{5}(-\mathrm{Cl}_2(3\pi/5) + a\,\mathrm{Cl}_2(\pi/5)),$$

from Lemma 11.7, and

$$S_2 = \tfrac{a}{4}(S_1 - S_4),$$
$$S_4 = \tfrac{b}{4}(S_2 - S_3),$$

from Lemma 11.6. Consequently one can express every $S_j$ $(j = 1, 2, 3, 4)$ as a $\mathbb{Q}(\sqrt{5})$-linear combination of $\mathrm{Cl}_2(\pi/5)$ and $\mathrm{Cl}_2(3\pi/5)$.

$m = 6$. From Lemma 11.7 we get

$$\begin{aligned} \mathrm{Cl}_2(\pi/6) &= T_1 + T_2 + S_3, \\ \mathrm{Cl}_2(3\pi/6) - \mathrm{Cl}_2(\pi/6) &= T_1 - T_2 - 2S_3, \\ \mathrm{Cl}_2(5\pi/6) - \mathrm{Cl}_2(3\pi/6) &= -T_1 - T_2 + 2S_3. \end{aligned} \tag{11.18}$$

Moreover in view of cases $m = 2$ and $m = 3$ considered above and of Lemma 11.8 we have

$$T_2 = T_{2,6}(\pi/6) = \tfrac{1}{4}T_{1,3}(\pi/3) = \tfrac{1}{4}\mathrm{Cl}_2(\pi/3),$$
$$S_3 = S_{3,6}(\pi/6) = \tfrac{1}{9}S_{1,2}(\pi/2) = \tfrac{1}{9}G.$$

Adding two first equalities (11.18) we get

$$T_1 = \tfrac{5}{9}G \tag{11.19}$$

and adding the first and the last one we get

$$\mathrm{Cl}_2(\pi/6) + \mathrm{Cl}_2(5\pi/6) = \tfrac{4}{3}G.$$

It is the formula given in Corollary 11.1(ii).

$m = 12$. We get analogously

$$\mathrm{Cl}_2(\pi/12) = T_1 + T_5 + T_2 + T_3 + T_4 + S_6,$$
$$\mathrm{Cl}_2(5\pi/12) - \mathrm{Cl}_2(3\pi/12) = T_1 + T_5 - T_2 - 2T_3 - T_4 + 2S_6.$$

Hence subtracting we obtain

$$\mathrm{Cl}_2(5\pi/12) - \mathrm{Cl}_2(3\pi/12) - \mathrm{Cl}_2(\pi/12) = -2T_2 - 3T_3 - 2T_4 + S_6. \tag{11.20}$$



Moreover applying Lemma 11.8 we get

$$T_2 = T_{2,12}(\pi/12) = \tfrac{1}{4}T_{1,6}(\pi/6) = \tfrac{5}{36}G,$$

in view of (11.19),

$$T_3 = T_{3,12}(\pi/12) = \tfrac{1}{9}T_{1,4}(\pi/4) = \tfrac{1}{9}\mathrm{Cl}_2(\pi/4) - \tfrac{1}{4}G,$$

in view of (11.17), and

$$T_4 = T_{4,12}(\pi/12) = \tfrac{1}{16}T_{1,3}(\pi/3) = \tfrac{1}{16}\mathrm{Cl}_2(\pi/3),$$

in view of (11.16). Similarly $S_6 = \frac{1}{36}G$.

Substituting these values into (11.20) we get finally

$$\begin{aligned} &24\,\mathrm{Cl}_2(5\pi/12) - 24\,\mathrm{Cl}_2(\pi/12) \\ &\quad = 16\,\mathrm{Cl}_2(\pi/4) - 3\,\mathrm{Cl}_2(\pi/3) - 4\,\mathrm{Cl}_2(\pi/2). \end{aligned} \tag{11.21}$$

Let us observe that in this formula all arguments belong to the interval $\langle 0, \pi/2\rangle$. In the next paragraphs we shall give another method of constructing analogous formulas.

11.2.4. *Dilogarithm of Wigner and Bloch.* We define the dilogarithm $D(z)$ of Wigner and Bloch by the formula

$$D(z) = -\mathrm{Im}\int_1^z \frac{\log(1-t)}{t}\,dt + \arg(1-z)\cdot\log|z|, \tag{11.22}$$

for $z \in \mathbb{C}\setminus\mathbb{R}$, and $D(z) = 0$ for $z \in \mathbb{R}\cup\{\infty\}$.

In other words, for $z \in \mathbb{C}\setminus\mathbb{R}$,

$$D(z) = \mathrm{Im}\,L(z) - G(z), \tag{11.23}$$

where

$$G(z) = \tfrac{1}{2}(\arg z\cdot\log|1-z| - \arg(1-z)\cdot\log|z|),$$

and $D(z) = 0$, for $z \in \mathbb{R}\cup\{\infty\}$.

From (11.22) or (11.23) it follows that, for $t \in \mathbb{R}$,

$$D(e^{it}) = \mathrm{Im}\,\mathrm{Li}_2(e^{it}) = \mathrm{Cl}_2(t). \tag{11.24}$$

The dilogarithm $D\colon \mathbb{C}\cup\{\infty\} \to \mathbb{R}$ is a real-analytic function on $\mathbb{C}\setminus\{0, 1\}$, continuous on $\mathbb{C}\cup\{\infty\}$. Its properties given in the next lemma can be easily deduced from Lemma 11.4.

LEMMA 11.10. *For* $z \in \mathbb{C}\cup\{\infty\}$, *we have*

(i) $D(z) + D(1-z) = 0$,
(ii) $D(z) + D(1/z) = 0$,
(iii) $D(z) + D(\bar{z}) = 0$,
(iv) $D(z^m) = m\sum_{j=0}^{m-1} D(\zeta_m^j z)$,
(v) *If* $(a, b, c, d, e)$ *is a* 5-*cycle, then*

$$D(a) + D(b) + D(c) + D(d) + D(e) = 0.$$

Thus $D$ vanishes on elements of $C(F)$ given in Lemmas 11.1 and 11.3.

**Theorem** 11.2. *For $z \in \mathbb{C}\setminus\mathbb{R}$, we have*

$$D(z) = \tfrac{1}{2}(\mathrm{Cl}_2(2\theta) + \mathrm{Cl}_2(2\omega) - \mathrm{Cl}_2(2\theta + 2\omega)), \tag{11.25}$$

*where $\theta = \arg z$, $\omega = \arg(1-\bar{z})$.*

**Proof.** It is easy to verify that, for $z \in \mathbb{C}\setminus\mathbb{R}$

$$\left(\bar{z}, z/\bar{z}, \frac{1-\bar{z}}{1-z}, 1-z, 1-\frac{z(1-\bar{z})}{\bar{z}(1-z)}\right)$$

is a nontrivial 5-cycle. Moreover

$$z/\bar{z} = \exp(2i\theta), \qquad \frac{1-\bar{z}}{1-z} = \exp(2i\omega), \qquad \frac{z(1-\bar{z})}{\bar{z}(1-z)} = \exp(2i(\theta+\omega)).$$

Consequently in view of (11.24) we get

$$D(z/\bar{z}) = \mathrm{Cl}_2(2\theta), \qquad D\left(\frac{1-\bar{z}}{1-z}\right) = \mathrm{Cl}_2(2\omega),$$
$$D\left(\frac{z(1-\bar{z})}{\bar{z}(1-z)}\right) = \mathrm{Cl}_2(2\theta + 2\omega).$$

Therefore Lemma 11.10(v) gives

$$D(\bar{z}) + D(1-z) + \mathrm{Cl}_2(2\theta) + \mathrm{Cl}_2(2\omega) - \mathrm{Cl}_2(2\theta + 2\omega) = 0,$$

and the result follows from Lemma 11.10(i) and (iii). □

It follows from (11.24) and (11.25) that every formula containing the function $\mathrm{Cl}_2$ can be transformed into a formula containing the dilogarithm $D$, and vice versa. Since Clausen's function has been tabulated (see e.g. Clausen [32], Ashour and Sabri [56], Lewin [58]) the formula (11.25) enables one to compute easily an approximate value of $D(z)$ for every $z \in \mathbb{C}$.

**Lemma** 11.11. (i) *For $0 < \theta < \pi$ the function $D(z)$ is positive and attains its maximum value on the half line $L_\theta = \{z : \arg z = \theta\}$ at $z = e^{i\theta}$.*

(ii) *The function $D(z)$ attains its maximum value on $C$ at $z = e^{\pi i/3}$.*

(iii) $\lim_{z\to\infty} D(z) = 0$.

**Proof.** (i) For $z \in L_\theta$, $0 < \theta < \pi$, the value of $D(z)$ is given by (11.25) with fixed $\theta$ and variable $\omega \in (0, \pi)$. Hence from (11.11) we get

$$\begin{aligned}\frac{d}{d\omega}D(z) &= \mathrm{Cl}_2'(2\omega) - \mathrm{Cl}_2'(2\theta + 2\omega)\\ &= -\log|2\sin\omega| + \log|2\sin(\theta+\omega)|\\ &= \log\frac{\sin(\theta+\omega)}{\sin\omega}.\end{aligned}$$

Therefore $\frac{d}{d\omega}D(z)$ is positive for $0 < \omega < \frac{1}{2}(\pi - \theta)$, and is negative for $\omega > \frac{1}{2}(\pi-\theta)$. Consequently $D(z)$ attains its maximum on $L_\theta$ for $\omega = \frac{1}{2}(\pi - \theta)$, i.e., for $z = e^{i\theta}$.

(ii) The claim follows from Lemma 11.6 and (i).



(iii) In view of Lemma 11.10(ii) we have

$$\lim_{z\to\infty} D(z) = -\lim_{z\to\infty} D(1/z) = -D(0) = 0. \quad \square$$

11.2.5. *Dedekind zeta function of a number field $F$.* For a number field $F$, its zeta function $\zeta_F(s)$ is defined by

$$\zeta_F(s) := \sum_{\mathfrak{a}} 1/(N\mathfrak{a})^s,$$

for $\mathrm{Re}\, s > 1$, where $\mathfrak{a}$ runs over all nonzero ideals of the ring of integers $O_F$ of the field $F$. In particular, for $F = \mathbb{Q}$, $\zeta_{\mathbb{Q}}(s)$ is the Riemann zeta function $\zeta(s) = \sum_{n=1}^{\infty} 1/n^s$.

The function $\zeta_F(s)$ can be extended to a meromorphic function defined on the whole complex plane with a simple pole at $s = 1$. It satisfies a functional equation

$$\Phi(s) = \Phi(1-s), \tag{11.26}$$

where

$$\Phi(s) = (d_F/4^{r_2}\pi^n)^{s/2}\Gamma(s/2)^{r_1}\Gamma(s)^{r_2}\zeta_F(s),$$

$r_1$ (resp. $r_2$) is the number of real (resp. complex) infinite primes of $F$, $n = (F : Q) = r_1 + 2r_2$, and $d_F$ is the discriminant of the field $F$. For details, see e.g. Narkiewicz [90, Chapter VII].

If $F$ is an abelian extension of $\mathbb{Q}$, there exists a more explicit formula for the zeta function of $F$. By the Kronecker-Weber theorem, every abelian extension $F$ of $\mathbb{Q}$ is a subfield of a cyclotomic field:

$$\mathbb{Q} \subset F \subset \mathbb{Q}(\zeta_m), \quad \text{for some } m.$$

Since the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$ is isomorphic to the multiplicative group $(\mathbb{Z}/m)^*$ of residues modulo $m$ (to an automorphism $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$ there corresponds such a residue $k \pmod m$ that $\sigma(\zeta_m) = \zeta_m^k$), the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_m)/F)$ can be considered as a subgroup $H$ of index $(F : \mathbb{Q})$ of $(\mathbb{Z}/m)^*$.

Then

$$\zeta_F(s) = \prod_\chi L(s, \chi'), \tag{11.27}$$

where $\chi$ runs over characters of the group $(\mathbb{Z}/m)^*$ trivial on $H$, $\chi'$ is the primitive Dirichlet character corresponding to $\chi$, and

$$L(s, \chi') = \sum_{n=1}^{\infty} \chi'(n)/n^s, \tag{11.28}$$

for $\mathrm{Re}\, s > 1$. (See, Narkiewicz [90, Theorem 8.2]). The factor $L(s, \chi')$ in (11.27) corresponding to the principal character is equal to $\zeta(s) = \zeta_{\mathbb{Q}}(s)$.

It is known (see e.g. Borevich and Shafarevich [66, Chapter V, §2]) that, for a primitive Dirichlet character $\chi$ of conductor $m$ we have

$$L(s,\chi)=\frac{\tau(\chi)}{m}\sum_{k=1}^{m-1}\overline{\chi(k)}\sum_{n=1}^{\infty}\zeta_m^{-nk}/n^s,$$

where $\tau(\chi)$ is the corresponding Gauss's sum.

Consequently, if the character $\chi$ is odd, $\chi(-1)=-\chi(1)$, we get

$$\begin{aligned}L(s,\chi)&=\frac{\tau(\chi)}{m}\sum_{k=1}^{m-1}\overline{\chi(k)}\sum_{n=1}^{\infty}(-\zeta_m^{nk}+\zeta_m^{-nk})/2n^s\\&=-i\frac{\tau(\chi)}{m}\sum_{k=1}^{m-1}\overline{\chi(k)}\sum_{n=1}^{\infty}\frac{\sin(2nk\pi/m)}{n^s}.\end{aligned}$$

Hence

$$L(2,\chi)=-i\frac{\tau(\chi)}{m}\sum_{k=1}^{m-1}\overline{\chi(k)}\mathrm{Cl}_2(2k\pi/m). \tag{11.29}$$

In particular if $\chi$ is an odd real primitive character, then $\tau(\chi)=i\sqrt{m}$, and consequently

$$L(2,\chi)=m^{-1/2}\sum_{k=1}^{m-1}\chi(k)\mathrm{Cl}_2(2k\pi/m).$$

It is known that every odd real primitive character $\chi$ is given by the Kronecker symbol $\chi(\bullet)=(\frac{d}{\bullet})$, where $d$ is a negative discriminant. Its conductor is equal to $|d|$.

Thus

$$L(2,\chi)=|d|^{-1/2}\sum_{k=1}^{|d|-1}\left(\frac{d}{k}\right)\mathrm{Cl}_2(2k\pi/|d|).$$

Since $(\frac{d}{k})=-(d/(|d|-k))$, and $\mathrm{Cl}_2(2(|d|-k)\pi/|d|)=-\mathrm{Cl}_2(2k\pi/|d|)$, the previous formula can be written in the form

$$L(2,\chi)=2|d|^{-1/2}\sum_{1\le k\le(|d|-1)/2}\left(\frac{d}{k}\right)\mathrm{Cl}_2(2k\pi|d|), \tag{11.30}$$

where $\chi(k)=(\frac{d}{k})$ is the Kronecker symbol and $d$ is a negative discriminant.

**11.3. $K$-theoretic background.** In algebraic $K$-theory there are functors $K_n$ $(n\ge 0)$ from the category of associative rings with 1 to the category of abelian groups. For a number field $F$, the groups $K_nO_F$ are finitely generated, and their ranks are described by the following



THEOREM 11.3. (Borel [72]). *For a number field $F$ and $n\ge 0$,*

$$\mathrm{rk}(K_nO_F)=\begin{cases}0, & \text{if } n \text{ is even},\ n>0,\\ r_2, & \text{if } n=3 \pmod 4,\\ r_1+r_2, & \text{if } n=1 \pmod 4,\ n>1,\\ r_1+r_2-1, & \text{if } n=1,\\ 1, & \text{if } n=0,\end{cases}$$

*where $r_1$ (resp. $r_2$) is the number of real, (resp. complex) places of $F$. Moreover $\mathrm{rk}(K_nF)=\mathrm{rk}(K_nO_F)$, for $n>1$.*

In particular, the group $K_2O_F$ is finite, and the groups $K_3F$ and $K_3O_F$ have ranks $r_2$.

We shall consider the mapping

$$\mathbb{D}\colon C(F)\to\mathbb{R}^{r_2}$$

defined as follows. For $b=\sum k_i[a_i]\in C(F)$, where $k_i\in\mathbf{Z}$, $a_i\in F^*\backslash\{1\}$, and for an injection $\sigma\colon F\to\mathbf{C}$, let $\sigma(b)=\sum k_i[\sigma(a_i)]$ be an element of $C(\sigma(F))$. Then we put

$$\mathbb{D}(b)=(D(\sigma_1(b)),\dots,D(\sigma_{r_2}(b))),$$

where $D$ is the dilogarithm of Bloch and $\sigma_1,\dots,\sigma_{r_2}$ are all complex places of $F$.

It is known that the mapping $\mathbb{D}$ can be decomposed as follows: $C(F)\to K_3F\xrightarrow{\rho}\mathbb{R}^{r_2}$, where $\rho$ is the regulator map (see e.g. Suslin [87]). Moreover $\mathbb{D}(C(F))$ is a lattice of rank $r_2$ in $\mathbb{R}^{r_2}$. Therefore the volume of $\mathbb{R}^{r_2}/\mathbb{D}(C(F))$ is finite. We denote it by $D_F$.

If, for some $b_1,\dots,b_{r_2}\in C(F)$, $\mathbb{D}(b_1),\dots,\mathbb{D}(b_{r_2})$ is a basis of the lattice $\mathbb{D}(C(F))$, then

$$D_F=|\det(D(\sigma_i(b_j)))_{1\le i,j\le r_2}|.$$

If $r_2=0$, we put $D_F=1$.

For a number field $F$, let

$$w_2(F)=\#H^0(G,\mu^{\otimes 2})=\#(\mu\otimes\mu)^G,$$

where $G$ is the Galois group of the algebraic closure $\overline{F}$ of $F$, and $\mu$ is the group of all roots of unity in $\overline{F}$.

It is easy to see that $w_2(F)=2\prod_p p^{m_p}$, where $p$ runs over rational primes, and $m_p$ is the maximal integer $m\ge 0$ such that $\zeta_{p^m}+\zeta_{p^m}^{-1}\in F$.

E.g. $w_2(F)=24$, if $F=\mathbf{Q}$ or $F$ is a quadratic extension of $\mathbf{Q}$, with the following exceptions:

$$w_2(Q(\sqrt{2}))=48,\qquad w_2(Q(\sqrt{5}))=120.$$

It is known that $\zeta_F(s)$ has at $s=-1$ a zero of multiplicity $r_2$. Therefore $\lim_{s\to-1}(s+1)^{-r_2}\zeta_F(s)$ exists and is $\ne 0$. We denote it by $\xi_F(-1)$. We have $\xi_F(-1)=\zeta_F(-1)\ne 0$, if $r_2=0$, i.e., if $F$ is totally real.

From the functional equation (11.26) it follows that

$$\zeta_F(-1) = (-1)^{r_1+r_2}\frac{|d|^{3/2}}{2^{r_1+3r_2}\cdot\pi^{2r_1+3r_2}}\zeta_F(2), \tag{11.31}$$

where $d$ is the discriminant of $F$.

Now we are ready to state a particular case of Lichtenbaum's conjecture.

CONJECTURE OF LICHTENBAUM. For every number field $F$,

$$\# K_2O_F = w_2(F)\frac{\pi^{r_2}|\zeta_F(-1)|}{D_F}. \tag{11.32}$$

In particular, if $F$ is totally real (i.e., $r_2 = 0$) we have $D_F = 1$, $\zeta_F(-1) \neq 0$, and the formula (11.32) takes the form:

CONJECTURE OF BIRCH AND TATE.

$$\# K_2O_F = w_2(F)\zeta_F(-1), \quad \text{for } F \text{ totally real.} \tag{11.33}$$

This last conjecture has been proved up to powers of 2 dividing both sides of (11.33). It has also been proved for infinitely many real quadratic fields.

If $r_2 = 1$, e.g. if $F$ is imaginary quadratic, there exists $b \in C(F)$ such that $D(b)$ generates the lattice $\mathbf{D}(C(F))$ in $\mathbf{R}^1$, and then $D_F = |D(b)|$.

In view of (11.31) Lichtenbaum's conjecture can be rewritten in the form

$$\# K_2O_F = w_2(F)\frac{|d|^{3/2}}{D_F\cdot 2^{r_1+3r_2}\cdot\pi^{2r_1+2r_2}}\zeta_F(2). \tag{11.34}$$

In particular, if $F$ is an abelian extension of $\mathbf{Q}$, then

$$\zeta_F(s) = \zeta(s)\cdot\prod_\chi L(s, \chi),$$

for appropriate characters $\chi$. Consequently, since $\zeta(2) = \pi^2/6$, from (11.34) we get

$$\# K_2O_F = w_2(F)\frac{|d|^{3/2}}{6\cdot D_F\cdot 2^{r_1+3r_2}\cdot\pi^{2r_1+2r_2-2}}\prod_\chi L(2, \chi).$$

Therefore, for an imaginary quadratic field $F$ of discriminant $d$ we get (since $r_1 = 0$, $r_2 = 1$, $w_2(F) = 24$)

$$\# K_2O_F = \frac{|d|^{3/2}}{2D_F}\cdot L(2, \chi),$$

where $\chi(a) = (\frac{d}{a})$ is the character corresponding to $F$.

Hence in view of (11.30) we have finally

$$\# K_2O_F = \frac{|d|}{D_F}\sum_{1\le k\le(|d|-1)/2}\left(\frac{d}{k}\right)\mathrm{Cl}_2\left(\frac{2k\pi}{|d|}\right), \tag{11.35}$$

for quadratic imaginary field $F$ of discriminant $d$.



**11.4. Examples.** Now we give several explicit numerical examples illustrating the general theory developed earlier in this chapter.

11.4.1. *Linear relations between the values of $L(x)$ at rational arguments.* In the present paragraph we consider sets $W_S$ of good $S$-units in the field $\mathbf{Q}$ of rational numbers for some sets $S$ consisting of small rational primes and the only infinite prime of $\mathbf{Q}$ which we denote by $\infty$. Applying results given in 11.1.2 and 11.1.3 we obtain some $Z$-linear relations between the numbers $L(a)$, $a \in W_S$.

In the set $W_S$ of good $S$-units in $\mathbf{Q}$ we define a relation of dependence. For a subset $A$ of $W_S$ and an element $a$ of $W_S$, we say that $a$ is dependent on $A$ if either

(i) $1 - a \in A$, or

(ii) $1/a \in A$, or

(iii) There is a 5-cycle $(a, b, c, d, e)$ such that $b, c, d, e \in A$.

Next we extend this relation by transitivity, i.e., we say that $a$ is dependent on $A$ if there exists a sequence of subsets of $W_S$

$$A_0 \subset A_1 \subset \cdots \subset A_k,$$

such that $A_0 = A$, $a \in A_k$, and for every $j = 1, 2, \dots, k$ every element of $A_j$ is dependent on $A_{j-1}$ in the above sense. We say that elements $a_1, a_2, \dots, a_t \in W_S$ are dependent, if one of them, say $a_j$, is dependent on the set $\{a_1, \dots, a_t\}\setminus\{a_j\}$.

LEMMA 11.12. *Let a good $S$-unit $a$ be dependent on a subset $\{a_1, \dots, a_t\}$ of $W_S$. Then*

$$L(a) = \sum_{j=1}^{t} r_j L(a_j) + r\frac{\pi^2}{6},$$

*for some $r, r_j \in \mathbf{Z}$, provided all $L(a)$, $L(a_j)$ are defined.*

PROOF. It is an immediate consequence of the definition of dependence and of Lemma 11.4. □

LEMMA 11.13. *If elements $a_1, a_2, \dots, a_t \in W_S$ are dependent, then elements $\lambda([a_1]), \lambda([a_2]), \dots, \lambda([a_t]) \in F^* \wedge F^*$ are linearly dependent in the sense that one of them is a $\mathbf{Z}$-linear combination of the others.*

PROOF. The lemma follows immediately from the definition of dependence in view of Lemmas 11.1 and 11.3. □

Let $a_1, a_2, \dots, a_t \in W_S$ be a minimal set such that the elements $\lambda([a_1])$, $\lambda([a_2]), \dots, \lambda([a_t])$ generate the group $\lambda(A(W_S))$. Then to every $a \in W_S$ which is dependent on the set $\{a_1, a_2, \dots, a_t\}$ we can apply Lemma 11.12, obtaining a representation of $L(a)$ as a $\mathbf{Z}$-linear combination of numbers $L(a_1), L(a_2), \dots, L(a_t)$, and $\pi^2/6$.

Since in general the set of such elements $a$ is large, we obtain in this way many linear relations. One may ask if every element of $W_S$ is dependent on a fixed minimal set $\{a_1, a_2, \dots, a_t\}$ (see 11.5.1).

We give below some numerical examples.

**Example** 11.1. Let $S = \{\infty, 2, 3\}$. It is well known that

$$W_S \cap (0, 1) = \{1/2, 1/3, 2/3, 1/4, 3/4, 1/9, 8/9\}.$$

Moreover we have

$$U_S \wedge U_S = \mathrm{lin}((-1) \wedge (-1), (-1) \wedge 2, (-1) \wedge 3, 2 \wedge 3).$$

From

$$\lambda([1/2]) = (-1) \wedge 2, \qquad \lambda([1/3]) = (-1) \wedge 3 + 2 \wedge 3,$$

we conclude that $\lambda([1/2])$ and $\lambda([1/3])$ generate a subgroup of index 4 in $U_S \wedge U_S$.

Since $(K_2\mathbf{Q})_S = \mathbf{Z}/2 \times \mathbf{F}_3^*$ has order 4 (see Milnor [71, Theorem 11.6]), it follows that $\lambda([1/2])$ and $\lambda([1/3])$ generate $\lambda(A(W_S))$, i.e., $\{1/2, 1/3\}$ is a minimal subset of $W_S$. Of course, one can also verify this directly.

There are two 5-cycles consisting of good $S$-units:

$$(1/2, 1/2, 2/3, 3/4, 2/3) \quad \text{and} \quad (1/3, 1/3, 3/4, 8/9, 3/4).$$

Hence it follows easily that every good $S$-unit is dependent on the set $\{1/2, 1/3\}$. Consequently in view of Lemma 11.12, for every $a \in W_S$, we have

$$L(a) = c_1 L(1/2) + c_2 L(1/3) + c_3 \pi^2/6,$$

for some $c_1, c_2, c_3 \in \mathbf{Z}$. Since $L(1/2) = \pi^2/12$, we can assume that $c_3 = 0$.

In particular applying Lemma 11.4 and the above 5-cycles we get

$$L(3/4) = 2L(1/3),$$
$$L(1/9) = 6L(1/3) - 4L(1/2) = 6L(1/3) - \pi^2/3.$$

**Example** 11.2. Let $S = \{\infty, 2, 3, 5\}$. We have (see Alex [76]):

$$W_S \cap (0, 1/2\rangle = \{1/2, 1/3, 1/4, 1/5, 1/6, 2/5, 3/8, 1/9, 4/9, 1/10, 1/16, 1/25, 9/25, 2/27, 5/32, 1/81, 3/128\}.$$

Moreover,

$$U_S \wedge U_S = \mathrm{lin}((-1) \wedge (-1), (-1) \wedge 2, (-1) \wedge 3, (-1) \wedge 5, 2 \wedge 3, 2 \wedge 5, 3 \wedge 5).$$

Since

$$\begin{aligned} &\lambda([1/2]) = (-1) \wedge 2, \\ &\lambda([1/3]) = (-1) \wedge 3 + 2 \wedge 3, \\ &\lambda([1/5]) = (-1) \wedge 5 + 2(2 \wedge 5), \\ &\lambda([1/6]) = (-1) \wedge 2 + (-1) \wedge 3 + 2 \wedge 5 + 3 \wedge 5, \end{aligned} \tag{11.36}$$

we see that the elements (11.36) generate a subgroup of index 16 in $U_S \wedge U_S$. The group $(K_2\mathbf{Q})_S = \mathbf{Z}/2 \times \mathbf{F}_3^* \times \mathbf{F}_5^*$ has order 16, and hence it follows



that $\lambda(A(W_S))$ is generated by elements (11.36). Obviously they are linearly independent.

Therefore $\{1/2, 1/3, 1/5, 1/6\}$ is a minimal subset of $W_S$.

Now we shall prove that every element of $W_S$ is dependent on this subset using the following 5-cycles consisting of good $S$-units. We shall replace freely $a$ by $1-a$ without mentioning it explicitly. The arrows indicate new good $S$-units which are dependent on earlier ones.

$$\begin{gathered}
(1/2, 1/3, 3/5, 5/6, 4/5) \to 2/5, \\
(1/2, 2/5, 5/8, 4/5, 3/4) \to 3/8, \\
(1/2, 1/2, 2/3, 3/4, 2/3) \to 1/4, \\
(1/3, 1/3, 3/4, 8/9, 3/4) \to 1/9, \\
(2/3, 2/3, 3/5, 5/9, 3/5) \to 4/9, \\
(1/4, 1/4, 4/5, 15/16, 4/5) \to 1/16, \\
(1/5, 1/5, 5/6, 24/25, 5/6) \to 1/25, \\
(3/5, 3/5, 5/8, 16/25, 5/8) \to 9/25, \\
(1/4, 2/3, 9/10, 5/6, 2/5) \to 1/10, \\
(1/6, 3/5, 25/27, 9/10, 4/9) \to 2/27, \\
(1/4, 4/9, 27/32, 8/9, 5/8) \to 5/32, \\
(1/9, 1/9, 9/10, 80/81, 9/10) \to 1/81, \\
(3/8, 1/16, 16/25, 125/128, 24/25) \to 3/128.
\end{gathered}$$

Thus we have proved that every good $S$-unit belonging to the interval $(0, 1/2\rangle$, and hence every good $S$-unit, is dependent on the set $\{1/2, 1/3, 1/5, 1/6\}$. This leads via Lemma 11.12 to an expression of $L(a)$, for every $a \in W_S \cap (0, 1)$, in the form

$$c_1 L(1/2) + c_2 L(1/3) + c_3 L(1/5) + c_4 L(1/6) + c_5 \pi^2/6,$$

with $c_j \in Z$. We can assume as above that $c_5 = 0$.

E.g. we have

$$\begin{aligned}
L(2/5) &= L(1/2) + L(1/3) - L(1/5) - L(1/6), \\
L(3/8) &= 2L(1/2) + L(1/3) - L(1/4) - 2L(1/5), \\
L(125/128) &= -7L(1/3) + 12L(1/5) + 3L(1/6).
\end{aligned}$$

**Example** 11.3. Let $S = \{\infty, 2, 3, 5, 7\}$. Alex [76] determined all good $S$-units in this case: The set $W_S \cap (0, 1/2\rangle$ is the sum of the set given in Example 11.2 and the set

$$\{1/7, 2/7, 3/7, 1/8, 2/9, 3/10, 5/12, 5/14, 1/15, 7/15, 7/16, 1/21, 5/21, 4/25, 7/25, 7/27, 1/28, 3/28, 7/32, 3/35, 8/35, 1/36, 1/49, 4/49, 9/49, 24/49, 1/50, 5/54, 1/64, 15/64, 25/81, 32/81, 27/125, 1/126, 7/135, 64/189, 1/225, 2/245, 81/256, 100/343, 32/375, 49/625, 5/1029, 1/2401, 1/4375\}.$$

We have

$$U_S \wedge U_S = \mathrm{lin}(\{(-1)\wedge q : q \in \{-1, 2, 3, 5, 7\}\} \cup \{p \wedge q : p, q \in \{2, 3, 5, 7\}, p < q\}).$$

One can verify analogously as in the above examples that

$$\{1/2, 1/3, 1/5, 1/6, 1/7, 1/8, 2/7\} \tag{11.37}$$

is a minimal subset of $W_S$. Moreover constructing corresponding 5-cycles (consisting of good $S$-units) one can verify that every good $S$-unit is dependent on the subset (11.37). Therefore applying Lemma 11.12 we can represent every element $L(a)$, for $a \in W_S$, as a $\mathbb{Z}$-linear combination of elements $L(a_j)$, for $a_j$ belonging to the set (11.37).

E.g. we have

$$L(1/64) = 2L(1/8) + 6L(1/4) - 2\frac{\pi^2}{6},$$
$$L(5/1029) = 10L(1/3) - 4L(1/5) - L(1/6) - 3L(2/7) - 9L(1/8),$$
$$L(1/4375) = 2L(1/8) + 7L(1/7) - 28L(1/6) - 12L(1/5) + 7\frac{\pi^2}{6}.$$

REMARKS. 1. One can define the relation of dependence in the set $W_S$ of good $S$-units for every number field $F$. However in this case it seems natural to consider in the definition one more condition:

If $\zeta_m \in F$ and $a \in W_S$, then $a^m$ is dependent on the set $\{a, \zeta_m a, \zeta_m^2 a, \ldots, \zeta_m^{m-1} a\}$ provided all these elements are good $S$-units.

For $F = \mathbb{Q}$ this condition is automatically satisfied, since $m = 2$, and the 5-cycle $(a, a, (1+a)^{-1}, 1-a^2, (1+a)^{-1})$ proves that $a^2$ is dependent on the set $\{a, -a\}$ provided $a$, $-a$ and $a^2$ are good $S$-units.

2. For $S = \{\infty, 2, 3, 5, 7, 11, 13\}$ the set of all good $S$-units in $\mathbb{Q}$ has been determined by de Weger [87]. Unfortunately, his paper does not contain the complete list of these $S$-units. Therefore one cannot verify if all of them are dependent on a minimal set.

3. From the theorem of Zsigmondy on primitive divisors (see e.g. Schinzel [62]) it follows that there are very few independent cyclotomic equations satisfied by rational numbers. In fact we have only

$$a^{-1}(1-a) = 1, \quad \text{for } a \mp 1/2,$$
$$a^{-1}(1-a)^3(1-a^2)^{-1} = 1, \quad \text{for } a = 1/3,$$
$$a^{-1}(1-a^2)^2(1-a^3)(1-a^6)^{-1} = 1, \quad \text{for } a = 1/2,$$

provided $a \in (0, 1)$. Multiplying both sides of these equations respectively by $a^2\wedge$, $a^2\wedge$ and $a^6\wedge$ we get

$$2[a] \in C(F), \quad \text{for } a = 1/2,$$
$$6[a] - [a^2] \in C(F), \quad \text{for } a = 1/3,$$
$$6[a^2] + 2[a^3] - [a^6] \in C(F), \quad \text{for } a = 1/2.$$



The corresponding linear relations for $L(x)$

$$2L(1/2) = \pi^2/6,$$
$$6L(1/3) - L(1/9) = 2\cdot\pi^2/6,$$
$$6L(1/4) + 2L(1/8) - L(1/64) = 2\cdot\pi^2/6$$

are well known, see Examples 11.1 and 11.3.

Therefore using cyclotomic equations we obtain very few linear relations between the numbers $L(a)$, $a \in \mathbb{Q} \cap (0, 1)$, and $\pi^2/6$.

11.4.2. *Linear relations between the values of* $L(x)$ *at arguments in a real quadratic field.*

11.4.2.1. Cyclotomic equations satisfied by quadratic algebraic integers. Denote by

$$c_n = c_n(x) = \prod_{\substack{k=0 \\ (n,k)=1}}^{n-1} (1 - \zeta_n^k x) \in \mathbb{Z}[x],$$

the $n$th cyclotomic polynomial, $n = 1, 2, \ldots$. Since $c_1^2, c_2^2, c_3, c_4, \ldots$ are symmetric polynomials of even degrees, they can be expressed as polynomials in $y = x + x^{-1}$ with integral coefficients multiplied by a power of $x$.

E.g. we have

$$c_1^2(x) = (1-x)^2 = x(y-2),$$
$$c_2^2(x) = (1+x)^2 = x(y+2),$$
$$c_3(x) = 1 + x + x^2 = x(y+1),$$
$$c_4(x) = 1 + x^2 = xy,$$
$$c_6(x) = 1 - x + x^2 = x(y-1),$$
$$c_5(x) = 1 + x + x^2 + x^3 + x^4 = x^2(y^2 + y - 1),$$
$$c_{10}(x) = c_5(-x) = x^2(y^2 - y - 1),$$
$$c_8(x) = c_4(x^2) = 1 + x^4 = x^2(y^2 - 2),$$
$$c_{12}(x) = c_6(x^2) = x^2(y^2 - 3), \quad \text{etc.}$$

From the definition of cyclotomic polynomials it follows that $1 - x^n = \prod_{d|n} c_d(x)$, and hence by the Möbius inversion formula

$$c_n(x) = \prod_{d|n} (1 - x^d)^{\mu(n/d)}, \tag{11.38}$$

where $\mu$ is the Möbius function.

Consequently for every number $a$ a multiplicative relation between the numbers $c_n(a)$, $n = 1, 2, \ldots$, and $a$ leads to a cyclotomic equation satisfied by $a$, and vice versa. Therefore we shall look for multiplicative relations between the numbers

$$a,\ c_1^2(a),\ c_2^2(a),\ c_3(a),\ c_4(a), \ldots. \tag{11.39}$$

Let $a \neq 0$ be an algebraic integer, which is not a root of unity. We say that $a$ is exceptional of order $n$, if every prime ideal dividing $a^n - 1$ divides $a^d - 1$ for some $d$, $1 \leq d < n$. In particular $a$ is exceptional of order 1 iff $a \neq 0, 1$ and $a - 1$ is a unit.

One can prove easily that

(i) If a prime ideal divides $a^r - 1$ and $a^s - 1$ for some natural numbers $r$ and $s$, then it divides also $a^{(r,s)} - 1$. Therefore in the above definition one can assume that $d$ divides $n$.

(ii) If $a^r$ is exceptional of order $n$, then $a$ is exceptional of order $rn$.

If an algebraic integer $a$, which is not a root of unity, satisfies the cyclotomic equation

$$a^A(1-a^{r_1})^{A_1}(1-a^{r_2})^{A_2}\cdots(1-a^{r_s})^{A_s} = 1,$$

where $0 < r_1 < r_2 < \cdots < r_s$ and $A_s \neq 0$, then evidently $a$ is exceptional of order $r_s$.

From a result of Rédei [58, Satz 5] it follows that if a quadratic algebraic integer $a$ is exceptional of order $n$, then $n \leq 4$ or $n = 6$, with the following exceptions (up to conjugation):

| | |
|---|---|
| $a = \pm\rho$, | where $\rho = \frac{1}{2}(-1+\sqrt{5})$, is exceptional of orders 12, 20 and 24, |
| $a = \rho^2$ | is exceptional of orders 10 and 12, |
| $a = -\rho^2$ | is exceptional of orders 5 and 12, |
| $a = \sqrt{2}$ | is exceptional of order 12, |
| $a = 1+i$ | is exceptional of order 10, |
| $a = -1+i$ | is exceptional of order 5. |

Applying this result one can determine all independent cyclotomic equations satisfied by a quadratic algebraic integer $a$. If $a$ is exceptional of order $n$, it is sufficient to find all independent multiplicative relations between the numbers $a$ and $1 - a^k$, for $k \mid n$.

Thus for numbers $a$ different from the values given explicitly above (and their conjugates) it is sufficient to look for multiplicative relations between the numbers $a$, $1-a$, $1-a^2$, $1-a^3$, $1-a^4$, and $1-a^6$, or equivalently (see (11.39)) between the numbers $a$, $1-a$, $1+a$, $1+a+a^2$, $1+a^2$, and $1-a+a^2$.

We give below some examples of this kind. We omit elementary details and state only the results (see Browkin [90]).

EXAMPLE 11.4. Let $a = \rho$. From the result of Rédei stated above it follows that $a$ is exceptional of orders $1, 2, 6, 12, 20$, and $24$ only. Hence there are only six independent cyclotomic equations satisfied by $a = \rho$. They are given below with the corresponding elements of $C(F)$ and corresponding **Z**-linear relations between the values of $L(a^k)$, $k = 1, 2, \ldots,$



and $D_F := \pi^2/30$. All these cyclotomic equations are known and the linear relations have been proved (see Lewin [84, (33a)–(33f) and (10a), (10b), (12a)–(12c), (34)], where all this is given in an equivalent form).

(1) $a^{-2}(1-a) = 1$, $\quad a\wedge$
$[a] \in C(F)$,
$L(a) = 3D_F$.

(2) $a^{-1}(1-a^2) = 1$, $\quad a^2\wedge$
$[a^2] \in C(F)$,
$L(a^2) = 2D_F$.

(3) $a^{-1}(1-a^3)^2(1-a^6)^{-1} = 1$, $\quad a^6\wedge$
$b_1 := 4[a^3] - [a^6] \in C(F)$,
$L(b_1) = 5D_F$.

(4) $a^{-2}(1-a^3)^3(1-a^4)(1-a^{12})^{-1} = 1$, $\quad a^{12}\wedge$
$b_2 := 12[a^3] + 3[a^4] - [a^{12}] \in C(F)$,
$L(b_2) = 19D_F$.

(5) $a^{-1}(1-a^4)^3(1-a^{10})(1-a^{20})^{-1} = 1$, $\quad a^{20}\wedge$
$b_3 := 15[a^4] + 2[a^{10}] - [a^{20}] \in C(F)$,
$L(b_3) = 14D_F$.

(6) $a^{-2}(1-a^3)^4(1-a^4)^{-1}(1-a^8)^2(1-a^{24})^{-1} = 1$, $\quad a^{24}\wedge$
$b_4 := 32[a^3] - 6[a^4] - 6[a^8] - [a^{24}] \in C(F)$,
$L(b_4) = 39D_F$.

11.4.2.2. Quadratic algebraic integers of norm 1. Let $a \in (0, 1)$ be a root of the polynomial $x^2 - sx + 1$, where $s > 2$ is an integer. Thus $a = \frac{1}{2}(s - \sqrt{s^2-4})$. Suppose that $a$ satisfies a cyclotomic equation. Then in view of the result of Rédei stated above either $a$ is an exceptional number of order $\leq 4$ or 6, or $a = \rho^2$ (which is an exceptional number also of orders 10 and 12).

Consequently, to find all cyclotomic equations satisfied by $a \neq \rho^2$ it is sufficient to look for independent multiplicative relations between the numbers $a$, $c_1^2(a)$, $c_2^2(a)$, $c_3(a)$, $c_4(a)$, $c_6(a)$. For $a = \rho^2$ we should consider also $c_{10}(a)$ and $c_{12}(a)$.

Since $a + a^{-1} = s$, we get

$$c_1^2(a) = a(s-2), \qquad c_2^2(a) = a(s+2), \qquad c_3(a) = a(s+1),$$
$$c_4(a) = as, \qquad c_6(a) = a(s-1).$$

Moreover $c_{10}(a) = a^2(s^2 - s - 1)$, and $c_{12}(a) = a^2(s^2-3)$.

Therefore we look for multiplicative relations between the five consecutive integers $s-2, s-1, s, s+1, s+2$, for $s > 2$. It is an easy exercise to prove that such a relation exists for $s = 3, 4, 5, 6, 7, 8$, and 10 only, since at least two of these five numbers should be powers of 2 or of 3.

In the example below we state results corresponding to these values of $s$ (see Lewin [**82**, **84**]).

EXAMPLE 11.5. Let $a = \frac{1}{2}(s - \sqrt{s^2-4})$, where $s = 3, 4, 5, 6, 7, 8$, and 10. The basic data are given in Table 11.1, where $\rho = \frac{1}{2}(-1+\sqrt{5}) = 2\cos(2\pi/5)$, $\tau = \sqrt{2}-1 = \tan(\pi/8)$, $\chi = 2-\sqrt{3} = \tan(\pi/12)$.

| $s$ | $a$ | $c_1^2$ | $c_2^2$ | $c_3$ | $c_4$ | $c_6$ | $c_{10}$ | $c_{12}$ |
|---|---|---|---|---|---|---|---|---|
| 3 | $\frac{1}{2}(3-\sqrt{5}) = \rho^2$ | $a$ | $5a$ | $4a$ | $3a$ | $2a$ | $5a^2$ | $6a^2$ |
| 4 | $2-\sqrt{3} = \chi$ | $2a$ | $6a$ | $5a$ | $4a$ | $3a$ | | |
| 5 | $\frac{1}{2}(5-\sqrt{21})$ | $3a$ | $7a$ | $6a$ | $5a$ | $4a$ | | |
| 6 | $3-2\sqrt{2} = \tau^2$ | $4a$ | $8a$ | $7a$ | $6a$ | $5a$ | | |
| 7 | $\frac{1}{2}(7-3\sqrt{5}) = \rho^4$ | $5a$ | $9a$ | $8a$ | $7a$ | $6a$ | | |
| 8 | $4-\sqrt{15}$ | $6a$ | $10a$ | $9a$ | $8a$ | $7a$ | | |
| 10 | $5-2\sqrt{6}$ | $8a$ | $12a$ | $11a$ | $10a$ | $9a$ | | |

TABLE 11.1

From the table we get the following multiplicative relations:

$s = 3$. $a^{-1}\cdot c_1^2 = 1$, $c_6^2 = c_3\cdot c_1^2$, $c_{10} = c_2^2\cdot c_1^2$, $c_{12} = c_6\cdot c_4$.

$s = 4$. $c_6\cdot c_1^2 = a\cdot c_2^2$, $a\cdot c_4 = c_1^4$.

$s = 5$. $c_6\cdot c_1^4 = a\cdot c_3^2$.

$s = 6$. $a\cdot c_2^4 = c_1^6$.

$s = 7$. $c_6^6 = a\cdot c_3^2\cdot c_2^6$.

$s = 8$. $a\cdot c_4^2\cdot c_3^3 = c_1^{12}$.

$s = 10$. $c_6^3\cdot c_1^8 = a\cdot c_2^{12}$.

We give below cyclotomic equations satisfied by these numbers $a$, which follow from the multiplicative relations given above. Then we give the elements of $C(F)$ corresponding to these cyclotomic equations, and $\mathbb{Z}$-linear relations (numerically determined in general) between the numbers $L(a^k)$, $k = 1, 2, \dots$, and $\pi^2/6$ which follow in the usual way.

We omit the cases $s = 3$ and $s = 7$ since then $a = \rho^2$, resp. $a = \rho^4$, and cyclotomic equations satisfied by these values of $a$ follow from those given in Example 11.4.



$s = 4$. $a = 2-\sqrt{3}$.

$$a^{-1}(1-a)^4(1-a^2)(1-a^4)^{-1} = 1, \quad a^4\wedge$$
$$a^{-1}(1-a)^5(1-a^2)^{-3}(1-a^3)^{-1}(1-a^6) = 1, \quad a^6\wedge$$
$$b_1 := 16[a] + 2[a^2] - [a^4] \in C(F),$$
$$b_2 := 30[a] - 9[a^2] - 2[a^3] + [a^6] \in C(F).$$

Using a calculator one can verify numerically that, to the accuracy employed,

$$L(b_1) = 5\cdot\pi^2/6, \qquad L(b_2) = 8\cdot\pi^2/6.$$

Both these equalities have been proved (see Lewin [**82**, formulas (91) and (92)]).

$s = 5$. $a = \frac{1}{2}(5-\sqrt{21})$.

$$a^{-1}(1-a)^7(1-a^2)^{-1}(1-a^3)^{-3}(1-a^6) = 1, \qquad a^6\wedge$$
$$b := 42[a] - 3[a^2] - 6[a^3] + [a^6] \in C(F),$$
$$L(b) = 10\cdot\pi^2/6 \quad \text{(numerically)}.$$

See Lewin [**84**, formula (73)].

$s = 6$. $a = 3-2\sqrt{2}$. Since $a = \tau^2$, for $\tau = \sqrt{2}-1$, we get a simpler multiplicative relation than that given above:

$$\tau\cdot c_2^2(a) = c_1^3(a),$$

and hence

$$\tau^{-1}(1-a)^5(1-a^2)^{-2} = 1.$$

Multiplying both sides of the last equality by $a\wedge$ and applying $a = \tau^2$ we get $5[a] - [a^2] \in C(F)$.

A calculation gives that, numerically, $5L(a) - L(a^2) = \pi^2/6$. This equality has been proved by Lewin [**82**, formula (97)].

$s = 8$. $a = 4-\sqrt{15}$.

$$a^{-1}(1-a)^{15}(1-a^2)^2(1-a^3)^{-3}(1-a^4)^{-2} = 1, \qquad a^2\wedge$$
$$b := 30[a] + 2[a^2] - 2[a^3] - [a^4] \in C(F),$$
$$L(b) = 5\cdot\pi^2/6 \quad \text{(numerically)}.$$

See Lewin [**84**, formula (76)].

$s = 10$. $a = 5-2\sqrt{6}$.

$$a^{-1}(1-a)^{23}(1-a^2)^{-15}(1-a^3)^{-3}(1-a^6)^3 = 1, \quad a^2\wedge$$
$$b := 46[a] - 15[a^2] - 2[a^3] + [a^6] \in C(F),$$
$$L(b) = 6\cdot\pi^2/6 \quad \text{(numerically)}.$$

See Lewin [**84**, formula (79)].

11.4.2.3. Algebraic numbers in the field $F = \mathbb{Q}(\sqrt{13})$. It is known that $\varepsilon = \frac{1}{2}(3+\sqrt{13})$ is the fundamental unit of the field $F$, the class number $h_F = 1$, 2 is prime in $F$, and 3 is decomposable: $3 = 3_1 \cdot 3_2$, where $3_1 = \frac{1}{2}(-1+\sqrt{13})$, $3_2 = \frac{1}{2}(1+\sqrt{13})$.

EXAMPLE 11.6. Let $a = \frac{1}{2}(-3+\sqrt{13}) = \varepsilon^{-1}$. One can easily verify that

$$c_1(a) = a\cdot 3_2,\qquad c_2(a) = 3_1,\qquad c_3(a) = a\cdot 2\cdot 3_2,$$
$$c_4(a) = a\cdot\sqrt{13},\qquad c_6(a) = a\cdot 2\cdot 3_1.$$

Consequently there is only one multiplicative relation between these numbers and $a$

$$c_6(a)\cdot c_1(a) = a\cdot c_3(a)\cdot c_2(a).$$

This leads to the unique cyclotomic equation satisfied by $a$

$$a^{-1}(1-a)^4(1-a^2)^{-2}(1-a^3)^{-2}(1-a^6) = 1.$$

Multiplying both sides of the equation by $a^6\wedge$ we get

$$b := 24[a] - 6[a^2] - 4[a^3] + [a^6] \in C(F).$$

One can compute numerically that

$$L(b) = 7\cdot\pi^2/6\quad \text{(cf. Lewin [84, equation (83)])}.$$

Before considering the next example we state the following obvious

LEMMA 11.14. *If $a \in F$ satisfies a cyclotomic equation*

$$a^A\prod_{r\in I}(1-a^r)^{A_r} = 1,$$

*then also the numbers $a^{-1}$, $-a$, and $\sigma(a)$, for every automorphism $\sigma$ of $F$, satisfy appropriate cyclotomic equations, namely:*

(i) $b = a^{-1}$ *satisfies* $b^B\prod_{r\in I}(1-b^r)^{2A_r} = 1$, *where* $B = -2(A+\sum_{r\in I} rA_r)$.

(ii) $c = -a$ *satisfies*

$$c^{2A}\prod_{\substack{r\in I,\\ r\text{ even}}}(1-c^r)^{2A_r}\cdot\prod_{\substack{r\in I,\\ r\text{ odd}}}\left(\frac{1-c^{2r}}{1-c^r}\right)^{2A_r} = 1.$$

(iii) $\sigma(a)$ *satisfies* $\sigma(a)^A\prod_{r\in I}(1-\sigma(a)^r)^{A_r} = 1$.

For every real number $a$, $a \neq 0, \pm 1$, exactly one of the elements of the set $V(a) := \{a, a^{-1}, -a, -a^{-1}\}$ belongs to the interval $(0, 1)$. For any $b \notin V(a)$ sets $V(a)$ and $V(b)$ are disjoint. Therefore if a real number $a$, $a \neq 0, \pm 1$, satisfies a cyclotomic equation, then we get a conjectural linear relation between the numbers $L(a_1^k)$, $k = 1, 2, \ldots$, and $\pi^2/6$, for $a_1 \in V(a)\cap(0, 1)$.

Moreover, for every automorphism $\sigma$ such that $\sigma(a)$ is real and $\sigma(a) \notin V(a)$, we get a linear relation between the numbers $L(a_2^k)$, $k = 1, 2, \ldots$, and $\pi^2/6$, for $a_2 \in V(\sigma(a))\cap(0, 1)$.



In particular, if $F$ is a real quadratic field, and $\sigma$ is its nontrivial automorphism, then, for $a \in F$, we have $\sigma(a) \in V(a)$ iff $a \in \mathbb{Q}$, $a^2 \in \mathbb{Q}$, or $Na = \pm 1$, where $N = N_{F/\mathbb{Q}}$ is the norm.

The numbers $a$ considered in Examples 11.4–11.6 satisfy $Na = \pm 1$. Therefore every cyclotomic equation satisfied by those $a$'s leads to only one linear relation between the numbers $L(a^k)$, $k = 1, 2, \ldots$, and $\pi^2/6$.

In the following example we have $\sigma(a) \notin V(a)$, and consequently a cyclotomic equation satisfied by $a$ will lead to two linear relations.

EXAMPLE 11.7. We use the notation of Example 11.6. Then $b := \frac{1}{2}(-1+\sqrt{13}) = 3_1$ satisfies $b^2+b-3 = 0$.

We have

$$c_1(b) = -\varepsilon^{-1},\qquad c_2(b) = 3_2,\qquad c_3(b) = 4,$$
$$c_4(b) = \tfrac{1}{2}(9-\sqrt{13}),\qquad Nc_4(b) = 17,\ c_6(b) = 2\cdot\varepsilon^{-1}\cdot 3_2.$$

There is only one multiplicative relation between these numbers: $c_6^2 = c_3c_2^2c_1^2$.

This leads to the cyclotomic equation

$$(1-b)^3(1-b^2)^{-4}(1-b^3)^{-3}(1-b^6)^2 = 1.$$

Unfortunately, $b$ does not belong to the interval $(0, 1)$, but $V(b)\cap(0, 1) = \{b^{-1}\}$. Therefore we apply Lemma 11.14 to get a cyclotomic equation satisfied by $c = b^{-1}$:

$$c^2(1-c)^3(1-c^2)^{-4}(1-c^3)^{-3}(1-c^6)^2 = 1,$$

where $c = b^{-1} = \frac{1}{6}(1+\sqrt{13})$. Multiplying both sides of the last cyclotomic equation by $c^3\wedge$ we obtain $9[c] - 6[c^2] - 3[c^3] + [c^6] \in C(F)$, and using a calculator we find numerically that

$$9L(c) - 6L(c^2) - 3L(c^3) + L(c^6) = 2\cdot\pi^2/6$$

(see Lewin and Abouzahra [90, formula (3.8)]).

Since $c' := \sigma(c) = \frac{1}{6}(1-\sqrt{13}) \notin V(c)$ and $V(c')\cap(0, 1) = \{-c'\}$, from the above cyclotomic equation satisfied by $c$, using Lemma 11.14, we get a cyclotomic equation satisfied by $d = -c'$:

$$d^{-2}(1-d)^3(1-d^2)(1-d^3)^{-3}(1-d^6) = 1.$$

Multiplying both sides of this equation by $d^6\wedge$ we obtain

$$18[d] + 3[d^2] - 6[d^3] + [d^6] \in C(F),$$

and using a calculator we find numerically that

$$18L(d) + 3L(d^2) - 6L(d^3) + L(d^6) = 8\cdot\pi^2/6,$$

(see Lewin [90, formula (3.7)]).

11.4.2.4. Algebraic numbers in the field $F = \mathbb{Q}(\sqrt{33})$. In $F = \mathbb{Q}(\sqrt{33})$ the class number $h_F = 1$, and the fundamental unit $\varepsilon = 23 + 4\sqrt{33}$ satisfies $N\varepsilon = 1$.

Denote

$$\begin{aligned} u &= \tfrac{1}{2}(5+\sqrt{33}), && \text{then } Nu = -2, \\ v &= 6+\sqrt{33}, && \text{then } Nv = 3, \\ u_1 &= \tfrac{1}{2}(1+\sqrt{33}) = u^3\varepsilon^{-1}, \\ u_2 &= \tfrac{1}{2}(-7-\sqrt{33}) = -\bar{u}^2\varepsilon, \\ u_3 &= \tfrac{1}{2}(17+\sqrt{33}) = u_1^2 = u^6\varepsilon^{-2}. \end{aligned}$$

We have $2 = -u\bar{u}$, $3 = v^2\varepsilon^{-1}$, i.e., $v = \bar{v}\varepsilon$.

EXAMPLE 11.8. Let $S = \{\infty_1, \infty_2, u, \bar{u}, v\}$. The following elements are good $S$-units (the list may be not complete):

$$\begin{aligned} -\varepsilon, &\quad \text{since } 1+\varepsilon = 24+4\sqrt{33} = u^2\bar{u}^2 v, \\ u, &\quad \text{since } 1-u = \tfrac{1}{2}(-3-\sqrt{33}) = \bar{u}v, \\ -u, &\quad \text{since } 1+u = \tfrac{1}{2}(7+\sqrt{33}) = \bar{u}^2\varepsilon, \\ v, &\quad \text{since } 1-v = -5-\sqrt{33} = u^2\bar{u}, \\ -v, &\quad \text{since } 1+v = 7+\sqrt{33} = -u\bar{u}^3\varepsilon, \\ u_1, &\quad \text{since } 1-u_1 = \tfrac{1}{2}(1-\sqrt{33}) = \bar{u}^3\varepsilon, \\ -u_1, &\quad \text{since } 1+u_1 = \tfrac{1}{2}(3+\sqrt{33}) = -\bar{u}v, \\ u_2, &\quad \text{since } 1-u_2 = \tfrac{1}{2}(9+\sqrt{33}) = u^2v\varepsilon^{-1}, \\ u_3, &\quad \text{since } 1-u_3 = \tfrac{1}{2}(-15-\sqrt{33}) = -\bar{u}^4v\varepsilon. \end{aligned}$$

The following good $S$-units are equivalent to the above under the action of $S_3$, but belong to the interval $(0, 1)$:

$$\begin{aligned} a_1 &:= (1+\varepsilon)^{-1} = \frac{6-\sqrt{33}}{12} = (u^2\bar{u}^2v)^{-1}, && 1-a_1 = \varepsilon\cdot(u^2\bar{u}^2v)^{-1}, \\ a_2 &:= u^{-1} = \frac{-5+\sqrt{33}}{4}, && 1-a_2 = -\bar{u}v\cdot u^{-1}, \\ a_3 &:= (1+u)^{-1} = \frac{7-\sqrt{33}}{8} = (\bar{u}^2\varepsilon)^{-1}, && 1-a_3 = u\cdot(\bar{u}^2\varepsilon)^{-1}, \\ a_4 &:= v^{-1} = \frac{6-\sqrt{33}}{3}, && 1-a_4 = -u^2\bar{u}\cdot v^{-1}, \\ a_5 &:= (1+v)^{-1} = \frac{7-\sqrt{33}}{16} = -(u\bar{u}^3\varepsilon)^{-1}, && 1-a_5 = -v\cdot(u\bar{u}^3\varepsilon)^{-1}, \\ a_6 &:= u_1^{-1} = \frac{-1+\sqrt{33}}{6} = \varepsilon\cdot u^{-3}, && 1-a_6 = -\bar{u}^3\varepsilon^2\cdot u^{-3}, \end{aligned}$$



$$\begin{aligned} a_7 &:= (1+u_1)^{-1} = \frac{-3+\sqrt{33}}{12} = -(\bar{u}v)^{-1}, && 1-a_7 = -u^3(\bar{u}v\varepsilon)^{-1}, \\ a_8 &:= (1-u_2)^{-1} = \frac{9-\sqrt{33}}{24} = \varepsilon\cdot(u^2v)^{-1}, && 1-a_8 = \bar{u}^2\varepsilon^2\cdot(u^2v)^{-1}, \\ a_9 &:= u_3^{-1} = \frac{17-\sqrt{33}}{128} = \varepsilon^2\cdot u^{-6}, && 1-a_9 = \bar{u}^4v\varepsilon^3\cdot u^{-6}. \end{aligned}$$

Applying these decompositions it is easy to express elements $\lambda([a_j])$, $1 \le j \le 9$, in the following form:

$$\begin{aligned} \lambda([a_j]) = {} & c_1(\varepsilon\wedge u) + c_2(\varepsilon\wedge\bar{u}) + c_3(\varepsilon\wedge v) + c_4(u\wedge\bar{u}) \\ & + c_5(u\wedge v) + c_6(\bar{u}\wedge v) + c((-1)\wedge x). \end{aligned}$$

The corresponding coefficients $c_1, c_2, \ldots, c_6, c$, and element $x$ are given in the Table 11.2.

| $j$ | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ | $c_6$ | $c$ | $x$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 2 | 1 | 0 | 0 | 0 | 1 | $v$ |
| 2 | 0 | 0 | 0 | −1 | −1 | 0 | 0 | – |
| 3 | −1 | 0 | 0 | 2 | 0 | 0 | 1 | $\varepsilon$ |
| 4 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | – |
| 5 | 0 | 0 | −1 | 0 | −1 | −3 | 1 | $-u\bar{u}\varepsilon v$ |
| 6 | 3 | 3 | 0 | 9 | 0 | 0 | 1 | $\varepsilon$ |
| 7 | 0 | −1 | −1 | 3 | 3 | 0 | 1 | $-u\bar{u}\varepsilon v$ |
| 8 | 2 | 2 | 1 | −4 | 0 | 2 | 1 | $v$ |
| 9 | 6 | 8 | 2 | −24 | −6 | 0 | 0 | – |

TABLE 11.2

From Table 11.2 it follows that $\lambda([a_j])$, for $1 \le j \le 6$, are linearly independent, and $\lambda([a_j])$, for $j = 7, 8, 9$, are their linear combinations:

$$\begin{aligned} \lambda([a_8]) &= \lambda([a_1]) + 4\lambda([a_2]) + 2\lambda([a_4]), \\ 11\lambda([a_7]) &= -10\lambda([a_1]) - 28\lambda([a_2]) - 11\lambda([a_3]) + 3\lambda([a_4]) \\ &\quad + \lambda([a_5]) + 3\lambda([a_6]), \\ 11\lambda([a_9]) &= 56\lambda([a_1]) + 236\lambda([a_2]) + 22\lambda([a_3]) + 102\lambda([a_4]) \\ &\quad + 34\lambda([a_5]) - 8\lambda([a_6]). \end{aligned}$$

One can compute numerically that

$$\begin{aligned} & 11L(a_7) + 10L(a_1) + 28L(a_2) + 11L(a_3) - 3L(a_4) - L(a_5) \\ & - 3L(a_6) = 9\cdot\pi^2/6, \end{aligned}$$

$$L(a_8) - L(a_1) - 4L(a_2) - 2L(a_4) = -\pi^2/6,$$
$$11L(a_9) - 56L(a_1) - 236L(a_2) - 22L(a_3) - 102L(a_4)$$
$$- 34L(a_5) + 8L(a_6) = -68 \cdot \pi^2/6.$$

Before considering the next example we prove the following general lemma on cyclotomic equations with a factor.

LEMMA 11.15. *Suppose that numbers $a$ and $b$ satisfy the following cyclotomic equation with a factor*

$$c^\alpha a^A \prod_{r\in I}(1 - a^r)^{A_r} = 1,$$
$$c^\beta b^B \prod_{s\in J}(1 - b^s)^{B_s} = 1,$$

*for some $A, B, A_r, B_s, \alpha, \beta \in Z$, $\alpha\beta \neq 0$, finite sets $I, J$ of natural numbers, and a complex number $c$. Suppose moreover that $a^u b^v c^w = 1$, for some nonzero integers $u, v, w$, and let $w\alpha\beta$ be even.*

*Then some nontrivial $\mathbb{Z}$-linear combination of elements $[a^r]$ and $[b^s]$ belongs to $C(F)$, where $r \in I$, $s \in J$.*

PROOF. From given cyclotomic equations with a factor it follows that, for natural numbers $M$ and $N$ satisfying $2|AM$, $2|BN$, $r|MA_r$, for $r \in I$, $s|NA_s$, for $s \in J$, we have

$$a^M \wedge c^{-\alpha} = \sum_{r\in I} k_r[a^r], \quad \text{where } k_r = \frac{MA_r}{r} \in \mathbb{Z},$$
$$b^N \wedge c^{-\beta} = \sum_{s\in J} l_s[b^s], \quad \text{where } l_s = \frac{NB_s}{s} \in \mathbb{Z}.$$

Consequently in view of $a^u b^v c^w = 1$ we get

$$N\beta u(a^M \wedge c^{-\alpha}) + M\alpha v(b^N \wedge c^{-\beta})$$
$$= -MN\alpha\beta(a^u b^v \wedge c) = MN\alpha\beta(c^w \wedge c)$$
$$= MN\alpha\beta w((-1) \wedge c) = 0,$$

since $\alpha\beta w$ is even. □

EXAMPLE 11.9. One can verify that the numbers

$$a = \tfrac{1}{8}(7 - \sqrt{33}) = \bar{u}^2 \cdot \varepsilon^{-1}, \qquad b = \tfrac{1}{2}(-5 + \sqrt{33}) = -\bar{u},$$

in the notation of Example 11.8, satisfy cyclotomic equations with a factor:

$$\varepsilon^2 a^5 (1 - a)^{-10}(1 - a^2)^4(1 - a^3)(1 - a^6)^{-1} = 1, \tag{11.40}$$

and

$$\varepsilon(1 - b)^{11}(1 - b^2)^{-8}(1 - b^3)^{-2}(1 - b^6)^2 = 1, \tag{11.41}$$

where $\varepsilon = 23 + 4\sqrt{33}$ is the fundamental unit of the field $F = \mathbb{Q}(\sqrt{33})$. Moreover we have $ab^2\varepsilon = 1$.



Therefore the assumptions of Lemma 11.15 are satisfied, and we obtain an element of $C(F)$ as follows. Multiplying both sides of (11.40) and (11.41) by $a^6\wedge$ and $b^3\wedge$ respectively, we get

$$\lambda([a^6] + 60[a] - 2[a^3] - 12[a^2]) = 12(a \wedge \varepsilon),$$

and

$$\lambda([b^6] + 33[b] - 2[b^3] - 12[b^2]) = -3(b \wedge \varepsilon).$$

Since $a = b^{-2} \cdot \varepsilon^{-1}$, we have $12(a \wedge \varepsilon) = -24(b \wedge \varepsilon)$.

Consequently

$$d := 8[b^6] + 264[b] - 16[b^3] - 96[b^2] - [a^6] - 60[a]$$
$$+ 2[a^3] + 12[a^2] \in C(F).$$

One can compute numerically that $L(d) = 74 \cdot \pi^2/6$.

11.4.3. *Linear relations between the values of $D(z)$ at imaginary quadratic arguments.* As we have remarked earlier every $\mathbb{Q}$-linear relation between the values of $D(z)$ is equivalent to a corresponding $\mathbb{Q}$-linear relation between the values of $\mathrm{Cl}_2(t)$. Therefore we shall use both functions equivalently.

Let $a$ be a root of the quadratic polynomial $x^2 - sx + 1$, where $s$ is a rational number, $0 < s < 2$. Let $s^2 - 4 = dr^2$, where $d$ is a squarefree integer, and $r \in \mathbb{Q}$. Since $0 < s < 2$, we have $d < 0$, i.e.,

$$a = \tfrac{1}{2}(s + r\sqrt{d}) \tag{11.42}$$

is an imaginary quadratic number. Let $F = \mathbb{Q}(a) = \mathbb{Q}(\sqrt{d})$.

To find cyclotomic equations satisfied by $a$ we apply the same method as in 11.4.2.2. In the present case there are known many more values of $s$ such that the numbers (11.39) are multiplicatively dependent. E.g. it is so for $s = m/n$, where $n = 2, 3, 4, 5$, $1 \leq m < 2n$, $(m, n) = 1$, and moreover for $s = 2/7, 4/7, 10/7$, and $11/7$.

We shall give below only one example of this kind. Others are similar and can be found in Boldy [89] and Browkin [89].

EXAMPLE 11.10. Let $s = 3/5$. Then $a = \frac{1}{10}(3 + \sqrt{-91})$. One can easily verify that

$$c_3(a) = a(s + 1) = a \cdot \frac{8}{5},$$
$$c_5(a) = a^2(s^2 + s - 1) = a^2 \cdot \frac{-1}{25},$$
$$c_6(a) = a(s - 1) = a \cdot \frac{-2}{5}.$$

Hence we get a multiplicative relation $c_6^3(a) = c_5(a)c_3(a)$, and the corresponding cyclotomic equation

$$(1 - a^6)^3(1 - a^5)^{-2}(1 - a^3)^{-4}(1 - a^2)^{-3}(1 - a)^5 = 1.$$

Multiplying both sides of this equation by $a^{30}\wedge$ we obtain an element of $C(F)$:

$$b := 15[a^6] - 6[a^5] - 40[a^3] - 45[a^2] + 150[a] \in C(F).$$

Since $|a| = 1$, and $t := \arg a = \arccos \frac{3}{10}$, from (11.24) we get $D(a^k) = \mathrm{Cl}_2(kt)$, for $k \in Z$.

Consequently

$$D(b) = 15\,\mathrm{Cl}_2(6t) - 6\,\mathrm{Cl}_2(5t) - 40\,\mathrm{Cl}_2(3t) - 45\,\mathrm{Cl}_2(2t) + 150\,\mathrm{Cl}_2(t).$$

Using tables (e.g. Ashour and Sabri [56]) we calculate numerically that $D(b) = 162.17617\ldots$.

On the other hand, for $F = \mathbf{Q}(\sqrt{d}) = \mathbf{Q}(\sqrt{-91})$, in view of formula (11.30), we have

$$\frac{1}{2}|d|^{3/2}L(2,\chi) = 91 \sum_{1\le j\le 45} \left(\frac{-91}{j}\right)\mathrm{Cl}_2\left(\frac{2\pi j}{91}\right) = 324.35236\ldots.$$

Consequently formula (11.35) gives the conjectural value of order of the group $K_2O_F$: $\# K_2O_F = 2$.

**Example** 11.11. Let $a = \frac{1}{2}(1+\sqrt{-7})$, $F = \mathbf{Q}(\sqrt{-7})$, $d = -7$. Since $a$ is a quadratic algebraic integer it follows from the paper of Rédei [58] that $a$ can be exceptional of orders $\le 4$ and 6 only. Therefore to find all cyclotomic equations satisfied by $a$ it is sufficient to find all multiplicative relations between the numbers

$$(11.43)\qquad \begin{array}{lll} c_1(a) = \bar{a}, & c_2(a) = -\bar{a}^2, & c_3(a) = \sqrt{-7}, \\ c_4(a) = -\bar{a}, & c_6(a) = -1. & \end{array}$$

There are only three independent multiplicative relations between these numbers (and $-1$):

$$c_1^2(a) = -c_2(a), \qquad c_4(a) = -c_1(a), \qquad c_6(a) = -1.$$

They lead to the cyclotomic equations:

$$\begin{aligned} &(1-a^2)(1-a)^{-3} = -1, \\ &(1-a^4)(1-a^2)^{-1}(1-a)^{-1} = -1, \\ &(1-a^6)(1-a^3)^{-1}(1-a^2)^{-1}(1-a) = -1. \end{aligned}$$

Multiplying them by $a^2\wedge$, $a^4\wedge$, and $a^6\wedge$ respectively we get elements of $C(F)$:

$$\begin{aligned} b_1 &:= [a^2] - 6[a] \in C(F), \\ b_2 &:= [a^4] - 2[a^2] - 4[a] \in C(F), \\ b_3 &:= [a^6] - 2[a^3] - 3[a^2] + 6[a] \in C(F). \end{aligned}$$

One can compute numerically that $D(b_1) = -D_F$, $D(b_2) = -D_F$, $D(b_3) = D_F$, where $D_F = 5.33348\ldots$.



On the other hand

$$\tfrac{1}{2}|d|^{3/2}L(2,\chi) = 7\cdot \sum_{1\le j\le 3}\left(\frac{-7}{j}\right)\mathrm{Cl}_2\left(\frac{2\pi j}{7}\right) = 10.66696\ldots.$$

Moreover it is known (see Tate [73]) that $\# K_2 0_F = 2$. Thus this example confirms the formula (11.35).

In fact we can prove that $D(b_1) = D(b_2) = -D(b_3)$. In view of (11.43) we have

$$1 + a^3 = c_2(a)\cdot c_6(a) = \bar{a}^2, \quad \text{and} \quad 1 + a^2 = c_4(a) = -\bar{a}.$$

Consequently

$$D(-a^3) = -D(1+a^3) = -D(\bar{a}^2) = D(a^2),$$

and

$$D(-a^2) = -D(1+a^2) = -D(-\bar{a}) = D(-a).$$

Therefore (see Lemma 11.10) we get

$$\begin{aligned} D(b_1) + D(b_3) &= (2D(-a) - 4D(a)) + (2D(-a^3) - 6D(-a)) \\ &= 2D(a^2) - 4D(a) - 4D(-a) = 0, \\ D(b_1) - D(b_2) &= (2D(-a) - 4D(a)) - (2D(-a^2) - 4D(a)) \\ &= 2D(-a) - 2D(-a^2) = 0. \end{aligned}$$

Now we give an example of an imaginary quadratic number satisfying generalized cyclotomic equations.

**Example** 11.12. Let $s = \frac{13}{7}$, then from (11.42) we get $a = \frac{1}{14}(13+3\sqrt{-3})$. One can verify that

$$\begin{aligned} c_2(a) &= 1 + a = 3\sqrt{-3}\cdot c_1(a), \\ c_3''(a) &= 1 - \zeta_3^2 a = 4\zeta_3\cdot c_1(a), \\ c_6'(a) &= 1 + \zeta_3^2 a = -2\sqrt{-3}\zeta_3\cdot c_1(a), \\ c_6''(a) &= 1 + \zeta_3 a = -\sqrt{-3}\zeta_3^2 c_1(a). \end{aligned}$$

Hence we get the following multiplicative relations:

$$c_6''(a)^3 = c_2(a)\cdot c_1(a)^2, \qquad c_6'(a)^4 = c_2(a)\cdot c_3''(a)^2\cdot c_6''(a),$$

i.e., we have the generalized cyclotomic equations:

$$\begin{aligned} &(1+\zeta_3 a)^3 = (1-a)^2(1+a), \\ &(1+\zeta_3^2 a)^4 = (1+a)(1-\zeta_3^2 a)^2(1+\zeta_3 a). \end{aligned}$$

Multiplying both sides of them by $(-a)\wedge$ and $(-a^3)\wedge$ respectively we get elements of $C(F)$:

$$\begin{aligned} b_1 &:= 3[-\zeta_3 a] - 2[a] - [-a] \in C(F), \\ b_2 &:= 12[-\zeta_3^2 a] - 3[-a] - 6[\zeta_3^2 a] - 3[-\zeta_3 a] \in C(F). \end{aligned}$$

Since $|a| = 1$ and $\arg a := t = \arccos\frac{13}{14}$, we have

$$D(b_1) = 3\,\mathrm{Cl}_2(t - \pi/3) - 2\,\mathrm{Cl}_2(t) + \mathrm{Cl}_2(\pi - t),$$

and

$$\begin{aligned} D(b_2) = {} & 12\,\mathrm{Cl}_2(t + \pi/3) - 3\,\mathrm{Cl}_2(t + \pi) - 6\,\mathrm{Cl}_2(t + 4\pi/3) \\ & - 3\,\mathrm{Cl}_2(t - \pi/3) \end{aligned} \tag{11.44}$$

One can compute numerically that

$$D(b_1) = -6\,\mathrm{Cl}_2(2\pi/3), \qquad D(b_2) = 30\,\mathrm{Cl}_2(2\pi/3).$$

Let us remark that after the substitution $t' = t + \pi/3$, in view of the triplication formula, (11.44) can be written in the form

$$D(b_2) = 15\,\mathrm{Cl}_2(t') + 6\,\mathrm{Cl}_2(\pi - t') - \mathrm{Cl}_2(3t').$$

Since $\cos t' = \cos(t + \pi/3) = 1/7$, we get the formula (7.45) of Browkin [89].

11.4.4. *Linear relations between the numbers* $\mathrm{Cl}_2(k\pi/m)$. We shall describe a method which gives some $\mathbf{Z}$-linear relations between the numbers $\mathrm{Cl}_2(k\pi/m)$, where $k \in \mathbf{Z}$, and $m$ is a fixed natural number. All these relations will be deduced from the multiplication formula:

$$\frac{1}{n}\mathrm{Cl}_2(nt) = \sum_{j=0}^{n-1}\mathrm{Cl}_2(t + 2\pi j/n). \tag{11.45}$$

For a fixed natural number $m$ we denote the sum

$$\sum_{j=1}^{r}\mathrm{Cl}_2(k_j\pi/m), \qquad k_j \in \mathbf{Z},$$

simply by $(k_1, k_2, \dots, k_r)$. We say that two such sums $(k_1, k_2, \dots, k_r)$ and $(l_1, l_2, \dots, l_s)$ are congruent (modulo the terms with the denominators less than $m$), if after omitting the terms with $(k_i, m) > 1$ and $(l_j, m) > 1$ the remaining terms in both sums are the same up to a permutation. We write then

$$(k_1, k_2, \dots, k_r) \equiv (l_1, l_2, \dots, l_s).$$

In particular, formula (11.45), for $n = m$ and $t = \pi/m$ gives $(1, 3, 5, \dots, 2m - 1) \equiv 0$.

If $m$ is even, for $n = 2$, $t = k\pi/m$, it follows from (11.45) that $(k, k + m) \equiv 0$, i.e., $(k) \equiv (m - k)$.

Let $p$ be a prime dividing $m$. Then, for $n = p$, $t = k\pi/m$, $1 \le k \le m/p$, the formula (11.45) takes the form

$$(k, k + 2m/p, k + 4m/p, \dots, k + (p-1)2m/p) \equiv 0 \tag{11.46}$$

In the left-hand side of the last congruence we can cancel all terms $k + j2m/p$ which are not prime to $m$. Thus we get a set of relations we were looking for.



We now explain the method on several examples. If $m$ is a prime power, we do not obtain anything more than the multiplication by $p$ formula. Therefore we shall consider only $m$'s which are not prime powers.

EXAMPLE 11.13. $m = 6$. Formula (11.46), for $p = 3$, gives $(1, 5, 9) \equiv 0$ i.e., $(1, 5) \equiv 0$, and we get the triplication formula:

$$\begin{aligned} \mathrm{Cl}_2(\pi/6) + \mathrm{Cl}_2(5\pi/6) &= \tfrac{1}{3}\,\mathrm{Cl}_2(\pi/2) - \mathrm{Cl}_2(3\pi/2) \\ &= \tfrac{4}{3}\,\mathrm{Cl}_2(\pi/2). \end{aligned}$$

(See Corollary 11.1.(ii)).

EXAMPLE 11.14. $m = 12$. The formula (11.46), for $p = 3$, gives

$$(1, 9, 17) \equiv (3, 11, 19) \equiv 0, \quad \text{i.e., } (1, 17) \equiv (11, 19) \equiv 0.$$

Hence in view of $(k) \equiv (m - k)$ we get

$$(1, -5) \equiv (1, -7) \equiv 0.$$

Performing necessary calculations we get once more formula (11.21), and

$$\mathrm{Cl}_2(\pi/12) - \mathrm{Cl}_2(7\pi/12) = \tfrac{1}{2}\,\mathrm{Cl}_2(\pi/2) - \tfrac{3}{2}\,\mathrm{Cl}_2(\pi/4).$$

EXAMPLE 11.15. $m = 24$. The formula (11.46), for $p = 3$, gives

$$(1, 17, 33) \equiv (3, 19, 35) \equiv 0, \quad \text{i.e., } (1, 7) \equiv (5, -11) \equiv 0.$$

Hence we get two relations

$$\mathrm{Cl}_2(\pi/24) + \mathrm{Cl}_2(7\pi/24) = \tfrac{1}{3}\,\mathrm{Cl}_2(\pi/8) + \mathrm{Cl}_2(5\pi/8) + \tfrac{1}{2}\,\mathrm{Cl}_2(7\pi/12),$$

and

$$\mathrm{Cl}_2(5\pi/24) - \mathrm{Cl}_2(11\pi/24) = \tfrac{1}{3}\,\mathrm{Cl}_2(5\pi/8) - \mathrm{Cl}_2(7\pi/8).$$

The second equation can be transformed to

$$\begin{aligned} \mathrm{Cl}_2(5\pi/24) - \mathrm{Cl}_2(11\pi/24) = {} & \tfrac{1}{12}\,\mathrm{Cl}_2(\pi/2) + \tfrac{1}{3}\,\mathrm{Cl}_2(\pi/4) \\ & - \mathrm{Cl}_2(\pi/8) + \tfrac{1}{3}\,\mathrm{Cl}_2(3\pi/8), \end{aligned}$$

where all arguments belong to the interval $\langle 0, \pi/2\rangle$.

EXAMPLE 11.16. $m = 14$. We have analogously, for $p = 7$,

$$(1, 5, 9, 13, 17, 21, 25) \equiv 0, \quad \text{i.e., } (1, 5, 5, 1, -3, -3) \equiv 0,$$

and hence $(1, -3, 5) \equiv 0$.

This leads to the relation

$$\begin{aligned} & \mathrm{Cl}_2(\pi/14) - \mathrm{Cl}_2(3\pi/14) + \mathrm{Cl}_2(5\pi/14) \\ & \quad = \tfrac{1}{4}(\mathrm{Cl}_2(\pi/7) - \mathrm{Cl}_2(3\pi/7) + \mathrm{Cl}_2(5\pi/7) + \tfrac{4}{7}\mathrm{Cl}_2(\pi/2)). \end{aligned}$$

EXAMPLE 11.17. $m = 15$. From (11.46) we get for $p = 3$,

$$(1, 11, -9) \equiv (2, 12, -8) \text{ i.e. } (1, 4) \equiv (2, -7) \equiv 0,$$

and, for $p = 5$,

$$(1, 7, 13, -11, -5) \equiv 0, \quad \text{i.e., } (1, 7, 2, -4) \equiv 0.$$

Consequently $(1, 7) \equiv (2, -4) \equiv 0$. Hence we get the following relations:

$$\mathrm{Cl}_2(\pi/15) + \mathrm{Cl}_2(7\pi/15) = \tfrac{2}{3}\mathrm{Cl}_2(\pi/5) + \tfrac{1}{3}\,\mathrm{Cl}_2(3\pi/5) + \tfrac{3}{5}\,\mathrm{Cl}_2(\pi/3),$$

$$\mathrm{Cl}_2(2\pi/15) - \mathrm{Cl}_2(4\pi/15) = -\tfrac{1}{3}\,\mathrm{Cl}_2(2\pi/5) - \tfrac{2}{3}\,\mathrm{Cl}_2(4\pi/5) + \tfrac{3}{5}\,\mathrm{Cl}_2(2\pi/3).$$

**11.5. Problems and conjectures.**

11.5.1. *The field of rational numbers* $\mathbb{Q}$. Let $S = \{\infty, 2, 3, \dots, p\}$ be the set of first $n$ prime numbers and $\infty$, i.e., $p = p_n$, and let $a_1, a_2, \dots, a_t \in \mathbb{Q}$ be a fixed minimal set of good $S$-units.

(a) For $a \in W_S \cap T$, prove that $\lambda([a])$ has a finite order iff $a = 1/2$.

(b) Is the index $(U_S \wedge U_S : \lambda(A(W_S)))$ finite? Equivalently, are free ranks of both groups equal, i.e. does $t - 1 = \binom{n}{2}$ hold?

(c) Does $\#(K_2\mathbb{Q})_S = (U_S \wedge U_S : \lambda(A(W_S)))$ hold? Equivalently, does $\lambda(A(W_S)) = \operatorname{Ker}\mathscr{K}$ hold?

It is known (see Milnor [71, Theorem 11.6]) that

$$\#(K_2\mathbb{Q})_S = 2\prod_{j=2}^{n}(p_j - 1).$$

(d) Does every good $S$-unit depend on the set $a_1, a_2, \dots, a_t$?

(e) For $1 \le k \le p/2$, let $q_k = k/(p-k)$. Evidently every $q_k$ is a good $S$-unit. Is it true that exactly $n - 1$ numbers among $q_k$ are multiplicatively independent?

(f) Is every number $L(b)$, for $b \in C(\mathbb{Q})$, an integral multiple of $\pi^2/6$, provided $L(b)$ is defined?

REMARKS. 1. Problem 1(a) is an easy exercise.

2. The answer to problems (b)–(f) is unknown. We conjecture that it is always in the affirmative.

3. We have verified that the answer to problem (d) is in the affirmative, for $S = \{\infty, 2, 3, 5, 7\}$. One can in principle extend this to $S = \{\infty, 2, 3, 5, 7, 11, 13\}$ applying results of de Weger [**87, 89**].

4. In problem (e) it is evident that the number of multiplicatively independent numbers $q_k$ does not exceed $n - 1$, since every $q_k$ has the form

$$q_k = p_1^{m_1} \cdots p_{n-1}^{m_{n-1}}, \quad \text{for some } m_1, \dots, m_{n-1} \in \mathbf{Z}.$$

We have verified (e), for all primes $p \le 47$ and for $p = 101$.

5. From the positive answer to the problem (e) follows the positive answer to problem (b). One can construct inductively a free basis of $\lambda(A(W_S))$ adding in the $n$th step the elements $\lambda([k_j/p])$, $(j = 1, \dots, n-1)$ such that the elements $q_{k_j}$ defined above are multiplicatively independent, for $j = 1, \dots, n-1$.

Since $\lambda([k_j/p]) = (k_j/p) \wedge (1 - k_j/p) = p \wedge q_{k_j} +$ (the sum of elements of the form $(-1) \wedge p$ and $p_k \wedge p_m$ with $k, m < n$), these elements $\lambda([k_j/p])$, for $j = 1, \dots, n-1$, are linearly independent.

6. One can consider analogous problems for any number field $F$ in place of $\mathbb{Q}$.

11.5.2. *Real quadratic fields.* (a) Let $F$ be a real quadratic field. Is it true that, for every $b \in C(F)$,

$$L(b) = r\frac{24}{w_2(F)}\frac{\pi^2}{6} = r\frac{4\pi^2}{w_2(F)}$$

holds, provided $L(b)$ is defined?



(b) Are the numerically-verified equalities obtained in Examples 11.5–11.9 true?

(c) Does there exist a real quadratic number $a$ which satisfies a cyclotomic equation, but the set $V(a) = \{a, -a, 1/a, -1/a\}$ does not contain any algebraic integer?

(d) For a number field $F$ let $C'(F)$ be the set of all $b \in C(F)$ such that, for every imbedding $\sigma : F \to \mathbf{R}$, $L(\sigma(b))$ is defined. Define the mapping $\mathbf{L} : C'(F) \to \mathbf{R}^{r_1}$, by the formula

$$\mathbf{L}(b) = (L(\sigma_1(b)), \dots, L(\sigma_{r_1}(b))),$$

where $\sigma_1, \dots, \sigma_{r_1}$ are all real places of $F$.

Is it true that $\mathbf{L}(C'(F))$ is a lattice in $\mathbf{R}^{r_1}$ of rank $r_1$? If so, is it possible to interpret the volume of $\mathbf{R}^{r_1}/\mathbf{L}(C'(F))$ in a similar way as the volume of $\mathbf{R}^{r_2}/\mathbf{D}(C(F))$ was interpreted in 11.3?

REMARKS. 1. The answer to problems 2(a)–2(d) is unknown. We conjecture that it is always in the affirmative (but cf. Remark 3. below).

2. Examples 11.4–11.9 confirm 2(a). Moreover, for $F = \mathbb{Q}(\sqrt{2})$, we have $w_2(F) = 48$, and

$$\lambda([\tfrac{1}{2}]) = \tfrac{1}{2} \wedge \tfrac{1}{2} = 2 \wedge 2 = (-1) \wedge 2 = 2((-1) \wedge \sqrt{2}) = 0.$$

Consequently $[\frac{1}{2}] \in C(F)$. From Lemma 11.4(i) it follows that $L(1/2) = \pi^2/12$. On the other hand $4\pi^2/w_2(F) = \pi^2/12$, since $w_2(F) = 48$. Thus this example also confirms 2(a).

3. It seems, however, that for the cubic field $F = \mathbb{Q}(\omega)$, where $\omega$ is the real root of $x^3 + x^2 - 1 = 0$, the set $\mathbf{L}(C'(F))$ is a lattice of rank 2 (see Lewin [**84**]). Here we have $r_1 = r_2 = 1$.

11.5.3. *Imaginary quadratic fields.* (a) Estimate from above and from below the volume $D_F$, where $F$ is an imaginary quadratic field of discriminant $d$; in particular, whether $\lim_{d\to-\infty} D_F = \infty$ holds.

(b) Is it true that $D_F \le 1$ iff $F$ is totally real?

(c) Prove the numerically-verified equalities given in Examples 11.10–11.12, and in the paper of Browkin [**89**].

REMARKS. 1. There is a geometric interpretation of the number $D(z)$, and consequently of $\mathrm{Cl}_2(t) = D(e^{it})$. Namely, for a complex number $z$, the volume of the asymptotic simplex with vertices $0, 1, z$ and $\infty$ in 3-dimensional hyperbolic space is equal to $|D(z)|$. Some formulas for the dilogarithm of Bloch can be interpreted as relations between volumes of the corresponding simplices. One may expect that some of the numerically-verified equalities for $D(z)$ or $\mathrm{Cl}_2(t)$ can be proved using this geometric interpretation.

For more details, see Milnor [**79, 83**], Dupont and Sah [**82**], Zagier [**86, 88, 90**].

11.5.4. *Clausen's function.* We state the following conjecture of Milnor (see Milnor [79, 83]): Every $\mathbf{Z}$-linear relation between the numbers $\mathrm{Cl}_2(j\pi/m)$ follows from the multiplication formula (11.45), or equivalently: For fixed number $m$, the numbers $\mathrm{Cl}_2(j\pi/m)$, where $0 \le j \le m/2$, $(j, m) = 1$, are linearly independent over $\mathbb{Q}$.

## REFERENCES


L. J. Alex, [76] *Diophantine equations related to finite groups*, Comm. Algebra **4** (1976), 77–100.

A. Ashour and A. Sabri, [56] *Tabulation of the function* $\psi(\theta) = \sum_{n=1}^{\infty} \sin n\theta/n^2$, Math. Tables and Other Aids of Computation **10** (1956), 57–65.

S. Bloch, [78] *Application of the dilogarithm function in algebraic K-theory and algebraic geometry*, Proc. of the Internat. Sympos. on Algebraic geometry, Kyoto, 1977 (M. Nagata, ed.), Kinokuniya Book-Store Co., Tokyo, 1978, pp. 103–114.

M. C. Boldy, [89] *The Bloch group and it's relation to K-theory*, Report 8917, Catholic University, Nijmegen, 1989.

A. Borel, [72] *Cohomologie réelle stable de groupes S-arithmétiques classiques*, C. R. Acad. Sci. Paris Sér. I Math. **274** (1972), 1700–1703.

Z. I. Borevich and I. R. Shafarevich, [66] *Number theory*, Academic Press, New York, 1966.

J. Browkin, [89] *Conjectures on the dilogarithm*, $K$-theory **3** (1989), 29–56.

——, [90] *Cyclotomic equations satisfied by quadratic integers* (to appear).

Th. Clausen, [32] *Über die Function* $\sin\varphi + (1/2^2)\sin 2\varphi + (1/3^2)\sin 3\varphi +$ *etc.*, J. Reine Angew. Math. **8** (1832), 298–300.

J. L. Dupont and C.-H. Sah, [82] *Scissors congruences*. II, J. Pure Appl. Algebra **25** (1982), 159–195.

J.-H. Evertse, [84] *On equations in S-units and the Thue-Mahler equation*, Invent. Math. **75** (1984), 561–584.

D. Grayson, [81] *Dilogarithm computations for* $K_3$, Lecture Notes in Math., vol. 854, Springer-Verlag, Berlin, 1981, pp. 168–177.

F. Keune, [81] *Another presentation for the* $K_2$ *of a local domain*, J. Pure Appl. Algebra **22** (1981), 131–141.

L. Lewin, [58] *Dilogarithms and associated functions*, Macdonald, London, 1958.

——, [81] *Polylogarithms and associated functions*, North Holland, New York, 1981.

——, [82] *The dilogarithm in algebraic fields*, J. Austral. Math. Soc. Ser. A **33** (1982), 302–330.

——, [84] *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), 345–373.

L. Lewin and M. Abouzahra, [90] *Supernumary polylogarithmic ladders and related functional equations*, Aequationes Math. **39** (1990), 210–253.

J. Milnor, [71] *Introduction to algebraic K-theory*, Princeton Univ. Press, Princeton, 1971.

——, [79] *Computation of Volume*, Chapter 7. Mimeographed lecture notes by W. P. Thurston, Princeton University, 1979.

——, [83] *On polylogarithms, Hurwitz zeta functions, and the Kubert identities*, Enseign. Math. **29** (1983), 281–322.

W. Narkiewicz, [90] *Elementary and analytic theory of algebraic numbers*, PWN, Warszawa, 1990.

L. Rédei, [58] *Über die algebraischzahlentheoretische Verallgemeinerung eines elementarzahlentheoretischen Satzes von Zsigmondy*, Acta Sci. Math. **19** (1958), 98–128.

A. Schinzel, [62] *On primitive prime factors of* $a^n - b^n$, Proc. Cambridge Phil. Soc. **58** (1962), 555–562.

——, [74] *Primitive divisors of the expression* $A^n - B^n$ *in algebraic number fields*, J. Reine Angew. Math. **268/9** (1974), 27–33.

A. A. Suslin, [87] *Algebraic K-theory of fields*, Proc. Internat. Congress of Math. 1986, Amer. Math. Soc., Providence, 1987, 222–241.





B. M. M. de Weger, [87] *Solving exponential diophantine equations using lattice basis reduction algorithms*, J. Number Theory **26** (1987), 325–367; Erratum ibid. **31** (1989), 88–89.

——, [89] *Algorithms for diophantine equations*, CWI Tract 65, Stichting Math. Centrum, Amsterdam, 1989.

D. Zagier, [86] *Hyperbolic manifolds and special values of Dedekind zeta-functions*, Invent. Math. **83** (1986), 285–301.

——, [88] *The dilogarithm function in geometry and number theory*, Number theory and related topics, Papers presented at the Ramanujan Colloquium, Bombay Tata Institute of Fundamental Research, 1988, pp. 231–249. (see also: *The remarkable dilogarithm*, J. Math. Phys. Sci. **22** (1988), 131–145).

——, [90] *Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields*, Max-Planck-Institut für Mathematik, Bonn, preprint MPI 90-44, 40 pages.

CHAPTER 12

# Function Theory of Polylogarithms

SPENCER BLOCH

The history of the polylogarithm functions

$$\mathrm{Li}_k(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^k} \qquad (k \geq 1) \tag{1}$$

is curious, and not, we think, typical of modern science. After considerable interest in the 19th century, there was a long period of neglect, during which the monograph of Lewin was the only major reference to appear. Then suddenly for a number of reasons; including relations with combinatorial definitions of Pontryagin classes, Hilbert's problem relating to scissors-congruence groups, volume calculations for hyperbolic three space, continuous cohomology classes for $\mathrm{GL}_n(\mathbf{C})$, and regulator mappings for algebraic $K$-groups of number fields; the *dilogarithm* $L_2(z)$ became chic.

Modern geometers' interest in $\mathrm{Li}_n(z)$ for general $n$ began with the work of Chen and others on iterated integrals and the nilpotent completion of $\pi_1(\mathbf{P}^1 - \{0, 1, \infty\})$. More recently (motivated by a conjecture of Zagier) Beilinson and Deligne have begun to understand the role of polylogarithms in the theories of mixed Hodge structures and mixed motives. Following Beilinson and Deligne, we want to discuss in concrete terms what one might call the classical function theory of the $\mathrm{Li}_n(z)$. There are two problems here; how to "understand" the multivalued nature of the polylogs, and how to interpret the veritable zoo of functional equations they satisfy. On the first point, modern theory offers definitive insights. On the second, the question of functional equations, one has a theoretical interpretation but still no effective way of enumerating them. Our main result in this direction, theorem (50), is a variant of a result of Zagier [**Z1**, §7, Proposition 1].

A *pure Hodge structure of weight* $n$ is a free abelian group $H_{\mathbf{Z}}$ together with a descending filtration (the Hodge filtration) $F^{\cdot}H_{\mathbf{C}}$ on $H_{\mathbf{C}} = H_{\mathbf{Z}} \otimes \mathbf{C}$. The Hodge filtration is required to be $n$-opposite to its complex conjugate in the following sense. Since $H$ is defined over $\mathbf{R}$ (in fact over $\mathbf{Z}$), $H_{\mathbf{C}}$ has a conjugation map "$\overline{\phantom{x}}$" leaving fixed $H_{\mathbf{R}}$. Write $\overline{F}^{\cdot}$ for the conjugate of $F^{\cdot}$.

The requirement is

(2) $$F^q H_{\mathbf{C}} \oplus \overline{F}^{n-q+1} H_{\mathbf{C}} \cong H_{\mathbf{C}}.$$

Write

(3) $$H^{p,q} = F^p \cap \overline{F}^q.$$

It follows easily that

(4) $$H_{\mathbf{C}} \cong \bigoplus_{p+q=n} H^{p,q}.$$

For example, the first cohomology group $H^1(X, \mathbf{Z})$ of a compact Riemann surface $X$ is a pure Hodge structure of weight 1. The Hodge filtration is given by

(5) $$\begin{aligned} &F^0 = H^1(X, \mathbf{C}), \\ &F^1 = H^{1,0} = H^0(X, \Omega^1) = \text{ the space of holomorphic 1-forms on } X. \end{aligned}$$

Another example more relevant to polylogs are the Tate Hodge structures $\mathbf{Z}(n)$ of weight $-2n$ defined by $\mathbf{Z}(n)_{\mathbf{Z}} \cong \mathbf{Z}$, and $\mathbf{Z}(n)_{\mathbf{C}} = H^{-n-n}$. Although this doesn't really have meaning unless one considers also the "de Rham realization" of $\mathbf{Z}(n)$ (which we will not do), it is customary to identify $Z(n)_{\mathbf{Z}} = \mathbf{Z} \cdot (2\pi i)^n \subset \mathbf{C}$.

To define $\mathbf{Q}$- (resp. $\mathbf{R}$-) Hodge structures, one proceeds as above, starting with a $\mathbf{Q}$ (resp. $\mathbf{R}$) vector space $H_{\mathbf{Q}}$ (resp. $H_{\mathbf{R}}$).

A *mixed Hodge structure* (MHS) $H$ is a free abelian group $H_{\mathbf{Z}}$ together with:

(1) an increasing filtration, the *weight filtration* $W.H_{\mathbf{Q}}$,
(2) a decreasing filtration $F^{\cdot} H_{\mathbf{C}}$, the *Hodge filtration.*

These data are required to be compatible in the sense that the filtration induced by $F^{\cdot}$ on $\mathrm{gr}_n^W(H_{\mathbf{Q}})$ is required to be a pure $\mathbf{Q}$-Hodge structure of weight $n$. (Exercise: show the two possible definitions of "filtration induced by $F^{\cdot}$ on $\mathrm{gr}_n^W(H_{\mathbf{Q}})$" coincide.) Definitions of $\mathbf{Q}$- and $\mathbf{R}$-mixed Hodge structures are analogous.

Here is a simple example. Choose $x \in \mathbf{C}^{\times}$. View $x$ as defining a map $\mathbf{Z} \to \mathbf{C}^{\times}$, $1 \mapsto x$; and consider the diagram

(6) $$\begin{tikzcd} 0 \arrow[r] & 2\pi i \cdot \mathbf{Z} \arrow[r] & \mathbf{C} \arrow[r, "\exp"] & \mathbf{C}^{\times} \arrow[r] & 0 \\ 0 \arrow[r] & 2\pi i \cdot \mathbf{Z} \arrow[r] \arrow[u, equal] & M_{1,x,\mathbf{Z}} \arrow[r] \arrow[u, "\varphi"] & \mathbf{Z} \arrow[r] \arrow[u, "x"] & 0, \end{tikzcd}$$



where $M_{1,x,\mathbf{Z}} = \{(a, b) \in \mathbf{C} \times \mathbf{Z} | \exp(a) = x^b\}$. Define

$$W_{-3}(M_{1,x,\mathbf{Z}}) = (0); \qquad W_{-2}(M_{1,x,\mathbf{Z}}) = W_{-1}(M_{1,x,\mathbf{Z}}) = 2\pi i \cdot \mathbf{Z};$$
$$W_0(M_{1,x,\mathbf{Z}}) = M_{1,x,\mathbf{Z}};$$
$$F^{-1}(M_{1,x,\mathbf{C}}) = M_{1,x,\mathbf{C}}; \qquad F^0(M_{1,x,\mathbf{C}}) = \mathrm{Ker}(\varphi \otimes \mathbf{C} : M_{1,x,\mathbf{C}} \to \mathbf{C});$$
$$F^1(M_{1,x,\mathbf{C}}) = (0).$$

One checks easily that these data give a mixed Hodge structure on $M_{1,x}$.

The category of $\mathbf{Q}$-mixed Hodge structures is an *abelian category*. This is an elementary but not a trivial fact. It means we can treat mixed Hodge structures rather like we do abelian groups. Homomorphisms between them have kernels and cokernels which are themselves mixed Hodge structures. The essential point is that if $f : H \to H'$ is a homomorphism of mixed Hodge structures, then image$(f)$ with weight and Hodge filtrations induced as quotients from those of $H$ coincides with coimage$(f) \subset H'$, with filtrations induced by restriction from those on $H'$. (The ol' philosopher: whenever a category defined via filtrations turns out to be an abelian category, one is onto something interesting.) For example, the mixed Hodge structure $M_{1,x}$ defined above fits into an exact sequence of mixed Hodge structures

(7) $$0 \to \mathbf{Z}(1) \to M_{1,x} \xrightarrow{p} \mathbf{Z}(0) \to 0.$$

Moreover, MHS has an internal tensor product and hom. (It is a *Tannakian* category.)

What does all this have to do with polylogarithms? Notice in the above sequence, $p^{-1}(1) = \{(y, x) \in \mathbf{C} \times \{x\} | e^y = x\}$, i.e.,

(8) $$p^{-1}(1) = \{(2n\pi i + \log x, x)\}.$$

In other words, $p^{-1}(1)$ is the *torseur* (principal homogeneous space) of all determinations of $\log x$. Here is how this is generalized to polylogs in **[D1]**.

Consider the matrix

(9) $$A(z) = \begin{pmatrix} 1 & & & & 0 \\ -\mathrm{Li}_1(z) & 2\pi i & & & \\ -\mathrm{Li}_2(z) & 2\pi i \cdot \log z & (2\pi i)^2 & & \\ -\mathrm{Li}_3(z) & \frac{2\pi i \cdot (\log z)^2}{2!} & (2\pi i)^2 \cdot \log z & (2\pi i)^3 & \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \quad\cdot \\ \cdot & \cdot & \cdot & & \quad\quad\cdot \end{pmatrix}.$$

Fix an integer $N \geq 1$, and restrict it to the first $N+1$ rows and columns to get an $(N+1) \times (N+1)$ matrix $A_N(z)$, which we view as acting on $\mathbf{C}^{[0,N]}$ written as columns. We define a $\mathbf{Q}$-mixed Hodge structure $M_{N,z}$ as follows. $M_{N,z,\mathbf{Q}}$ is the $\mathbf{Q}$-span of the columns of $A_N(z)$. The weight filtration is given by

(10) $$W_{-2k} M_{N,z,\mathbf{C}} = \mathbf{C}^{[k,N]}.$$

Note this coincides with the $\mathbb{C}$-span of columns $k+1$ through $n$, so $W.$ is defined over $\mathbb{Q}$. The Hodge filtration is

$$F^{-k}M_{N,z,\mathbb{C}} = \mathbb{C}^{[0,k]}. \tag{11}$$

Although log and the $\mathrm{Li}_n$ are multiple valued functions, the MHS $M_{N,z,\mathbb{Q}}$ is independent of the choice of branch. Indeed, if we fix a real number $t$ with $0 < t < 1$ and choose a path $\gamma$ from $t$ to $z$ then we have fixed a matrix $A(z)_\gamma$. If we preceed $\gamma$ by a small loop (counterclockwise) $\gamma_0$ about 0 (resp. $\gamma_1$ about 1) one computes

$$A(z)_{\gamma\gamma_0} = A(z)_\gamma \cdot \exp(e_0); \qquad A(z)_{\gamma\gamma_1} = A(z)_\gamma \cdot \exp(e_1), \tag{12}$$

where $e_1 = \varepsilon_{2,1}$ and $e_0 = \varepsilon_{3,2}+\varepsilon_{4,3}+\varepsilon_{5,4}+\cdots$. Here $\varepsilon_{i,j}$ is the matrix with 1 in the $(i,j)$-th place and zeroes elsewhere. In particular, the $\mathbb{Q}$-structure on $\mathbb{C}^{[0,n]}$ given by the span of the columns is unchanged. Formulas (10) and (11) show the weight and Hodge filtrations are independent of the choice of branch, so the Hodge structure itself is independent.

Another way of saying this is to view $z$ as a parameter on $\mathbb{P}^1 - \{0, 1, \infty\}$. The $\mathbb{Q}$-span of the columns of $A_N(z)$ embeds (locally) in the sheaf $\mathscr{O}^{\oplus N+1}$ and the $N+1$ dimensional $\mathbb{Q}$-vector space thus obtained is independent of the choice of branch. The fundamental group $\pi_1(\mathbb{P}^1 - \{0, 1, \infty\}, t)$ acts on this vector space as in (12), so we obtain a *local system* on $\mathbb{P}^1 - \{0, 1, \infty\}$. For each point $z$ in the space, the fibre of this local system is the underlying rational structure $M_{N,z,\mathbb{Q}}$ for a MHS $M_{N,z}$, so we get a *variation of Hodge structure* over $\mathbb{P}^1 - \{0, 1, \infty\}$.

The submixed Hodge structure $W_{-2}M_{N,z} \subset M_{N,z}$ is given by rows and columns 2 to $N+1$ of (9). In the following lemma, we use the fact that the category of mixed Hodge structures has a tensor product, so mixed Hodge structures like $\mathrm{Sym}^n H$ exist. We write $H(r) = H \otimes \mathbb{Z}(r)$.

LEMMA 13. *$W_{-2}M_{N,z} \cong \mathrm{Sym}^{N-1}(M_{1,z})(1)$ as $\mathbb{Q}$-mixed Hodge structures.*

PROOF. Write $v_0$ and $v_1$ for the obvious basis of $\mathbb{C}^{[0,1]}$. $M_{1,z,\mathbb{Q}}$ is the $\mathbb{Q}$-span of $v_0 + \log(z)\cdot v_1$, and $2\pi i \cdot v_1$. $\mathrm{Sym}^{N-1}(M_{1,z,\mathbb{Q}})$ is the span of

(13a)
$$(2\pi i)^{N-1}v_1^{N-1},\ (2\pi i)^{N-2}v_1^{N-2}\cdot(v_0+\log(z)\cdot v_1),\ \dots,\ (v_0+\log(z)\cdot v_1)^{N-1}.$$

Write $w_1, \dots, w_N$ for the standard basis of $\mathbb{Q}^{[1,N]}$ and map

$$v_1^{N-r}v_0^{r-1} \mapsto (r-1)!\,w_{N-r+1}; \qquad 1 \le r \le N. \tag{14}$$

One has

$$\begin{aligned} &[(r-1)!]^{-1}(2\pi i)^{N-r}v_1^{N-r}\cdot(v_0+\log(z)\cdot v_1)^{r-1} \\ &\mapsto (2\pi i)^{N-r}[w_{N-r+1} + \log(z)w_{N-r+2} + (\log(z)^2/2!)w_{N-r+3} \\ &\qquad + \cdots + (\log(z)^{r-1}/(r-1)!)w_N], \end{aligned} \tag{15}$$



so the $\mathbb{Q}$-span of the elements (13a) coincides with the $\mathbb{Q}$-span of the right-hand side of (15), which is $[2\pi i]^{-1}$ times the $\mathbb{Q}$-span of the columns in $W_{-2}M_{N,z,\mathbb{Q}}$. The rest of the argument is straightforward. Q.E.D.

As a consequence of (13), we have an extension of mixed Hodge structures, the *polylogarithm extension* of $\mathbb{Q}$-mixed Hodge structures

$$0 \to \mathrm{Sym}^{N-1}(M_{1,z})(1) \to M_{N,z} \to \mathbb{Q}(0) \to 0, \tag{16}$$

which encodes the monodromy of the polylogarithms. Recall (7) that $M_{1,z}$ is an extension of $\mathbb{Z} = \mathbb{Z}(0)$ by $\mathbb{Z}(1)$. This implies

$$\mathrm{gr}^W_{-2k} M_{N,z} \cong \mathbb{Z}(k); \qquad 0 \le k \le N. \tag{17}$$

I want to exhibit the class of extension (16) in terms of the polylogs. To simplify, write (16) as $0 \to E \to M \xrightarrow{p} \mathbb{Q} \to 0$. Choose a map of abelian groups $f: \mathbb{Q} \to M_{\mathbb{Q}}$ splitting $p$. The map $f$ is determined by $f(1)$, and we may modify $f(1)$ by an element of $E_{\mathbb{Q}}$. $f$ is a map of Hodge structures if and only if $f(1) \in M_{\mathbb{Q}} \cap F^0M_{\mathbb{C}}$. But $F^0M_{\mathbb{C}} \cap E_{\mathbb{C}} = F^0E_{\mathbb{C}} = (0)$, so

$$p: F^0M_{\mathbb{C}} \cong \mathbb{Q}(0)_{\mathbb{C}} = \mathbb{C}. \tag{18}$$

There exists, therefore, a unique $e \in E_{\mathbb{C}}$ such that $f(1) + e \in F^0M_{\mathbb{C}}$. If $e \in E_{\mathbb{Q}}$ then $f'$ defined by $f'(1) = f(1) + e$ splits $p$ as a map of Hodge structures. In other words,

$$\mathrm{Ext}^1_{\mathbb{Q}\text{-MHS}}(\mathbb{Q}(0), E) \cong E_{\mathbb{C}}/E_{\mathbb{Q}}. \tag{19}$$

(A similar analysis shows

$$\mathrm{Ext}^1_{\mathbb{R}\text{-MHS}}(\mathbb{R}(0), E) \cong E_{\mathbb{C}}/E_{\mathbb{R}}.) \tag{20}$$

In particular, the polylog extension (16) corresponds to a canonical element

$$\mathrm{polylog}_N(z) \in \mathrm{Sym}^{N-1}(M_{1,z})(1)_{\mathbb{C}}/\mathrm{Sym}^{N-1}(M_{1,z})(1)_{\mathbb{Q}}. \tag{21}$$

We will exhibit this element. The discussion involving the diagram (6) can be reinterpreted as saying that the exponential sequence $(\otimes\mathbb{Q})$,

$$0 \to \mathbb{Q}\cdot 2\pi i \to \mathbb{C} \to \mathbb{C}^\times \otimes \mathbb{Q} \to 0, \tag{22}$$

is itself an enormous extension of mixed Hodge structures, where $\mathbb{Q}\cdot 2\pi i$ is identified with $\mathbb{Q}(1)_{\mathbb{Q}}$, and $\mathbb{C}^\times \otimes \mathbb{Q} = \oplus\mathbb{Q}(0)_{\mathbb{Q}}$. Write $M_1$ for the term in the middle. $M_{1,\mathbb{Q}} \cong \mathbb{C}$, and the Hodge filtration is given by the kernel of multiplication

$$F^0M_{1,\mathbb{C}} = \mathrm{Ker}(M_{1,\mathbb{C}} = \mathbb{C}\otimes_{\mathbb{Q}}\mathbb{C} \to \mathbb{C}). \tag{23}$$

We have from (6) that $M_{1,z} \to M_1$ (the map is an inclusion unless $z$ is a root of 1), so we can view $\mathrm{polylog}_N(z)$ as an element of

$$\begin{aligned} \mathrm{Sym}^{N-1}_{\mathbb{Q}}(M_1)_{\mathbb{C}}/\mathrm{Sym}^{N-1}_{\mathbb{Q}}(M_1)(1)_{\mathbb{Q}} &\cong \mathrm{Sym}^{N-1}_{\mathbb{Q}}(\mathbb{C})\otimes_{\mathbb{Q}}(\mathbb{C}/2\pi i\mathbb{Q}) \\ &\cong \mathrm{Sym}^{N-1}_{\mathbb{Q}}(\mathbb{C})\otimes_{\mathbb{Z}}\mathbb{C}^\times. \end{aligned} \tag{24}$$

The image of $M_{1,z,\mathbf{Q}}$ is the $\mathbf{Q}$-span of $2\pi i$ and $\log z$ in $M_{1\mathbf{Q}} = \mathbf{C}$, so $\mathrm{Sym}_{\mathbf{Q}}^{N-1}(M_{1,z})_{\mathbf{Q}} \subset \mathrm{Sym}_{\mathbf{Q}}^{N-1}(\mathbf{C})$ is spanned by

$$(25) \qquad (2\pi i)^{\odot a} \odot (\log z)^{\odot(N-1-a)}, \qquad 0 \le a \le N-1,$$

where $\odot$ means product in Sym (as opposed to product in $\mathbf{C}$).

To calculate $\mathrm{polylog}_N(z) \in \mathrm{Sym}_{\mathbf{Q}}^{N-1}(\mathbf{C}) \otimes_{\mathbf{Z}} \mathbf{C}^\times$, let $C_i(z)$ be the $i+1$st column of the matrix $A(z)$ in (9). Let $\mathbf{1}_1$ be the column vector with 1 in the first row and zeroes below, and write

$$(26) \qquad \mathbf{1}_1 = C_0 + a_1(z)C_1(z) + a_2(z)C_2(z) + \cdots.$$

One shows easily

$$(27) \qquad a_p = (2\pi i)^{-p}[\mathrm{Li}_p(z) - \cdots + (-1)^j \mathrm{Li}_{p-j}(z)\cdot \log(z)^j/j! + \cdots + (-1)^{p-1}\mathrm{Li}_1(z)\cdot\log(z)^{p-1}/(p-1)!].$$

One has

(28)

$$\mathrm{polylog}_N(z) = \sum_{\nu=0}^{\nu=N-1} (N-1-\nu)!^{-1}(2\pi i)^{\odot\nu}\odot(\log z)^{\odot(N-1-\nu)}\otimes\exp(2\pi i\cdot a_{\nu+1}(z)).$$

For example

(29)

$$\mathrm{polylog}_1(z) = \frac{1}{(1-z)} \in \mathbf{C}^\times;$$
$$\mathrm{polylog}_2(z) = \log(z)\otimes\exp(\mathrm{Li}_1(z)) + 2\pi i\otimes\exp((2\pi i)^{-1}\{\mathrm{Li}_2(z) - \log(z)\cdot\mathrm{Li}_1(z)\}) \in \mathbf{C}\otimes_{\mathbf{Z}}\mathbf{C}^\times.$$

The expressions (28) are independent of the choice of branch, but they are rather complicated tensors. We may simplify things a bit by considering the extension (16) as an extension of $\mathbf{R}$-Hodge structures. The resulting class, $\mathrm{polylog}_{N,\mathbf{R}}(z)$, is simply the image of $\mathrm{polylog}_N(z)$ in

$$(30) \qquad \mathrm{Sym}_{\mathbf{R}}^{N-1}(\mathbf{C})\otimes_{\mathbf{R}}(\mathbf{C}/2\pi i\mathbf{R}) \cong \mathbf{R}^N.$$

For example

(31)

$$\begin{aligned}\mathrm{polylog}_{2,\mathbf{R}}(z) = {} & -\log|z|\cdot\log|1-z| \\ & + i\cdot\{-\arg(z)\cdot\log|1-z| + \mathrm{Im}(\mathrm{Li}_2(z)) \\ & + \arg(z)\cdot\log|1-z| + \log|z|\cdot\arg(1-z)\} \\ = {} & -\log|z|\cdot\log|1-z| + i\cdot\{\log|z|\cdot\arg(1-z) + \mathrm{Im}(\mathrm{Li}_2(z))\}.\end{aligned}$$

What is new here is the single-valued function

$$(32) \qquad \log|z|\cdot\arg(1-z) + \mathrm{Im}(\mathrm{Li}_2(z)).$$



More generally, a basis of $\mathrm{Sym}_{\mathbf{R}}^{N-1}(\mathbf{C})$ is given by $i^{\odot\mu}$, $0 \le \mu \le N-1$. It follows from (28) that

$$(33) \qquad \mathrm{polylog}_{N,\mathbf{R}}(z) = \sum_{\mu=0}^{N-1}(N-1-\mu)!^{-1}E_{\mu+1}\log|z|^{N-1-\mu}\cdot i^{\odot\mu},$$

where

$$(34) \qquad E_{\mu+1}(z) = -2\pi\sum_{\nu=0}^{\mu}\mathrm{Im}(a_{\nu+1}(z))(2\pi)^\nu\arg(z)^{\mu-\nu}/(\mu-\nu)!.$$

A tedious calculation shows

(35)

$$\pm E_\mu(z) = \begin{cases} \mathrm{Re}(\mathrm{Li}_\mu(z)) - \mathrm{Re}(\mathrm{Li}_{\mu-1}(z))\cdot\log|z|/1! + \cdots \\ \quad + (-1)^{\mu-1}\mathrm{Re}(\mathrm{Li}_1(z))\cdot(\log|z|)^{\mu-1}/(\mu-1)!; & \mu \text{ odd} \\ \mathrm{Im}(\mathrm{Li}_\mu(z)) - \mathrm{Im}(\mathrm{Li}_{\mu-1}(z))\cdot\log|z|/1! + \cdots \\ \quad + (-1)^{\mu-1}\mathrm{Im}(\mathrm{Li}_1(z))\cdot(\log|z|)^{\mu-1}/(\mu-1)!; & \mu \text{ even}. \end{cases}$$

(The sign is $+$ for $\mu \equiv 1$ or $2 \pmod 4$; $-$ otherwise.)

I next want to consider the question of functional equations for the polylogs. Define the map

(36)

$$\partial : (\mathrm{Sym}_{\mathbf{Q}}^n\mathbf{C})\otimes_{\mathbf{Q}}\mathbf{C}^\times \to (\mathrm{Sym}_{\mathbf{Q}}^{n-1}\mathbf{C})\otimes_{\mathbf{Q}}\mathbf{C}^\times\otimes_{\mathbf{Q}}\mathbf{C}^\times,$$
$$\partial(a_1\odot a_2\odot\cdots\odot a_n\otimes b) = \sum a_1\odot a_2\odot\cdots\odot\hat{a}_i\odot\cdots\odot a_n\otimes b\otimes\exp(a_i).$$

One checks that the sequence

$$(37) \qquad 0 \to \mathbf{C}^\times\otimes\mathbf{Q} \xrightarrow{\iota} (\mathrm{Sym}_{\mathbf{Q}}^n\mathbf{C})\otimes_{\mathbf{Q}}\mathbf{C}^\times \xrightarrow{\partial} (\mathrm{Sym}_{\mathbf{Q}}^{n-1}\mathbf{C})\otimes_{\mathbf{Q}}\mathbf{C}^\times\otimes_{\mathbf{Q}}\mathbf{C}^\times$$

is exact, where $\iota(x) = (2\pi i)^{\odot n}\otimes x$.

LEMMA (38). *For $N \ge 2$, we have*

$$(39) \qquad \partial(\mathrm{polylog}_N(z)) = \mathrm{polylog}_{N-1}(z)\otimes z.$$

*In particular,*

$$(40) \qquad \partial(\mathrm{polylog}_2(z)) = (1-z)\otimes z.$$

PROOF. This is straightforward from (28). Q.E.D.

The key to finding functional equations is the following sort of *rigidity*. Let $\mathcal{O}$ be the ring of germs of analytic functions at the origin in $\mathbf{C}^q$ for some $q$, and let $t = (t_1, \ldots, t_q)$ be coordinates. Suppose $z = z(t) \in \mathcal{O}$. The functions defining $\mathrm{polylog}_N$ are locally analytic, so

$$(41) \qquad \mathrm{polylog}_N(z(t)) \in \mathrm{Sym}_{\mathbf{Q}}^{N-1}\mathcal{O}\otimes_{\mathbf{Q}}\mathcal{O}^\times.$$

Let $\Omega = \mathcal{O}\cdot dt_1 + \cdots + \mathcal{O}\cdot dt_q$ be the module of germs of analytic differential 1-forms at the origin in $\mathbf{C}^q$. Consider the map

(42) $$\mathcal{D} : \mathrm{Sym}_{\mathbf{Q}}^{N-1}\mathcal{O}\otimes_{\mathbf{Q}}\mathcal{O}^\times \to \Omega;$$
$$\mathcal{D}(a_1(t)\circ a_2(t)\circ\cdots\circ a_n(t)\otimes b(t)) = a_1a_2\cdots a_n\, db/b.$$

Note that $\mathcal{D}\circ\iota : \mathcal{O}^\times \to \Omega$ is given by

(43) $$\mathcal{D}\circ\iota(x) = (2\pi i)^{N-1}dx/x.$$

LEMMA (44). $\mathcal{D}(\mathrm{polylog}_N(z(t))) = 0$.

PROOF. Using (27) and induction, the reader can verify that

(45) $$da_p(z) = (-1)^{p-1}\log(z)^{p-1}\cdot dz/((p-1)!(z-1)); \qquad p\geq 1.$$

Substituting in (28), one is then reduced to showing

(46) $$0 = \Sigma_0^{N-1}(-1)^\nu[\nu!\cdot(N-1-\nu)!]^{-1},$$

which is straightforward. Q.E.D.

PROPOSITION (47). *Let $z_1(t), \ldots, z_r(t)\in\mathcal{O}^\times$, and suppose $1-z_i\in\mathcal{O}^\times$ as well, for $1\leq i\leq r$. Suppose further that for suitable $a_i\in\mathbf{Q}^\times$*

(48) $$\Sigma a_i\cdot\mathrm{polylog}_{N-1}(z_i)\otimes z_i = 0 \quad \textit{in}\ \mathrm{Sym}_{\mathbf{Q}}^{N-2}\mathcal{O}\otimes_{\mathbf{Q}}\mathcal{O}^\times\otimes\mathcal{O}^\times.$$

*Then $\Sigma a_i\cdot\mathrm{polylog}_N(z_i)$ is constant independent of $t$. In fact*

(49) $$\Sigma a_i\cdot\mathrm{polylog}_N(z_i)\in\mathbf{C}^\times\subset\mathcal{O}^\times\hookrightarrow\mathrm{Sym}_{\mathbf{Q}}^{N-1}\mathcal{O}\otimes_{\mathbf{Q}}\mathcal{O}^\times.$$

PROOF. Equations (38) and (48) imply that $\Sigma a_i\cdot\mathrm{polylog}_N(z_i)\in\mathcal{O}^\times$. To show it is constant, we note that (43) and (44) imply the logarithmic derivative of this expression is $0$. Q.E.D.

THEOREM (50) (rigidity). *Let $z_1(t), \ldots, z_r(t)\in\mathcal{O}^\times$, and suppose $1-z_i\in\mathcal{O}^\times$ as well, for $1\leq i\leq r$. Suppose further that for suitable $a_i\in\mathbf{Q}^\times$*

(51) $$\Sigma(1-z_i)^{a_i}\otimes z_i^{\otimes N-1} = 0\ \textit{in}\ \mathcal{O}^\times\otimes_{\mathbf{Q}}\mathcal{O}^{\times\otimes N-1}.$$

*Choose local analytic branches $\log(z_i)$. Then there exist constants $c_{i,\nu}\in\mathbf{C}^\times$, $0\leq\nu\leq N-1$ such that*

(52) $$\sum_i\left[a_i\ \mathrm{polylog}_N(z_i)+\sum_\nu(2\pi i)^{\odot\nu}\odot(\log z_i)^{\odot(N-1-\nu)}\otimes c_{i,\nu}\right]$$
$$=0\in\mathrm{Sym}^{N-1}\mathcal{O}\otimes\mathcal{O}^\times.$$

*If the $z_i(t)$ are algebraic functions of $t$ then*

(53) $$\sum_i a_i\cdot\mathrm{polylog}_N(z_i)\in\ \textit{Image}\ (\mathrm{Sym}^{N-1}\mathbf{C}\otimes\mathbf{C}^\times\to\mathrm{Sym}^{N-1}\mathcal{O}\otimes\mathcal{O}^\times),$$

*i.e., this expression is constant in $t$.*



PROOF. Consider an expression for some $M$, $1\leq M\leq N-1$:

(53a)
$$A=\sum_i\left[a_i\cdot\mathrm{polylog}_M(z_i)+\sum_\nu(2\pi i)^{\odot\nu}\odot(\log z)^{\odot(M-1-\nu)}\otimes c_{i,\nu}\right]\otimes z_i^{N-M},$$

in $\mathrm{Sym}^{M-1}\mathcal{O}\otimes\mathcal{O}^\times\otimes\mathcal{O}^{\times\otimes N-M}$, where the $c_{i,\nu}$ are constant. Note that (51) implies the existence of such an expression for $M=1=c_{i,0}$. One has

(54) $$A=(\partial\otimes(\mathrm{id})^{\otimes N-M-1})B,$$

with $\partial$ as in (36), and

(55) $$B=\sum_i\left[a_i\cdot\mathrm{polylog}_{M+1}(z_i)+\sum_\nu(M-\nu)^{-1}(2\pi i)^{\odot\nu}\odot(\log z)^{\odot(M-\nu)}\otimes c_{i,\nu}\right]\otimes z_i^{N-M-1}$$
$$\in\mathrm{Sym}^M\mathcal{O}\otimes\mathcal{O}^\times\otimes\mathcal{O}^{\times\otimes N-M-1}$$

Consider as in (42) the map

(56) $$\mathcal{D}\otimes(\mathrm{id})^{\otimes N-M-1}:\mathrm{Sym}^M\mathcal{O}\otimes\mathcal{O}^\times\otimes\mathcal{O}^{\times\otimes N-M-1}\to\Omega\otimes\mathcal{O}^{\times\otimes N-M-1}.$$

I claim

(57) $$\mathcal{D}\otimes(\mathrm{id})^{\otimes N-M-1}(B)=0.$$

Indeed, looking at (55), the terms involving $\mathrm{polylog}_{M+1}(z_i)$ die by (44), while the other terms die because the $c_{i,\nu}$ are constant.

Assume now $A=0$. It follows from the analogue of (37) that

(58) $$B\in\mathcal{O}^\times\otimes\mathcal{O}^{\times\otimes N-M-1}\hookrightarrow\mathrm{Sym}^M\mathcal{O}\otimes\mathcal{O}^\times\otimes\mathcal{O}^{\times N-M-1};$$
$$a\otimes\alpha\mapsto(2\pi i)^{\odot M}\otimes a\otimes\alpha.$$

From (57) and (43) one has $B\in\mathbf{C}^\times\otimes\mathcal{O}^{\times\otimes N-M-1}$. In fact

(59) $$B\in(\mathbf{C}^\times\otimes\mathcal{O}^{\times\otimes N-M-1})\cap(\mathcal{O}^\times\otimes\langle z^{\otimes N-M-1}\rangle)=\mathbf{C}^\times\otimes\langle z^{\otimes N-M-1}\rangle,$$

where $\langle z^{\otimes N-M-1}\rangle\subset\mathbf{Q}\otimes\mathcal{O}^{\times\otimes N-M-1}$ is the $\mathbf{Q}$-span of the elements $z_i^{\otimes N-M-1}$. In the definition of $B$ we are free to add terms $(2\pi i)^{\odot M}\otimes c_{i,M}\otimes z_i^{\otimes N-M-1}$, so we can arrange $B=0$.

Now replace $A$ by $B$ and repeat the above argument. In the end, it follows that for suitable $c_{i,\nu}\in\mathbf{C}^\times$, $0\leq\nu\leq N-1$, equation (52) holds.

Note that (52) implies

(60) $$\sum_\nu(2\pi i)^{\odot\nu}\odot(\log z_i(t))^{\odot(N-1-\nu)}\otimes c_{i,\nu}\in\mathrm{Sym}^{N-1}\mathcal{O}\otimes\mathbf{C}^\times$$

is independent of path in $t$. If this expression is nonconstant, i.e., does not lie in the image of $\mathrm{Sym}^{N-1}\mathbf{C}\otimes\mathbf{C}^\times\to\mathrm{Sym}^{N-1}\mathcal{O}\otimes\mathbf{C}^\times$, it is not hard to show there exist constants $b_{i,\nu} \in \mathbb{C}$ such that the function

$$\sum_{\nu,i} b_{i,\nu} \log z_i(t)^{N-1-\nu} \tag{61}$$

is single valued, nonconstant and analytic in $t$. Assuming the $z_i$ are algebraic functions of $t$, this function has at worst logarithmic poles, a contradiction. Q.E.D.

**Corollary** (62). *Let $E_{\mu+1}$ be as in (35), and let $z_i(t)$ be as in (50) for some $N > \mu$. Then there exist constants $d_{i,\nu} \in \mathbb{R}$, $0 \le \nu \le \mu$ such that*

$$\sum_i \log|z_i|^{N-1-\mu} \cdot \left[ a_i \cdot E_{\mu+1}(z_i(t)) + \sum_{0\le\nu\le\mu} d_{i,\nu} \cdot \arg(z_i)^{\mu-\nu} \right] = 0. \tag{63}$$

*If the $z_i(t)$ are algebraic, one has*

$$\sum_i a_i \cdot E_{\mu+1}(z_i(t)) = \text{constant (depending on } \mu). \tag{64}$$

**Proof.** Using (30) we can reinterpret (52) as an equation of functions of $t$ with values in $\mathbb{R}^N = \mathrm{Sym}_{\mathbb{R}}^{N-1}(\mathbb{C}) = \bigoplus_{\nu\in N-1} \mathbb{R} \cdot i^{\odot\nu}$. The coefficient of $i^{\odot\mu}$ in this expression is easily seen to be

$$\begin{aligned}\sum_i \Bigg[ & a_i \cdot (N-1-\mu)!^{-1} E_{\mu+1}(z_i) \cdot \log|z_i|^{N-1-\mu} \\ & + \log|z_i|^{N-1-\mu} \sum_{0\le\nu\le\mu} \binom{N-1-\mu}{\mu-\nu} (2\pi)^\nu \arg(z_i)^{\mu-\nu} \cdot \log|c_{i,\nu}| \Bigg].\end{aligned} \tag{65}$$

Equation (63) follows by collecting constants. Finally, (64) follows from equation (53). Q.E.D.

**Example** (66). Let $t_1, \dots, t_5$ be 5 points on $\mathbb{P}^1$, and let $z_i(t)$ be the cross-ratio of $t_1, \dots, \hat{t}_i, \dots, t_5$. One checks that

$$\Sigma(-1)^i(1-z_i) \otimes z_i + \Sigma(-1)^i(1-z_i^{-1}) \otimes z_i^{-1} = 1 \text{ in } \mathcal{O}^\times \otimes \mathcal{O}^\times. \tag{67}$$

The function $E_2$ is odd, and $\Sigma(-1)^i E_2(z_i) = 0$.

## Bibliography


[Be1] A. A. Beilinson, *Polylogarithm and cyclotomic elements*, Preprint.

[Be2] ——, *Notes on absolute Hodge cohomology*, Contemporary Math., Vol. 55, Part 1, (1986)

[BGSV] A. A. Beilinson, A. B. Goncharov, V. V. Schechtman, and A. N. Varchenko, *Aomoto dilogarithms, mixed Hodge structures, and motivic cohomology of pairs of triangles on the plane*, The Grothendieck Festschrift, vol. 1, Prog. Math., vol. 86, Birkhäuser, Boston, 1990.

[BMS] A. A. Beilinson, R. MacPherson, and V. V. Schechtman, *Notes on motivic cohomology*, Duke Math. J. **54** (1987), 679–710.

[Bl1] S. Bloch, *Applications of the dilogarithm function in algebraic K-theory and algebraic geometry*, Internat. Sympos. on Algebraic Geometry, Kyoto, 1977, Kinokuniya Book-Store Co. Ltd., Tokyo, 1978.





[Bl2] ——, *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*, (unpublished preprint).

[D1] P. Deligne, *Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs*, Preprint.

[D2] ——, *Le groupe fondamental de la droite projective moins trois points*, Galois groups over $\mathbb{Q}$, Math. Sci. Res. Inst. Publ., vol. 16, Springer-Verlag, Berlin and New York, 1989.

[Du] J. Dupont, *On polylogarithms*, Nagoya Math. J. **114** (1989), 1–20.

[Ra] D. Ramakrishnan, *On the monodromy of higher logarithms*, Proc. Amer. Math. Soc. **85** (1982), 596–599.

[Z1] D. Zagier, *Polylogarithms, Dedekind zeta functions, and the algebraic K-theory of fields*, Preprint.

[Z2] ——, *The Bloch-Wigner-Ramakrishnan polylogarithm function*, Math. Ann. **286** (1990), 613–624.

[Z3] ——, *The remarkable dilogarithm*, J. Math. Phys. Sci. **22** (1988), 131–145.

CHAPTER 13

# Partition Identities and the Dilogarithm

J. H. LOXTON

## 13.1. Introduction.

The theme of the earlier chapters is that the curious linear relations between the values of polylogarithm functions should occur by specialisation of functional equations. This follows the basic principle of transcendence theory, wherein special values of functions are related to the behaviour of their differential equations or functional equations. The dilogarithm is the simplest of the so-called $G$-functions of Siegel. Both in general and in this particular case, the arithmetic behaviour of $G$-functions remains mysterious adding special interest to the study of the dilogarithm.

We begin with a catalogue of relations for the dilogarithm function. They are written most economically in terms of Rogers' dilogarithm function

$$L(z) = L_2(z) = \mathrm{Li}_2(z) + \tfrac{1}{2}\log z \cdot \log(1-z).$$

(Throughout this paper, we play for safety by restricting the argument of the dilogarithm to the interval between 0 and 1. The functions $\mathrm{Li}_2(z)$ and $L(z)$ have analytic continuations to the complex plane cut along the real axis from $-\infty$ to 0 and 1 to $\infty$, but care is required. For this reason, it may sometimes be preferable to replace the logarithmic term in the definition of $L(z)$ by $\frac{1}{2}\log|z| \cdot \log(1-z)$. However, the resulting function is not analytic so that functional equations such as (13.3) and (13.4) will need some qualification. Of course, for values of $z$ between 0 and 1, there is no difference.)

Euler obtained the values

$$L(1) = \pi^2/6, \quad \text{and} \quad L(\tfrac{1}{2}) = \pi^2/12,$$

and Landen evaluated

$$L(\tfrac{1}{2}(\sqrt{5}-1)) = \pi^2/10, \quad \text{and} \quad L(\tfrac{1}{2}(3-\sqrt{5})) = \pi^2/15.$$

No other elementary evaluations of the dilogarithm are known. However, there are many strange relations such as those found by Watson and involving the roots of the cubic $x^3+2x^2-x-1$. If the roots of this cubic are denoted by $\alpha$, $-\beta$ and $-1/\gamma$, so that $\alpha = \frac{1}{2}\sec\frac{2\pi}{7}$, $\beta = \frac{1}{2}\sec\frac{\pi}{7}$ and $\gamma = 2\cos\frac{3\pi}{7}$

all lie between 0 and 1, then

$$L(\alpha) - L(\alpha^2) = \pi^2/42$$
$$2L(\beta) + L(\beta^2) = 5\pi^2/21$$
$$2L(\gamma) + L(\gamma^2) = 4\pi^2/21 .$$

Lewin has observed that these relations for values of the dilogarithm follow a pattern. In each case, the algebraic number on which the relation is based satisfies a polynomial equation which Lewin has called a *cyclotomic equation*. This has the shape

$$\prod_r (1 - x^r)^{e_r} = x^{e_0}, \tag{13.1}$$

and the corresponding relation for the dilogarithm is

$$\sum_r \frac{e_r}{r} L(x^r) = c\pi^2, \tag{13.2}$$

where $c$ is a rational number. Usually, the range of the summation index $r$ is over the positive divisors of an integer $n$, but there are exceptions and no such restriction on $r$ is needed in the analysis. The term 'cyclotomic equation' is perhaps unfortunate since (13.1) is usually not the cyclotomic equation for a root of unity. However, the name is firmly entrenched in this context and cannot be avoided. The *index* of the cyclotomic equation (13.1) is the maximum $r$ occurring in the product.

Table 13.1 sets out the relations known to the author which have index at most 6 and which involve algebraic numbers of degree at most 3. In the final column of the table, the note $A$ indicates that the relation can be deduced from simple functional equations as in §13.3 below and **A** that it follows from more complicated functional equations, $P$ denotes that the relation can be obtained by asymptotic analysis of partition identities as in §13.4, and ? indicates that the relation has been verified numerically to high precision but that no analytic proof is known to the author. There are many further relations of all three types for algebraic numbers of higher degree or of higher index. (The table needs some qualification. I have omitted some simple relations which follow immediately from the functional equations (13.3) and (13.4). Thus, $\theta$ satisfies $1 - \theta = \theta^3$ which gives 2 simple relations and the entry in Table 13.1 leads to a relation for $\theta^2$. Similarly, $\omega$ satisfies $1 - \omega^2 = \omega^3$ and $1 - \omega = \omega^5$ which gives 3 simple relations and the entry in Table 1 leads to a relation for $\omega^3$.)

The discovery of the relations involving the 3 bases $\frac{1}{2}\sec\frac{\pi}{9}$, $\frac{1}{2}\sec\frac{2\pi}{9}$ and $2\cos\frac{4\pi}{9}$ attached to the cubic $x^3 + 3x^2 - 1$ is amusing. The first was found by the asymptotic analysis to be described later. Lewin observed the parallel with Watson's identities and conjectured the second and third relations of the triple. The second was immediately proved by the asymptotic method. At this stage, the third was a numerical fact and for some time had no analytic



| Base | Coefficients | | | | | | | Discoverer | Proof |
|---|---|---|---|---|---|---|---|---|---|
| $x$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_4$ | $e_6$ | $c$ | | (see text) |
| $1$ | 1 | | | | | | $\frac{1}{6}$ | Euler | *A*, *P* |
| $\frac{1}{2}$ | 1 | | | | | | $\frac{1}{12}$ | Euler | *A*, *P* |
| $\frac{1}{3}$ | 3 | −1 | | | | | $\frac{1}{6}$ | Lewin [2] | *A*, *P* |
| $\rho = \frac{1}{2}(\sqrt{5} - 1)$ | 1 | | | | | | $\frac{1}{10}$ | Landen | *A*, *P* |
| $\rho^2 = \frac{1}{2}(3 - \sqrt{5})$ | 1 | | | | | | $\frac{1}{15}$ | Landen | *A*, *P* |
| $\rho^3 = \sqrt{5} - 2$ | 2 | −1 | | | | | $\frac{1}{12}$ | Landen | **A** |
| $\rho^4 = \frac{1}{2}(7 - 4\sqrt{5})$ | 7 | −6 | 4 | | | −3 | $\frac{7}{36}$ | Coxeter | **A** |
| $\sqrt{2} - 1$ | 2 | −1 | | | | | $\frac{1}{8}$ | Lewin [2] | *A*, *P* |
| " | 1 | 2 | | −1 | | | $\frac{5}{48}$ | Lewin [2] | *A*, *P* |
| $3 - 2\sqrt{2}$ | 5 | −2 | | | | | $\frac{1}{6}$ | Lewin [2] | **A** |
| $\frac{1}{2}(\sqrt{3} - 1)$ | 2 | 1 | −1 | | | | $\frac{5}{36}$ | Loxton [6] | **A**, *P* |
| $\sqrt{3} - 1$ | 2 | −3 | −1 | | | 1 | $\frac{1}{12}$ | Lewin [3] | **A** |
| $2 - \sqrt{3}$ | 4 | 1 | | −1 | | | $\frac{5}{24}$ | Lewin [2] | **A** |
| " | 5 | −3 | −1 | | | 1 | $\frac{2}{9}$ | Lewin [2] | **A** |
| $5 - 2\sqrt{6}$ | 23 | −15 | −3 | | | 3 | $\frac{1}{2}$ | Lewin [3] | ? |
| $\frac{1}{2}(\sqrt{13} - 3)$ | 4 | −2 | −2 | | | 1 | $\frac{7}{36}$ | Lewin [3] | ? |
| $\frac{1}{6}(\sqrt{13} - 1)$ | 3 | 1 | −3 | | | 1 | $\frac{2}{9}$ | Browkin [5] | ? |
| $\frac{1}{6}(\sqrt{13} + 1)$ | 3 | −4 | −3 | | | 2 | $\frac{1}{9}$ | Browkin [5] | ? |
| $4 - \sqrt{15}$ | 15 | 2 | −3 | −2 | | | $\frac{5}{12}$ | Lewin [3] | ? |
| $\frac{1}{2}(5 - \sqrt{21})$ | 7 | −1 | −3 | | | 1 | $\frac{5}{18}$ | Lewin [3] | ? |
| $\frac{1}{2}\sec\frac{2\pi}{7}$ | 1 | −2 | | | | | $\frac{1}{42}$ | Watson | *A*, *P* |
| $\frac{1}{2}\sec\frac{\pi}{7}$ | 1 | 1 | | | | | $\frac{5}{42}$ | Watson | *A*, *P* |
| $2\cos\frac{3\pi}{7}$ | 1 | 1 | | | | | $\frac{2}{21}$ | Watson | *A*, *P* |
| $\frac{1}{2}\sec\frac{\pi}{9}$ | 1 | 2 | −1 | | | | $\frac{7}{54}$ | Loxton [6] | *P* |
| $\frac{1}{2}\sec\frac{2\pi}{9}$ | 1 | −3 | −1 | | | 1 | $-\frac{1}{54}$ | Lewin [2] | *P* |
| $2\cos\frac{4\pi}{9}$ | 1 | −3 | −1 | | | 1 | $\frac{1}{54}$ | Lewin [2] | |
| $2x + x^3 = 1$ | 1 | 5 | | −4 | | | $\frac{1}{6}$ | Lewin [2] | **A** |
| " | 3 | 1 | 12 | | | −6 | $\frac{1}{3}$ | Lewin [2] | **A** |
| $x + 2x^3 = 1$ | 2 | 1 | 3 | −2 | | | $\frac{1}{4}$ | Lewin [2] | **A** |
| $\theta: (\theta + \theta^3 = 1)$ | 2 | 6 | 3 | | | −3 | $\frac{1}{2}$ | Lewin [2] | **A** |
| $\theta^3$ | 5 | −9 | −6 | | | 6 | $\frac{1}{6}$ | Lewin [2] | **A** |
| $\omega^2: (\omega^2 + \omega^3 = 1)$ | 1 | 6 | 6 | | | −6 | $\frac{1}{3}$ | Lewin [2] | **A** |
| $x + x^2 + x^3 = 1$ | 1 | 1 | −3 | | | | $\frac{1}{12}$ | Lewin [2] | **A** |
| " | 2 | 3 | | −2 | | | $\frac{1}{4}$ | Lewin [2] | **A** |

TABLE 13.1. $\sum_r \frac{e_r}{r} L(x^r) = c\pi^2$.

derivation. It has recently been proved by H. Gangl, a student of D. Zagier in Bonn, using methods based on Bloch groups. As this story illustrates, there are a number of methods available for finding and proving these relations, but an ultimate straightforward uniform procedure has yet to be discovered.

**13.2. Cyclotomic equations.**

Although it has not been proved, it appears that the base $x$ of each dilogarithm relation of the shape (13.2) must satisfy a cyclotomic equation. Such an equation imposes a considerable restriction on the base of the relation, but it is not easy to see just which algebraic numbers are the roots of cyclotomic equations. The following remarks may be of interest, but they do not go very far towards resolving this problem. Similar results have recently been obtained by G. A. Ray [7] and more detailed proofs can be found there.

Probably, there are only finitely many algebraic numbers of each degree which satisfy cyclotomic equations. Towards this, we can show that the index of the cyclotomic equation is bounded in terms of the degree of the algebraic number.

Suppose that $x$ is an algebraic number of degree $d$ and satisfies the cyclotomic equation

$$\prod_r (1-x^r)^{e_r} = x^{e_0}.$$

Write $x = A/B$, where $A$ and $B$ are relatively prime algebraic integers. (The integers $A$ and $B$ may be in a larger field.) Thus

$$\prod_r (B^r - A^r)^{e_r} = A^{e_0} B^{\sum r e_r - e_0}.$$

But, $B^n - A^n$ has a primitive divisor when $n$ is sufficiently large relative to $d$, that is, a divisor which does not divide $A^m - B^m$ for $0 < m < n$. (See Stewart [10].) Consequently, the previous equation is impossible when its index, $n$ say, is sufficiently large relative to $d$ because the factor $B^n - A^n$ has a prime divisor which cannot be cancelled by any of the other factors in the equation. Stewart [10] gives an explicit form of the primitive divisor theorem which makes this bound effective. Moreover, for $d = 1$, the explicit bound can be made small enough to carry the argument further.

The only rational numbers in the interval $(0, 1)$ which satisfy a cyclotomic equation are $\frac{1}{2}$ and $\frac{1}{3}$. Suppose $x = A/B$ satisfies a cyclotomic equation as above and $A$ and $B$ are relatively prime rational integers. In this case, $B^n - A^n$ has a primitive divisor for $n > 6$. The cases with $n \le 6$ can be analysed case by case by examining the prime factors of the various factors. Of course, this does not prove that there are no other nice relations for the dilogarithm at rational points, but it does provide some further evidence for this assertion.



**13.3. Accessible relations.**

13.3.1. *Elementary functional equations.* The function $L(z)$ satisfies the functional equations

$$L(z) + L(1-z) = \pi^2/6, \tag{13.3}$$

and

$$L(z) = L(z/(1+z)) + \tfrac{1}{2}L(z^2), \tag{13.4}$$

both of which can be readily verified by differentiation. It may be conjectured that all the relations between values of the dilogarithm should be explained by specialisation of functional equations. Indeed, this line of thinking has led Lewin [4] to several very complicated functional equations, but not even these seem to give all the known relations.

The relations which have been deduced from the two functional equations (13.3) and (13.4) are marked by $A$ in the preceding table. These might be called 'accessible', following Lewin. It is not at all easy to see what distinguishes the 'accessible' relations and the derivations can be quite complicated. We illustrate such an argument by deriving Watson's three relations cited in §13.1.

13.3.2. *Watson's relations.* Let $\alpha$, $-\beta$, and $-1/\gamma$ be the three roots of the cubic $x^3 + 2x^2 - x - 1 = 0$, so that $\alpha$, $\beta$ and $\gamma$ all lie between 0 and 1. Then, Watson's relations are

$$L(\alpha) - L(\alpha^2) = \pi^2/42, \qquad 2L(\beta) + L(\beta^2) = 5\pi^2/21,$$

and

$$2L(\gamma) + L(\gamma^2) = 4\pi^2/21.$$

To prove these relations, first note that

$$\alpha = \frac{1}{2}\sec\frac{2\pi}{7}, \qquad \beta = \frac{1}{2}\sec\frac{\pi}{7} = \frac{1}{1+\alpha},$$

and

$$\gamma = 2\cos\frac{3\pi}{7} = \frac{\alpha}{1+\alpha}.$$

Now, $\alpha$ satisfies the equation $\alpha^3 + 2\alpha^2 - \alpha - 1 = 0$, so we compute

$$\frac{1}{1+\beta} = \frac{1+\alpha}{2+\alpha} = \alpha^2,$$

$$\frac{\gamma}{1+\gamma} = \frac{\alpha}{1+2\alpha} = \frac{1}{(1+\alpha)^2} = \beta^2.$$

From the second functional equation (13.4),

$$L(\alpha) = L(\gamma) + \tfrac{1}{2}L(\alpha^2), \qquad L(\beta) = L(1-\alpha^2) + \tfrac{1}{2}L(\beta^2),$$

and

$$L(\gamma) = L(\beta^2) + \tfrac{1}{2}L(1-\alpha),$$

and Watson's relations follow by using the first of the functional equations and some easy elimination.

Despite the similarity with the relations involving the trigonometric functions of multiples of $\pi/9$, I have been unable to deduce the relations involving these quantities from the functional equations.

The functional equation armoury contains a number of more complicated equations such as Abel's two-variable equation and the recently discovered equations of Lewin [4] and Ray [7]. These have been used to obtain the entries marked **A** in Table 13.1. In all these cases, the proofs yield the relevant cyclotomic equations and the conjugate relations, if any. (Watson's 3 relations are an example of a set of conjugate relations.) A proof that this natural property holds in general is given in Chapter 7.

**13.4. Partition identities.**

13.4.1. *Some identities of Slater.* Richmond and Szekeres [8] observed that values of the dilogarithm appear in the course of asymptotic analysis of partition identities. Sometimes, this leads to relations of the type considered above and several of the entries in the table can be verified by this method. These are marked by $P$ in Table 1. We illustrate the method by sketching the derivation of the identities involving $\frac{1}{2}\sec\frac{\pi}{9}$ and $\frac{1}{2}\sec\frac{2\pi}{9}$.

The asymptotic analysis is based on certain identities which express basic hypergeometric series as infinite products. Slater [9] gives a list of 130 identities of this type, on which the analysis in [6] was based. In stating these, we use the standard abbreviations

$$(a)_n = (a;q)_n = (1-a)(1-aq)\cdots(1-aq^{n-1}),$$

for $n$ a positive integer, and

$$(a)_\infty = (a;q)_\infty = \lim_{n\to\infty}(a;q)_n .$$

Identity 92 from Slater's list is

$$\sum_{n=0}^{\infty}\frac{(q^3;q^3)_n q^{n(n+1)}}{(q)_n(q)_{2n+1}} = \frac{(q^9;q^{27})_\infty(q^{18};q^{27})_\infty(q^{27};q^{27})_\infty}{(q)_\infty}. \tag{13.5}$$

Similar methods yield another identity not in Slater's list:

(13.6)
$$1+\sum_{n=1}^{\infty}\frac{(-q^3;q^3)_{n-1}q^{n(n+1)}}{(-q)_{n-1}(q)_{2n}} = \frac{(q;q^9)_\infty(q^8;q^9)_\infty(q^9;q^9)_\infty(q^7;q^{18})_\infty(q^{11};q^{18})_\infty}{(q)_\infty}.$$

The derivation of this last identity and the other relevant identities are given in [6].

All these identities involving basic hypergeometric series are descendants of the famous Rogers-Ramanujan identity

$$1+\sum_{n=1}^{\infty}\frac{q^{n^2}}{(q)_n} = \frac{1}{(q;q^5)_\infty(q^4;q^5)_\infty}.$$



They have combinatorial interpretations, some more transparent than others. For example, the Rogers-Ramanujan identity amounts to the following statement: the number of partitions of $n$ whose parts have minimal difference 2 is equal to the number of partitions into parts congruent to 1 or 4 modulo 5.

13.4.2. *Statement of the theorem.* The power series coefficients of the basic hypergeometric series occurring in these identities can be found by applying the circle method in the following way. Consider the power series

$$\sum_{n=0}^{\infty} q^{n(an+b)/2} \Big/ \prod_{j=1}^{r}(q^{c_j};q^{d_j})_n^{e_j} = \sum_{k=0}^{\infty} a_k q^k , \tag{13.7}$$

where $a$, $b$, $c_j$, $d_j$ and $e_j$ are integers satisfying the conditions $a \geq 0$, $b > 0$ if $a = 0$, $a \equiv b \pmod 2$, and $d_j > 0$. Suppose also that the power series expansion of each of the products $\prod_{j=1}^{r}(q^{c_j};d^{d_j})_n^{-e_j}$ has positive coefficients and that the equation

$$\prod_{j=1}^{r}(1-x^{d_j})^{e_j} = x^a \tag{13.8}$$

has a unique root $\xi$ between 0 and 1. Then, as $k\to\infty$,

$$\frac{(\log a_k)^2}{4k} \to \sum_{j=1}^{r}\frac{e_j}{d_j}L(1-\xi^{d_j}). \tag{13.9}$$

13.4.3. *The circle method.* We sketch the proof of the above assertion. The argument follows the method developed by Szekeres [11]. Consider first the power series

$$q^{n(an+b)/2} \Big/ \prod_{j=1}^{r}(q^{c_j};q^{d_j})_n^{e_j} = \sum_{k=0}^{\infty} a_{kn} q^k .$$

The $a_{kn}$ can be obtained by Cauchy's formula applied to a suitably chosen circle, say $q = e^{-\beta+i\theta}$ $(-\pi \leq \theta \leq \pi)$, with radius $e^{-\beta}$ less than 1. Thus

$$\begin{aligned} a_{kn} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp\Big\{ &-\sum_{j=1}^{r}\sum_{\nu=0}^{n-1} e_j \log(1-e^{-(\beta-i\theta)(d_j\nu+c_j)}) \\ &+(k-\tfrac{1}{2}n(an+b))(\beta-i\theta)\Big\}\, d\theta . \end{aligned} \tag{13.10}$$

The radius, $e^{-\beta}$, of the circle of integration is determined by the saddle-point condition which is obtained by setting the logarithmic derivative of the integrand with respect to $\beta$ equal to 0:

$$\sum_{j=1}^{r}\sum_{\nu=0}^{n-1}\frac{e_j(d_j\nu+c_j)}{e^{\beta(d_j\nu+c_j)}-1} = k-\tfrac{1}{2}n(an+b). \tag{13.11}$$

Split the integral (13.10) for $a_{kn}$ into two pieces: the major arc described by $|\theta| \le \theta_0 = k^{-5/7}$ and the complementary minor arc. On the major arc, the logarithm, $f(\theta)$ say, of the integrand can be approximated by a Taylor expansion in powers of $\theta$. The saddle point equation (13.11) means that the coefficient of $\theta$ is 0. The result is

$$f(\theta) = (k - \tfrac{1}{2}n(an+b))\beta - \sum_{j=1}^{r}\sum_{\nu=0}^{n-1} e_j \log(1 - e^{-\beta(d_j\nu + c_j)}) - \frac{A\theta^2}{2\beta^3} + O\left(\frac{\theta^3}{\beta^4}\right),$$

where

$$A = \sum_{j=1}^{r} \frac{e_j}{d_j} \int_0^{d_j u} \frac{t^2 e^t}{(e^t - 1)^2}\, dt$$

and $u = \beta n$. The integral representation for $A$ comes about by applying the Euler-Maclaurin formula to the sum

$$\frac{d^2}{d\theta^2}\left\{-\sum_{\nu=0}^{n-1} \log(1 - e^{-(\beta - i\theta)(d_j\nu + c_j)})\right\} = -\sum_{\nu=0}^{n-1} \frac{(d_j\nu + c_j)^2 e^{\beta(d_j\nu + c_j)}}{(e^{\beta(d_j\nu + c_j)} - 1)^2}.$$

Consequently, the contribution to $\log a_{kn}$ from the major arc is

$$\begin{aligned}
&\log \frac{1}{2\pi}\int_{-\theta_0}^{\theta_0} \exp f(\theta)\, d\theta \\
&= \left(k - \frac{1}{2}n(an+b)\right)\beta - \sum_{j=1}^{r}\sum_{\nu=0}^{n-1} e_j \log(1 - e^{-\beta(d_j\nu + c_j)}) \\
&\quad + \log \frac{1}{2\pi}\int_{-\theta_0}^{\theta_0} \exp\left(-\frac{A\theta^2}{2\beta^3} + O\left(\frac{\theta^3}{\beta^4}\right)\right) d\theta \\
&= \left(k - \frac{1}{2}n(an+b)\right)\beta - \sum_{j=1}^{r}\sum_{\nu=0}^{n-1} e_j \log(1 - e^{-\beta(d_j\nu + c_j)}) \\
&\quad + \frac{1}{2}\log(\beta^3/2\pi A) + o(1),
\end{aligned} \tag{13.12}$$

since $\beta \approx k^{-1/2}$ (as shown below) and so $\theta_0^2/\beta^3 \approx k^{1/14} \to \infty$ and $\theta_0^3/\beta^4 \approx k^{-1/7} \to 0$ as $k \to \infty$. The minor arc contributes $o(1)$ to $\log a_{kn}$ as shown in [**11**]. By the Euler-Maclaurin formula applied to the saddle-point condition (13.11),

$$\frac{1}{\beta^2}\sum_{j=1}^{r} \frac{e_j}{d_j}\int_0^{d_j u} \frac{t}{e^t - 1}\, dt = k - \frac{au^2}{2\beta^2} + O\left(\frac{1}{\beta}\right),$$

that is

$$k = \frac{1}{\beta^2}\left\{\sum_{j=1}^{r} \frac{e_j}{d_j}\int_0^{d_j u} \frac{t}{e^t - 1}\, dt + \frac{1}{2}au^2\right\} + O\left(\frac{1}{\beta}\right), \tag{13.13}$$



so that $\beta$ is of order $k^{-1/2}$. Again, applying the Euler-Maclaurin formula, an integration by parts, and Stirling's formula,

$$\begin{aligned}
\sum_{\nu=1}^{n-1} \log(1 - e^{-\beta(d_j\nu + c_j)}) &= \sum_{\nu=1}^{n-1}\left(\log \frac{1 - e^{-\beta(d_j\nu + c_j)}}{\beta d_j \nu} + \log \beta d_j \nu\right) \\
&= \frac{1}{\beta d_j}\int_0^{d_j u} \log \frac{1 - e^{-t}}{t}\, dt + n\log \beta d_j + \log n! + o(\beta^{-1}) \\
&= n\log \frac{1 - e^{-d_j u}}{d_j u} + n - \frac{1}{\beta d_j}\int_0^{d_j u} \frac{t}{e^t - 1}\, dt - n\log d_j u - n + o(\beta^{-1}),
\end{aligned}$$

and, by (13.11) again, the formula (13.12) for $\log a_{kn}$ simplifies to give

$$\log a_{kn} = \frac{1}{\beta}\sum_{j=1}^{r}\left\{\frac{2e_j}{d_j}\int_0^{d_j u} \frac{t}{e^t - 1}\, dt - e_j u \log(1 - e^{-d_j u})\right\} + o(k^{1/2}). \tag{13.14}$$

The next step is to determine $n$ so that this expression for $\log a_{kn}$ is maximal. Let the difference operator $\Delta$ denote differences formed by moving from $n$ to $n+1$. Applying $\Delta$ to the saddle point equation (13.11) leads to

$$\sum_{j=1}^{r} \frac{e_j(d_j n + c_j)}{e^{\beta(d_j n + c_j)} - 1} + \sum_{j=1}^{r}\sum_{\nu=0}^{n-1} \frac{e_j(d_j\nu + c_j)^2 e^{d_j\nu + c_j}}{(e^{\beta(d_j\nu + c_j)} - 1)^2}\Delta\beta \sim -an,$$

and, after another application of Euler-Maclaurin, this gives an expression for $\Delta\beta$, namely

$$\Delta\beta = \beta^3\left\{\sum_{j=1}^{r} \frac{e_j(d_j n + c_j)}{e^{\beta(d_j n + c_j)} - 1} + an\right\} \Big/ \left\{\sum_{j=1}^{r} \frac{e_j}{d_j}\int_0^{d_j u} \frac{t^2 e^t}{(e^t - 1)^2}\, dt\right\} + o(k^{-1}).$$

Thus, $\Delta\beta$ is of order $\beta^2$ and this justifies the various approximations made in the calculations. By applying $\Delta$ to the formula (13.12) for $\log a_{kn}$ and invoking (13.11), we get

$$\Delta \log a_{kn} = -\sum_{j=1}^{r} e_j \log(1 - e^{-d_j u}) - au + o(1).$$

The condition for $\log a_{kn}$ to be maximal is that this last expression should vanish, that is

$$\sum_{j=1}^{r} e_j \log(1 - e^{-d_j u}) = -au. \tag{13.15}$$

Set $\xi = e^{-u}$. Since $0 < \xi < 1$, the hypotheses and (13.15) fix $\xi$ as the root of the cyclotomic equation (13.8). For $n$ at the maximum, (13.13) and (13.14) yield

$$\log a_{kn} = 2k^{1/2}\left\{\sum_{j=1}^{r} \frac{e_j}{d_j}\int_0^{d_j u} \frac{t}{e^t - 1}\, dt - \frac{1}{2}e_j u \log(1 - e^{-d_j u})\right\}^{1/2} + o(k^{1/2})$$

in which

$$\int_0^{d_j u} \frac{t}{e^t - 1}\,dt - \frac{1}{2} d_j u \log(1 - e^{-d_j u}) = L(1 - \xi^{d_j}).$$

Finally, $a_k = \sum a_{kn}$, where the sum over $n$ has at most $k$ terms, all of them positive, so that

$$\log a_k = 2k^{1/2} \left\{ \sum_{j=1}^{r} \frac{e_j}{d_j} L(1 - \xi^{d_j}) \right\}^{1/2} + o(k^{1/2}),$$

which is the assertion (13.9).

13.4.4. *An application of the theorem.* By way of application, we carry out the asymptotic analysis of Slater's identity 92 cited earlier in this section. The identity can be rewritten in the form

$$\text{(13.15a)} \quad \sum_{n=0}^{\infty} \frac{(q^3; q^3)_n q^{n(n+1)}}{(q)_n (q; q^2)_n (q^2; q^2)_n} = \frac{(q^9; q^{27})_\infty (q^{18}; q^{27})_\infty (q^{27}; q^{27})_\infty}{(q)_\infty}.$$

Let $\sum a_k q^k$ be the power series expansion side of either side of this identity. The cyclotomic equation corresponding to the left-hand side of (13.15a) according to (13.8) is

$$\frac{(1-x)(1-x^2)^2}{(1-x^3)} = x^2,$$

that is, $x^3 + 3x^2 - 1 = 0$. Let $\kappa$ be the real root between 0 and 1 of this equation, so that $\kappa = \frac{1}{2}\sec \pi/9$. Our result gives

$$\text{(13.16)} \quad \begin{aligned}(\log a_k)^2 &\sim 4k \left\{ L(1-\kappa) + L(1-\kappa^2) - \frac{1}{3} L(1-\kappa^3) \right\} \\ &\sim -4k \left\{ L(\kappa) + L(\kappa^2) - \frac{1}{3} L(\kappa^3) - \frac{5\pi^2}{18} \right\},\end{aligned}$$

as $k \to \infty$. The estimate for $\log a_k$ from the right-side of the identity follows easily from the well-known estimate for the partition function; namely, if

$$\prod_{n=1}^{\infty} (1 - q^n)^{-1} = \sum_{k=0}^{\infty} p(k) q^k,$$

then

$$(\log p_k)^2 \sim \frac{2\pi^2 k}{3},$$

as $k \to \infty$. The right side of the identity (13.15a) is obtained by omitting 3 residue classes modulo 27 from the last infinite product, whence

$$\text{(13.17)} \quad (\log a_k)^2 \sim \frac{24}{27} (\log p(k))^2 \sim \frac{16\pi^2 k}{27},$$

and, finally, by equating the two estimates (13.16) and (13.17),

$$\text{(13.18)} \quad L(\kappa) + L(\kappa^2) - \frac{1}{3} L(\kappa^3) = \frac{7\pi^2}{54}.$$



The working for the relation involving $\frac{1}{2}\sec\frac{2\pi}{9}$ is given in [6].

Asymptotic analysis of partition identities in this fashion has a number of useful consequences. For example, we were able to pick up 2 misprints in Slater's list in [9] because the identities stated led to false equations for the dilogarithm. Unfortunately, there is no automatic path in the opposite direction. A cyclotomic equation and the putative dilogarithm relation associated with it contain some hints about the shape of a possible partition identity, but there seems to be rather too much uncertainty, foiling an attempt to discover the required partition identity by computer algebra. Some luck is required. For example, consider the 3 conjugate bases $\kappa = \frac{1}{2}\sec\frac{\pi}{9}$, $\lambda = \frac{1}{2}\sec\frac{2\pi}{9}$ and $\mu = 2\cos\frac{4\pi}{9}$. Now, $\kappa$ is a root of the equation $x^3 + 3x^2 - 1 = 0$ and we have derived its dilogarithm relation (13.18) from the partition identity (13.5). The equation for $\lambda$ is $x^3 - 3x^2 + 1 = 0$, obtained by substituting $-x$ for $x$ in the previous equation. Substituting $-q$ for $q$ in (13.5) gives (13.6), after a certain amount of manipulation, and this leads successfully to the dilogarithm relation

$$6L(\lambda) - 9L(\lambda^2) - 2L(\lambda^3) = -\frac{\pi^2}{9}.$$

The equation for $\mu$ is $x^3 - 3x + 1 = 0$, obtained by substituting $-1/x$ for $x$ in the equation for $\kappa$. The obvious attempt, substituting $-1/q$ for $q$ in (5) leads to the partition identity

$$1 - \sum_{n=1}^{\infty} \frac{(-q^3; q^3)_{n-2} q^{n-1} (1 - q^{n-1} - q^n - q^{n+1} + q^{3n})}{(-q)_{n-2} (q)_{2n}} = \frac{(q; q^6)_\infty (q^5; q^6)_\infty}{(q)_\infty}$$

of a similar type to the others, but the mixed signs in the numerator on the left-hand side prevent an application of Theorem 13.4.2. In fact, the expected leading term from the left-hand side cancels and the asymptotic analysis is insufficient to give the next term.

**13.5. Generalisations and extensions.**

The identities considered in §13.4 form one family of generalisations of the Rogers-Ramanujan identities. Another notable generalisation is the multidimensional identity of Andrews and Gordon,

$$\sum_{n_1, n_2, \ldots, n_r \geq 0} \frac{q^{\sum_{i=1}^{r} N_i^2}}{\prod_{i=1}^{r} \prod_{j=1}^{n_i} (1 - q^j)} = \prod_{n \not\equiv 0, \pm(r+1) \bmod 2r+3} (1 - q^n)^{-1}.$$

This is the starting point for the analysis of Richmond and Szekeres [8], referred to in §13.4. Comparison of the asymptotic expansions for the power series coefficients of the two sides of the identity as in the previous section yields the following theorem. Define $\delta_j$ by the recurrence

$$\delta_1 = 2\left(1 - \cos\frac{\pi}{2r+3}\right),$$

$$\delta_{j+1} = \delta_j / D_j^2, \qquad j = 1, 2, \ldots, r-1,$$

where

$$D_j = \prod_{i=1}^{j}(1-\delta_i), \qquad j = 1, 2, \ldots, r.$$

Then

$$\sum_{j=1}^{r} L(\delta_j) = \frac{\pi^2 r}{3(2r+3)}.$$

If we let $r \to \infty$ in this identity, we obtain the amusing formula

$$\sum_{k=2}^{\infty} L\left(\frac{1}{k^2}\right) = \frac{\pi^2}{6}.$$

Lewin [2] has given direct proofs of these identities based on the functional equations for the dilogarithm. Another family of identities of this type is given in [8] and some more complicated ones based on further identities of Andrews and others were found by D. Acreman in his thesis.

The asymptotic method has many other applications. For example, consider the following remarkable assertion from Ramanujan's first notebook:

$$\begin{aligned}
&\frac{1}{(1-x^2)(1-x^3)(1-x^5)(1-x^7)(1-x^{11})(1-x^{13})(1-x^{17})\&c} \\
&\quad = 1 + \frac{x^2}{1-x} + \frac{x^5}{(1-x)(1-x^2)} + \frac{x^{10}}{(1-x)(1-x^2)(1-x^3)} \\
&\qquad + \frac{x^{17}}{(1-x)(1-x^2)(1-x^3)(1-x^4)} + \&c.
\end{aligned}$$

It would be astounding if the sequence of primes fitted such an identity and, in fact, the purported identity fails. It is still surprising that the coefficients agree up to the terms in $x^{20}$. The identity occurs again in Ramanujan's second notebook and is crossed out, so that Ramanujan must have known that it was false. Andrews raised the question of finding 'Ramanujan pairs', that is pairs of sequences $\{a_n\}$ and $\{b_n\}$ such that

$$\prod_{n\geq 1}(1-q^{a_n})^{-1} = 1 + \sum_{n\geq 1}\frac{q^{b_1+b_2+\cdots+b_n}}{(1-q)(1-q^2)\cdots(1-q^n)}.$$

With Acreman in [1], we showed that such an identity imposes tough restrictions on the sequences $\{a_n\}$ and $\{b_n\}$. For example, if $a_n \sim b_n$ and $a_n$ and $b_n$ are regularly varying, then $a_n$ and $b_n$ must be asymptotic to arithmetic progressions. In particular, a 'Ramanujan pair' in which the sequences both grow like the prime numbers is impossible.

There are more interesting and more delicate applications of the method, particularly by Richmond and Szekeres. In the present context, it would be very interesting to discover some of the relations for the higher order polylogarithms from partition identities, but so far none have been forthcoming.



## REFERENCES


1. D. Acreman and J. H. Loxton, *Asymptotic analysis of Ramanujan pairs*, Aequationes Math. **30** (1986), 106–117.
2. L. Lewin, *The dilogarithm in algebraic fields*, J. Austral. Math. Soc. Ser. A **33** (1982), 302–330.
3. ——, *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), 345–373.
4. ——, *The order independence of the polylogarithmic ladder structure—implications for a new category of functional equations*, Aequationes Math. **30** (1986), 1–20.
5. L. Lewin and M. Abouzahra, *Supernumary polylogarithmic ladders and related functional equations*, Aequationes Math. **39** (1990), 210–253.
6. J. H. Loxton, *Special values of the dilogarithm function*, Acta Arith. **43** (1984), 155–166.
7. G. A. Ray, *Multivariable polylogarithm identities*, see Chapter 7 of this volume.
8. B. Richmond and G. Szekeres, *Some formulas related to dilogarithms, the zeta function and the Rogers-Ramanujan identities*, J. Austral. Math. Soc. Ser. A **31** (1981), 362–373.
9. L. J. Slater, *Further identities of the Rogers-Ramanujan type*, Proc. London Math. Soc. (2) **54** (1951), 147–167.
10. C. L. Stewart, *Primitive divisors of Lucas and Lehmer numbers*, Transcendence theory: advances and applications (A. Baker and D. W. Masser eds.) Academic Press, New York, 1977, 79–92.
11. G. Szekeres, *Some asymptotic formulae in the theory of partitions* (II), Quart. J. Math. Oxford Ser. (2) **4** (1953), 96–111.

CHAPTER 14

# The Dilogarithm and Volumes of Hyperbolic Polytopes

RUTH KELLERHALS

**14.0. Introduction.** In this chapter, we consider the Euler dilogarithm $\mathrm{Li}_2(z)$ in connection with the problem of calculating volumes of noneuclidean polytopes. In contrast to the euclidean case, where the volume of an arbitrary simplex $S \subset E^n$, $n \geq 3$, spanned by vectors $p_0, \ldots, p_n$, is given by the "elementary" formula

$$\mathrm{vol}_n(S) = \frac{1}{n!} |\det(p_0, \ldots, p_n)|, \tag{14.1}$$

the corresponding volume problem for noneuclidean $n$-simplexes is considerably more difficult, and—in full generality—an unsolved problem.

However, in 1852, Schläfli proved a very beautiful formula for the volume differential $d\mathrm{vol}_n$ on the set of spherical $n$-simplexes $S$ [27, p. 227]:

$$d\mathrm{vol}_n(S) = \frac{1}{n-1} \sum_{1 \leq j < k \leq n} \mathrm{vol}_{n-2}(S_j \cap S_k)\, d\alpha_{jk}, \qquad \mathrm{vol}_0 := 1, \tag{14.2}$$

where $\alpha_{jk}$ is the dihedral angle formed by the facets $S_j$, $S_k$ of $S$. This formula, the three-dimensional hyperbolic version of which was already known to Lobachevsky, was very elegantly reproved and extended to the hyperbolic case by H. Kneser (see [16] and [4, §5.1]). However, the remaining single integration cannot be carried out even in the simplest case of a three-dimensional simplex. Consequently, one has to look for polyhedral objects whose geometry allows one to simplify the last but most difficult step.

We consider the class of $d$-truncated orthoschemes $R_d \subset H^n$, $0 \leq d \leq 2$, $n \geq 2$, which are convex polytopes bounded by $n + d + 1$ hyperplanes $H_0, \ldots, H_{n+d}$ such that

$$H_j \perp H_k \quad \text{for } 2 \leq |j - k| \leq n,$$

(for $d = 2$, indices are taken modulo $n + 3$). For $d = 0$, these polytopes are the ordinary orthoschemes $R$, first introduced by Schläfli. They are determined (up to isometry) by their $n$ nonright dihedral angles.

For spherical orthoschemes, Schläfli derived a variety of results concerning his volume function $f_n$, which is proportional to $\mathrm{vol}_n$ in such a way that

$f_n = 1$ for a totally right-angled $n$-orthoscheme. Independently of him, in 1836, Lobachevsky expressed the volume of a three-dimensional hyperbolic orthoscheme $R = R(\alpha_1, \alpha_2, \alpha_3)$ in the form (see [4, (18), p. 250])

(14.3)

$$\begin{aligned}\operatorname{vol}_3(R) = \frac{1}{4}\Big\{&Л(\alpha_1+\theta) - Л(\alpha_1-\theta) + Л\left(\frac{\pi}{2}+\alpha_2-\theta\right) + Л\left(\frac{\pi}{2}-\alpha_2-\theta\right)\\ &+Л(\alpha_3+\theta) - Л(\alpha_3-\theta) + 2Л\left(\frac{\pi}{2}-\theta\right)\Big\},\end{aligned}$$

where $Л(\omega) := -\int_0^\omega \log|2\sin t|dt$ denotes the Lobachevsky function, related to the dilogarithm by

$$Л(\omega) = \tfrac{1}{2}\operatorname{Im}(\operatorname{Li}_2(e^{2i\omega})), \qquad i^2 = -1,$$

and where

$$0 \le \theta := \arctan\frac{\sqrt{\cos^2\alpha_2 - \sin^2\alpha_1\sin^2\alpha_3}}{\cos\alpha_1\cos\alpha_3} < \frac{\pi}{2}.$$

About 1935, Coxeter [8] reformulated and combined the results of Lobachevsky for hyperbolic and of Schläfli for spherical orthoschemes by introducing the function

(14.4)

$$S(\alpha_1, \alpha_2, \alpha_3) := \sum_{r=1}^{\infty}\frac{(-X)^r}{r^2}(\cos 2r\alpha_1 - \cos 2r\alpha_2 + \cos 2r\alpha_3 - 1) - \alpha_1^2 + \alpha_2^2 - \alpha_3^2,$$

where

$$X = \frac{\sin\alpha_1\sin\alpha_3 - D}{\sin\alpha_1\sin\alpha_3 + D} \quad \text{with } D = \sqrt{\cos^2\alpha_1\cos^2\alpha_3 - \cos^2\alpha_2}.$$

He showed that

$$S\left(\frac{\pi}{2}-\alpha_1, \alpha_2, \frac{\pi}{2}-\alpha_3\right) = \frac{\pi^2}{2}f_3(\alpha_1, \alpha_2, \alpha_3),$$

and, for hyperbolic orthoschemes $R = R(\alpha_1, \alpha_2, \alpha_3)$, for which $D^2 < 0$,

$$\frac{i}{4}S\left(\frac{\pi}{2}-\alpha_1, \alpha_2, \frac{\pi}{2}-\alpha_3\right) = \operatorname{vol}_3(R).$$

In 1962, Böhm [4, §5] analyzed Coxeter's method and generalized it to spaces of nonzero constant curvature of arbitrary dimension without, however, solving the higher dimensional volume problem. We shall see that the Böhm-Coxeter method is even applicable to $d$-truncated orthoschemes in $H^3$ and that the volume formula (14.3) remains valid (up to a minor modification in one case).

This chapter is organized as follows: In §14.1, we collect some basic material about real hyperbolic space. Then, we introduce the notion of schemes due to Schläfli and Vinberg [29, 30] to describe polytopes with many right dihedral angles and, in particular, orthoschemes of degree (of truncation) $d$. In §14.2, we discuss the volume problem for $d$-truncated orthoschemes in $H^3$ (cf. [14]). For this, a "schematic" version of Schläfli's differential formula is presented for the polytopes under consideration. By Schläfli's differential formula, one can see that there is a fundamental difference in the volume problems of even and odd dimensions; the Reduction formula in §14.2.2 shows that the first problem can be reduced to the second one (cf. [15])! Next, we generalize the integration method of Böhm and Coxeter to derive explicit volume formulae for $d$-truncated orthoschemes that are analogous to (14.3) (Theorem 14.5 and Theorem 14.6). In §14.3, we discuss some applications; in particular, the volumes of all Coxeter orthoschemes of degree $d$, forming fundamental polyhedra for hyperbolic Coxeter groups, are determined. We append the corresponding list for the ten ordinary Coxeter orthoschemes. By means of dissection into (truncated) orthoschemes, we calculate the volumes of the totally asymptotic regular simplexes in $\overline{H^n}$ for $n = 2, 3, 4, 6$, as well as the volumes of the four totally asymptotic regular polyhedra, the tetrahedron $S^\infty_{\text{reg}}(\frac{\pi}{3})$, the hexahedron $H^\infty_{\text{reg}}(\frac{\pi}{3})$, the octahedron $O^\infty_{\text{reg}}(\frac{\pi}{4})$ and the dodecahedron $D^\infty_{\text{reg}}(\frac{\pi}{3})$. Also by dissection, we derive some interesting functional equations for the Lobachevsky function $Л(\omega)$.

In §14.4, a few further aspects are considered. We survey results concerning small elements in the volume spectrum of hyperbolic 3-space forms. By a result of Borel, volumes of arithmetic 3-folds are computable in terms of Dedekind zeta functions, which we demonstrate with two examples. Finally, in §14.4.2, we give a very lay introduction to the fascinating circle of ideas around Hilbert's Third Problem concerning scissors congruence. Following the paper [9] of Dupont and Sah, we summarize definitions and properties of the different scissors congruence groups in hyperbolic space and describe how these groups admit a more general homological treatment. By the works of Bloch, Wigner and Dupont, Sah, we see at the end how geometrical notions such as volume and Dehn's invariant can be unified on the level of scissors congruence.

The author would like to thank Prof. Don Zagier for reading a preliminary version of this chapter and making some helpful suggestions.

**14.1. A particular class of hyperbolic polytopes.**

14.1.1. *Hyperbolic space.* Let $X^n$ denote either the $n$-dimensional euclidean space $E^n$, the $n$-sphere $S^n$ or the $n$-dimensional hyperbolic space $\overline{H^n}$. Let $S^n$ be embedded in $E^{n+1}$, and use for $H^n$ the model in the Lorentz space $E^{n,1}$ of signature $(n, 1)$, i.e., if $E^{n,1}$ denotes the $(n+1)$-dimensional real vector space $\mathbb{R}^{n+1}$, together with the bilinear form

$$\langle x, y\rangle := -x_0y_0 + x_1y_1 + \cdots + x_ny_n, \quad \forall x, y \in \mathbb{R}^{n+1},$$

of signature $(n, 1)$, then $H^n$ can be interpreted as

$$H^n = \{x \in E^{n,1} | \langle x, x\rangle = -1, x_0 > 0\}.$$

Or, in the projective model, $H^n$ is the interior $IQ_{n,1}$ of real projective space

$P^n$ with respect to the quadric

$$Q_{n,1} := \{[x] \in P^n | \langle x, x\rangle = 0\}.$$

The closure $\overline{H^n}$ of $H^n$ in $P^n$ represents the natural compactification of $H^n$. Points of the boundary $\partial H^n = \overline{H^n} - H^n$ are called *points at infinity* of $H^n$. Points in $P^n$ lying outside $\overline{H^n}$ are said to be *ideal points* of $H^n$ relative to $Q_{n,1}$, and the set of all such points is denoted by $AQ_{n,1}$.

To every point in $P^n$ corresponds a hyperplane in $P^n$ and vice versa: Let $P = [x] \in P^n$. A point $[y] \in P^n$ is said to be *conjugate to* $[x]$ *relative to* $Q_{n,1}$ if $\langle x, y\rangle = 0$ holds. The set of all points which are conjugate to $P = [x]$ form a projective hyperplane

$$\Pi_P := \{[y] \in P^n | \langle x, y\rangle = 0\},$$

the *polar hyperplane to* $P$. $P$ is called the *pole to* $\Pi_P$, and is denoted by $\mathrm{Pole}(\Pi_P)$. The map *pole* $\mapsto$ *polar hyperplane* is a bijection between the points and hyperplanes of $P^n$ known as the duality principle of the projective space $P^n$ (see [8, §4E]), and has the following properties (see [8, §4]):

(a) $P \in AQ_{n,1}$, $P \in Q_{n,1}$ or $P \in IQ_{n,1}$ if and only if $\Pi_P$ intersects, touches or avoids the quadric $Q_{n,1}$.

(b) If two lines $g, h$ in $P^2$ intersect at $S := g \cap h$, then $\Pi_S$ is the line determined by $\mathrm{Pole}(g)$, $\mathrm{Pole}(h)$.

(c) If a line $g$ in $P^2$ contains the point $\mathrm{Pole}(h)$ of the line $h$, then $g \perp h$ holds.

14.1.2. *The scheme of a polytope.* Let $P \subset \overline{H^n}$ denote a convex polytope bounded by finitely many hyperplanes $H_i$, $i \in I$, which are characterized by unit normal vectors $e_i \in E^{n,1}$ directed outwards with respect to $P$, say, i.e. (for basic notations and properties, see [29, §1]):

$$H_i = e_i^\perp := \{x \in H^n | \langle x, e_i\rangle = 0\} \quad \text{with } \langle e_i, e_i\rangle = 1.$$

We always assume that $P$ is acute-angled (i.e., all dihedral angles $\neq \frac{\pi}{2}$ are of measure strictly less than $\frac{\pi}{2}$) and of finite volume. The Gram matrix $G(P) := (\langle e_i, e_j\rangle)_{i,j\in I}$ of the vectors $e_i$, $i \in I$, associated to $P$ is an indecomposable symmetric matrix of signature $(n, 1)$ with entries $\langle e_i, e_i\rangle = 1$ and $\langle e_i, e_j\rangle \leq 0$ for $i \neq j$, having the following geometrical meaning (see [29, §1]):

$$-\langle e_i, e_j\rangle = \begin{cases} 0 & \text{if } H_i \perp H_j, \\ \cos\alpha_{ij} & \text{if } H_i, H_j \text{ intersect on } P \text{ at the angle } \alpha_{ij}, \\ 1 & \text{if } H_i, H_j \text{ are parallel}, \\ \cosh l_{ij} & \text{if } H_i, H_j \text{ admit a common perpendicular} \\ & \text{of length } l_{ij}. \end{cases}$$

On the other hand, if $G = (g_{ij})$ is an indecomposable symmetric $m \times m$ matrix of rank $n+1$ with $g_{ii} = 1$ and $g_{ij} \leq 0$, for $i \neq j$, then $G$ can be realized as Gram matrix $G(P)$ of an acute-angled polytope $P \subset X^n$ of finite volume in the following way [29, §2]:

1. If $G$ is positive definite ($G$ is *elliptic*), then $m = n+1$, and $G$ is the Gram matrix of a simplex in $S^n$ uniquely defined up to a motion.
2. If $G$ is positive semidefinite ($G$ is *parabolic*), then $m = n+2$, and $G$ is the Gram matrix of a simplex in $E^{n+1}$ uniquely defined up to a similarity.
3. If $G$ is of signature $(n, 1)$ ($G$ is *hyperbolic*), then $G$ is the Gram matrix of a convex polytope with $m$ facets (faces of codimension 1) in $\overline{H^n}$ uniquely defined up to a motion.

The Gram matrix $G(P)$ reflects combinatorial and metrical properties of an acute-angled polytope $P \subset \overline{H^n}$ [29, §3-4]. In particular, every ordinary vertex $p$ of $P$ is characterized by an elliptic principal submatrix of $G(P)$ of order $n$ describing the spherical vertex polytope $P_p$ (intersection of $P$ with the surface of a sufficiently small ball around $p$) of dimension $n - 1$ associated to $p$. To every vertex $q$ at infinity corresponds a parabolic principal submatrix of $G(P)$ of rank $n-1$ indicating that the vertex polytope $P_q$ is euclidean of dimension $n-1$.

For the geometric description of polytopes with many right dihedral angles, the language of schemes is much more convenient (see [30, §3]). A scheme $\Sigma$ is a weighted graph [30, §2] whose nodes $n_i$, $n_j$ are joined by an edge with positive weight $c_{ij}$ or are not joined at all; the last fact will be indicated by $c_{ij} = 0$. A subscheme of $\Sigma$ is a subgraph of $\Sigma$ with each pair of nodes connected by the same weighted edge as in $\Sigma$. The number $|\Sigma|$ of nodes is called the order of $\Sigma$. To every scheme $\Sigma$ of order $m$ corresponds a symmetric matrix $A(\Sigma) = (a_{ij})$ of order $m$ with $a_{ii} = 1$ on the diagonal and nonpositive entries $a_{ij} = -c_{ij} \leq 0$, $i \neq j$, off the diagonal. $\Sigma$ is connected if and only if $A(\Sigma)$ is indecomposable. Rank, determinant, permanent and character of definiteness of $\Sigma$ are defined to be the corresponding ones of $A(\Sigma)$. Furthermore, $\Sigma$ is said to be either elliptic, or parabolic, or hyperbolic if either all its components are elliptic, or apart from elliptic components there is at least one parabolic component, or exactly one component is hyperbolic.

In particular, if $\Sigma(1, \ldots, m)$ denotes a linear scheme with nodes $1, \ldots, m$ and weights $c_i := c_{i,i+1}$, $1 \leq i \leq m-1$, then the following useful recursion formulae hold for $\det\Sigma(1, \ldots, m)$ and $\mathrm{per}\Sigma(1, \ldots, m)$:

LEMMA 14.1. *Let* $m \geq 3$. *Then*

$$\det\Sigma(1, \ldots, m) = \det\Sigma(1, \ldots, m-1) - c_{m-1}^2 \det\Sigma(1, \ldots, m-2), \tag{14.5}$$

$$\mathrm{per}\Sigma(1, \ldots, m) = \mathrm{per}\Sigma(1, \ldots, m-1) + c_{m-1}^2 \mathrm{per}\Sigma(1, \ldots, m-2). \tag{14.6}$$

PROOF. Equation (14.5) is well known (cf. [27, p. 258]). To prove (14.6), denote by $C := A(\Sigma(1, \ldots, m)) = (-c_{ij})$ the symmetric $m \times m$ matrix associated to $\Sigma(1, \dots, m)$ with diagonal elements $-c_{ii} := 1$. Then, by definition,

$$\operatorname{per}\Sigma(1, \dots, m) = \operatorname{per} C = (-1)^m \sum c_{1\sigma(1)} \cdots c_{m\sigma(m)},$$

where $\sigma$ runs through all permutations of $\{1, \dots, m\}$. Since the permanent of every symmetric $m \times m$ matrix $M = (\mu_{ij})$ satisfies (cf. [23, Theorem 1.1(c) and (1.4)])

$$\operatorname{per} M = \sum_{k=1}^{m} \mu_{km} \operatorname{per} M(k|m) = \sum_{k=1}^{m} \mu_{mk} \operatorname{per} M(m|k), \tag{14.7}$$

where $M(i|j)$ is the matrix which is obtained from $M$ by deleting row $i$ and column $j$, we obtain

$$\operatorname{per} C = -c_{m-1} \operatorname{per} C(m-1|m) + \operatorname{per} C(m|m). \tag{14.8}$$

Applying (14.7) to the $(m-1) \times (m-1)$ matrix $C_m := C(m-1|m)$ we get

$$\operatorname{per} C_m = -c_{m-1} \operatorname{per} C_m(m-1|m-1). \tag{14.9}$$

Since $C_m(m-1|m-1) = A(\Sigma(1, \dots, m-2))$, (14.8) together with (14.9) imply (14.6). Q.E.D.

The scheme $\Sigma(P)$ of an acute-angled polytope $P \subset X^n$ is defined to be the scheme whose matrix $A(\Sigma)$ coincides with the Gram matrix $G(P)$, i.e., whose nodes $i$ correspond to the bounding hyperplanes $H_i = e_i^{\perp}$ (or equivalently to their normal vectors $e_i$) of $P$ and whose weights equal $-\langle e_i, e_j \rangle_{X^n}$, $i, j \in I$.

Two acute-angled polytopes $P_1, P_2 \subset \overline{H^n}$ are said to be *of the same schematic type* if their schemes $\Sigma(P_1)$, $\Sigma(P_2)$ are of the same graphical type (i.e., their underlying graphs as one-dimensional simplicial complexes are homeomorphic) and if corresponding weights $c_{ij}^1$ of $\Sigma(P_1)$ and $c_{ij}^2$ of $\Sigma(P_2)$ satisfy:

$$c_{ij}^1 \begin{cases} > \\ = 1 \\ > \end{cases} \Leftrightarrow \quad c_{ij}^2 \begin{cases} > \\ = 1. \\ < \end{cases}$$

It follows that polytopes of the same schematic type are of the same combinatorial type [3].

For the schemes of Coxeter polytopes $P_C \subset X^n$ (all dihedral angles are of the form $\frac{\pi}{p}$, $p \in \mathbf{N}$, $p \geq 2$) the usual conventions are adopted and—for convenience—used sometimes even in the non-Coxeter case: If two nodes are related by the weight $\cos\frac{\pi}{p}$, then they are joined by a $(p-2)$-fold line for $p = 3, 4$ and by a single line marked $p$ (or $\alpha = \frac{\pi}{p}$) for $p \geq 5$. If two bounding hyperplanes of $P_C \subset X^n$, $X^n \neq S^n$, are parallel, then the corresponding nodes are joined by a line marked $\infty$; if they are divergent (occurring at most in the hyperbolic case), then their nodes are joined by a dotted line and the weight $> 1$ is dropped.



14.1.3. *Orthoschemes of degree* $d$. The simplest examples of schemes are the linear and cyclic ones. One class of acute-angled hyperbolic polytopes described by such schemes is the following (see [13, 14]):

DEFINTION. An *n-dimensional orthoscheme of degree* $d$, $0 \leq d \leq 2$, is a convex polytope in $\overline{H^n}$, $n \geq 2$, denoted by $R_d$ such that its scheme $\Sigma_d := \Sigma(R_d)$ is connected and linear of length $n+d+1$ for $d = 0, 1$ or cyclic of order $n+3$ for $d = 2$.

Hence, orthoschemes of degree $d$ in $H^n$ are bounded by $n+d+1$ hyperplanes $H_0, \dots, H_{n+d}$ such that

$$H_i \perp H_j \quad \text{for } j \neq i-1, i, i+1, \tag{14.10}$$

where, for $d = 2$, indices are taken modulo $n+3$. Orthoschemes of degree $d$ allow the following geometrical description. For $d = 0$, they constitute the class of (ordinary) orthoschemes introduced by Schläfli (see [27, p. 243]): An orthoscheme in $X^n$ $(n \geq 1)$ is a simplex bounded by $n+1$ hyperplanes $H_0, \dots, H_n$ such that $H_i \perp H_j$ for $2 \leq |i-j| \leq n$. Or, equivalently, it has vertices $P_0, \dots, P_n$ numbered in such a way that $\operatorname{span}(P_0, \dots, P_i) \perp \operatorname{span}(P_i, \dots, P_n)$ for $1 \leq i \leq n-1$. The initial and final vertices $P_0$, $P_n$ of the orthogonal edge-path $P_0P_1, \dots, P_{n-1}P_n$ are called principal vertices, since they are distinguished in several ways. E.g. in $\overline{H^n}$, at most the principal vertices may be points at infinity (see [4, Satz 15, p. 188]).

In the projective model for $H^n$ (see 14.1.1), orthoschemes of degree $d > 0$ can be derived from ordinary ones by allowing $d$ of its principal vertices (and with them possibly further vertices) to lie outside the quadric $Q_{n,1}$, and then by cutting off the ideal vertices by means of the polar hyperplanes $H_{n+1} := \Pi_{P_n}$ resp. $H_{n+2} := \Pi_{P_0}$ (inasmuch as they lie outside $Q_{n,1}$). Hence, orthoschemes of degree $d$ are $d$-times (polarly) truncated orthoschemes bounded by hyperplanes $H_0, \dots, H_{n+d}$ with the property (14.10) (cf. 14.1.1).

REMARK. By adjoining to the bounding hyperplanes $H_0, \dots, H_n$ the polar hyperplanes associated to the principal vertices of an orthoscheme, viewed as an object in projective space, the configuration of the corresponding $n+3$ outer normal vectors $E^{n,1}$ form a Napier cycle in $E^{n,1}$. These were introduced and in the crystallographic case classified by Im Hof [13].

By construction, orthoschemes $R_d$ of degree $d$ are of finite volume (cf. also [29, Theorem 4.1]). Furthermore, they have at most $n+3$ nonright dihedral angles (or *essential* angles) $\alpha_1, \dots, \alpha_m$, $m \leq n+3$, and all of them are acute, i.e., $\alpha_i < \frac{\pi}{2}$ for $i = 1, \dots, m$ (see [4, §4.8, Hilfssatz 2] and the definition).

Let $d > 0$, and denote by $\Sigma_d$ the scheme of $R_d$. Then, removing $d$ nonconnected nodes in $\Sigma_d$ leaves two disjoint components $\sigma_1, \sigma_2$ of $\Sigma_d$, which satisfy (see [13, Proposition 1.4])

$$(14.11)\qquad \sigma_1 \begin{cases} \text{elliptic,} \\ \text{parabolic,} \\ \text{hyperbolic,} \end{cases} \Leftrightarrow \quad \sigma_2 \begin{cases} \text{hyperbolic,} \\ \text{parabolic,} \\ \text{elliptic.} \end{cases}$$

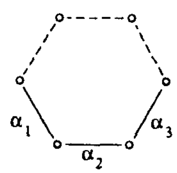
Therefore, if $n \geq 4$, $\Sigma_d$ always contains a connected hyperbolic subscheme of order $n+1$ all of whose weights $c_1, \dots, c_n$ are of the form $0 < c_i = \cos\alpha_i < 1$, $1 \leq i \leq n$ (see [13, Proposition 1.8]). Such schemes are said to be of *type A*.

In dimension three, however, the situation is principally different. First, for $R_d \subset H^3$ is compact, then, by the Euler equation for compact polyhedra, the number $m$ of essential angles equals three. Moreover, all but one of the different schemes $\Sigma_d$ are of type $A$; according to the degree $d$, the schemes of type A are of the form and likewise characterized by angle inequalities as follows (see (14.11) and 14.1.2):

(14.12) $\Sigma_0 : \circ\overset{\alpha_1}{—}\circ\overset{\alpha_2}{—}\circ\overset{\alpha_3}{—}\circ \qquad \alpha_1+\alpha_2 > \frac{\pi}{2},\ \alpha_2+\alpha_3 > \frac{\pi}{2},$

(14.13) $\Sigma_1 : \circ\overset{\alpha_1}{—}\circ\overset{\alpha_2}{—}\circ\overset{\alpha_3}{—}\circ\cdots\circ \qquad \alpha_1+\alpha_2 < \frac{\pi}{2},\ \alpha_2+\alpha_3 > \frac{\pi}{2},$

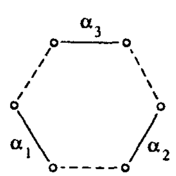
(14.14) $\Sigma_2 :$




$\alpha_1+\alpha_2 < \frac{\pi}{2},\ \alpha_2+\alpha_3 < \frac{\pi}{2}.$

The exceptional scheme is said to be of *type B*. By means of (14.11), it is given by

(14.15)




$0 < \alpha_1, \alpha_2, \alpha_3 < \frac{\pi}{2}.$

The corresponding polyhedron is a Lambert cube, i.e., a cube bounded by pairs of opposite Lambert quadrangles with equal angle $\alpha_j$, $1 \leq j \leq 3$, (see Figure 14.1; all figures are projective, and unlabelled dihedral angles are right angles).



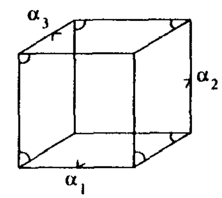



FIGURE 14.1

**14.2. The volume of a $d$-truncated orthoscheme.**

14.2.1. *The Schläfli differential formula.* For $n \geq 1$, let $\Sigma$ denote the elliptic linear scheme of order $n+1 \geq 2$ associated to a spherical $n$-orthoscheme $R$. The normalized volume function

$$(14.16)\qquad f_n(\Sigma) = f_n := c_n \operatorname{vol}_n(R) \quad \text{with } c_n = \frac{2^{n+1}}{\operatorname{vol}_n(S^n)} = \frac{2^n}{\pi^{(n+1)/2}}\Gamma((n+1)/2),\ f_0 := 1$$

is called the function of Schläfli (see [27, Nr. 23, p. 238]). The function $f_n$ is proportional to $\operatorname{vol}_n(R)$ such that $f_n = 1$ for the orthoscheme with all dihedral angles equal to $\frac{\pi}{2}$. Moreover, the function of Schläfli satisfies the following factorization property (see [27, Nr. 23, p. 238]):

LEMMA 14.2. *Let $\Sigma$ denote an elliptic linear scheme of order $n+1 \geq 2$ consisting of disjoint components $\sigma_1, \dots, \sigma_r$ of orders $n_1+1, \dots, n_r+1 \geq 1$. Then*

$$(14.17)\qquad f_n(\Sigma) = f_{n_1}(\sigma_1)\cdots f_{n_r}(\sigma_r).$$

For spherical Coxeter orthoschemes, Schläfli determined explicitly all possible values of $f_n$ (cf. [27, Nr. 30, p. 268 ff]); in particular, he found

$$(14.18)\qquad f_n(A_{n+1}) = \frac{2^{n+1}}{(n+2)!},$$

where $A_{n+1}$ denotes the scheme

$$\circ—\circ— \quad \cdots \quad —\circ—\circ$$

of order $n+1$. Interpreting hyperbolic $n$-space $H^n$ as upper half of the pseudosphere of radius $i = \sqrt{-1}$ in $E^{n+1}$, the notion of Schläfli's function can be carried over to orthoschemes $R_d \subset H^n$ of degree $d$, $0 \leq d \leq 2$, with graph $\Sigma_d$:

The function

$$(14.19)\qquad F_n(\Sigma_d) := i^n c_n \operatorname{vol}_n(R_d) \quad \text{with } i^2 = -1, \quad F_0 := 1,$$

where the constant $c_n$ is defined as in (14.16), is called the Schläfli function of $R_d$. Thus, for even dimensions,

$$F_{2n}(\Sigma_d) = (-1)^n \left(\frac{2}{\pi}\right)^n \cdot \prod_{p=1}^{n}(2p-1)\cdot \operatorname{vol}_{2n}(R_d), \qquad n \geq 1,$$

is a real-valued function.

Denote by $\mathscr{R}_\varsigma$ the set of compact $d$-truncated orthoschemes in $H^n$ of schematic type $\varsigma$ (cf. 14.1.2). Since every element of $\mathscr{R}_\varsigma$ is acute-angled (see 14.1.2), its congruence class is uniquely determined by its dihedral angles (see [3, §3, Uniqueness Theorem]). Therefore, Schläfli's volume function $F_n = F_n|\mathscr{R}_\varsigma$ restricted on $\mathscr{R}_\varsigma$ may be expressed as a function of the dihedral angles. The differential of $F_n$ depending on the dihedral angles can be represented by Schläfli's formula as follows (see [15, §2]):

**Theorem 14.3.** (Schläfli's differential formula). *Let $F_n$, $n \geq 2$, be the Schläfli function on the set $\mathscr{R}_\varsigma$ of compact $d$-truncated orthoschemes in $H^n$ of schematic type $\varsigma$ with essential angles $\alpha_1, \dots, \alpha_{m(\varsigma)}$ and with scheme $\Sigma_\varsigma$. Denote by $F_{n-2}(k)$ the Schläfli function of the apex of codimension 2 associated to the essential angle $\alpha_k$ of measure $f_1(k) := \frac{2}{\pi}\alpha_k$, $1 \leq k \leq m(\varsigma)$. Then*

$$dF_n(\Sigma_\varsigma) = \sum_{k=1}^{m(\varsigma)} F_{n-2}(k)\, df_1(k). \tag{14.20}$$

Schläfli discovered this formula for spherical simplexes, and separately for the more basic orthoschemes. Much later, H. Kneser found a different, very elegant proof for both, spherical and hyperbolic simplexes [16]. As Schläfli already remarked (cf. [27, Nr. 25, p. 246 ff, Nr. 32, p. 272 ff], and [29, Corollary, p. 48]), the differential formula for orthoschemes can be generalized to arbitrary acute-angled polytopes by means of subdivision into orthoschemes [15, §2.2].

14.2.2. *The reduction formula.* As can be read off from Schläfli's differential formula, there is a fundamental difference in calculating volumes of polytopes of even and odd dimensions. In fact, as the two-dimensional case already indicates, the volume of an even-dimensional simplex can be expressed in terms of the volumes of certain lower dimensional spherical ones. This reduction principle was proved by Schläfli in the spherical case and extended by Hopf to the hyperbolic case by means of analytic continuation (cf. [12, p. 134ff]). In principle, for every class of even-dimensional polytopes, an appropriate formula can be derived as soon as their schematic type is known, namely by an inductive argument using Schläfli's differential formula. In terms of Schläfli's function, the reduction formula for orthoschemes $R_d \subset \overline{H^{2n}}$, $n \geq 1$, of degree $d$, $0 \leq d \leq 2$, reads as follows (see [15, §3]):

**Theorem 14.4.** (Reduction formula). *Denote by $R_d \subset \overline{H^{2n}}$, $0 \leq d \leq 2$, $n \geq 1$, a $2n$-dimensional orthoscheme of degree $d$ with scheme $\Sigma_d$. Then*

$$F_{2n}(\Sigma_d) = \sum_{k=0}^{n} \frac{(-1)^k}{k+1}\binom{2k}{k}\sum_\sigma f_{2n-(2k+1)}(\sigma), \qquad \sum f_{-1} := 1, \tag{14.21}$$

*where $\sigma$ runs through all elliptic subschemes of order $2(n-k)$ of $\Sigma_d$ all of whose components are of even order.*



Thus, in order to calculate volumes of noneuclidean polytopes, it is sufficient to consider the volume problem for polytopes of odd dimensions.

14.2.3. *The principal parameter and the fundamental relations.* Let $R_d \subset H^3$, $0 \leq d \leq 2$, denote a compact $d$-truncated orthoscheme with essential angles $\alpha_1, \alpha_2, \alpha_3$, and scheme $\Sigma_d$. For the integration of its Schläfli differential, the principal parameter $\theta$ is of fundamental importance. To introduce this notion, we begin with the following

**Definition.** A connected subscheme $\sigma$ of order 4 of $\Sigma_d$ is called *maximal* if its number of weights having the form $\cos\alpha$ is maximal.

Hence, if $\Sigma_d$ is of type $A$, then

$$\circ\overset{\alpha_1}{\text{——}}\circ\overset{\alpha_2}{\text{——}}\circ\overset{\alpha_3}{\text{——}}\circ$$

is the unique maximal subscheme of $\Sigma_d$, whereas in case $B$, exactly the schemes

$$\circ\overset{\alpha_i}{\text{——}}\circ\cdots\circ\overset{\alpha_j}{\text{——}}\circ, \qquad 1 \leq i < j \leq 3,$$

are maximal. Furthermore, since every connected subscheme of order 4 of $\Sigma_d$ is hyperbolic, a maximal subscheme of $\Sigma_d$ has negative determinant (see [13, Proposition 1.2]).

**Definition.** The *principal parameter* $\theta$ of $R_d$ is given by

$$0 \leq \theta := \arccos\left(\frac{\operatorname{per}\sigma + \det\sigma - 2}{\operatorname{per}\sigma - \det\sigma - 2}\right)^{1/2} \leq \frac{\pi}{2}, \tag{14.22}$$

where $\sigma$ is a maximal subscheme of $\Sigma_d$.

In the nondegenerate case, $\det\sigma < 0$ and $\operatorname{per}\sigma + \det\sigma > 2$. Consequently, the quotient in (14.22) is positive and less than 1. Thus, if $\Sigma_d$ is of type $A$, $\theta$ is well defined and takes the form

$$\cos^2\theta = \frac{\cos^2\alpha_1\cos^2\alpha_3}{\cos^2\alpha_1 - \sin^2\alpha_2 + \cos^2\alpha_3}, \quad \text{or} \tag{14.23}$$
$$\cos^2\theta(\cos^2\theta - \sin^2\alpha_2) = (\cos^2\theta - \cos^2\alpha_1)(\cos^2\theta - \cos^2\alpha_3).$$

In case $B$, it remains to check that $\theta$ is independent of the chosen maximal subscheme of $\Sigma_d$:

**Lemma 14.3.** *Let $\Sigma$ denote the scheme in Figure 14.2 of type $B$ describing a Lambert cube in $H^3$. For every maximal subscheme $\sigma_k$, $1 \leq k \leq 3$, of $\Sigma$*

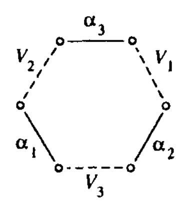



Figure 14.2

*with weights* $\cos\alpha_{k-1}$, $\cosh V_k$, $\cos\alpha_{k+1}$ *(indices are taken modulo* 3), *set*

$$\theta_k := \arccos\left(\frac{\operatorname{per}\sigma_k + \det\sigma_k - 2}{\operatorname{per}\sigma_k - \det\sigma_k - 2}\right)^{1/2}, \quad 0 \le \theta_k \le \frac{\pi}{2}.$$

*Then,* $\theta_k = \theta_l$ *for* $1 \le k < l \le 3$.

PROOF. We prove that $\cos^2\theta_1 = \cos^2\theta_3$, i.e.,

$$\frac{\cos^2\alpha_1\cos^2\alpha_2}{\cos^2\alpha_1 + \sinh^2 V_3 + \cos^2\alpha_2} = \frac{\cos^2\alpha_2\cos^2\alpha_3}{\cos^2\alpha_2 + \sinh^2 V_1 + \cos^2\alpha_3}, \quad \text{or} \tag{14.24}$$

$$\cos^2\alpha_1(\cos^2\alpha_2 + \sinh^2 V_1) = \cos^2\alpha_3(\cos^2\alpha_2 + \sinh^2 V_3).$$

Since $\Sigma$ is of signature (3,1), the extensions

$$\overline{\sigma_1}:\ \circ\overset{\alpha_3}{\text{---}}\circ\overset{V_1}{\cdots\cdots}\circ\overset{\alpha_2}{\text{---}}\circ\overset{V_3}{\cdots\cdots}\circ, \qquad \overline{\sigma_3}:\ \circ\overset{\alpha_1}{\text{---}}\circ\overset{V_3}{\cdots\cdots}\circ\overset{\alpha_2}{\text{---}}\circ\overset{V_1}{\cdots\cdots}\circ$$

of $\sigma_1$ and $\sigma_3$ in $\Sigma$ have vanishing determinant. Hence, by Lemma 14.1, (14.5), we obtain

$$\det(\overline{\sigma_1}) = \sinh^2 V_3(\cosh^2 V_1 - \sin^2\alpha_3) - \sin^2\alpha_3\cos^2\alpha_2 = 0,$$
$$\det(\overline{\sigma_3}) = \sinh^2 V_1(\cosh^2 V_3 - \sin^2\alpha_1) - \sin^2\alpha_1\cos^2\alpha_2 = 0,$$

and therefore

$$\cos^2\alpha_1\cos^2\alpha_2 + \sinh^2 V_1\sinh^2 V_3 = \cos^2\alpha_2 - \cos^2\alpha_1\sinh^2 V_1,$$
$$\cos^2\alpha_2\cos^2\alpha_3 + \sinh^2 V_1\sinh^2 V_3 = \cos^2\alpha_2 - \cos^2\alpha_3\sinh^2 V_3.$$

Subtraction yields

$$\cos^2\alpha_2(\cos^2\alpha_1 - \cos^2\alpha_3) = \cos^2\alpha_3\sinh^2 V_3 - \cos^2\alpha_1\sinh^2 V_1,$$

which proves (14.24). By the same procedure, one derives the remaining equalities. Q.E.D.

This lemma implies that the principal parameter $\theta$ of a Lambert cube satisfies

$$\cos^2\theta = \frac{\cos^2\alpha_{k-1}\cos^2\alpha_{k+1}}{\cos^2\alpha_{k-1} + \sinh^2 V_k + \cos^2\alpha_{k+1}},$$

and

$$\cos\theta^2(\cos^2\theta + \sinh^2 V_k) = (\cos^2\theta - \cos^2\alpha_{k-1})(\cos^2\theta - \cos^2\alpha_{k+1}), \tag{14.25}$$

where indices are taken modulo 3.

In the asymptotic cases, which form the transitions from one degree to another, the principal parameter $\theta$ is realizable as a dihedral angle in $R_d \subset \overline{H^3}$. E.g., if $R_0$ is an ordinary orthoscheme, then

(14.26) $\quad R_0$ simply asymptotic $\quad\Leftrightarrow\quad \theta = \frac{\pi}{2} - \alpha_2,$

(14.27) $\quad R_0$ doubly asymptotic $\quad\Leftrightarrow\quad \theta = \alpha_1 = \frac{\pi}{2} - \alpha_2 = \alpha_3.$



Let $R_0$ denote a compact ordinary orthoscheme. Then, the following relationship holds between the measures of the essential angles $\alpha_k$ and the corresponding apices $V_k$, $k = 1, 2, 3$, (see [19] or [4, (11), p. 229]):

$$\tanh V_k = \tan\theta\cdot\tan\bar\alpha_k \quad\text{or}\quad V_k = \frac{1}{2}\log\frac{\cos(\theta - \bar\alpha_k)}{\cos(\theta + \bar\alpha_k)}, \quad k = 1, 2, 3, \tag{14.28}$$

where

$$\bar\alpha_k := \begin{cases} \alpha_2, & k = 2, \\ \frac{\pi}{2} - \alpha_k, & k = 1, 3. \end{cases} \tag{14.29}$$

Now, similar relations hold for orthoschemes $R_d$ of degree $d > 0$. In fact, in the projective model of $H^3$, a maximal subscheme of $\Sigma_d$ describes an ordinary orthoscheme in $P^3$ with $d$ ideal principal vertices lying outside the quadric $Q_{3,1}$. Since hyperbolic geometry admits a complex continuation to the space $AQ_{3,1}$ of ideal points of $H^3$ such that the distance between pole and polar line equals $i\frac{\pi}{2}$ (cf. [25, §5]), the fundamental relations for $d$-truncated orthoschemes can be summarized as follows (see [14, §3.3]):

PROPOSITION 14.1. *Let $R_d$ denote a compact orthoscheme of degree $d$, $0 \le d \le 2$, with essential angles $\alpha_k$, $0 < \alpha_k < \frac{\pi}{2}$, corresponding apices of lengths $V_k$, $k = 1, 2, 3$, and principal parameter $\theta$. Then,*

$$V_k = \frac{1}{2}\log\left|\frac{\cos(\theta - \bar\alpha_k)}{\cos(\theta + \bar\alpha_k)}\right|, \qquad k = 1, 2, 3, \quad \text{with} \tag{14.30}$$

$$\bar\alpha_k := \begin{cases} \alpha_2, & \text{if } R_d \text{ of type A and } k = 2, \\ \frac{\pi}{2} - \alpha_k, & \text{otherwise.} \end{cases} \tag{14.31}$$

Proposition 14.1 induces the following important identity

$$\frac{\partial V_k}{\partial\theta} = \frac{\sin\alpha_k\cos\alpha_k}{\cos^2\theta - \sin^2\bar\alpha_k} \quad \text{for } k = 1, 2, 3, \tag{14.32}$$

where $\bar\alpha_k$ satisfies (14.31).

The principal parameter characterizes also degenerations of $R_d$ in the following way:

PROPOSITION 14.2. *Let $R_d$, $0 \le d \le 2$, denote a $d$-truncated orthoscheme with graph $\Sigma_d$, essential angles $0 \le \alpha_1, \alpha_2, \alpha_3 \le \frac{\pi}{2}$ and principal parameter $\theta$. If $\theta = 0$ or $\theta = \frac{\pi}{2}$, then* $\operatorname{vol}_3(R_d) = 0$.

PROOF. (i) Let $\theta = 0$. Then, by (14.22), $\det\sigma = 0$ for every maximal subscheme $\sigma$ of $\Sigma_d$. Let $\Sigma_d$ be of type $A$. First, if $0 < \alpha_1, \alpha_2, \alpha_3 < \frac{\pi}{2}$, Proposition 14.1 implies that the corresponding apices are of lengths zero. Hence, $R_d$ is point-shaped. Second, the condition $\theta = 0$ implies that $\cos\alpha_2 = \sin\alpha_1\sin\alpha_3$. Thus

$$\alpha_2 = 0 \quad\Leftrightarrow\quad \alpha_1 = \alpha_3 = \frac{\pi}{2} \quad\text{or}\quad \alpha_2 = \frac{\pi}{2} \quad\Leftrightarrow\quad \alpha_1 \text{ (and/or } \alpha_3) = 0.$$

In both cases, at least one vertex triangle of $R_d$ degenerates, and therefore $\mathrm{vol}_3(R_d)=0$. If $\Sigma_d$ is of type $B$, $\theta=0$ reads as

$$\cosh V_k = \sin\alpha_{k-1}\sin\alpha_{k+1} = 1 \quad \text{for } k \bmod 3,$$

where $V_k$ denotes the length of the apex associated to $\alpha_k$. Hence, $\alpha_k = \frac{\pi}{2}$ and $V_k = 0$ for $k=1,2,3$, and $R_d$ is point-shaped.

(ii) Let $\theta = \frac{\pi}{2}$. By (14.22), $\mathrm{per}\,\sigma + \det\sigma - 2 = 0$ implies that at least one essential angle equals $\frac{\pi}{2}$. Let $\Sigma_d$ be of type $A$. If $\alpha_1$ and/or $\alpha_3 = \frac{\pi}{2}$, then $\mathrm{vol}_3(R_d)=0$, since at least one facet degenerates in dimension. Hence, it remains to consider the case $\alpha_2 = \frac{\pi}{2}$. If $d>0$, then at least one vertex triangle degenerates, and therefore $\mathrm{vol}_3(R_d)=0$. If $d=0$, and if $0<\alpha_1, \alpha_3<\frac{\pi}{2}$, then Proposition 14.1 yields $V_1 = V_3 = 0$ and therefore $\mathrm{vol}_3(R_d)=0$. In the other case, where e.g. $\alpha_1 = 0$ (or $\frac{\pi}{2}$), one vertex triangle (or one facet) degenerates. If $\Sigma_d$ is of type $B$, at least two opposite Lambert quadrangles of the cube $R_d$ degenerate to a point, from which $\mathrm{vol}_3(R_d)=0$ follows. Q.E.D.

14.2.4. *The Euler dilogarithm and Lobachevsky's function.* As will be seen later, the volume of a hyperbolic polyhedron is expressible in terms of the Euler dilogarithm and the Lobachevsky function related to it.

Let $z\in\mathbf{C}$, $|z|\le 1$. Denote by

$$\mathrm{Li}_2(z) := \sum_{n=1}^{\infty}\frac{z^n}{n^2} = -\int_0^z \frac{\log(1-t)}{t}\,dt$$

the Euler dilogarithm function. Splitting $\mathrm{Li}_2(z)$ up into real and imaginary parts, one deduces for

$$\mathrm{Li}_2(r,\phi) := \mathrm{Re}(\mathrm{Li}_2(re^{i\phi})) = \sum_{n=1}^{\infty}\frac{r^n\cos(n\phi)}{n^2} = -\frac{1}{2}\int_0^r \frac{\log(1-2t\cos\phi+t^2)}{t}\,dt$$

the following properties (see [18, §5]):

$$\mathrm{Li}_2(r,0) = \mathrm{Li}_2(r), \qquad \mathrm{Li}_2(r,\pi) = \mathrm{Li}_2(-r); \tag{14.33}$$

$$\mathrm{Li}_2(r,\pi+\phi) = \mathrm{Li}_2(-r,\phi), \qquad \mathrm{Li}_2(r, 2n\pi\pm\phi) = \mathrm{Li}_2(r,\phi); \tag{14.34}$$

$$\mathrm{Li}_2\left(\frac{1}{r},\phi\right) + \mathrm{Li}_2(r,\phi) = -\frac{1}{2}\log^2 r + \frac{1}{2}(\pi-\phi)^2 - \frac{\pi^2}{6}; \tag{14.35}$$

$$\mathrm{Li}_2\left(r,\frac{\pi}{3}\right) = \frac{1}{6}\mathrm{Li}_2(-r^3) - \frac{1}{2}\mathrm{Li}_2(-r). \tag{14.36}$$

For $z = e^{i\phi}$, $0\le\phi\le 2\pi$,

$$\begin{aligned}\mathrm{Li}_2(e^{i\phi}) &= \sum_{n=1}^{\infty}\frac{\cos(n\phi)}{n^2} + i\sum_{n=1}^{\infty}\frac{\sin(n\phi)}{n^2}\\ &= \frac{\pi^2}{6} - \frac{\phi(2\pi-\phi)}{4} - i\int_0^{\phi}\log\left|2\sin\frac{t}{2}\right|\,dt.\end{aligned}$$



The *Lobachevsky function* is defined by

$$Л(z) := -\int_0^z \log|2\sin t|\,dt.$$

Thus, for $-\frac{\pi}{2} < \mathrm{Re}(z) \le \frac{\pi}{2}$, one obtains

$$Л(z) = \frac{i}{2}\left\{\mathrm{Li}_2(e^{2iz}) + z(\pi - z) - \frac{\pi^2}{6}\right\}. \tag{14.37}$$

Furthermore, if $0<\alpha,\beta,\gamma<\frac{\pi}{2}$ are such that $\sin\alpha\sin\gamma > \cos\beta$, then (14.37) together with (14.33)–(14.35) yield the following formula, which will be of use later, for the function $S$ defined in (14.4):

(14.38)

$$\begin{aligned} S\left(\frac{\pi}{2}-\alpha, \beta, \frac{\pi}{2}-\gamma\right) &= \mathrm{Li}_2(r,2\alpha) - \mathrm{Li}_2(r,2\beta) + \mathrm{Li}_2(r,2\gamma) - \mathrm{Li}_2(-r)\\ &\quad - \left(\frac{\pi}{2}-\alpha\right)^2 + \beta^2 - \left(\frac{\pi}{2}-\gamma\right)^2\\ &= \mathrm{Re}\left(\frac{1}{i}\left\{Л(\alpha+i\tau) - Л(\alpha-i\tau) - Л\left(\frac{\pi}{2}-\beta+i\tau\right) + Л\left(\frac{\pi}{2}-\beta-i\tau\right)\right.\right.\\ &\qquad \left.\left. + Л(\gamma+i\tau) - Л(\gamma-i\tau) + 2Л\left(\frac{\pi}{2}-i\tau\right)\right\}\right), \end{aligned}$$

where

$$\rho := \frac{\sqrt{\sin^2\alpha\sin^2\gamma - \cos^2\beta}}{\cos\alpha\cos\gamma}, \qquad r := \frac{1-\rho}{1+\rho}, \qquad \tau := -\frac{1}{2}\log r. \tag{14.39}$$

Let $\omega\in\mathbf{R}$, Then

(14.40)

$$\begin{aligned} Л(\omega) &= \frac{1}{2}\mathrm{Im}(\mathrm{Li}_2(e^{2i\omega})) = \frac{1}{2}\sum_{n=1}^{\infty}\frac{\sin(2n\omega)}{n^2}\\ &= -\int_0^{\omega}\log|2\sin t|\,dt = \left(\frac{\pi}{2}-\omega\right)\log 2 + \int_0^{(\pi/2)-\omega}\log|\cos t|\,dt, \end{aligned}$$

which is closely related to the Clausen function (see [18, §4])

$$\mathrm{Cl}_2(\omega) := \sum_{n=1}^{\infty}\frac{\sin(n\omega)}{n^2} = -\int_0^{\omega}\log\left|\sin\frac{t}{2}\right|\,dt$$

according to

$$Л(\omega) = \tfrac{1}{2}\mathrm{Cl}_2(2\omega), \quad \forall\omega\in\mathbf{R},$$

and which satisfies the following

PROPERTIES. (see [18, §4]).

(a) $Л(\omega)$ is well-defined and continuous for all $\omega\in\mathbf{R}$. $Л(\omega)$ is odd and $\pi$-periodic. $Л(\omega)$ assumes its maximum value at $\omega_k = \frac{\pi}{6} + k\pi$, $k\in\mathbf{Z}$.

(b) $Л(\omega)$ satisfies the following distribution law:

$$Л(n\omega) = n \sum_{k \bmod n} Л\left(\omega + \frac{k\pi}{n}\right), \quad \forall n \in \mathbf{N}, \ \forall \omega \in \mathbf{R}.$$

In particular for $n = 2$, this relation yields the duplication law

$$Л(2\omega) = 2Л(\omega) + 2Л\left(\tfrac{\pi}{2} + \omega\right), \quad \forall \omega \in \mathbf{R}.$$

(c) For actual computation, the following representation of $Л(\omega)$ is very useful (cf. [32, Appendix 1, p. 294])

$$Л(\pi t) = \pi t \left\{9 - \log|2 \sin \pi t| - \sum_{k=1}^{4}\left(c_k t^{2k+1} + k \log \frac{k+t}{k-t}\right) + \varepsilon\right\}$$

with

$c_1 = 0.14754863716$, $c_2 = 0.00142852188$,
$c_3 = 0.00002919407$, $c_4 = 0.00000076258$,

and $|\varepsilon| < 1.2 \times 10^{-11}$ for $|t| \leq 1/2$.

14.2.5. *The volume formula for $R_d$.* Let $R_d$ denote a compact orthoscheme of degree $d$, $0 \leq d \leq 2$, with scheme $\Sigma_d$, with essential angles $\alpha_k$ and corresponding apices of length $V_k$, $k = 1, 2, 3$. Then, by Schläfli's volume differential formula (see Theorem 14.3, 14.2.1),

$$d\mathrm{vol}_3(R_d) = -\frac{1}{2}\sum_{k=1}^{3} V_k \, d\alpha_k, \tag{14.41}$$

where, for $k = 1, 2, 3$, the coefficients $V_k$ are given by (see Proposition 14.1, 14.2.3)

$$V_k = \frac{1}{2} \log\left|\frac{\cos(\theta - \bar{\alpha}_k)}{\cos(\theta + \bar{\alpha}_k)}\right| \quad \text{with } \bar{\alpha}_k := \begin{cases} \alpha_2, & \text{if } R_d \text{ of type A and } k = 2, \\ \frac{\pi}{2} - \alpha_k, & \text{otherwise.}\end{cases} \tag{14.42}$$

Here, $\theta$ denotes the principal parameter of $R_d$ given by

$$0 \leq \theta = \arccos\left(\frac{\operatorname{per}\sigma + \det\sigma - 2}{\operatorname{per}\sigma - \det\sigma - 2}\right)^{1/2} \leq \frac{\pi}{2},$$

where $\sigma$ is a maximal subscheme of $\Sigma_d$; $\theta$ depends on $\alpha_1, \alpha_2, \alpha_3$. Thus, $V_k$ are very complicated expressions in the essential angles of $R_d$. However, by regarding $\alpha_1, \alpha_2, \alpha_3, \theta =: \alpha_4$ as four parameters independent from each other, the differential (14.41) suitably extended by $d\alpha_4$ can be integrated and thereafter identified with the volume of $R_d$. In this context, set

$$\widehat{V}_k(\alpha_1, \ldots, \alpha_4) := \frac{1}{2}\log\left|\frac{\cos(\alpha_4 - \bar{\alpha}_k)}{\cos(\alpha_4 + \bar{\alpha}_k)}\right|, \quad k = 1, 2, 3. \tag{14.43}$$



Then

$$\widehat{V}_k|_{\alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3)} = V_k(\alpha_1, \alpha_2, \alpha_3), \quad k = 1, 2, 3.$$

Consider the region

$$G := \{(\alpha_1, \ldots, \alpha_4) \in \mathbf{R}^4 \mid 0 \leq \alpha_1, \ldots, \alpha_4 \leq \tfrac{\pi}{2};\ \alpha_4 \neq \tfrac{\pi}{2} - \bar{\alpha}_k,\ k = 1, 2, 3\},$$

and on $G$ the following differential form

$$\Omega := \sum_{k=1}^{4} W_k \, d\alpha_k, \quad \text{with} \tag{14.44}$$

$$W_k(\alpha_1, \ldots, \alpha_4) := -\frac{1}{2}\widehat{V}_k(\alpha_1, \ldots, \alpha_4), \quad k = 1, 2, 3, \tag{14.45}$$

and where $W_4 \in C^1(G)$ is determined by

(I) $W_4$ satisfies the integrability conditions $\partial W_i/\partial \alpha_k = \partial W_k/\partial \alpha_i$ for $1 \leq i < k \leq 4$.

(II) $W_4 = 0$ for $\alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3)$.

By definition, $W_k(\alpha_1, \ldots, \alpha_4)$ depends only on $\alpha_k$ and $\alpha_4$, i.e.,

$$\frac{\partial W_k}{\partial \alpha_i} = \frac{\partial W_i}{\partial \alpha_k} = 0 \quad \text{for } 1 \leq i < k \leq 3.$$

On the other hand, 14.2.3, (14.32), yields

$$\frac{\partial W_k}{\partial \alpha_4} = -\frac{1}{2}\frac{\sin\alpha_k \cos\alpha_k}{\cos^2\alpha_4 - \sin^2\bar{\alpha}_k} \quad \text{for } k = 1, 2, 3,$$

where $\bar{\alpha}_k$ is determined by (14.42). Hence, according to type A or B of $\Sigma_d$, $W_4$ is given as follows (see [4, (11), p. 237] for the case $d = 0$)

(14.46)

$$\mathrm{A}: \ W_4 = \frac{1}{4}\log\frac{(\cos^2\alpha_4 - \sin^2\alpha_2)\cos^2\alpha_4}{(\cos^2\alpha_4 - \cos^2\alpha_1)(\cos^2\alpha_4 - \cos^2\alpha_3)},$$

$$\mathrm{B}: \ W_4 = \frac{1}{4}\log\frac{\sin^2\alpha_4\cos^4\alpha_4}{(\cos^2\alpha_1 - \cos^2\alpha_4)(\cos^2\alpha_2 - \cos^2\alpha_4)(\cos^2\alpha_3 - \cos^2\alpha_4)}.$$

It is obvious that $W_4 \in C^1(G)$, and that it satisfies (I). By means of 14.2.3, (14.23) and (14.25), one easily derives (II). Therefore, the differential form $\Omega$ restricted to the hypersurface

$$\alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3) \quad \text{in } \mathbf{R}^4$$

is identical with Schläfli's volume differential (14.41). Since $\Omega$ satisfies (I), it is exact and path-independently integrable in every connected component of $G$. To perform the integration of $\Omega$, we distinguish between $R_d$ with scheme $\Sigma_d$ of type A or B.

A. Let $R_0$ denote an ordinary orthoscheme with essential angles $0 < \alpha_1, \alpha_2, \alpha_3 < \frac{\pi}{2}$; by 14.1.3, (14.12), these satisfy

$$\alpha_1 + \alpha_2 > \tfrac{\pi}{2}, \qquad \alpha_2 + \alpha_3 > \tfrac{\pi}{2},$$

and thus, by 14.2.3, (14.22),

$$0 \le \theta < \alpha_1, \tfrac{\pi}{2} - \alpha_2, \alpha_3 < \tfrac{\pi}{2}.$$

Consider the convex region

$$G_0 := \{(\alpha_1, \dots, \alpha_4) \in G \mid 0 \le \alpha_4 < \tfrac{\pi}{2} - \alpha_2 < \alpha_1, \alpha_3 < \tfrac{\pi}{2}\},$$

and on $G_0$, the differential form

$$\Omega = \sum_{k=1}^{4} W_k \, d\alpha_k,$$

where, for $k = 1, 2, 3$,

$$W_k = -\frac{1}{4} \log \left| \frac{\cos(\alpha_4 - \overline{\alpha}_k)}{\cos(\alpha_4 + \overline{\alpha}_k)} \right|,$$

$$W_4 = \frac{1}{4} \log \frac{(\cos^2\alpha_4 - \sin^2\alpha_2)\cos^2\alpha_4}{(\cos^2\alpha_4 - \cos^2\alpha_1)(\cos^2\alpha_4 - \cos^2\alpha_3)}.$$

Take an arbitrary point $P := (\alpha_1, \dots, \alpha_4) \in G_0$ and integrate $\Omega$ along the line from $(\alpha_1, \alpha_2, \alpha_3, 0)$ to $P$. Since $W_k(\alpha_1, \alpha_2, \alpha_3, 0) = 0$ for $k = 1, 2, 3$, the integral

$$\begin{aligned} \widehat{V} &:= \int_0^{\alpha_4} W_4(\alpha_1, \alpha_2, \alpha_3, t)\, dt \\ &= \sum_{k=1}^{3} \frac{(-1)^k}{4} \left\{ Л\left(\frac{\pi}{2} + \alpha_4 + \overline{\alpha}_k\right) + Л\left(\frac{\pi}{2} + \alpha_4 - \overline{\alpha}_k\right) \right\} \\ &\quad + \frac{1}{2} Л\left(\frac{\pi}{2} - \alpha_4\right) \end{aligned} \tag{14.47}$$

is an antiderivative of $\Omega$ in $G_0$. Restricted to $\alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3)$, $\widehat{V}$ represents the volume $\operatorname{vol}_3(R_0)$ of $R_0$, since:

(i) For $\alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3)$, Leibniz' Rule yields together with (I), (II), and Proposition 14.2, 14.2.3

$$\begin{aligned} \frac{\partial \widehat{V}}{\partial \alpha_k} &= \frac{\partial}{\partial \alpha_k} \int_0^{\theta} W_4(\alpha_1, \alpha_2, \alpha_3, t)\, dt \\ &= W_4(\alpha_1, \alpha_2, \alpha_3, \theta) \cdot \frac{\partial \theta}{\partial \alpha_k} + \int_0^{\theta} \frac{\partial W_4}{\partial \alpha_k}(\alpha_1, \alpha_2, \alpha_3, t)\, dt \\ &= W_k(\alpha_1, \alpha_2, \alpha_3, \theta) - W_k(\alpha_1, \alpha_2, \alpha_3, 0) = W_k(\alpha_1, \alpha_2, \alpha_3, \theta) \\ &= -\frac{1}{2} V_k(\alpha_1, \alpha_2, \alpha_3) = \frac{\partial \operatorname{vol}_3(R_0)}{\partial \alpha_k}, \qquad k = 1, 2, 3. \end{aligned}$$



(ii) For $\alpha_4 = \theta = 0$, both $\widehat{V}$ and $\operatorname{vol}_4(R_0)$ vanish according to (14.47) and Proposition 14.2, 14.2.3.

Furthermore, the antiderivative $\widehat{V}$ of $\Omega$ certainly extends to the cube

$$\{(\alpha_1, \dots, \alpha_4) \in \mathbf{R}^4 \mid 0 \le \alpha_1, \dots, \alpha_4 \le \tfrac{\pi}{2}\}$$

in $\mathbf{R}^4$, still satisfying

$$\frac{\partial \widehat{V}}{\partial \alpha_k} = W_k(\alpha_1, \alpha_2, \alpha_3, \theta) \quad \text{for } \alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3), \ 1 \le k \le 3,$$

$$\widehat{V}(\alpha_1, \alpha_2, \alpha_3, 0) = 0.$$

Hence, the restriction $\widehat{V}(\alpha_1, \alpha_2, \alpha_3, \alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3))$ to the domain of definition of $R_d$, $d > 0$, is identical to $\operatorname{vol}_3(R_d)$ (see 14.1.3, (14.13), (14.14)). Thus, we have proved the following

THEOREM 14.5. *Let $R_d$, $0 \le d \le 2$, denote a $d$-truncated orthoscheme with scheme $\Sigma_d$ of type $A$ and with essential angles $0 \le \alpha_1, \alpha_2, \alpha_3 \le \frac{\pi}{2}$. If $\sigma$ denotes the maximal subscheme of $\Sigma_d$, then*

(14.48)

$$\begin{aligned} \operatorname{vol}_3(R_d) = \frac{1}{4} \Big\{ & Л(\alpha_1 + \theta) - Л(\alpha_1 - \theta) + Л\left(\frac{\pi}{2} + \alpha_2 - \theta\right) + Л\left(\frac{\pi}{2} - \alpha_2 - \theta\right) \\ & + Л(\alpha_3 + \theta) - Л(\alpha_3 - \theta) + 2Л\left(\frac{\pi}{2} - \theta\right) \Big\}, \end{aligned}$$

*where*

$$\begin{aligned} 0 \le \theta &= \arccos \left( \frac{\operatorname{per}\sigma + \det\sigma - 2}{\operatorname{per}\sigma - \det\sigma - 2} \right)^{1/2} \\ &= \arctan \left( \frac{\cos^2\alpha_2 - \sin^2\alpha_1 \sin^2\alpha_3}{\cos^2\alpha_1 \cos^2\alpha_3} \right)^{1/2} \le \frac{\pi}{2}. \end{aligned}$$

B. Let $R$ be a Lambert cube with essential angles $0 < \alpha_k < \frac{\pi}{2}$ and corresponding apices of length $V_k$, $k = 1, 2, 3$. By 14.2.3, (14.25), the principal parameter $\theta$ of $R$ satisfies

$$\tan^2\theta = \frac{\cosh^2 V_k - \sin^2\alpha_{k-1} \sin^2\alpha_{k+1}}{\cos^2\alpha_{k-1} \cos^2\alpha_{k+1}}, \qquad k \bmod 3,$$

which implies that $\theta > \alpha_1, \alpha_2, \alpha_3$. Hence, consider the convex region

$$H := \{(\alpha_1, \dots, \alpha_4) \in \mathbf{R}^4 \mid 0 \le \alpha_k < \alpha_4 \le \tfrac{\pi}{2}, k = 1, 2, 3\},$$

choose an arbitrary point $P = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in H$ and integrate the differential form $\Omega$ from $(\alpha_1, \alpha_2, \alpha_3, \frac{\pi}{2})$ to $P$. Since $W_k(\alpha_1, \alpha_2, \alpha_3, \frac{\pi}{2}) = 0$

for $k = 1, 2, 3$, one obtains

(14.49)

$$\begin{aligned}\widehat{V} : &= \int_{\pi/2}^{\alpha_4} W_4(\alpha_1, \alpha_2, \alpha_3, t)\,dt \\ &= \frac{1}{4}\int_{\pi/2}^{\alpha_4} \log \frac{\sin^2 t \cos^4 t}{(\cos^2\alpha_1 - \cos^2 t)(\cos^2\alpha_2 - \cos^2 t)(\cos^2\alpha_3 - \cos^2 t)}\,dt \\ &= \frac{1}{4}\sum_{k=1}^{3}\{Л(\alpha_k + \alpha_4) - Л(\alpha_k - \alpha_4)\} + Л\left(\frac{\pi}{2} - \alpha_4\right) - \frac{1}{2}Л(\alpha_4).\end{aligned}$$

Restricting to the hypersurface $\alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3)$ in $\mathbf{R}^4$, we can identify $\widehat{V}$ with the volume $\mathrm{vol}_3(R)$ of the Lambert cube $R$, since again (cf. A.):

(i) For $\alpha_4 = \theta(\alpha_1, \alpha_2, \alpha_3)$:

$$\frac{\partial \widehat{V}}{\partial \alpha_k} = \frac{\partial \mathrm{vol}_3(R)}{\partial \alpha_k}, \qquad k = 1, 2, 3.$$

(ii) For $\alpha_4 = \theta = \frac{\pi}{2}$, $\widehat{V}$ and $\mathrm{vol}_3(R)$ vanish according to (14.49) and Proposition 14.3, 14.2.3.

Thus, we derived the following

THEOREM 14.6. *Let $R$ denote a Lambert cube with essential angles $0 \leq \alpha_k \leq \frac{\pi}{2}$ and corresponding apices of length $V_k$, $k = 1, 2, 3$. Then*

$$\text{(14.50)}\quad \mathrm{vol}_3(R) = \frac{1}{4}\sum_{k=1}^{3}\{Л(\alpha_k + \theta) - Л(\alpha_k - \theta)\} - \frac{1}{4}Л(2\theta) + \frac{1}{2}Л\left(\frac{\pi}{2} - \theta\right),$$

*where*

$$0 \leq \theta = \arctan\left(\frac{\cosh^2 V_k - \sin^2\alpha_{k-1}\sin^2\alpha_{k+1}}{\cos^2\alpha_{k-1}\cos^2\alpha_{k+1}}\right)^{1/2} \leq \frac{\pi}{2}, \qquad k \bmod 3.$$

**Remarks** (a) By means of hyperbolic trigonometry, the quantities $\cosh^2 V_k$ in Theorem 14.6 can be expressed as functions of the essential angles $\alpha_1$, $\alpha_2$, $\alpha_3$ according to

$$\text{(14.51)}\quad \begin{aligned}\cosh^2 V_k &= 1 + \frac{1}{2}\left(\sqrt{A_k^2 + (2B_k \sin\alpha_k)^2} - A_k\right) \quad \text{with} \\ A_k &= \cos^2\alpha_{k-1} + \cos^2\alpha_{k+1} - B_k^2, \\ B_k &= \frac{\cos\alpha_{k-1}\cos\alpha_{k+1}}{\cos\alpha_k}, \qquad k \bmod 3.\end{aligned}$$

(b) In the limiting case of an asymptotic orthoscheme $R_2$ of degree 2 with scheme given in Figure 14.3 the two formulae for $\mathrm{vol}_3(R_2)$ of Theorem 14.5 and Theorem 14.6 coincide.

Apart from this exceptional case, these two formulae are conceptually different and cannot be related to each other by means of suitable functional equations for $Л(\omega)$. This can be seen by evaluating both expressions for the values $\alpha_1 = \alpha_2 = \alpha_3 = \frac{\pi}{4}$ using (14.51).



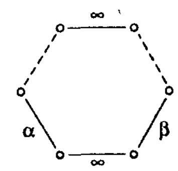



FIGURE 14.3

**14.3. Applications.**

14.3.1. *Volumes of Coxeter polytopes.* By means of Theorems 14.5, 14.6 and the reduction formula 14.3, the volumes of all Coxeter orthoschemes of degree $d$, $0 \leq d \leq 2$, of dimension three and of even dimensions (existing only up to dimension eight) can be explicitly calculated. In dimension three, there are infinitely many of such Coxeter polyhedra. For $d = 0$, however, there are exactly 10 realizations with schemes and volumes according to the following list in Table 14.1.

TABLE 14.1

| Scheme | Volume |
|---|---|
| $\circ - \circ - \circ \overset{6}{-} \circ$ | $\frac{1}{8}Л\left(\frac{\pi}{3}\right) \simeq 0.0423$ |
| $\circ - \circ = \circ = \circ$ | $\frac{1}{6}Л\left(\frac{\pi}{4}\right) \simeq 0.0763$ |
| $\circ - \circ \overset{5}{-} \circ - \circ$ | $\simeq 0.0391$ |
| $\circ - \circ \overset{6}{-} \circ - \circ$ | $\frac{1}{2}Л\left(\frac{\pi}{3}\right) \simeq 0.1692$ |
| $\circ = \circ - \circ \overset{5}{-} \circ$ | $\simeq 0.0359$ |
| $\circ = \circ - \circ \overset{6}{-} \circ$ | $\frac{5}{24}Л\left(\frac{\pi}{6}\right) \simeq 0.1057$ |
| $\circ = \circ = \circ = \circ$ | $\frac{1}{2}Л\left(\frac{\pi}{4}\right) \simeq 0.2290$ |
| $\circ \overset{5}{-} \circ - \circ \overset{5}{-} \circ$ | $\simeq 0.0933$ |
| $\circ \overset{5}{-} \circ - \circ \overset{6}{-} \circ$ | $\simeq 0.1715$ |
| $\circ \overset{6}{-} \circ - \circ \overset{6}{-} \circ$ | $\frac{1}{2}Л\left(\frac{\pi}{6}\right) \simeq 0.2537$ |

For $d = 1$, the polyhedron with scheme

$$\circ \overset{\infty}{-} \circ = \circ = \circ \overset{\infty}{-} \circ$$

and with volume $Л(\frac{\pi}{4}) \simeq 0.4580$ is the simply truncated Coxeter orthoscheme of maximal volume. For $d = 2$, the volume is maximal and equal to $2Л(\frac{\pi}{4}) \simeq 0.9160$ for the totally asymptotic Lambert cube with scheme given in Figure 14.4.

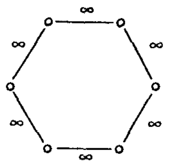



FIGURE 14.4

The values for the even-dimensional cases may be found in [15, Appendix]. By dissection into orthoschemes, R. Meyerhoff determined the volumes of all Coxeter simplexes of dimension three (see [21, Appendix]).

14.3.2. *Volumes of totally asymptotic and regular hyperbolic simplexes.* Every simplex $S \subset X^n$ with acute dihedral angles can be dissected into orthoschemes in several ways. The dissection process $\chi_p := \chi_p(S)$ consists in taking a point $p \in S$ and in drawing successively the perpendiculars to the faces of lower dimensions. If $p$ is an interior point of $S$, then $S$ is subdivided into $(n+1)!$ orthoschemes. If $p$ coincides with a vertex $v \in S$, then $S$ is dissected into $n!$ orthoschemes.

Consider a totally asymptotic simplex $S_\infty \subset \overline{H^3}$ with angles $\alpha_1, \alpha_2, \alpha_3$ along edges intersecting in a vertex and thus satisfying $\alpha_1 + \alpha_2 + \alpha_3 = \pi$. By $\chi_v(S_\infty)$, where $v$ denotes a vertex of $S_\infty$, $S_\infty$ is subdivided into three pairs of congruent doubly asymptotic orthoschemes characterized by the angle $\alpha_i$ and, by Theorem 14.5, with volume $\frac{1}{2}\text{Л}(\alpha_i)$, $1 \le i \le 3$, respectively. Hence, one obtains (cf. also [22, Lemma 2, p. 18])

$$\text{(14.52)} \qquad \operatorname{vol}_3(S_\infty) = \text{Л}(\alpha_1) + \text{Л}(\alpha_2) + \text{Л}(\alpha_3).$$

Denote by $S_{\text{reg}}$ a regular simplex (implying that all facets and vertex simplexes are regular), which is therefore parametrized by the dihedral angle $2\alpha$, say. Then, the dissecting orthoschemes with respect to $\chi_c$, where $c$ is the center of $S$, and with respect to $\chi_v$ are all congruent, and described by the characteristic scheme

$$\sigma_{n+1} = \sigma_{n+1}(\alpha): \quad \circ \overset{\alpha}{\text{——}} \circ \text{——} \circ \text{——} \cdots \text{——} \circ \text{——} \circ$$

of order $n+1$ for $\chi_c$ (cf. [4, Satz 1, p. 271]), and by

$$\nu_{n+1} = \nu_{n+1}(\alpha): \quad \circ \overset{2\alpha}{\text{——}} \circ \overset{\alpha}{\text{——}} \circ \text{——} \circ \text{——} \cdots \text{——} \circ \text{——} \circ$$

of order $n+1$ for $\chi_v$; this last result follows inductively from the fact that $\chi_v(S)$ induces the process $\chi_{c'}(S(v))$ for the regular facet simplex $S(v)$ opposite to $v$, where $c'$ is the center of $S(v)$. Then, $S(v)$ and, simultaneously, the vertex simplex $S_v$ of $v \in S$ are dissected each into $n!$ congruent orthoschemes. Hence, the $(n+1)!$ congruent orthoschemes subdividing $S_{\text{reg}}$ have principal vertices whose vertex orthoschemes are described by $\sigma_n$ and $\nu_n$, respectively. In particular, if $S^\infty_{\text{reg}}(2\alpha) \subset \overline{H^n}$, $n \ge 2$, denotes a totally asymptotic regular simplex with scheme $\Sigma^\infty_{n+1}$, then



$$\text{(14.53)} \qquad F_n(\Sigma^\infty_{n+1}) = (n+1)!F_n(\sigma_{n+1}) = n!F_n(\nu_{n+1}).$$

Since the subschemes $\sigma_n$, $\nu_n$ are parabolic with $\det\sigma_n = \det\nu_n = 0$, it follows by Lemma 14.1, (14.5), using (14.18) that

$$\text{(14.54)} \qquad \cos 2\alpha = \frac{1}{n-1}.$$

Thus, up to motion, there is only one totally asymptotic regular simplex in $\overline{H^n}$, $n \ge 2$. Using Theorem 14.5 for the three-dimensional case, and for $n = 2m \ge 2$ even, the reduction formula 14.3

$$\text{(14.55)} \qquad \begin{aligned} F_{2m}(\Sigma^\infty_{2m+1}) &= (2m+1)!F_{2m}(\sigma_{2m+1}) \\ &= (2m+1)!\sum_{k=0}^{m} \frac{(-1)^k}{k+1}\binom{2k}{k}\sum_\sigma f_{2m-(2k+1)}(\sigma), \end{aligned}$$

where $\sigma$ runs through all elliptic subschemes of order $2(m-k)$ of $\sigma_{2m+1}$ all of whose components are of even order, we obtain the following results:

(a) For $n = 2$, the area of a totally asymptotic triangle equals $\pi \simeq 3.1416$.

(b) For $n = 3$, the condition (14.54) implies that $\alpha = \frac{\pi}{6}$. If $R_\infty$ denotes the doubly asymptotic orthoscheme given by

$$\nu_4(\pi/6): \quad \circ \text{——} \circ \overset{6}{\text{——}} \circ \text{——} \circ,$$

one gets by (14.53) (cf. [22, Corollary, p. 20])

$$\text{(14.56)} \qquad \operatorname{vol}_3(S^\infty_{\text{reg}}(\pi/3)) = 3!\operatorname{vol}_3(R_\infty) = 3\text{Л}(\pi/3) \simeq 1.0149.$$

(c) For $n = 4$, we have $\cos 2\alpha = \frac{1}{3}$. Using (14.55) and results from 14.2.1. we deduce

$$\begin{aligned} \frac{1}{5!}F_4(\Sigma^\infty_5) = F_4(\sigma_5) &= f_3(A_4) + f_1(\circ\text{——}\circ)f_1(\circ\overset{\alpha}{\text{——}}\circ) \\ &\quad - (3f_1(\circ\text{——}\circ) - f_1(\circ\overset{\alpha}{\text{——}}\circ)) + 2 \\ &= \frac{2}{3}\left(\frac{1}{5} - \frac{\alpha}{\pi}\right). \end{aligned}$$

Thus,

$$\text{(14.57)} \qquad \operatorname{vol}_4(S^\infty_{\text{reg}}(2\alpha)) = \frac{\pi^2}{12}F_4(\Sigma^\infty_5) = \frac{4\pi}{3}(\pi - 5\alpha) \simeq 0.2689.$$

(d) For $n = 6$, we have $\cos 2\alpha = 1/5$, and

$$\begin{aligned} \frac{1}{7!}F_6(\Sigma^\infty_7) &= F_6(\sigma_7) \\ &= f_5(A_6) + f_3(A_4)f_1(\sigma_2) + f_1(A_2)f_3(\sigma_4) \\ &\quad - (3f_3(A_4) + f_3(\sigma_4) + 3f_1(A_2)f_1(\sigma_2) + 3(f_1(A_2))^2) \\ &\quad + 2(5f_1(A_2) + f_1(\sigma_2)) - 5 \\ &= -\frac{1}{3}\left\{f_3(\sigma_4) - \frac{4}{5\pi}\alpha + \frac{17}{105}\right\}, \end{aligned}$$

from which it follows that

$$\mathrm{vol}_6(S^{\infty}_{\mathrm{reg}}(2\alpha)) = 2\pi^3\left\{7f_3(\sigma_4)+\frac{17}{15}\right\} - \frac{56\pi^2}{5}\alpha. \tag{14.58}$$

By Schläfli's result for volumes of spherical orthoschemes (see 14.0, (14.4)), which can also be deduced from Theorem 14.5, (14.48), by means of analytic continuation using 14.24, (14.38), we obtain

(14.59)
$$f_3(\sigma_4) = \frac{2}{\pi^2}\left\{\mathrm{Li}_2(a,2\alpha)+\frac{1}{6}(\mathrm{Li}_2(a^3)-\mathrm{Li}_2(-a^3))-\frac{1}{2}(\mathrm{Li}_2(a)+\mathrm{Li}_2(-a))\right\}$$
$$-\frac{1}{2}\left(1-\frac{2}{\pi}\alpha\right)^2+\frac{1}{6}$$

where

$$\alpha = \arccos(\sqrt{3/5}), \qquad a = \frac{\sqrt{3}-1}{\sqrt{3}+1}, \qquad \mathrm{Li}_2(a,2\alpha) = \sum_{n=1}^{\infty}\frac{a^n}{n^2}\cos(2n\alpha).$$

This yields

$$\mathrm{vol}_6(S^{\infty}_{\mathrm{reg}}(2\alpha)) \simeq 0.0102. \tag{14.60}$$

By a result of Haagerup and Munkholm [11], a simplex in $\overline{H^n}$, $n \geq 2$, is of maximal volume if and only if it is totally asymptotic and regular. Hence, for $n = 2, 3, 4, 6$, (a)–(d) provide an upper bound for the volume of an arbitrary simplex in $\overline{H^n}$, and we see that $\mathrm{vol}_n(S^{\infty}_{\mathrm{reg}})$ is a decreasing function with respect to $n$ verifying the inequality

$$\frac{n-1}{n^2} \leq \frac{\mathrm{vol}_{n+1}(S^{\infty}_{\mathrm{reg}})}{\mathrm{vol}_n(S^{\infty}_{\mathrm{reg}})} \leq \frac{1}{n}, \tag{14.61}$$

due to Haagerup and Munkholm (see [11, Proposition 2, p. 4]).

14.3.3. *Geometrical functional equations for* $Л(\omega)$. Using different methods (e.g., cutting and pasting, limiting processes) for calculating the volume of a given convex polyhedron, one simultaneously obtains functional equations for the characteristic volume function $Л(\omega)$. In the following paragraph, we will present some examples which are derived by dissection into orthoschemes of degree $d = 0, 1, 2$.

EXAMPLES. (a) Every orthoscheme $R_d \subset H^3$ of degree $d > 0$ admits a dissection into exactly three orthoschemes $R^i_{d-1}$, $i = 1, 2, 3$, of degree $d-1$. In the simplest case of a simply truncated asymptotic orthoscheme given by the graph

$$\circ \cdots\cdots \circ \overset{\frac{\pi}{2}-\alpha}{\text{——}} \circ \overset{\alpha}{\text{——}} \circ \overset{\beta}{\text{——}} \circ, \quad \text{where } 0 < \alpha+\beta < \frac{\pi}{2},$$



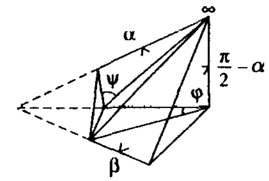



FIGURE 14.5

this dissection (see Figure 14.5) implies the identity

(14.62)
$$\begin{aligned}
&2Л(\alpha)+Л\left(\frac{\pi}{2}-\alpha+\beta\right)-Л\left(\frac{\pi}{2}+\alpha+\beta\right)\\
&=2\left\{Л(\alpha+\psi)-Л\left(\frac{\pi}{2}+\alpha\right)-Л\left(\frac{\pi}{2}+\psi\right)\right\}\\
&\quad+Л(\psi+\phi)+Л(\psi-\phi)+Л\left(\frac{\pi}{2}-\phi+\alpha\right)+Л\left(\frac{\pi}{2}+\phi+\alpha\right)\\
&\quad+Л\left(\frac{\pi}{2}+\beta-\alpha-\psi\right)-Л\left(\frac{\pi}{2}+\beta+\alpha+\psi\right),
\end{aligned}$$

where $\phi, \psi \in (0, \pi/2)$ with

$$\phi = \arccos(\sin\alpha/\cos\beta), \quad \psi = \arctan((\cos^2\beta-\sin^2\alpha)/\sin\alpha\cos\alpha).$$

Lewin (private communication) showed that the second part of (14.62) can be reduced to the first part of (14.62) by two applications of Kummer's equation for $Л(\omega) = \frac{1}{2}\mathrm{Cl}(2\omega)$ (see [18, (4.68) and (4.69)]).

(b) Let $R$ denote the totally asymptotic Lambert cube with dihedral angles $\frac{\pi}{4}$. Since $\mathrm{vol}_3(R) = 2\cdot Л(\frac{\pi}{4}) = 4\cdot\frac{1}{2}Л(\frac{\pi}{4})$ (cf. 14.3.1.), one could expect a dissection of $R$ into 2 congruent simply truncated orthoschemes with graph

$$\circ \overset{\infty}{\text{——}} \circ = \circ = \circ \overset{\infty}{\text{——}} \circ,$$

and into 4 congruent orthoschemes with graph

$$\circ = \circ = \circ = \circ.$$

In fact, Figure 14.6 shows such a dissection. More obvious is the subdivision of $R$ into 2 congruent simplexes $S$ with scheme given in Figure 14.7. Hence, $\mathrm{vol}_3(S) = Л(\frac{\pi}{4})$.

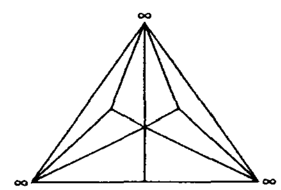



FIGURE 14.6

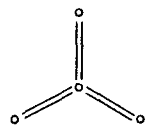


Figure 14.7

(c) Denote by $H_{\mathrm{reg}}^{\infty}$ the totally asymptotic regular hexahedron (or "cube") in $\overline{H^3}$ all of whose dihedral angles have to be equal to $\frac{\pi}{3}$. $H_{\mathrm{reg}}^{\infty}$ may be viewed as totally asymptotic regular tetrahedron $S_{\mathrm{reg}}^{\infty}(\frac{\pi}{3})$ to each of whose facets has been adjoined another tetrahedron. On the other hand, the dissection $\chi_c(H_{\mathrm{reg}}^{\infty})$ (cf. 14.3.2), where $c$ denotes the center of $H_{\mathrm{reg}}^{\infty}$, yields a subdivision into 48 orthoschemes with graph

$$\circ\!=\!=\!\circ\!-\!-\!\circ\overset{6}{-\!-}\circ.$$

Thus, by 14.3.1 and (14.56), one gets

$$9Л\left(\frac{\pi}{3}\right) = 12\left(Л\left(\frac{\pi}{12}\right) - Л\left(\frac{5\pi}{12}\right)\right).$$

By 14.2.4, (a), this equation is equivalent to the distribution law

$$Л\left(4\cdot\frac{\pi}{12}\right) = 4\cdot\sum_{1\le k\le 3} Л\left(\frac{\pi}{12}+\frac{k\pi}{4}\right).$$

By another geometrical construction, Thurston proved the distribution law for $Л(3\alpha)$, $0<\alpha<\frac{\pi}{3}$, (cf. [**26**, Proposition 4.12, p. 201]).

(d) Denote by $g_m = g(P_1, \dots, P_m)$, $m \ge 3$, a plane regular hyperbolic $m$-gon with vertices $P_1, \dots, P_m$ and consider the pyramid $C := C_m(\alpha, 2\beta) \subset H^3$ over $g_m$ with apex $A$ such that the dihedral angles formed by $g_m$ and the laterals equal $\alpha$ and the angles between two intersecting laterals equal $2\beta$. With respect to the dissection $\chi_A(C)$, $C$ admits a simplicial subdivision into $4m$ congruent orthoschemes given by

$$\circ\overset{\alpha}{-\!-}\circ\overset{\beta}{-\!-}\circ\overset{\frac{\pi}{m}}{-\!-}\circ.$$

If all $m+1$ vertices of $C$ are points are infinity, then $\alpha = \frac{\pi}{2} - \beta = \frac{\pi}{m}$, and for the volume of $C_m^{\infty} := C_m(\frac{\pi}{m}, \frac{m-2}{m}\pi)$, Theorem 14.5 yields

$$\mathrm{vol}_3(C_m^{\infty}) = mЛ\left(\frac{\pi}{m}\right). \tag{14.63}$$

In particular, for $m=3$, we obtain again the formula (14.56) for the volume of $S_{\mathrm{reg}}^{\infty}$. Now assume that the vertices $P_1, \dots, P_m$ of the basis $g_m$ are ideal in such a way that all edges $P_iP_{i+1}$, $1 \le i \le m-1$, *intersect* the absolute quadric $Q_{3,1}$. Furthermore, let the apex $A \in AQ_{3,1}$ be an ideal point; $A$ can be chosen "near" the quadric $Q_{3,1}$ such that the edges $AP_i$, $i = 1, \dots, m$, are *secants* with respect to $Q_{3,1}$ *of equal hyperbolic length.*



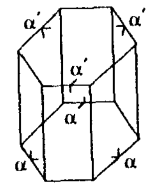



FIGURE 14.8

Thus, polar truncation of each vertex of $C = C_m(\alpha, 2\beta)$ yields a $2m$-gonal prism $P_{2m}(\alpha, \alpha')$ whose mutually orthogonal laterals form by turns the angles $\alpha$, $\pi/2$ resp. $\alpha'$, $\pi/2$ with one resp. the other totally rectangular $2m$-gon (see Figure 14.8 for the case $m=3$). The dissection $\chi_A(C)$ of the underlying pyramid $C$ in $P^3$ into orthoschemes induces a subdivision of $P_{2m}(\alpha, \alpha')$ into $2m$ congruent Lambert cubes with essential angles $\alpha$, $\alpha'$, $\pi/m$. (See Figure 14.8.)

In the limiting case $\alpha = \alpha' = 0$, Theorem 14.6 yields for the volume of $P_{2m}^{\infty} := P_{2m}(0, 0)$

(14.64)
$$\mathrm{vol}_3(P_{2m}^{\infty}) = 2mЛ(\theta_m) + \frac{m}{2}\left\{Л\left(\frac{\pi}{m}+\theta_m\right) - Л\left(\frac{\pi}{m}-\theta_m\right) - Л(2\theta_m) + 2Л\left(\frac{\pi}{2}-\theta_m\right)\right\},$$

$$0 \le \theta_m := \operatorname{arccot}\left(\cos\frac{\pi}{m}\right) \le \frac{\pi}{2}.$$

In particular, if $m=3$, the prism $P_6^{\infty}$ is the totally asymptotic regular octahedron $O_{\mathrm{reg}}^{\infty}$ with dihedral angles $\frac{\pi}{4}$, which can be dissected by means of the "central" dissection $\chi_c(O_{\mathrm{reg}}^{\infty})$ into 16 congruent orthoschemes given by

$$\circ\!=\!=\!\circ\!=\!=\!\circ\!=\!=\!\circ$$

as well as into 4 congruent simplexes $S$ with scheme given in Figure 14.9.

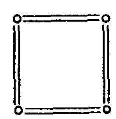
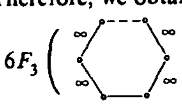
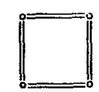


FIGURE 14.9

Therefore, we obtain the relations

$6F_3\Big($  $\Big) = 16F_3(\circ\!=\!=\!\circ\!=\!=\!\circ\!=\!=\!\circ) = 4F_3\Big($  $\Big)$, which, by 14.3.1 and (14.64), imply that $\mathrm{vol}_3(O^{\infty}_{\mathrm{reg}}) = 8\text{Л}(\frac{\pi}{4})$, $\mathrm{vol}_3(S) = 2\text{Л}(\frac{\pi}{4})$, and

$$(14.65)\qquad \begin{aligned} &6\text{Л}(\theta_3) + \frac{3}{2}\left\{\text{Л}\left(\frac{\pi}{3}+\theta_3\right) - \text{Л}\left(\frac{\pi}{3}-\theta_3\right) - \text{Л}(2\theta_3) + 2\text{Л}\left(\frac{\pi}{2}-\theta_3\right)\right\} \\ &\quad = 8\text{Л}\left(\frac{\pi}{4}\right) \simeq 3.6639, \end{aligned}$$

where $\theta_3 = \arctan 2$.

(e) Finally, consider the remaining totally asymptotic regular polyhedron, the dodecahedron $D^{\infty}_{\mathrm{reg}}$ with dihedral angles $\frac{\pi}{3}$. The dissection $\chi_c(D^{\infty}_{\mathrm{reg}})$ starting from the center $c \in D^{\infty}_{\mathrm{reg}}$ yields a subdivision into 120 congruent orthoschemes

$$\circ \overset{5}{—} \circ — \circ \overset{6}{—} \circ\,,$$

which leads to the equality

$$(14.66)\qquad \mathrm{vol}_3(D^{\infty}_{\mathrm{reg}}) = 30\left\{20\text{Л}\left(\frac{\pi}{3}\right) + \text{Л}\left(\frac{\pi}{5}+\frac{\pi}{6}\right) - \text{Л}\left(\frac{\pi}{5}-\frac{\pi}{6}\right)\right\}.$$

### 14.4. Further aspects.

14.4.1. *Volumes of hyperbolic 3-folds and Dedekind zeta functions.* Denote by $M^n$, $n \geq 2$, an $n$-dimensional complete hyperbolic space form $H^n/\Gamma$ of finite volume (orientable or nonorientable), where $\Gamma$ is a discrete group of isometries of $H^n$. If $\Gamma$ acts without fixpoints on $H^n$, then $M^n$ is a manifold; otherwise it is an orbifold locally modelled on $\mathbf{R}^n$ modulo a finite linear group action. One of the most important (topological) invariants of $M^n$ is its volume. By a result of Wang [31], the volumes $\mathrm{vol}_n(M^n)$, where $M^n$ runs through all $n$-manifolds, form a discrete subset of $\mathbf{R}_+$ if $n \neq 3$ (for $n$ even, this is a consequence of the Gauss-Bonnet formula). For $n = 3$, however, Jørgensen and Thurston [28] proved that the volume spectrum

$$\mathbf{Vol} := \{\mathrm{vol}_3(M) \mid M \text{ hyperbolic 3-fold}\}$$

is a closed, nondiscrete subset of $\mathbf{R}_+$, which is well-ordered and of order type $\omega^{\omega}$. In particular, there is a manifold (resp. orbifold) of minimum volume $v_1$ (resp. $v_1'$), and a 1-cusped manifold (resp. orbifold) of minimum volume $v_\omega$ (resp. $v_c'$). In the special case of orientable 3-manifolds, Neumann and Zagier [24] determined the metric structure of the corresponding volume spectrum near its accumulation points. In particular, they proved that if $\{M_k\}$ is the set of all manifolds obtained from a given manifold $M$ by performing Dehn surgery on a single cusp on $M$, then (cf. [24, Corollary, p. 308])

$$\#\{k \mid \mathrm{vol}_3(M_k) < \mathrm{vol}_3(M) - \tfrac{1}{x}\} = 6\pi x + O(x^{1/2}) \quad \text{as } x \to \infty.$$

However, in contrast to the manifold case, there are noncompact 3-orbifolds whose volumes are isolated (cf. [20, p. 278]). Up to now, very little is known about the smallest elements of **Vol**. In the sequel, we shall collect what is



known to us (the lower volume bounds are due mostly to Meyerhoff [20]):

$$(14.67)\qquad 0.00115 < v_1 \leq \mathrm{vol}_3(W) \simeq 0.9427, \qquad (v_1)$$

where $W$ denotes the orientable Weeks manifold which is obtained by (5,1), (5,2) Dehn surgery on the complement of the Whitehead link in $S^3$ (cf. [7]). Here, the lower bound is due F. Gehring and G. Martin (cf. [F. Gehring, G. Martin, Inequalities for Möbius transformations and discrete groups, Preprint]) and improves the earlier result of 0.00082 found by Meyerhoff.

$$(14.68)\qquad 0.0000013 < v_1' \leq \mathrm{vol}_3(Q) \simeq 0.0391, \qquad (v_1')$$

where $Q$ is the following orientable tetrahedral orbifold (cf. [21]): Consider the reflection group $\Gamma$ associated to the Coxeter orthoscheme

$$\circ — \circ \overset{5}{—} \circ — \circ.$$

This simplex admits an inner twofold symmetry induced by a rotation of $\pi$. Denote by $\Gamma'$ the isometry group generated by $\Gamma$ and this rotation, and by $\Gamma'_+$ its subgroup of orientation-preserving isometries. Then, $Q$ is the quotient $H^n/\Gamma'_+$ of volume $\simeq 0.0391$ (see 14.3.1).

$$(14.69)\qquad 0.5074 = v/2 \leq v_\omega = \mathrm{vol}_3(G) = v \simeq 1.0149, \qquad (v_\omega)$$

where $v = \mathrm{vol}_3(S^{\infty}_{\mathrm{reg}})$, and where $G$ denotes the Gieseking manifold; this is the unique noncompact (nonorientable) 3-manifold of minimal volume (cf. [1]). $G$ is obtained from $S^{\infty}_{\mathrm{reg}}(\pi/3)$ by identifying two faces by means of a rotation of $2\pi/3$ about a common vertex and by identifying the opposite two faces by a rotation about a common vertex.

$$(14.70)\qquad \sqrt{3}/24 \leq v_c' \leq \mathrm{vol}_3(Q_c) \simeq 0.0846, \qquad (v_c')$$

where $Q_c$ denotes the 1-cusped orientable doublecover of the (nonorientable) tetrahedral orbifold with Coxeter scheme

$$\circ — \circ — \circ \overset{6}{—} \circ.$$

$(v_\vartheta)$ Kojima and Miyamoto [17] considered compact 3-manifolds with nonempty geodesic boundary and showed that the one of minimum volume $v_\vartheta$, which is necessarily orientable but not unique, admits a subdivision into two regular truncated tetrahedra with dihedral angles $\pi/6$ or, equivalently, into 6 doubly truncated Coxeter orthoschemes with graph given in Figure 14.10.

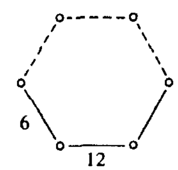



FIGURE 14.10

Hence, by Theorem 14.5, $v_\theta \simeq 6.4520$.

($v_{\omega^N}$) Adams [2] investigated $N$-cusped 3-manifolds $M_{\omega^N}$ and proved that $\mathrm{vol}_3(M_{\omega^N}) \geq N \cdot v$, where again $v = \mathrm{vol}_3(S^\infty_{\mathrm{reg}})$. These lower bounds are—in a unique way—realizable for $N = 1$ (by the Gieseking manifold) and for $N = 2$, whereas for $N > 2$, $\mathrm{vol}_3(M_{\omega^N}) > N \cdot v$.

For arithmetically definable hyperbolic space forms, the volumes can be related to values of Dedekind zeta functions at 2 (cf. [5] and [22]):

Denote by $F = \mathbf{Q}(\sqrt{-d})$, $d \geq 1$ square-free, an imaginary quadratic number field with discriminant $d > 0$, and by $\mathscr{O}_d$ its ring of integers. Then, the group $\Gamma = \mathrm{PSL}(2, \mathscr{O}_d)$ is a discrete subgroup of $\mathrm{PSL}(2, \mathbf{C})$ and acts therefore on $H^3$ by isometries.

Let $\zeta_F$ be the Dedekind zeta function associated to $F$, which can be written in the form

$$\zeta_F(s) = \zeta(s) \cdot \sum_{r=1}^{\infty} \left(\frac{-d}{r}\right) r^{-s}, \tag{14.71}$$

where $(\frac{-d}{r})$ is the Kronecker symbol with values 0 or $\pm 1$ associated to $F$. By a result of Humbert, the volume of a fundamental domain $D$ of $\Gamma$ is given by (cf. [22, p. 20])

$$\mathrm{vol}_3(D) = d^{3/2}\zeta_F(2)/4\pi^2. \tag{14.72}$$

Borel [5, 6] generalized Humbert's formula (14.72) to arbitrary number fields $F$ having exactly $b$ complex places and at least $a$ real places, where $a, b$ are nonnegative integers with $a + b \geq 1$ (see Example (b) for $a = 2$, $b = 1$).

Consider Humbert's formula (14.72) and use (see [22, p. 21])

$$\left(\frac{-d}{r}\right)\sqrt{d} = \sum_{0<k<d} \left(\frac{-d}{k}\right) \sin\left(\frac{2\pi i k r}{d}\right).$$

Then, one obtains the following expression using 14.2.4, (14.40),

$$\mathrm{vol}_3(D) = \frac{d}{12} \sum_{0<k<d} \left(\frac{-d}{k}\right) Л\left(\frac{k\pi}{d}\right), \tag{14.73}$$

which together with (14.72) implies

$$\zeta_{\mathbf{Q}(\sqrt{-d})}(2) = \frac{\pi^2}{3\sqrt{d}} \sum_{0<k<d} \left(\frac{-d}{k}\right) Л\left(\frac{k\pi}{d}\right). \tag{14.74}$$

In [32], Zagier extended this result for $\zeta_F(2)$ to arbitrary number fields $F$ first in terms of $A(x) = 2Л(\mathrm{arccot} x)$. This motivated his conjecture about the representation of $\zeta_F(m)$, $m > 1$, in terms of (modified) polylogarithms (cf. [33]).

Examples. (a) Consider the field $F = \mathbf{Q}(\sqrt{-3})$. Then, by a result of Meyerhoff (cf. [21]),

$$H^3/\mathrm{PGL}(2, \mathscr{O}_3) = Q_c.$$



Since $[\mathrm{PGL}(2, \mathscr{O}_3) : \mathrm{PSL}(2, \mathscr{O}_3)] = 2$, the generalized Humbert formula combined with $(v_c)$ yields

$$3^{3/2}\zeta_{\mathbf{Q}(\sqrt{-3})}(2)/8\pi^2 = \mathrm{vol}_3(Q_c) = \frac{1}{4}Л\left(\frac{\pi}{3}\right) \simeq 0.0846, \quad \text{and}$$

$$\zeta_{\mathbf{Q}(\sqrt{-3})}(2) = \frac{2\pi^2}{3\sqrt{3}} Л\left(\frac{\pi}{3}\right).$$

(b) Consider the field $K = \mathbf{Q}(\sqrt{3 + 2\sqrt{5}})$ of discriminant $-275$ with $a = 2$, $b = 1$. Choose a maximal order $\mathscr{D}$ (all are mutually conjugate) of the Hamilton quaternion algebra over $K$. Then, one can associate to it a discrete subgroup $\Gamma_{\mathscr{D}}$ of $\mathrm{SL}(2, \mathbf{C})$ such that $H^3/\Gamma_{\mathscr{D}} = Q$ (cf. [5, p. 30] and [21, Remark (2), p. 186]). Thus, the formula of Borel (cf. [6, §2]) and $(v_1')$ yield

$$\frac{275^{3/2}}{2^7\pi^6}\zeta_K(2) = \mathrm{vol}_3(Q) \simeq 0.0391.$$

Therefore, we obtain

$$\begin{aligned}\zeta_K(2) = \frac{2^5\pi^6}{275^{3/2}} \Big\{ & 2Л\left(\frac{\pi}{3} + \theta\right) - 2Л\left(\frac{\pi}{3} - \theta\right) - Л\left(\frac{\pi}{6} + \theta\right) \\ & + Л\left(\frac{\pi}{6} - \theta\right) + 2Л\left(\frac{\pi}{2} - \theta\right)\Big\} \\ \simeq 1.0537, &\end{aligned}$$

where $\theta = \arctan(\sqrt{2\sqrt{5} - 3})$.

14.4.2. *Scissors congruence and Dehn invariants.* By cutting and pasting of polytopes in a space $X^n$ of constant curvature, one immediately gets in touch with the notion of scissors congruence groups $\mathscr{P}(X^n)$ and Hilbert's Third Problem concerning scissors congruence (equidecomposability of polytopes):

Let $G$ denote the group of isometries of $X^n$. Then, the scissors congruence group $\mathscr{P}(X^n)$ is defined to be the free abelian group generated by symbols $[P]$, one for each polytope $P \subset X^n$, modulo the relations

(i) $[P] = [P_1] + [P_2]$, if $P = P_1 + P_2$ in the sense of elementary geometry.
(ii) $[gP] = [P]$, where $g \in G$.

The scissors congruence problem consists in finding a complete system of invariants for the classes of polytopes in $\mathscr{P}(X^n)$. By means of the Dehn invariants (suitably defined and including volume) the scissors congruence problem was solved for $n \leq 2$ (this is a classical result), in $E^3$ by Sydler and in $E^4$ by Jessen (for references, see [9, p. 159]). We are mainly interested in $\mathscr{P}(H^n)$ and, in particular, in the case $n = 3$. For the group $\mathscr{P}(H^3)$, one conjectures that the scissors congruence class of a hyperbolic polyhedron is determined by its volume and its Dehn invariant; recall that the classical Dehn invariant is given by

$$\Psi : \mathscr{P}(H^3) \to \mathbf{R} \otimes_{\mathbf{Z}} \mathbf{R}/2\pi\mathbf{Z},$$

associating to a class of $\mathcal{P}(H^3)$, represented by a polyhedron $P \subset H^3$ with dihedral angles $\alpha_A$ along edges $A$ of length $l(A)$, the expression

$$\Psi(P) = \sum_A l(A) \otimes \alpha_A. \tag{14.75}$$

**Remark.** The Dehn invariant looks very similar to Schläfli's volume differential (cf. 14.2.1., or [**15**, §2.2])

$$d\mathrm{vol}_3(P) = -\frac{1}{2}\sum_A l(A)d\alpha_A.$$

Like this Schläfli differential, the Dehn invariant can be extended to the set of asymptotic polyhedra representing the elements of $\mathcal{P}(\overline{H^3})$ (cf. [**9**, p. 168]).

Besides $\mathcal{P}(H^n)$ and $\mathcal{P}(\overline{H^n})$ there are other notions of hyperbolic scissors congruence groups, e.g., the group $\mathcal{P}(\partial H^n)$ generated by classes of totally asymptotic polytopes. Some of these groups are identical; by a result of Dupont (cf. [**9**, Theorem 2.1, p. 162]), the groups $\mathcal{P}(H^n)$ and $\mathcal{P}(\overline{H^n})$ are isomorphic for $n \geq 2$. If $\mathcal{P}(\overline{H^n})_\infty$ denotes the group associated to all totally asymptotic simplexes in $\overline{H^n}$, then $\mathcal{P}(\overline{H^{2k+1}})$ equals $\mathcal{P}(\overline{H^{2k+1}})_\infty$ for $k > 1$ (cf. [**26**, Proposition 3.7, p. 195]). Moreover, one can show that $\mathcal{P}(H^n)$ (resp. $\mathcal{P}(\overline{H^n})$) can be generated by classes of orthoschemes (resp. doubly asymptotic orthoschemes) (cf. [**26**, §4]), which implies that $\mathcal{P}(H^n)$, $n \geq 0$, (resp. $\mathcal{P}(\overline{H^n})$ for $n > 1$) is 2-divisible. In particular, $\mathcal{P}(\overline{H^3}) = \mathcal{P}(\overline{H^3})_\infty \cong \mathcal{P}(H^3)$ is 2-divisible.

Divisibility questions about $\mathcal{P}(\overline{H^3})$ can be reduced to corresponding problems for doubly asymptotic orthoschemes $R_\infty = R_\infty(\alpha)$ with graph

$$\circ \overset{\alpha}{\text{———}} \circ \overset{\frac{\pi}{2}-\alpha}{\text{———}} \circ \overset{\alpha}{\text{———}} \circ,$$

whose volumes equal $\frac{1}{2}Л(\alpha)$ (cf. Theorem 14.5): Consider the map

$$L: \mathbf{R} \to \mathcal{P}(\overline{H^3})$$

given by

$$L(\alpha) := [R_\infty(\alpha)], \tag{14.76}$$

$$L(\alpha) = -L(-\alpha), \qquad L(\alpha + \pi) = L(\alpha).$$

Then, one can show that $L$ satisfies a distribution law analogous to 14.2.4, (a), for the Lobachevsky function $Л(\alpha)$ (cf. [**26**, pp. 200–202]), which therefore admits a geometrical interpretation in terms of cutting and pasting of totally asymptotic simplexes (see also 14.3.3, (c)), and from which the divisibility of $\mathcal{P}(\overline{H^3})$ follows.

All these results can be brought into a more general context which allows a very elegant description of volume and scissors congruence problems in



$\overline{H^3}$. Consider the group $\mathcal{P}(\partial H^3)$ homologically defined by Dupont as follows (cf. [**9**, pp. 165–166]). Let $\mathcal{P}(\partial H^n)$ be the abelian group generated by $(a_0, \ldots, a_n)$, $a_i \in \partial H^n$, satisfying

(i) $(a_0, \ldots, a_n) = 0$, if $a_0, \ldots, a_n$ lie in a subspace of dimension $< n$.
(ii) $\sum_{0\leq i\leq n+1}(-1)^i(a_0, \ldots, \hat{a}_i, \ldots, a_{n+1}) = 0$ for $a_i \in \partial H^n$ arbitrary.
(iii) $(ga_0, \ldots, ga_n) = \det g \cdot (a_0, \ldots, a_n)$ for $a_i \in \partial H^n$ and $g$ an isometry of $H^n$.

This group is closely related to Thurston's group $\mathcal{P}'(\partial H^n)$ which is obtained from $\mathcal{P}(\partial H^n)$ by replacing (i) and (iii) by

(i)′ $(a_0, \ldots, a_n) = 0$, if $a_i = a_j$ for $i \neq j$.
(iii)′ $(ga_0, \ldots, ga_n) = (a_0, \ldots, a_n)$ for $a_i \in \partial H^n$ and $g$ an orientation preserving isometry of $H^n$.

For $n = 3$, the group $\mathcal{P}'(\partial H^3)$ can be viewed as a special case of a group $\mathcal{P}_F$ associated to an arbitrary field $F$, studied independently by Bloch, Wigner and Thurston: $\mathcal{P}_F$ is defined to be the abelian group generated by 4-tuples $(x_0, x_1, x_2, x_3)$, $x_i \in P^1(F) = F \cup \{\infty\}$ and $x_i \neq x_j$ for $i \neq j$, subject to the relations

(i) $(gx_0, gx_1, gx_2, gx_3) = (x_0, x_1, x_2, x_3)$ for $g \in \mathrm{PGL}(2, F)$.
(ii) $\sum_{0\leq i\leq 4}(-1)^i(x_0, \ldots, \hat{x}_i, \ldots, x_4) = 0$ for distinct $x_i \in P^1(F)$.

Consider the case $F = \mathbf{C}$, and use for $H^3$ the upper half space model $\mathbf{C} \times \mathbf{R}_+$ bounded by the Riemann sphere $\partial H^3 = \mathbf{C} \cup \{\infty\} = P^1(\mathbf{C})$. Then, $\mathrm{PGL}(2, \mathbf{C}) = \mathrm{PSL}(2, \mathbf{C})$ is the group of orientation preserving isometries of $H^3$ acting on $\partial H^3$ by

$$g(z) = \frac{az+b}{cz+d} \quad \text{for } z \in \mathbf{C} \cup \{\infty\}, \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PGL}(2, \mathbf{C}). \tag{14.77}$$

Recall that the cross-ratio $\{a_0 : a_1 : a_2 : a_3\}$ of four distinct points $a_0, a_1, a_2, a_3 \in P^1(\mathbf{C})$ is defined by

$$\{a_0 : a_1 : a_2 : a_3\} = (a_0 - a_2)(a_1 - a_3)/(a_0 - a_3)(a_1 - a_2) \in \mathbf{C}\backslash\{0, 1\}, \tag{14.78}$$

and that one has

(a) $\{\infty : 0 : 1 : z\} = z$.
(b) $\mathrm{PSL}(2, \mathbf{C})$ acts 3-transitively on $P^1(\mathbf{C})$.
(c) For two 4-tuples $(a_0, a_1, a_2, a_3)$, $(b_0, b_1, b_2, b_3)$ of distinct points in $P^1(\mathbf{C})$, one has
$\{a_0 : a_1 : a_2 : a_3\} = \{b_0 : b_1 : b_2 : b_3\} \quad \Leftrightarrow$
$\exists g \in \mathrm{PSL}(2, \mathbf{C})$ with $(b_0, b_1, b_2, b_3) = (ga_0, ga_1, ga_2, ga_3)$.

Using this, one observes that $\mathcal{P}'(\partial H^3)$ is the abelian group generated by $[z] := [(\infty, 0, 1, z)]$ for $z \in \mathbf{C}\backslash\{0, 1\}$, satisfying

$$\sum_{0\leq i\leq 4}(-1)^i[\{a_0 : \cdots : \hat{a}_i : \cdots : a_4\}] = 0 \quad \text{for arbitrary } a_i \in P^1(\mathbf{C}). \tag{14.79}$$

If all components $a_i$ in (14.79) are distinct, one can use (a)–(c) to express (14.79) in the form

(14.80) $[z_1]-[z_2]+[z_1/z_2]-[(1-z_1)/(1-z_2)]+[(1-z_2)z_1/(1-z_1)z_2]=0$,

where $z_1, z_2 \in \mathbf{C}\backslash\{0, 1\}$, $z_1 \neq z_2$. In fact, Dupont and Sah showed that $\mathscr{P}'(\partial H^3)$ is completely characterized by this relation, which implies that $\mathscr{P}'(\partial H^3) = \mathscr{P}_{\mathbf{C}}$ (see [**9**, Remark, p. 171]).

Denote by $S_\infty \subset \overline{H^3}$ a totally asymptotic simplex. By the above, its vertices can be brought into the form $\infty, 0, 1, z$, where $z \in P^1(\mathbf{C})$, and the volume of $S_\infty(z) = (\infty, 0, 1, z)$ becomes an expression in $z$: Since each euclidean vertex triangle of $S_\infty$ is similar to the triangle in $\mathbf{C}$ with vertices $0, 1, z$, the angles of $S_\infty$ along edges meeting at a vertex equal $\alpha := \arg z$, $\beta := \arg(1-(1/z))$, and $\gamma := \arg(1/(1-z)) = \pi - (\alpha+\beta)$ (cf. [**24**, p. 311]). Hence by (14.52), we obtain that

(14.81) $$\operatorname{vol}_3(S_\infty(z)) = Л(\arg z) + Л\left(\arg\left(1-\frac{1}{z}\right)\right) + Л\left(\arg\left(\frac{1}{1-z}\right)\right).$$

Now, denote by

$$D(z) := \operatorname{Im}(\operatorname{Li}_2(z)) + \arg(1-z)\log|z|,$$
$$z \in \mathbf{C}\backslash\{0, 1\},\ -\pi < \arg(1-z) < \pi,$$

the Bloch-Wigner Dilogarithm which has the following properties (cf. [**9**, p. 172]):

$$D(e^{2i\alpha}) = 2Л(\alpha);$$
$$-D(z) = D(\overline{z}) = D(1/z) = D(1-z);$$
$$D(z^n) = n \cdot \sum_{\mu^n=1} D(\mu z);$$

$$D(z_1) - D(z_2) + D(z_1/z_2) - D((1-z_1)/(1-z_2)) + D((1-z_2)z_1/(1-z_1)z_2) = 0$$
$$\text{for } z_1, z_2 \in \mathbf{C}\backslash\{0, 1\},\ z_1 \neq z_2.$$

Then, by a result of Bloch and Wigner (cf. [**9**, (4.13), p. 173]),

(14.82) $$D(z) = \operatorname{vol}_3(S_\infty(z)),$$

which together with (14.81) yields Kummer's equation (cf. [**18**, (5.5)]

(14.83) $$D(z) = Л(\arg z) + Л(\arg(1-\overline{z})) - Л(\arg z + \arg(1-\overline{z})).$$

Bloch considered $D$ as ("imaginary") part of a more general function, which, slightly modified by Dupont and Sah (cf. [**9**, (4.14), p. 173]), is of the following form:

$$\varrho: \mathbf{C}\backslash\{0, 1\} \to \wedge^2_{\mathbf{Z}}(\mathbf{C}),$$



(14.84)
$$\varrho(z) := \frac{1}{2\pi i}\log(z) \wedge \frac{1}{2\pi i}\log(1-z) + 1 \wedge \frac{1}{4\pi^2}\operatorname{Li}_2(z) - 1 \wedge \frac{1}{4\pi^2}\operatorname{Li}_2(1-z),$$

is well-defined for $0 < \operatorname{Re}(z) < 1$ satisfying $\varrho(z) + \varrho(1-z) = 0$ and can be analytically continued to $\mathbf{C}\backslash\{0, 1\}$. Here, $\wedge^2_{\mathbf{Z}}(\mathbf{C})$ is the second exterior power written additively (i.e., it consists of formal sums of symbols $a \wedge b$, $a, b \in \mathbf{C}$, which are bimultiplicative and satisfy $a \wedge a = 0$). Consider the map

$$\lambda: \mathscr{P}_{\mathbf{C}} \to \wedge^2_{\mathbf{Z}}(\mathbf{C}^\times)$$

defined by $\lambda[z] := z \wedge (1-z)$, where $\mathbf{C}^\times$ is the mulitplicative group $\mathbf{C}\backslash\{0\}$. Then by a result of Dupont and Sah, the following diagram commutes (cf. [**9**, p. 171–172]):

$$\begin{array}{ccc} \mathscr{P}_{\mathbf{C}} & \xrightarrow{2\lambda} & \wedge^2_{\mathbf{Z}}(\mathbf{C}^\times) \\ \downarrow & & \downarrow{\scriptstyle -\log^-} \\ \mathscr{P}(\partial H^3) & \xrightarrow{\Psi_e} & \mathbf{R} \otimes_{\mathbf{Z}} \mathbf{R}/2\pi\mathbf{Z} \end{array},$$

where the left vertical arrow is given by the surjective map sending $[z]$ onto $(\infty, 0, 1, z)$; on $\mathscr{P}(\partial H^3)$, mapped canonically onto $\mathscr{P}(\overline{H^3})$, the extended Dehn invariant $\Psi_e$ is given by

$$\Psi_e(S_\infty) = 2\sum_{k=1}^{3} \log(2\sin\alpha_k) \otimes \alpha_k$$

for a totally asymptotic simplex $S_\infty$ with angles $\alpha_1, \alpha_2, \alpha_3$ along edges intersecting in a vertex (cf. [**9**, p. 168–169]). The function $-\log^-$ sends $r \wedge e^{2\pi i\alpha}$ onto $-\log|r| \otimes \alpha$. Now, the function $\lambda[z]$ satisfies the 5 term equation (cf. [**9**, (4.15), p. 173])

(14.85) $$\lambda[z_1] - \lambda[z_2] + \lambda[z_1/z_2] - \lambda[(1-z_1)/(1-z_2)] + \lambda[(1-z_2)z_1/(1-z_1)z_2] = 0.$$

Consider, the induced additive homomorphism

$$\varrho: \mathscr{P}_{\mathbf{C}} \to \wedge^2_{\mathbf{Z}}(\mathbf{C}),$$

and denote by $e$ the exponential map defined by $e(z \wedge w) = \exp(2\pi i z) \wedge \exp(2\pi i w)$ for $z, w \in \mathbf{C}$. Then, $e \circ \varrho = \lambda$, and, as to (14.85), one could expect to $\varrho(z)$ has an analogous property. In fact, $\varrho$ even satisfies (cf. [**9**, (4.18), p. 174])

(14.86) $$\operatorname{Im}\varrho[z] = \frac{D(z)}{2\pi^2}!$$

## References


1. C. C. Adams, *The noncompact hyperbolic 3-manifold of minimal volume*, Proc. Amer. Math. Soc. **100** (1987), 601–606.
2. ——, *Volumes of N-cusped hyperbolic 3-manifolds*, J. London Math. Soc. (2) **38** (1988), 507–565.

3. E. M. Andreev, *On convex polyhedra in Lobacevskii spaces*, Math. USSR Sb. **10** No. 3 (1970), 413–440.
4. J. Böhm and E. Hertel, *Polyedergeometrie in n-dimensionalen Räumen konstanter Krümmung*, Birkhäuser, Basel, 1981.
5. A. Borel, *Commensurability classes and volumes of hyperbolic 3-manifolds*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **VIII** (1981), 1–33.
6. T. Chinburg and E. Friedman, *The smallest arithmetic hyperbolic three-orbifold*, Invent. Math. **86** (1986), 507–527.
7. T. Chinburg, *A small arithmetic hyperbolic three-manifold*, Proc. Amer. Soc. **100** (1987), 140–144.
8. H. S. M. Coxeter, *The functions of Lobatschefsky and Schläfli*, Quart. J. Math. Oxford **6** (1935), 13–29.
9. J. L. Dupont and C. H. Sah, *Scissors congruences*, II, J. Pure Appl. Algebra **25** (1982), 159–195.
10. O. Giering, *Vorlesungen über höhere Geometrie*, Vieweg, Braunschweig, Wiesbaden, 1982.
11. U. Haagerup and H. J. Munkholm, *Simplices of maximal volume in hyperbolic n-space*, Acta Math. **147** (1981), 1–12.
12. H. Hopf, *Die Curvatura integra Clifford-Kleinscher Raumformen*, Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. (1925), 131–141.
13. H.-C. Im Hof, *Napier cycles and hyperbolic Coxeter groups*, Bull. Soc. Math. Belg. (to appear).
14. R. Kellerhals, *On the volume of hyperbolic polyhedra*, Math. Ann. **285** (1989), 541–569.
15. ——, *On Schläfli's reduction formula*, Math. Z. **206** (1991), 193–210.
16. H. Kneser, *Der Simplexinhalt in der nichteuklidischen Geometrie*, Deutsche Math. **1** (1936), 337–340.
17. S. Kojima and Y. Miyamoto, *The smallest hyperbolic 3-manifolds with totally geodesic boundary*, Preprint Tokyo Inst. Tech., 1990.
18. L. Lewin, *Polylogarithms and associated functions*, North Holland, New York and Oxford, 1981.
19. W. Maier, *Inhaltsmessung im $R_3$ fester Krümmung*, Arch. Math. **5** (1954), 266–273.
20. R. Meyerhoff, *Sphere-packing and volume in hyperbolic 3-space*, Comment. Math. Helv. **61** (1986), 271–278.
21. ——, *A lower bound for the volume of hyperbolic 3-orbifolds*, Duke Math. J. **57** (1988), 185–203.
22. J. Milnor, *Hyperbolic geometry: the first 150 years*, Bull. Amer. Math. Soc. **6** (1982), 9–24.
23. H. Minc, *Permanents*, Encyclopedia of Mathematics and its Applications, Addison-Wesley, Reading, Massachusetts, 1978.
24. W. D. Neumann and D. B. Zagier, *Volumes of hyperbolic three-manifolds*, Topology **24** (1985), 307–331.
25. E. Roeser, *Die nicht-euklidischen Geometrien und ihre Beziehungen untereinander*, Oldenburg, München, 1957.
26. C. H. Sah, *Scissors congruences*. I, *Gauss-Bonnet map*, Math. Scand. **49** (1981), 181–210.
27. L. Schläfli, *Theorie der vielfachen Kontinuität*, Gesammelte Mathematische Abhandlungen vol. 1, Birkhäuser, Basel, 1950.
28. W. P. Thurston, *Three dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. **6** (1982), 357–381.
29. È. B. Vinberg, *Hyperbolic reflection groups*, Russian Math. Surveys **40** (1985), 31–75.
30. ——, *The absence of crystallographic groups of reflections in Lobachevsky spaces of large dimension*, Trans. Moscow Math. Soc. **47** (1985), 75–112.
31. H. C. Wang, *Topics in totally discontinuous groups*, Symmetric Spaces, (Boothby-Weiss ed.), New York, 1972.
32. D. B. Zagier, *Hyperbolic manifolds and special values of Dedekind zeta-functions*, Invent. Math. **83** (1986), 285–301.
33. ——, *Polylogarithms, Dedekind zeta functions, and the algebraic K-theory of fields*, Arithmetic algebraic geometry, (G. van der Geer, et al., ed.), Birkhäuser, Boston, Basel, Berlin, 1991.

CHAPTER 15

# Introduction to Higher Logarithms

RICHARD M. HAIN AND ROBERT MACPHERSON

Few mathematicians would disagree with the assertion that the logarithm is one of the most important functions in mathematics. During the nineteenth century an analogous function, the *dilogarithm*, was the subject of much research. First defined by Leibnitz in 1696, the dilogarithm was subsequently studied by Euler, Spence, Abel, Hill, Jonquière Kummer, Lindelöf, Lobachevsky, and many others [L]. Recently there has been a resurgence of interest in this remarkable function, due in large part to the work of Bloch [**B1, B2, B3**] in number theory and $K$-theory; Gabrielov, Gelfand and Losik [**GGL**] on the combinatorial formula for the first Pontrjagin class; Wigner (see [**B2**] and [**Dp1**]) on group cohomology; and the work of Lewin [L]. Other recent work includes [**A1, A2, A3, Be1, BGSV, D2, Dp1, Dp2, DS, G, GM, Li, Lo, M1, M2, R1, R2, R3, Y, Z1, Z2**].

The dilogarithm has properties analogous to those of the logarithm. It has been widely believed, both in the nineteenth century and more recently, that these two functions should be the first two elements of an infinite sequence of higher logarithms which share analogous properties. Several sequences of such functions have been proposed, but until recently, no function beyond the dilogarithm in any of these sequences was known to possess all of the desired properties.

In [**HM**] we proposed a new approach to constructing higher logarithms $\{L_p\}$ which we hope will produce the true generalizations of the logarithm and the dilogarithm. The difficulty in this approach lies in constructing the functions $L_p$; once existence is established, the function will automatically possess the desired properties. This is to be contrasted with the classical approach where the difficulty lies in establishing that given functions possess the sought after properties. As evidence for our program, we have constructed the first four functions.

**15.1. The problem of generalizing the logarithm and the dilogarithm.** The logarithm $\log x$, may be defined as the analytic continuation of the power series

$$(15.1) \qquad -\log(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n}, \qquad |x| < 1$$

to $\mathbb{C}^*$. It possesses three fundamental properties: one analytic, one topological, and one algebraic in nature.

ANALYTIC PROPERTY. The logarithm may be written as a line integral

$$\log x = \int_1^x \frac{dz}{z},$$

of a logarithmic 1-form on $\mathbb{C}^*$.

TOPOLOGICAL PROPERTY. The logarithm is a multivalued function on $\mathbb{C}^*$. Let $\sigma_0$ be homotopy class of loops based at (say) $1/2$ in $\mathbb{C}$ with winding number 1 about 0 (Figure 1). Let $M(\sigma_0)$ be the monodromy operator whose value on a function is its analytic continuation along $\sigma_0$. Then

$$M(\sigma_0)\log x = \log x + 2\pi i.$$

In other words, $M(\sigma_0)$ acts on the two dimensional vector space of germs of functions at $z = 1/2$ with basis $\log x$, 1 through the matrix

$$M(\sigma_0) = \begin{pmatrix} 1 & 2\pi i \\ 0 & 1 \end{pmatrix}.$$

The monodromy group is the discrete, 1-step unipotent group

$$\begin{pmatrix} 1 & \mathbb{Z}(1) \\ 0 & 1 \end{pmatrix},$$

where $\mathbb{Z}(p)$ denotes the subgroup $(2\pi i)^p\mathbb{Z}$ of $\mathbb{C}$.

ALGEBRAIC PROPERTY. The logarithm satisfies the three-term functional equation

$$\log x - \log xy + \log y = 0,$$

provided that appropriate branches have been chosen. The logarithm owes much of its utility to this functional equation which can be thought of as a 2-cocycle condition. When suitably interpreted, the logarithm represents the universal first Chern class.

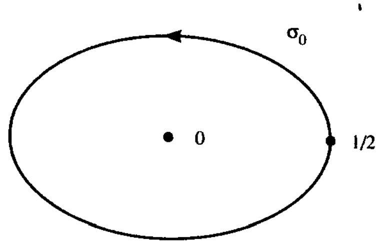



FIGURE 1



The classical dilogarithm function $\ln_2(x)$, may be defined as the analytic continuation of the power series

$$\ln_2(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2}. \tag{15.2}$$

It has three properties analogous to those of the logarithm.

ANALYTIC PROPERTY. The dilogarithm may be written as an iterated line integral of logarithmic 1-forms of length two

$$\ln_2(x) = \int_0^x \frac{dz}{1-z}\frac{dz}{z} := \int_0^x \int_0^z \frac{dw}{1-w}\frac{dz}{z},$$

of the type studied systematically by K.-T. Chen (see [C], for example). This expression shows that the dilogarithm can be analytically continued to any point of $\mathbb{C} - \{0, 1\}$.

TOPOLOGICAL PROPERTY. The dilogarithm is a multivalued function on $\mathbb{C} - \{0, 1\}$. Let $\sigma_0$ be the homotopy class of loops in $\mathbb{C} - \{0, 1\}$ based at $1/2$ that encircle 0, and $\sigma_1$ the homotopy class of loops based at $1/2$ that encircle 1 (Figure 2). Denote the corresponding monodromy operators by $M(\sigma_0)$ and $M(\sigma_1)$. Then

$$M(\sigma_0)\ln_2(x) = \ln_2(x), \qquad M(\sigma_1)\ln_2(x) = \ln_2(x) - 2\pi i \log x.$$

In other words, $M(\sigma_0)$ and $M(\sigma_1)$ act on the three dimensional vector space of germs of functions at $z = 1/2$ with basis $\ln_2(x)$, $\log x$, and 1 via the matrices

$$M(\sigma_0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2\pi i \\ 0 & 0 & 1 \end{pmatrix}, \qquad M(\sigma_1) = \begin{pmatrix} 1 & -2\pi i & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The monodromy group associated to $\ln_2(x)$ is the discrete, 2-step unipotent group

$$\begin{pmatrix} 1 & \mathbb{Z}(1) & \mathbb{Z}(2) \\ 0 & 1 & \mathbb{Z}(1) \\ 0 & 0 & 1 \end{pmatrix}.$$

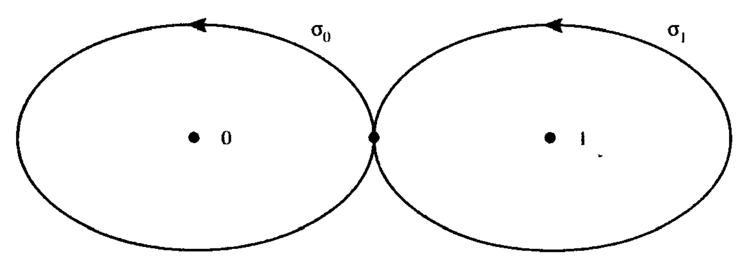



FIGURE 2

ALGEBRAIC PROPERTY. If we define $\phi(x)$, a form of the dilogarithm due to Rogers [Ro], by

$$\phi(x) = \frac{1}{2}(\ln_2(x) - \ln_2(1-x)),$$

then $\phi(x)$ satisfies the 5-term functional equation

$$\phi(x) - \phi(y) + \phi(y/x) - \phi\left(\frac{y-1}{x-1}\right) + \phi\left(\frac{x(y-1)}{y(x-1)}\right) = \frac{\pi^2}{6},$$

provided that we choose the branches of each of the five-terms carefully. The five functions $x$, $y$, $y/x$, $(y-1)/(x-1)$ and $x(y-1)/y(1-x)$ arise naturally as the cross ratios of the four element subsets of the configuration $(y, x, 1, 0, \infty)$ of five points on the projective line. A functional equation, equivalent to the one above, was discovered by Spence in 1809, rediscovered by Abel in 1828, and then again by many others. (The form above is due to Rogers [Ro].) When suitably interpreted, this five-term equation is a 4-cocycle condition, and the cocycle associated to $\phi(x)$ represents the second Chern class in certain cases [GGL, B2, GM, Dp1, DS, Be1]).

**15.2. The quest for higher logarithms.** No red blooded mathematician could compare the properties of the logarithm and dilogarithm above without wondering if they were the first two terms of an infinite sequence of higher logarithms possessing the:

> *Analytic property*: the $p$th logarithm should be defined by integrating a closed iterated integral of logarithmic 1-forms of length $\leq p$.
>
> *Topological property*: the $p$th logarithm should be a multivalued function whose associated monodromy group is discrete, and unipotent of length exactly $p$. (The last condition will imply, in particular, that the $p$th logarithm cannot be expressed as a polynomial of functions obtained by integrating iterated integrals of length $\leq (p-1)$.)
>
> *Algebraic property*: the $p$th logarithm should satisfy a natural functional equation. Since it is expected that the $p$th logarithm will be naturally associated with the $p$th Chern class, this equation should be a $2p$-cocycle condition, and therefore be of the form
>
> $$\sum_{j=0}^{2p}(-1)^j \mathscr{L}_p(A_j(x)) = 0,$$
>
> where the $A_j(x)$ are algebraic functions from a variety into the domain of $\mathscr{L}_p$.

The classical higher logarithms, or *polylogarithms* as we shall call them, are the naive generalizations of the logarithm and dilogarithm obtained by



extrapolating from (15.1) and (15.2):

$$\ln_p(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^p}, \qquad |x| < 1.$$

These first appeared in the literature in the late eighteenth century (cf. [L]). The integral formula

$$\ln_p(x) = \int_0^x \ln_{p-1}(z)\frac{dz}{z} = \int_0^x \frac{dz}{1-z}\overbrace{\frac{dz}{z}\cdots\frac{dz}{z}}^{p-1}$$

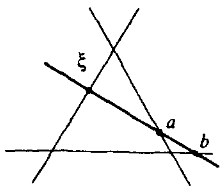
shows that $\ln_p(x)$ can be analytically continued to a multivalued function on $\mathbb{C} - \{0, 1\}$ and that it is an iterated integral of length $p$. Its monodromy group is a discrete, unipotent group of length exactly $p$ [R2]. The polylogarithms therefore possess the desired analytic and topological properties. As for the algebraic property, the lower polylogarithms satisfy many functional equations, but the number of terms in the functional equation of each considered most natural by Lewin is

| $p$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| No. of terms | 3 | 5 | 9 | 20 | 33 |

[L; p. 239]. No pattern in these equations is discernable and no generalization has been found for the higher polylogarithms.

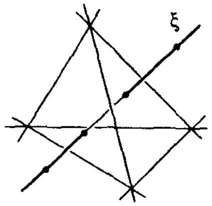
Another definition of real higher logarithms was proposed in [GM] for even $p$. They are real-valued functions defined on a real algebraic variety. These satisfy the algebraic property; however, it is not clear that these extend to complex valued functions defined on the complex points of their domains, nor that these functions would possess the analytic and topological properties. We hope that these functions will turn out to be the analogues of the Bloch-Wigner-Ramakrishnan functions ([R3, Z2]) of the higher logarithms we have proposed.

**15.3. Higher logarithms.** The higher logarithms that we propose are functions, not of a single complex variable, but of a point in a complex algebraic manifold $G^p_{p-1}$ of dimension $p^2$, which is an open subset of the self dual Grassmann manifold of $(p-1)$-dimensional linear subspaces of $\mathbf{P}^{2p-1}$. More precisely, $G^p_q$ is the open subset of the Grassmann manifold $G(q, \mathbf{P}^{p+q})$ of $q$ dimensional linear subspaces $\xi$ of $\mathbf{P}^{p+q}$ which are transverse to the configuration of coordinate hyperplanes. The following figures depict elements of the real $G^1_1$, $G^2_1$ and $G^2_0$ (Figures 3–5). In the first illustration, the line $\xi$ is required to avoid the vertices of the coordinate simplices, so $G^1_1 = \mathbb{C}^* \times \mathbb{C}^*$ with coordinates $(a, b)$. In the second, $\xi$ is required to avoid the edges of the coordinate simplices. In the third, $\xi$ is required to avoid the coordinate axes, so that $G^2_0 = \mathbb{C}^* \times \mathbb{C}^*$.



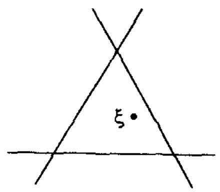

$G_1^1$

FIGURE 3




$G_1^2$

FIGURE 4




$G_0^2$

FIGURE 5

Since the transversality requirement defining $G_q^p$ is a generic condition, we call $G_q^p$ the *generic part* of the grassmannian. The manifold $G_{p-1}^p$ on which we expect the $p$-logarithm to be defined is "self dual".

The idea that higher logarithms should be functions on the $G_q^p$ is due to Gelfand and MacPherson [GM]; their higher logarithms are defined on the real points of $G_q^{2p}$. Damiano [Da] showed that the only possibly nonzero Gelfand-MacPherson higher logarithms occur when $q = 2p - 1$.

When $q > 0$, there are $p+q+1$ *face maps* $A_i : G_q^p \to G_{q-1}^p$; $A_i$ takes the element $\xi \subseteq \mathbb{P}^{p+q}$ of $G_q^p$ to its intersection with the $i$th coordinate hyperplane $\approx \mathbb{P}^{p+q-1}$ of $\mathbb{P}^{p+q}$. The face maps $A_i : G_1^1 \to G_0^1$ and $A_i : G_1^2 \to G_0^2$ are illustrated in Figures 6 and 7. In our approach, the existence of a $p$-logarithm function $L_p$ will guarantee that it satisfies the functional equation

$$\sum_{j=0}^{2p} (-1)^j A_j^* L_p = 0, \tag{15.3}$$

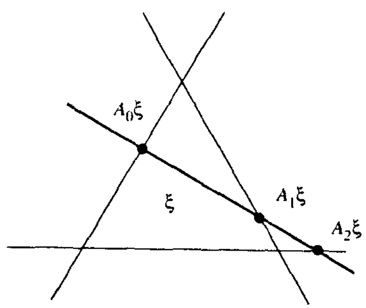



FIGURE 6

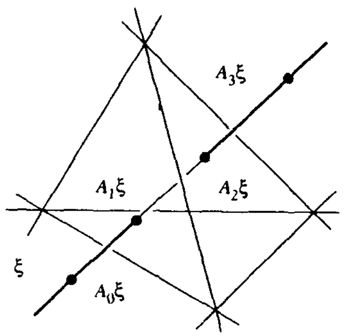



FIGURE 7



where the $A_j$ are the $2p+1$ face maps $G_p^p \to G_{p-1}^p$. Because the function $L_p$ will be multivalued, one has to be very careful with branches. (Such issues are dealt with in §2 of [HM].)

**15.4. The higher logarithm bicomplex.** The $p$-logarithm function is a component of a cochain in a certain double complex which we now describe briefly.

For fixed $p$, the face maps $A_i : G_q^p \to G_{q-1}^p$ satisfy the usual identities dual to those that hold between the faces of a simplex. This means that $\{G_q^p\}_{q=0,\dots,p}$ is a truncated simplicial variety $G_\bullet^p$. It is natural to put $G_q^p$ in dimension $p+q$ as there are $p+q+1$ face maps emanating from it. If we apply a contravariant, abelian group-valued functor to $G_\bullet^p$, we will obtain a cochain complex with differential

$$A^* = \sum_{i=0}^{p} (-1)^i A_i^* .$$

More generally, if we apply a contravariant, cochain complex valued functor, we will obtain a double complex. We will apply the multivalued de Rham complex functor $\tilde{\Omega}^\bullet$.

Briefly, the complex of multivalued differential forms on a complex algebraic manifold $X$ is

$$\tilde{\Omega}^\bullet(X) = \tilde{\mathscr{O}}(X) \otimes \Omega^\bullet(X),$$

where $\Omega^\bullet(X)$ denotes the holomorphic forms on $X$ with logarithmic singularities at infinity, and $\tilde{\mathscr{O}}(X)$ consists of all multivalued functions on $X$ obtained by integrating a relatively closed iterated integral[1] of elements of $\Omega^1(X)$. Both $\tilde{\mathscr{O}}(X)$ and $\tilde{\Omega}^\bullet(X)$ come equipped with a canonical filtration, called the *weight filtration*[2]; if $X = G_q^p$, then $W_{2l}\tilde{\mathscr{O}}(X)$ consists of those functions obtained by integrating iterated integrals of length not exceeding $l$. Thus, for example,

$$\log x = \int_1^x \frac{dz}{z} \in W_2\tilde{\mathscr{O}}(\mathbb{C}^*),$$

$$\ln_p(x) = \int_0^x \frac{dz}{1-z} \overbrace{\frac{dz}{z} \cdots \frac{dz}{z}}^{p-1} \in W_{2p}\tilde{\mathscr{O}}(\mathbb{C} - \{0, 1\}).$$

Combining these filtrations defines a weight filtration on $\tilde{\Omega}^\bullet(X)$.

Neglecting the problem of choosing branches, (dealt with in [HM, §5]) we obtain a double complex $(W_{2p}\tilde{\Omega}^\bullet(G_\bullet^p), D)$ by applying $W_{2p}\tilde{\Omega}^\bullet$ to the simplicial space $G_\bullet^p$. The differential $D$ is the total differential $d \pm A^*$. For

[1] An iterated line integral $I$ is *relatively closed* if its value on a path $\gamma$ depends on the homotopy class of $\gamma$ relative to its endpoints.

[2] This is the weight filtration in the sense of Deligne [D1].

example, the double complex for $p = 3$ is

$$
(*)\quad
\begin{array}{ccccccc}
W_6\tilde{\mathcal{O}}(G_3^3) & \longrightarrow & W_6\tilde{\Omega}^1(G_3^3) & \longrightarrow & W_6\tilde{\Omega}^2(G_3^3) & \longrightarrow & \Omega^3(G_3^3) \\
\uparrow & & \uparrow & & \uparrow & & \uparrow \\
W_6\tilde{\mathcal{O}}(G_2^3) & \longrightarrow & W_6\tilde{\Omega}^1(G_2^3) & \longrightarrow & W_6\tilde{\Omega}^2(G_2^3) & \longrightarrow & \Omega^3(G_2^3) \\
\uparrow & & \uparrow & & \uparrow & & \uparrow \\
W_6\tilde{\mathcal{O}}(G_1^3) & \longrightarrow & W_6\tilde{\Omega}^1(G_1^3) & \longrightarrow & W_6\tilde{\Omega}^2(G_1^3) & \longrightarrow & \Omega^3(G_1^3) \\
\uparrow & & \uparrow & & \uparrow & & \uparrow \\
W_6\tilde{\mathcal{O}}(G_0^3) & \longrightarrow & W_6\tilde{\Omega}^1(G_0^3) & \longrightarrow & W_6\tilde{\Omega}^2(G_0^3) & \longrightarrow & \Omega^3(G_0^3).
\end{array}
$$

The manifold $G_0^p$ is just $\mathbb{P}^p$ minus the union of the coordinate hyperplanes, and is isomorphic to $(\mathbb{C}^*)^p$. On this there is a canonical $p$-form, the "*volume*" *form*:

$$\mathrm{vol}_p = \frac{dx_1}{x_1} \wedge \cdots \wedge \frac{dx_p}{x_p} \in \Omega^p(G_0^p).$$

**Definition.** A *p-logarithm* is a $2p-1$ cochain $Z_p$ in the Grassmann bicomplex $W_{2p}\tilde{\Omega}^\bullet(G_\bullet^p)$ that satisfies the equation $DZ_p = \mathrm{vol}_p$.

Note that it is not at all clear that such a cochain $Z_p$ exists. The first obstruction is that $\mathrm{vol}_p$ be closed in the double complex as

$$A^*\mathrm{vol}_p = D\,\mathrm{vol}_p = D^2(Z_p) = 0.$$

This can be verified by direct computation for small values of $p$.

**Theorem A** [HM, §9]. *For all $p$, $A^*\mathrm{vol}_p = 0$.*

This we prove by taking residues and exploiting the action of the symmetric group on $G_1^p$. P. Cartier has given a very elegant proof of the vanishing of $A^*\mathrm{vol}_p$ using Cartan's theory of basic forms.

Note that the component $L_p$ of a $p$-logarithm $Z_p$ in $W_{2p}\tilde{\mathcal{O}}(G_{p-1}^p)$ satisfies

(i) $A^*L_p = \sum_{i=0}^{2p}(-1)^i L_p = 0$.
(ii) $L_p$ is obtained by integrating a relatively closed iterated integral of logarithmic 1-forms of length not exceeding $p$.

We shall call $L_p$ a *p-logarithm function.* It clearly possesses the desired analytic and algebraic properties. As for the topological property, the analytical property implies that $L_p$ has unipotent monodromy group of length $\leq p$. To say that the group has length exactly $p$ is equivalent to showing that $L_p$ is *indecomposable*; that is, $L_p$ cannot be expressed as a polynomial of elements of $W_{2p-2}\tilde{\mathcal{O}}(G_{p-1}^p)$.



Neither the cochain $Z_p$ nor the function $L_p$ is unique. The cochain can be adjusted by coboundaries and the function by functions of the form $A^*G$, where $G \in W_{2p}\tilde{\mathcal{O}}(G_{p-2}^p)$, which satisfy the functional equation trivially.

**Theorem B** [HM, §9]. *For $p = 1, 2, 3$, there is a p-logarithm which is unique modulo coboundaries.*[3] *In each case, the associated p-logarithm function is indecomposable and nontrivial in*

$$W_{2p}\tilde{\mathcal{O}}(G_{p-1}^p)/A^*W_{2p}\tilde{\mathcal{O}}(G_{p-2}^p).$$

When $p = 1$, the cochain $\log x \in W_2\tilde{\mathcal{O}}(\mathbb{C}^*)$ represents $Z_1$ and the functional equation $A^*\log x = 0$ is the usual one.

At first glance it seems that the 2-logarithm function cannot be the classical dilogarithm or Rogers' function $\phi(x)$ as $L_2$ is defined on $G_1^2$ while $\phi(x)$ is defined on $\mathbb{C} - \{0, 1\}$. However, an element of $G_1^2$ is a line in $\mathbb{P}^3$ that intersects the four coordinate hyperplanes in four distinct points. Taking $\xi$ to be the cross ratio of these four points defines a function $\pi\colon G_1^2 \to \mathbb{C} - \{0, 1\}$. There is a representative of $Z_2$ where

$$L_2 = \pi^*\phi - \pi^2/6.$$

Similarly, there is a projection of $G_2^2$ onto the domain of the functional equation of $\phi(x)$, and the functional equation $A^*L_2 = 0$ is just the pullback of Rogers' functional equation.

Both the logarithm and the dilogarithm have single, real-valued cousins which also satisfy natural functional equations. In the case of the logarithm, this function is the logarithm of the absolute value $D_1\colon \mathbb{C}^* \to \mathbb{R}$ defined by $z \mapsto \log|z|$. It satisfies the functional equation

$$D_1(x) - D_1(xy) + D_1(y) = 0,$$

and possesses the *symmetry property*

$$\sigma^*D_1 = -D_1,$$

for all $\sigma$ in the symmetric group on 2 letters, $\Sigma_2$, which acts on $\mathbb{C}^*$ by letting the generator take $z$ to $z^{-1}$. The functional equation implies that $D_1$ represents a cohomology class in $H^1(\mathrm{GL}(\mathbb{C}), \mathbb{R})$. This class is the first Cheeger-Simons Chern class $\hat{c}_1$ of the universal flat bundle over $\mathrm{BGL}(\mathbb{C})^\delta$, the classifying space of stable, flat complex vector bundles. It also defines the first regulator mapping $r_1\colon K_1(\mathbb{C}) = \mathbb{C}^* \to \mathbb{R}$.

The single-valued cousin of the dilogarithm is the *Bloch-Wigner function*

$$D_2\colon \mathbb{C} - \{0, 1\} \to \mathbb{R},$$

**[B1]** (see also **[Z1]**) which is defined by

$$D_2(z) = \operatorname{Im}\ln_2(z) + \arg(1-z)\log|z|.$$

[3] We have also constructed a 4-logarithm and proved its uniqueness, but have not verified the indecomposability and nontriviality.

It satisfies the 5-term functional equation

$$D_2(x) - D_2(y) + D_2\left(\frac{y}{x}\right) - D_2\left(\frac{y-1}{x-1}\right) + D_2\left(\frac{x(y-1)}{y(x-1)}\right) = 0.$$

Viewing $\mathbb{C} - \{0, 1\}$ as the space of ordered 4-tuples of distinct points on the projective line modulo projective equivalence, we see that $\Sigma_4$, the symmetric group on 4 letters, acts naturally on $\mathbb{C} - \{0, 1\}$. The Bloch-Wigner function satisfies the symmetry condition

$$\sigma^* D_2 = \operatorname{sgn}(\sigma) D_2,$$

for all $\sigma \in \Sigma_4$. Just as $D_1$ represents $\hat{c}_1$, $D_2$ represents the second Cheeger-Simons Chern class,

$$\hat{c}_2 \in H^3(\mathrm{GL}(\mathbb{C}), \mathbb{R}),$$

of the universal flat bundle over $\mathrm{BGL}(\mathbb{C})^\delta$ **[Dp1]**, and, by composition with the Hurewicz homomorphism $K_3(\mathbb{C}) \to H_3(\mathrm{GL}(\mathbb{C}))$, it defines the regulator mapping $r_2 \colon K_3(\mathbb{C}) \to \mathbb{R}$ **[B2]**.

One can also construct the single-valued cousin of the 3-logarithm function **[HM**, §11**]**. It is a single-valued function $D_3 \colon Y_2^3 \to \mathbb{R}$, where

$$Y_q^3 = \left\{ \begin{matrix} \text{ordered } (4+q)\text{-tuples of} \\ \text{points in } \mathbb{P}^2, \text{ no 3 on a line} \end{matrix} \right\} \Big/ \text{projective equivalence}.$$

The seven ways of forgetting a point give 7 maps $A_j \colon Y_3^3 \to Y_2^3$. In **[HM**, §11**]**, we show that $D_3$ satisfies the functional equation

$$\sum_{j=0}^{6} (-1)^j A_j^* D_3 = 0,$$

and that it possesses the symmetry property

$$\sigma^* D_3 = \operatorname{sgn}(\sigma) D_3,$$

for all $\sigma \in \Sigma_6$, where $\Sigma_6$ acts on $Y_2^3$ by permuting the points.

Recently, Goncharov **[G]** has given an explicit formula for $D_3$ in terms of Zagier's single-valued cousin of the classical trilogarithm. He and Yang **[Y]** have independently shown that $D_3$ defines a class in $H^5_{\mathrm{cts}}(\mathrm{GL}_3(\mathbb{C}), \mathbb{R})$ which is a nonzero rational multiple of Borel's class, and that the third Borel regulator can be expressed in terms of $D_3$. It follows from **[DHZ]** that $D_3$ also represents the third Cheeger-Simons Chern class $\hat{c}_3 \in H^5(\mathrm{GL}(\mathbb{C}), \mathbb{R})$.

**15.5. Multivalued Deligne cohomology.** The definition of $p$-logarithms sketched above is only part of a more complicated definition of $p$-logarithms as Deligne cohomology classes of the simplicial variety $G^p_\bullet$. We now describe this in more detail. An exposition of the role of these higher logarithms in algebraic $K$-theory is given in **[BMS]**.



Suppose that $X$ is a complex algebraic manifold. The most refined version of the Deligne cohomology $H^\bullet_{\mathscr{D}}(X, \mathbb{Q}(p))$ of $X$ is Beilinson's *absolute Hodge cohomology* **[Be2]** which is defined to be the cohomology of the complex

$$D(X, \mathbb{Q}(p)) = \operatorname{cone}[W_{2p} A^\bullet_{\mathbb{Q}}(X) \to W_{2p} A^\bullet_{\mathbb{C}}(X) / F^p W_{2p} A^\bullet_{\mathbb{C}}(X)][-1],$$

where $(A^\bullet_{\mathbb{Q}}(X), W_\bullet) \to (A^\bullet_{\mathbb{C}}(X), F^\bullet, W_\bullet)$ is a natural mixed Hodge complex[4] whose cohomology is Deligne's natural mixed Hodge structure on $H^\bullet(X)$ **[D1]**.

We now introduce the multivalued Deligne complex of $X$: Let $\mathfrak{g}(X)$ be the Malcev Lie algebra associated to $\pi_1(X)$, topologized by its weight filtration (or equivalently, by its lower central series). The complex of multivalued forms $\widetilde{\Omega}^\bullet(X)$ on $X$ is a continuous $\mathfrak{g}(X)$ module. Let $\mathscr{C}(\mathfrak{g}(X), \Omega^\bullet(X))$ be the complex of continuous $\widetilde{\Omega}^\bullet(X)$ valued cochains on $\mathfrak{g}(X)$. That is,

$$\mathscr{C}(\mathfrak{g}, \widetilde{\Omega}^\bullet) = \varinjlim \mathscr{C}(\mathfrak{g}/W_{-l-1}, W_l \widetilde{\Omega}^\bullet),$$

where, for a module $V$ over the Lie algebra $\mathfrak{h}$, $\mathfrak{g}(\mathfrak{h}, V)$ denotes the Chevalley-Eilenberg complex $\Lambda(\mathfrak{h}^*) \otimes V$ with the usual differential.

Define a Hodge filtration on $\mathscr{C}(\mathfrak{g}, \widetilde{\Omega}^\bullet)$ by

$$F^p \mathscr{C}(\mathfrak{g}, \widetilde{\Omega}^\bullet) = \bigoplus_{q \geq p} \mathscr{C}(\mathfrak{g}, \widetilde{\Omega}^q).$$

The weight filtrations on $\mathfrak{g}$ and $\widetilde{\Omega}^\bullet$ induce a weight filtration on $\mathscr{C}(\mathfrak{g}, \widetilde{\Omega}^\bullet)$.

Let $\widetilde{\mathscr{A}}$ denote the category whose objects are universal coverings of complex algebraic manifolds, and whose morphisms are lifts of morphisms between varieties to universal coverings. The assignment of $(\mathscr{C}(\mathfrak{g}(X), \widetilde{\Omega}^\bullet(X)), W_\bullet, F^\bullet)$ to $X$ defines a functor from $\widetilde{\mathscr{A}}$ into the category of bifiltered differential graded algebras.

Let $X$ be an object of $\widetilde{\mathscr{A}}$. Denote the $\mathbb{Q}$-form of the Malcev Lie algebra associated to $\pi_1(X)$ by $\mathfrak{g}_{\mathbb{Q}}(X)$ and the continuous $\mathbb{Q}$ valued cochains on it by $\mathscr{C}(\mathfrak{g}_{\mathbb{Q}}(X), \mathbb{Q})$.

Consider the complex

$$M(X, \mathbb{C}/\mathbb{Q}(p)) := \operatorname{cone}[W_{2p}\mathscr{C}(\mathfrak{g}_{\mathbb{Q}}(X), \mathbb{Q}) \to W_{2p}\mathscr{C}(\mathfrak{g}(X), \widetilde{\Omega}^\bullet(X))][-1].$$

This may be viewed as a double complex ( $W_1 H^1(X) = 0$ case). (See Diagram A, page 348.)

The *multivalued Deligne complex of* $X$ is defined to be the quotient complex

$$MD(X, \mathbb{Q}(p)) = M(X, \mathbb{C}/\mathbb{Q}(p))/F^p.$$

This can be viewed as a double complex ( $W_1 H^1(X) = 0$ case). (See Diagram B, page 349.)

[4] Note that the weight filtration on $A^\bullet(X)$ is the *filtration décalée* of the commonly used weight filtration (see **[D1]**).

$$
\begin{tikzcd}
 & 0 & 0 & & & & 0 & & \\
p & \Lambda^p H^1(\mathfrak{g}_{\mathbb{Q}}) \arrow[u] \arrow[r] & \Lambda^p H^1(\mathfrak{g}) \arrow[u] & & & & & & \\
\vdots & \vdots \arrow[u] & \vdots \arrow[u] & & & & & & \\
r & W_{2p}\Lambda^r \mathfrak{g}^*_{\mathbb{Q}} \arrow[u] \arrow[r] & \cdots & {} \arrow[r] & W_{2p}\Lambda^r(\mathfrak{g}^*)\otimes\tilde{\Omega}^{s-1} \arrow[r] & {} & & & \\
\vdots & \vdots \arrow[u] & \vdots & & \vdots \arrow[u] & & 0 & & \\
1 & W_{2p}\mathfrak{g}^*_{\mathbb{Q}} \arrow[u] \arrow[r] & {} & & & \cdots \arrow[r] & W_{2p}\mathfrak{g}^*\otimes\tilde{\Omega}^{p-1} \arrow[u] \arrow[r] & 0 & \\
0 & \mathbb{Q} \arrow[u] \arrow[r, "(2\pi i)^p"] & W_{2p}\tilde{\Omega}^0 \arrow[u] \arrow[r] & \cdots \arrow[r] & W_{2p}\tilde{\Omega}^{s-1} \arrow[u] \arrow[r] & \cdots \arrow[r] & W_{2p}\tilde{\Omega}^{p-1} \arrow[u] \arrow[r] & \Omega^p \arrow[r] & 0. \\
 & 0 & 1 & & s & & p & p+1 &
\end{tikzcd}
$$

DIAGRAM A



$$
\begin{tikzcd}
 & 0 & 0 & & & & 0 & \\
p & \Lambda^p H^1(\mathfrak{g}_{\mathbb{Q}}) \arrow[u] \arrow[r] & \Lambda^p H^1(\mathfrak{g}) \arrow[u] & & & & & \\
\vdots & \vdots \arrow[u] & \vdots \arrow[u] & & & & & \\
r & W_{2p}\Lambda^r \mathfrak{g}^*_{\mathbb{Q}} \arrow[u] \arrow[r] & \cdots & {} \arrow[r] & W_{2p}\Lambda^r(g^*)\otimes\tilde{\Omega}^{s-1} \arrow[r] & {} & & \\
\vdots & \vdots \arrow[u] & \vdots & & \vdots \arrow[u] & & 0 & \\
1 & W_{2p}\mathfrak{g}^*_{\mathbb{Q}} \arrow[u] \arrow[r] & {} & & & \cdots \arrow[r] & W_{2p}\mathfrak{g}^*\otimes\tilde{\Omega}^{p-1} \arrow[u] \arrow[r] & 0 \\
0 & \mathbb{Q} \arrow[u] \arrow[r, "(2\pi i)^p"] & W_{2p}\tilde{\Omega}^0 \arrow[u] \arrow[r] & \cdots \arrow[r] & W_{2p}\tilde{\Omega}^{s-1} \arrow[u] \arrow[r] & \cdots \arrow[r] & W_{2p}\tilde{\Omega}^{p-1} \arrow[u] \arrow[r] & 0. \\
 & 0 & 1 & & s & & p &
\end{tikzcd}
$$

DIAGRAM B

Define the *multivalued Deligne cohomology* of the object $X$ of $\widetilde{\mathscr{A}}$ by

$$H^\bullet_{\mathscr{M}\mathscr{D}}(X, \mathbb{Q}(p)) = H^\bullet(MD(X, \mathbb{Q}(p))).$$

This defines a functor from $\widetilde{\mathscr{A}}$ into the category of graded rings. (Actually, one should sheafify this construction and define $H^\bullet_{\mathscr{M}\mathscr{D}}(X, \mathbb{Q}(p))$ to be the Čech hypercohomology of the sheaf $\mathscr{M}\mathscr{D}_X(\mathbb{Q}(p))$. This complication will not concern us here.)

There is a natural homomorphism $H^\bullet_{\mathscr{M}\mathscr{D}}(X, \mathbb{Q}(p)) \to H^\bullet_{\mathscr{D}}(X, \mathbb{Q}(p))$.

If $X_\bullet$ is a simplicial object of $\widetilde{\mathscr{A}}$, one can define a triple complex $MD(X_\bullet, \mathbb{Q}(p))$ in the obvious way. The multivalued Deligne cohomology of $X_\bullet$ is defined to be the cohomology of this complex. One has a canonical homomorphism

$$H^\bullet_{\mathscr{M}\mathscr{D}}(X_\bullet, \mathbb{Q}(p)) \to H^\bullet_{\mathscr{D}}(X_\bullet, \mathbb{Q}(p)).$$

THEOREM C. *When $p = 1, 2, 3, 4$ the canonical homomorphism*

$$H^k_{\mathscr{M}\mathscr{D}}(G^p_\bullet, \mathbb{Q}(p)) \to H^k_{\mathscr{D}}(G^p_\bullet, \mathbb{Q}(p))$$

*is an isomorphism when $k \le 2p$.*

**15.6. Higher logarithms as Deligne cohomology classes.** Applying the multivalued Deligne complex functor $MD(-, \mathbb{Q}(p))$ to $G^p_\bullet$, we obtain a triple complex in which the double complex $(*)$ is imbedded in the ground floor (Figure 8).

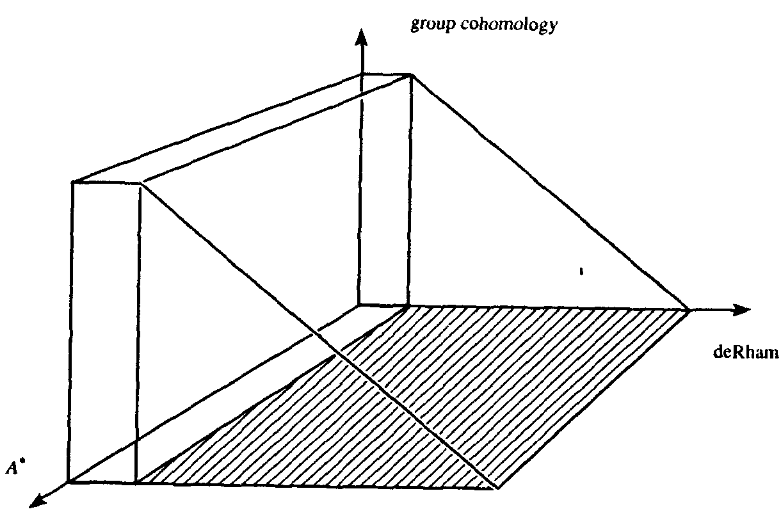



FIGURE 8



We also have the triple complex $M(G^p_\bullet, \mathbb{C}(\mathbb{Q}(p)))$. The short exact sequence of complexes

$$0 \to \Omega^p(G^p_\bullet)[-p-1] \to M(G^p_\bullet, \mathbb{C}/\mathbb{Q}(p)) \to MD(G^p_\bullet, \mathbb{Q}(p)) \to 0,$$

gives rise to a long exact sequence of cohomology groups. Consider the connecting homomorphism

$$\delta\colon H^{2p}_{\mathscr{M}\mathscr{D}}(G^p_\bullet, \mathbb{Q}(p)) \to \Omega^p(G^p_0).$$

DEFINITION. A *generalized $p$-logarithm* is an element $C_p$ of

$$H^{2p}_{\mathscr{M}\mathscr{D}}(G^p_\bullet, \mathbb{Q}(p))$$

satisfying $\delta(C_p) = \mathrm{vol}_p$.

Equivalently, $C_p$ is represented by a $2p-1$ cochain $Z_p$ in the triple complex $M(G^p_\bullet, \mathbb{C}/\mathbb{Q}(p))$ which satisfies

$$DZ_p = \mathrm{vol}_p. \tag{15.4}$$

If $Z_p$ represents a generalized $p$-logarithm, then the components of $Z_p$ which lie in the ground floor of the triple complex $M(G^p_\bullet, \mathbb{C}/\mathbb{Q}(p))$ is a $p$-logarithm as defined earlier. Observe also that the conditions imposed on the component

$$L_p \in W_{2p}\widetilde{\mathscr{O}}(G^p_{p-1})$$

of $Z_p$ by (15.4) correspond to the three essential properties of the $p$-logarithm function discussed earlier; the analytic property corresponds to the de Rham differential

$$d\colon W_{2p}\widetilde{\mathscr{O}}(G^p_{p-1}) \to \widetilde{\mathscr{O}}(G^p_{p-1}) \otimes \Omega^1(G^p_{p-1});$$

the topological property to the group cohomology differential

$$\partial\colon W_{2p}\widetilde{\mathscr{O}}(G^p_{p-1}) \to \mathscr{C}(\mathfrak{g}(G^p_{p-1}), \widetilde{\mathscr{O}});$$

and the algebraic property to the combinatorial differential

$$A^\bullet\colon W_{2p}\widetilde{\mathscr{O}}(G^p_{p-1}) \to \widetilde{\mathscr{O}}(G^p_p).$$

The next result is the generalization of Theorem B to the current setting.

THEOREM D. *When $p = 1, 2, 3$,*

$$\delta\colon H^{2p}_{\mathscr{M}\mathscr{D}}(G^p_\bullet, \mathbb{Q}(p)) \to \Omega^p(G^p_0)$$

*is an isomorphism. In particular, there is a canonical generalized $p$-logarithm.*

ACKNOWLEDGEMENTS. The research reported here was supported in part by grants from the National Science Foundation.

## REFERENCES

[A1] K. Aomoto, *A generalization of Poincaré normal functions on a polarized manifold*, Proc. Japan Acad. Ser. A Math. Sci. **55** (1979), 353–358.
[A2] ——, *Addition theorems of Abel type for hyperlogarithms*, Nagoya Math. J. **88** (1982), 55–71.
[A3] ——, *Special values of hyperlogarithms and linear difference schemes*, Illinois J. Math. **34** (1990), 191–216.
[Be1] A. Beilinson, *Higher regulators and values of L-functions*, J. Soviet Math. **30** (1985), 2036–2070. Translated from Sovr. Probl. Mat. **24**, Mosc. VINITI (1984), 181–238.
[Be2] ——, *Notes on absolute Hodge cohomology*, Applications of Algebraic $K$-Theory to Algebraic Geometry and Number Theory, Contemp. Math. vol. 55, Part I, Amer. Math. Soc., Providence, 1986, pp. 35–68.
[BGSV] A. Beilinson, A. B. Goncharov, V. V. Schechtman, and A. N. Varchenko, *Aomoto integrals, mixed Hodge structures and motivic cohomology of pairs of triangles on the plane*, Preprint, 1988.
[BMS] A. Beilinson, R. MacPherson, and V. Schechtman, *Notes on motivic cohomology*, Duke Math. J. **54** (1987), 679–710.
[B1] S. Bloch, *Applications of the dilogarithm function in algebraic K-theory and algebraic geometry*, Internat. Sympos. on Algebraic Geometry, Kyoto, 1977, pp. 103–114.
[B2] ——, *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*, unpublished manuscript, 1978.
[B3] ——, *The Dilogarithm and extensions of Lie algebras*, Algebraic $K$-Theory (M. Stein and E. Friedlander eds.), Lecture Notes in Math., vol. 854, Springer-Verlag, Berlin, Heidelberg, New York, 1984, pp. 1–23.
[C] K.-T. Chen, *Iterated path integrals*, Bull. Amer. Math. Soc. 83 (1977), 831–879.
[Da] D. Damiano, *Vanishing of the characteristic forms on the orbits of the Cartan subgroup*, Adv. in Math. **44** (1982), 313–317.
[D1] P. Deligne, *Theorie de Hodge* II, Inst. Hautes Études Sci. Publ Math. **40** (1971), 5–57.
[D2] ——, Letter to Spencer Bloch, April 3, 1984.
[Dp1] J. Dupont, *The dilogarithm as a characteristic class for flat bundles*, J. Pure Appl. Algebra **44** (1987), 137–164.
[Dp2] ——, *On polylogarithms*, Nagoya Math. J. **114** (1989), 1–20.
[DHZ] J. Dupont, R. Hain, and S. Zucker, *Regulators and characteristic classes of flat bundles*, Preprint, 1990.
[DS] J. Dupont and C. H. Sah, *Scissors congruences* II, J. Pure Appl. Algebra **25** (1982), 159–195.
[GGL] A. M. Gabrielov, I. M. Gelfand, and M. H. Losik, *Combinatorial calculation of characteristic classes*, Funktsional Anal. i Prilozhen. **9** (1975), no. 1, 54–55; no. 2, 12–28; no. 3, 5–26.
[GM] I. M. Gelfand and R. MacPherson, *Geometry in grassmannians and a generalization of the dilogarithm*, Adv. in Math. **44** (1982), 279–312.
[HM] R. Hain and R. MacPherson, *Higher logarithms*, Illinois J. Math. **34** (1990), 392–475.
[G] A. Goncharov, *The classical threelogarithm, algebraic K-theory of fields and Dedekind zeta functions*, Preprint, 1990.
[L] L. Lewin, *Polylogarithms and associated functions*, North-Holland, New York, 1981.
[Li] S. Lichtenbaum, *The construction of weight-two arithmetic cohomology*, Invent. Math. **88** (1987), 183–215.
[Lo] J.-L. Loday, *Comparaison des homologies du groupe linéaire et de son algèbre de Lie*, Ann. Inst. Fourier (Grenoble) **37** (1987), 167–190.
[M1] J. Milnor, *Hyperbolic geometry: the first* 150 *years*, Bull. Amer. Math. Soc. **6** (1982), 9–24.
[M2] ——, *On polylogarithms, Hurwitz zeta functions, and the Kubert identities*, Enseign. Math. **29** (1983), 281–322.
[R1] D. Ramakrishnan, *A regulator for curves via the Heisenberg group*, Bull. Amer. Math. Soc. **5** (1981), 191–195.



[R2] ——, *On the monodromy of higher logarithms*, Proc. Amer. Math. Soc. **85** (1982), 596–599.
[R3] ——, *Analogs of the Bloch-Wigner function for higher polylogarithms*, Applications of Algebraic $K$-Theory to Algebraic Geometry and Number Theory, Part I, Contemp. Math., vol. 55, Amer. Math. Soc., Providence, (1986), pp. 371–376.
[Ro] L. J. Rogers, *On function sum theorems connected with the series* $\sum_1^\infty x^n/n^2$, Proc. London Math. Soc. **4** (1907), 169–189.
[Y] J. Yang, *The Hain-MacPherson trilogarithm, the third Borel regulator and the values of Dedekind zeta functions at* 3, Preprint, 1990.
[Z1] D. Zagier, *The remarkable dilogarithm*, Max-Planck Institute preprint, 1988.
[Z2] ——, *The Bloch-Wigner-Ramakrishnan polylogarithm function*, Math. Ann. **286** (1990), 613–624.

CHAPTER 16

# Some Miscellaneous Results

L. LEWIN

**16.1. Clausen's function and the di-gamma function for rational arguments.**

16.1.1. de Doelder [1] has developed a relation between $\mathrm{Cl}_2(p\pi/q)$ and the second derivative of the logarithm of the gamma function, $\Psi'(z)$, at rational points. As is well known, $\Psi(p/q)$, for integral $p$ and $q$, can be expressed as a finite summation of logarithmic and circular functions. The same method can be used for the higher derivatives, the expression involving integrals which can be expressed in terms of polylogarithms. For $\Psi'$ the integrals can be put in terms of Clausen's function. The inverse problem of the expression of Clausen's function in terms of $\Psi'$ was given in de Doelder's 1984 paper, outlined here.

16.1.2. Starting with the expression of $\zeta(2)$ as $\mathrm{Li}_2(1)$ and changing the variable of integration and the path of integration yields the following sequence:

$$\begin{aligned}\pi^2/6 &= -\int_0^1 \frac{\log(1-x)}{x}dx = -\int_0^{e^{i\theta}} \frac{\log(1-ye^{-i\theta})}{y}dy \\ &= -\int_0^1 \frac{\log(1-ye^{-i\theta})}{y}dy - \int_1^{e^{i\theta}} \frac{\log(1-ye^{-i\theta})}{y}dy.\end{aligned} \tag{16.1}$$

Integrating the first integral by parts, and putting $y = e^{i(\theta-\psi)}$ in the second gives

$$\pi^2/6 = \int_0^1 \frac{\log y}{y-e^{i\theta}}dy - i\int_0^\theta \log(1-e^{-i\psi})\,d\psi. \tag{16.2}$$

The real and imaginary parts of this equation yield

$$\mathrm{Cl}_2(\theta) = -\sin\theta \int_0^1 \frac{\log y\,dy}{1-2y\cos\theta+y^2}, \tag{16.3}$$

$$\frac{\pi^2}{6} + \frac{\theta^2-2\pi\theta}{4} = \int_0^1 \frac{[(y-\cos\theta)\log y]\,dy}{1-2y\cos\theta+y^2}; \qquad 0 \le \theta \le 2\pi. \tag{16.4}$$

We give here de Doelder's derivation for the case $p$ odd and $q$ even. Then $z = \exp(ip\pi/q)$ is a solution of $z^q = -1$, an equation whose $q$ roots are

(16.5) $$z_k = \exp[i(2k+1)\pi/q], \qquad k = 0, 1, \dots, (q-1).$$

Hence, with $\theta = p\pi/q$ one can factorize $\rho^q + 1$ as

(16.6) $$\rho^q + 1 = (\rho^2 - 2\rho\cos\theta + 1)\cdot P,$$

with

(16.7) $$P = \prod_{m=0}^{(q-2)/2}{}' \left(\rho^2 - 2\rho\cos\left(\frac{2m+1}{q}\pi\right) + 1\right).$$

The prime on the product means that the term with either $m = (p-1)/2$ or $(2q-p-1)/2$ is omitted. Clearly $P$ is a polynomial of degree $q-2$ in $\rho$, with coefficients $A_k$ defined by

(16.8) $$P = A_0\rho^{q-2} + A_1\rho^{q-3} + \cdots + A_{q-2}.$$

Inserting this into (16.7) and identifying coefficients gives $A_0 = 1$, $A_1 = 2\cos\theta$, together with the recursion relation

(16.9) $$A_{k+1} = 2\cos\theta A_k - A_{k-1}, \qquad k = 1, 2, \dots, (q-3).$$

This is readily shown to have the solution

(16.10) $$A_k = \frac{\sin(k+1)\theta}{\sin\theta}, \qquad k = 0, 1, \dots, (q-2).$$

Hence

(16.11) $$(\rho^2 - 2\rho\cos\theta + 1)^{-1} = \frac{(\rho^q+1)^{-1}}{\sin\theta}\sum_{k=0}^{q-2}\rho^{q-k-2}\sin(k+1)\theta.$$

The integral

$$I(\alpha) = \int_0^1 \frac{\rho^\alpha}{\rho^2 - 2\rho\cos\theta + 1}\,d\rho$$

can therefore be expanded as

(16.12) $$I(\alpha) = \sum_{k=0}^{q-2}\frac{\sin(k+1)\theta}{\sin\theta}\int_0^1 \frac{\rho^{q+\alpha-k-2}}{\rho^q+1}\,d\rho.$$

With the substitution $x = \rho^q$ this integral becomes

$$\frac{1}{q}\int_0^1 \frac{x^{(\alpha-k-1)/q}}{1+x}\,dx = \frac{1}{2q}\left\{\Psi\left(1 + \frac{\alpha-k-1}{2q}\right) - \Psi\left(\frac{1}{2} + \frac{\alpha-k-1}{2q}\right)\right\},$$

on using the formula $\Psi(\beta) - \Psi(1) = \int_0^1((1-x^{\beta-1})/(1-x))\,dx$. Hence

(16.13) $$I(\alpha) = \frac{1}{2q}\sum_{k=1}^{q-1}\frac{\sin k\theta}{\sin\theta}\left[\Psi\left(1 + \frac{\alpha-k}{2q}\right) - \Psi\left(\frac{1}{2} + \frac{\alpha-k}{2q}\right)\right].$$



Differentiating with respect to $\alpha$, taking $\alpha = 0$ and using (16.3) gives

(16.14)
$$\begin{aligned}\mathrm{Cl}_2(\pi p/q) &= -\sin(p\pi/q)I'(0)\\ &= -\frac{1}{4q^2}\sum_{k=1}^{q-1}\left[\Psi'\left(1 - \frac{k}{2q}\right) - \Psi'\left(\frac{1}{2} - \frac{k}{2q}\right)\right]\sin(kp\pi/q),\end{aligned}$$

($p$ odd, $q$ even).

An analysis, slightly more complicated in detail, gives the identical result for $p$ and $q$ both odd, $q > 1$; while the sign of the second term is reversed when $p$ is even. All forms are covered by

(16.15)
$$\mathrm{Cl}_2(\pi p/q) = -\frac{1}{4q^2}\sum_{k=1}^{q-1}\left[\Psi'\left(1 - \frac{k}{2q}\right) + \cos p\pi\,\Psi'\left(\frac{1}{2} - \frac{k}{2q}\right)\right]\sin(kp\pi/q),$$

($q > 1$; $p$, $q$ not both even).

In a subsequent paper [2] de Doelder notes that, when $p$ is even and $q$ odd, (16.15) can be simplified by using the duplication formula for the gamma function, to give

(16.16) $$\mathrm{Cl}_2(\pi p/q) = -\frac{1}{q^2}\sum_{k=1}^{q-1}\Psi'\left(1 - \frac{k}{q}\right)\sin(kp\pi/q),$$

($p$ even; $q > 1$, odd).

Two interesting trigonometric series can also be deduced from the same material [2]:

(16.17) $$\sum_{k=1}^{q-1}\frac{\pi^2\cos k\theta}{\sin^2(\pi k/2q)} = \frac{\pi^2}{3} + 4q^2\left(\frac{\pi^2}{6} + \frac{\theta^2}{4} - \frac{\pi\theta}{2}\right); \qquad \theta = p\pi/q,$$

($p$ odd, $q$ even; $p < 2q$).

(16.18)
$$\sum_{k=1}^{q-1}\frac{\pi^2\cos k\theta}{\sin^2(\pi k/(2r+1))} = \frac{-\pi^2}{6} + (2r-1)^2\left(\frac{\pi^2}{6} + \frac{\theta^2}{4} - \frac{\pi\theta}{2}\right); \quad \theta = \frac{p\pi}{2r+1},$$

($p$ even, $q$ odd $= 2r+1$; $p \le r$).

16.1.3. Although the inverse solution of $\Psi'$ in terms of $\mathrm{Cl}_2$ is implicit in the above material a direct derivation is simpler. With $\theta$ now standing for $2\pi/q$, a partial-fraction devolution gives, for the case $q$ odd, $= 2n+1$:

(16.19) $$\frac{qy^{p-1}}{1-y^q} = \frac{1}{1-y} + 2\sum_{m=1}^{n}\frac{\cos m(p-1)\theta - y\cos mp\theta}{y^2 - 2y\cos m\theta + 1}.$$

In (16.3) and (16.4) replace $\theta$ by $m\theta$ and take the sum of the second equation times $\cos mp\theta$ and the first times $\sin mp\theta$:

$$\begin{aligned}(16.20)\qquad &\left[\frac{\pi^2}{6}+\frac{m^2\theta^2-2m\pi\theta}{4}\right]\cos mp\theta+\mathrm{Cl}_2(m\theta)\sin mp\theta\\ &=\int_0^1\frac{[y\cos mp\theta-\cos m(p-1)\theta]\log y\,dy}{y^2-2y\cos m\theta+1}.\end{aligned}$$

Multiply (16.19) by $\log y$ and integrate with respect to $y$ from 0 to 1, utilizing

$$(16.21)\qquad \Psi'(p/q)=-\int_0^1\frac{q^2y^{p-1}\log y}{1-y^q}\,dy.$$

Then, summing (16.20) from $m=1$ to $n$ gives

$$\begin{aligned}(16.22)\qquad &\frac{1}{q}\Psi'(p/q)\\ &=-\int_0^1\frac{\log y}{1-y}\,dy\\ &\quad+2\sum_{m=1}^{n}\left\{\left[\frac{\pi^2}{6}+\frac{m^2\theta^2-2m\pi\theta}{4}\right]\cos(mp\theta)+\sin(mp\theta)\mathrm{Cl}_2(m\theta)\right\}.\end{aligned}$$

The trigonometric series and the first integral are readily evaluated, leading to

$$(16.23)\qquad \Psi'(p/q)=\frac{\pi^2}{2\sin^2(\pi p/q)}+2q\sum_{m=1}^{[(q-1)/2]}\sin\left(\frac{2mp\pi}{q}\right)\mathrm{Cl}_2\left(\frac{2m\pi}{q}\right).$$

In this form, the equation can be shown to hold whether $q$ is even or odd.

A particular case of the above results, due to Fettis [3], is

$$(16.24)\qquad \mathrm{Cl}_2(\pi/3)=[\Psi'(1/3)-2\pi^2/3]/2\sqrt{3}.$$

From this, and the relation between the inverse tangent integral and the Claussen function, de Doelder deduced that

$$(16.25)\qquad \mathrm{Ti}_2(1/\sqrt{3})=\frac{\pi}{12}\log 3+5[\Psi'(1/3)-2\pi^2/3]/12\sqrt{3}.$$

A number of other examples are given in [1], including

$$(16.26)\qquad \mathrm{Cl}_2(\pi/6)=\frac{1}{24}[\sqrt{3}\Psi'(1/3)+16G-2\pi^2/\sqrt{3}],$$

where $G=\mathrm{Cl}_2(\pi/2)$ is Catalan's constant.

16.1.4. A more direct way of achieving similar results, which has the advantage of applying also to higher orders, stems from a consideration of the series. From the definition of $\mathrm{Li}_n$ as a series,

$$(16.27)\qquad \mathrm{Li}_n(e^{i2\pi p/q})=\sum_{m=1}^{\infty}\frac{e^{im2\pi p/q}}{m^n}.$$



Put $m=qr+s$ with $r=0,1,\dots,\infty$; $s=0,1,\dots,(q-1)$, and $r$ and $s$ not simultaneously zero. Then

$$\begin{aligned}\mathrm{Li}_n(e^{i2\pi p/q})&=\sum_{r=0}^{\infty}\sum_{s=0}^{q-1}{}'\frac{e^{i(2\pi p/q)(qr+s)}}{(qr+s)^n}\\ (16.28)\qquad &=\sum_{r=1}^{\infty}\frac{1}{(qr)^n}+\sum_{r=0}^{\infty}\sum_{s=1}^{q-1}\frac{e^{is2\pi p/q}}{(qr+s)^n}\\ &=\frac{1}{q^n}\left[\zeta(n)+\sum_{r=0}^{\infty}\sum_{s=1}^{q-1}\frac{e^{is2\pi p/q}}{(r+s/q)^n}\right].\end{aligned}$$

Now $\Psi'(\alpha)=\sum_{r=0}^{\infty}(r+\alpha)^{-2}$. Differentiating $n-2$ times gives

$$\Psi^{(n-1)}(\alpha)=(-1)^n(n-1)!\sum_{r=0}^{\infty}(r+\alpha)^{-n}.$$

Hence

$$(16.29)\qquad \sum_{r=0}^{\infty}(r+s/q)^{-n}=\frac{(-1)^n}{(n-1)!}\Psi^{(n-1)}(s/q).$$

Inserting in (16.29) gives

$$(16.30)\qquad \mathrm{Li}_n(e^{i2\pi p/q})=\frac{1}{q^n}\left[\zeta(n)+\frac{(-1)^n}{(n-1)!}\sum_{s=1}^{q-1}e^{is2\pi p/q}\Psi^{(n-1)}(s/q)\right].$$

Taking the imaginary part with $n$ even, and the real part with $n$ odd, gives

*n even*

$$(16.31)\qquad q^n\mathrm{Cl}_n(2\pi p/q)=\frac{1}{(n-1)!}\sum_{s=1}^{q-1}\sin(s2\pi p/q)\Psi^{(n-1)}(s/q),$$

*n odd*

$$(16.32)\qquad q^n\mathrm{Cl}_n(2\pi p/q)=\zeta(n)-\frac{1}{(n-1)!}\sum_{s=1}^{q-1}\cos(s2\pi p/q)\Psi^{(n-1)}(s/q).$$

For $n=2$ this form is somewhat simpler than (16.15), though the argument of $\mathrm{Cl}_2$ is doubled.

**16.2. An infinite integral of a product of two polylogarithms.**

16.2.1. Adamchik and Kölbig [4] have used residue techniques coupled with Mellin transforms to evaluate an infinite integral depending on several parameters and containing a product of two polylogarithms. The integral in question is

$$(16.33)\qquad I_{n,m}(\alpha,\sigma,\omega,r)=\int_0^{\infty}x^{\alpha-1}\mathrm{Li}_n(-\sigma x)\mathrm{Li}_m(-\omega x^r)\,dx,$$

for positive integers $n$ and $m$, complex $\alpha$, $\sigma$, $\omega$, and real $r \neq 0$ such that the integral exists. It can be readily shown that the integral converges if

$$\begin{rcases} -1 - r < \operatorname{Re}\alpha < 0 & \text{if } r > 0 \\ -1 < \operatorname{Re}\alpha < -r & \text{if } r < 0 \end{rcases}, \tag{16.34}$$

and $|\arg\sigma|, |\arg\omega| < \pi$.

By taking the change of variable $x = \xi^{1/r}$, (16.33) gives the symmetry relation

$$|r| I_{n,m}(\alpha, \sigma, \omega, r) = I_{m,n}(\alpha/r, \omega, \sigma, 1/r). \tag{16.35}$$

16.2.2. An integral representation for $\operatorname{Li}_k(z)$ is

$$\operatorname{Li}_k(-z) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{\pi}{\sin \pi s} \cdot \frac{z^{-s}}{(-s)^k}\, ds, \tag{16.36}$$

where $-1 < a < 0$, $|\arg z| < \pi$, $k = 0, 1, 2, \ldots$. With this representation, (16.33) can be written, taking $rs = -t$:

$$\begin{aligned} I_{n,m} = \frac{1}{\sigma^\alpha |r|} \int_0^\infty \frac{du}{u} &\left( \frac{u^\alpha}{2\pi i} \int_{a-i\infty}^{a+i\infty} \frac{\pi}{\sin \pi s} \cdot \frac{u^s}{(-s)^n}\, ds \right) \\ &\cdot \left( \frac{1}{2\pi i} \int_{a'-i\infty}^{a'+i\infty} \frac{\pi}{\sin(\pi t/r)} \cdot \frac{(\zeta/u)^{-t}}{-(t/r)^m}\, dt \right) \end{aligned} \tag{16.37}$$

with $\zeta = \sigma\omega^{-1/r}$.

Applying well-known theorems for the Mellin transform, in particular the product theorem, there is obtained

(16.38)

$$I_{n,m}(\alpha, \sigma, \omega, r) = \frac{(-1)^{n+1}\pi^2}{\sigma^\alpha |r|} \cdot \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} \frac{(s+\alpha)^{-n}(s/r)^{-m}\zeta^{-s}}{\sin[\pi(s+\alpha)]\sin(\pi s/r)}\, ds.$$

This integral is the basis for the subsequent development. It can be put in terms of a very general integral known as Fox's $H$-function, and also as a special case of Meijer's $G$-function. The integral is evaluated by the residue technique and gives rise, in the general case, to an exceedingly complicated formula, for which the reader is referred to the original paper. The special



case $\alpha = 0$ (which implies $r < 0$) simplifies somewhat, and gives

(16.39)

$$\begin{aligned} &\int_0^\infty \operatorname{Li}_n(-\sigma x)\operatorname{Li}_m(-\omega x^r)\frac{dx}{x} \\ &= -(\mp 1)^{m+n} \left\{ \frac{\pi}{|r|^n} \sum_k \frac{(-1)^k \zeta^{\mp |r| k}}{k^{m+n} \sin(\pi |r| k)} + |r|^{m-1}\pi \sum_l \frac{(-1)^l \zeta^{\mp l}}{l^{(m+n)} \sin(\pi l/|r|)} \right. \\ &\qquad \left. - |r|^{-n} \sum_K \frac{\zeta^{\mp |r| K}\left[\frac{m+n}{|r|K} \pm \log\zeta\right]}{K^{m+n}(-1)^{K(1+r)}} \right\} \\ &\quad + H(\pm 1)(-1)^{n+1}\pi^{m+n+1} r^m \sum_{m_1=0}^{m+n+1} \frac{(-\frac{1}{\pi}\log\zeta)^{m_1}}{m_1!} \\ &\qquad \cdot \sum_{m_2=0}^{m+n+1-m_1} r^{-m_2} B^*_{m_2} B^*_{m+n+1-m_1-m_2}, \end{aligned}$$

where $H(1) = 1$, $H(-1) = 0$; $B^*_j = (|2^j - 2|/j!)|B_j|$, $B_j$ are the Bernoulli numbers, the upper signs correspond to $|\zeta| > 1$, the lower to $|\zeta| < 1$. The summation indices are restricted to those values of $k, l \in \mathbb{N}$ for which

$$|r|k \pm \alpha \notin \mathbb{N}, \qquad (l \mp \alpha)/|r| \notin \mathbb{N} \tag{16.40}$$

and

$$K \in \mathbb{N}, \qquad |r|K \pm \alpha \in \mathbb{N}; \qquad \mathbb{N} = \{1, 2, 3, \ldots\}, \tag{16.41}$$

(The numbers $B^*_j$ are here given by $x \operatorname{cosec} x = \sum_{j=0}^\infty B^*_j x^j$.)

16.2.3. For some special cases the series can be summed in terms of Lerch's function

$$\Phi(z, u, v) = \sum_{k=0}^\infty z^k/(k+v)^u, \qquad |z| < 1. \tag{16.42}$$

For $r = -2$, (16.39) becomes

(16.43)

$$\begin{aligned} &\int_0^\infty \operatorname{Li}_n(-\sigma x)\operatorname{Li}_m(-\omega x^{-2})\frac{dx}{x} \\ &= (\mp)^{m+n}\{2^{-(n+1)}\pi\zeta^{\mp 1}\Phi(-\zeta^{\mp 2}, m+n, 1/2) \\ &\qquad - 2^{n+1}(m+n)\operatorname{Li}_{m+n+1}(-\zeta^{\mp 2}) \pm 2^{-n}\log\zeta \operatorname{Li}_{m+n}(-\zeta^{\mp 2})\} \\ &\quad + H(\pm 1)(-\pi)^{m+n+1} 2^m \sum_{m_1=0}^{m+n+1} \frac{1}{m_1!}\left(\frac{-\log\zeta}{\pi}\right)^{m_1} \\ &\qquad \cdot \sum_{m_2=0}^{m+n+1-m_1} (-2)^{-m_2} B^*_{m_2} B^*_{m+n+1-m_1-m_2}. \end{aligned}$$

For $n = m = 1$,

$$\begin{aligned}&\int_0^\infty \log(1+\sigma x)\log(1+\omega x^{-2})\frac{dx}{x}\\ &= \pi \mathrm{Ti}_2(\zeta^{\mp 1}) + \frac{1}{2}\mathrm{Li}_3(-\zeta^{\mp 2}) \pm \frac{1}{2}\log\zeta\,\mathrm{Li}_2(-\zeta^{\mp 2})\\ &\quad + \begin{cases} \frac{1}{3}\log^3\zeta + \frac{5}{12}\pi^2\log\zeta, & (|\zeta|>1),\\ 0, & (|\zeta|<1).\end{cases}\end{aligned} \tag{16.44}$$

For $n = 1$, $m = 2$,

(16.45)

$$\begin{aligned}&\int_0^\infty \log(1+\sigma x)\mathrm{Li}_2(-\omega x^{-2})\frac{dx}{x}\\ &= \pm\frac{1}{4}\pi\zeta^{\mp 1}\Phi\left(-\zeta^{\mp 2}, 3, \frac{1}{2}\right) \pm \frac{3}{4}\mathrm{Li}_4(-\zeta^{\mp 2})\\ &\quad + \frac{1}{2}\log z\,\mathrm{Li}_3(-\zeta^{\mp 2}) - \begin{cases} \frac{1}{6}\log^4\zeta + \frac{5}{12}\pi^2\log^2\zeta + \frac{53}{480}\pi^4, & (|\zeta|>1).\\ 0, & (|\zeta|<1).\end{cases}\end{aligned}$$

With $\sigma = \omega = 1$ and $n + m$ odd,

$$\begin{aligned}&\int_0^\infty \mathrm{Li}_n(-x)\mathrm{Li}_m(-x^2)\frac{dx}{x}\\ &= \frac{2^{-(n+2)}\pi^{n+m-1}}{(n+m-1)!}\left\{|E_{n+m-1}| - \frac{2^{n+m+1}}{n+m+1}|B_{n+m+1}|\right\},\end{aligned} \tag{16.46}$$

a form containing both the Euler and Bernoulli numbers.

For $\sigma = \omega = 1$, (16.44) gives the special case

$$\int_0^\infty \log(1+x)\log(1+x^{-2})\frac{dx}{x} = \pi G - \frac{3}{8}\zeta(3), \tag{16.47}$$

where $G$ is Catalan's constant. A further interesting case is

$$\int_0^\infty x^{-3}\mathrm{Li}_n(-x)\mathrm{Li}_n(-x^2)\,dx = 2^{n-2}\pi - \frac{2^{-n-1}}{(n-1)!}S, \tag{16.48}$$

where

(16.49)

$$\begin{aligned}S &= \pi^2\sum_{k=0}^{[(n-1)/2]}\frac{(2n-2k-2)!}{(n-2k-1)!(2k)!}\pi^{2k}|E_{2k}|\\ &\quad + (-1)^{n+1}\left[\sum_{k=0}^{n-1}[(-1)^n - (-1)^k](k-n)\frac{(n+k-1)!}{k!}(1-2^{k-n})\zeta(n-k+1)\right.\\ &\qquad \left. -(2\pi)^{n+1}\sum_{k=0}^{n}(-1)^k\frac{(k+n-1)!}{k!}(2\pi)^{-k}\sum_{j=0}^{n+1-k}2^{-j}B_j^*B_{n+1-k-j}^*\right].\end{aligned}$$

A short proof that $S$ is zero for all $n$ is given in the next section.



A large number of other special cases are given in the original paper, including results for $r = 2$, $\alpha = -2$, and $r = 1$, $\alpha = -1/2$. For the latter, for example, with $\sigma = \omega = 1$,

$$\int_0^\infty x^{-3/2}[\mathrm{Li}_2(-x)]^2\,dx = \frac{8\pi}{3}(\pi^2 + 24\log 2). \tag{16.50}$$

This result is interesting if only because it breaks the rather common pattern of expressions homogeneous in powers of $\pi$ and the logarithm.

Some further results, presented at the Fourth International Conference on Computer Algebra in Physical Research [5] include the following interesting examples:

$$\int_0^\infty x^{-5/2}[\mathrm{Li}_2(-x)]^2\,dx = \frac{8\pi}{27}(-\pi^2 - 8\log 2 + 20), \tag{16.51}$$

$$\int_0^\infty \log(1+x)\log(1+1/x^2)\,dx = \frac{5\pi^2}{24} - \frac{\pi}{2}(2-\log 2) + 2G, \tag{16.52}$$

$$\int_0^\infty \mathrm{Li}_n(-x)\mathrm{Li}_m(-1/x)\,dx/x = (m+n)\zeta(m+n+1), \tag{16.53}$$

(16.54)

$$\int_0^\infty x^{-3/4}\log(1+x)\mathrm{Li}_2(-1/x)\,dx = -2\pi\sqrt{2}[5\pi^2/3 + 16(3\log 2 + G - 4)],$$

$$\int_0^\infty x^{-3/4}\mathrm{Li}_2(-x)\mathrm{Li}_2(-1/x)\,dx = 256\pi\sqrt{2}(3 - 3\log 2 - G), \tag{16.55}$$

$$\int_0^\infty x^{-1/2}\mathrm{Li}_2(-x)\mathrm{Li}_2(-1/x)\,dx = 16\pi(3 - 4\log 2), \tag{16.56}$$

$$\int_0^\infty x^{-2}\log(1+x)\mathrm{Li}_4(-x)\,dx = -\left[\frac{\pi^4}{40} + \frac{\pi^2}{3} + 2\zeta(3) + 4\zeta(5)\right], \tag{16.57}$$

$$\int_0^\infty x^{-2}\mathrm{Li}_3(-x)\mathrm{Li}_4(-x)\,dx = \frac{2\pi^4}{15} + \frac{10\pi^2}{3} + 20\zeta(3) + 4\zeta(5), \tag{16.58}$$

and many others in the same vein.

16.2.4. We consider the case $n$ odd, $= 2m + 1$. The case for $n$ even is similar. Define modified Euler and Bernoulli numbers $E_j^*$ and $B_j^*$ by

$$\theta\operatorname{cosec}\theta = \sum_{j=0}^\infty B_j^*\theta^j; \qquad \sec\theta = \sum_{j=0}^\infty E_j^*\theta^j. \tag{16.59}$$

Then

$$\zeta(2k) = \pi^{2k}B_{2k}^*/(2 - 2^{2-2k}). \tag{16.60}$$

The second series in (16.49) has a nonzero coefficient of $\zeta(n-k+1)$ only if $n$ and $k$ have opposite parity. Hence only zeta functions of even argument are involved, and these can all be put in terms of $\pi$ via (16.60). All the terms in (16.49) involve only powers of $\pi$, and collecting terms of the same power gives, after some minor rearrangements,

(16.61)
$$S=\sum_{k=0}^{m}\pi^{2k+2}\frac{(4m-2k)!}{(2m-2k)!}\{E_{2k}^{*}-2(2k+1)B_{2k+2}^{*}-2^{2k+2}\sum_{j=0}^{2k+2}2^{-j}B_j^{*}B_{2k+2-j}^{*}\}.$$

Putting $2k+2-2j$ for $j$ in the summation gives, for the terms in curly braces, (and recalling that $B_j^{*}$ is zero for $j$ odd),

$$E_{2k}^{*}-2(2k+1)B_{2k+2}^{*}-\sum_{j=0}^{k+1}2^{2j}B_{2j}^{*}B_{2k+2-2j}^{*}. \tag{16.62}$$

Consider the expansion

$$\frac{\theta}{\sin\theta}\cdot\frac{2\theta}{\sin 2\theta}=\sum_{r=0}^{\infty}B_{2r}^{*}\theta^{2r}\sum_{s=0}^{\infty}B_{2s}^{*}\theta^{2s}2^{2s}. \tag{16.63}$$

The coefficient of $\theta^{2k+2}$ is $\sum_{j=0}^{k+1}B_{2j}^{*}2^{2j}B_{2k+2-2j}^{*}$. However, the expression on the left of (16.63) can be written as

$$\frac{1}{\cos\theta}\left(\frac{\theta^2}{\sin^2\theta}\right)=\frac{\theta^2}{\cos\theta}+\frac{\theta^2\cos\theta}{\sin^2\theta}=\frac{\theta^2}{\cos\theta}+\frac{2\theta}{\sin\theta}-\frac{d}{d\theta}\frac{\theta^2}{\sin\theta}.$$

Hence the coefficient of $\theta^{2k+2}$ is $E_{2k}^{*}+B_{2k+2}^{*}[2-2(2k+2)]$. Accordingly, (16.62) is zero, and all terms in $S$ in (16.61) vanish identically.

**16.3. Cyclotomic and polylogarithmic equations for a Salem number.**

16.3.1. The Salem number $x$, satisfying the base equation

$$x^{10}+x^9-x^7-x^6-x^5-x^4-x^3+x+1=0 \tag{16.64}$$

was studied by Ray [6] in relation to its dilogarithmic properties. From a rather involved functional equation he was able to derive the analytic result

(16.65)
$$\begin{aligned}&\tfrac12\mathrm{Li}_2(x^{18})-\mathrm{Li}_2(x^{17})+\tfrac12\mathrm{Li}_2(x^{16})-\mathrm{Li}_2(x^{13})+2\mathrm{Li}_2(x^{11})\\&\quad+\mathrm{Li}_2(x^{10})-2\mathrm{Li}_2(x^{9})-\tfrac12\mathrm{Li}_2(x^{8})+\mathrm{Li}_2(x^{7})-\mathrm{Li}_2(x^{5})-\mathrm{Li}_2(x^{4})\\&\quad+\mathrm{Li}_2(x^{3})+\tfrac32\mathrm{Li}_2(x^{2})-\mathrm{Li}_2(x)-\tfrac12\log^2(x)=0,\end{aligned}$$

where $x=1.1762808\ldots$ is a real root of (16.64). Because the equation is symmetric it possesses roots in inverse pairs, the other real root of (16.64) being $x=0.8501371\ldots$; and it may be verified, using the inversion formula for the dilogarithm, that (16.65), with no change, is also satisfied for the smaller root. All the other roots occur in conjugate pairs on the unit circle, the one with the smallest argument being

$$u=e^{i\theta};\qquad \theta=62.81495\ldots^{\circ}. \tag{16.66}$$



It is a conjecture, with no known exceptions, that, under these circumstances, if the dilogarithm $\mathrm{Li}_2(x^m)$ is replaced by $\mathrm{Cl}_2(m\theta)$, an equation of the form (16.65), with all non-Claussen terms suppressed, will also be satisfied. It is straightforward to verify numerically that this is indeed the case for $\theta$ given by (16.66).

16.3.2. Corresponding to (16.65) a cyclotomic equation can be written down, containing fourteen factors. Manipulations of (16.64) might be expected to yield some additional results, and in fact a total of fifty cyclotomic equations were initially thus determined. This is by far and away the largest number known for any base, and they are listed in §16.3.7. They have indices $N$ ranging from 6 to 132, the distribution thinning out considerably at the upper end. Subsequently, D. Zagier, using a more systematic approach, has located a further twenty-one indices, the largest being 360, and making for a total of seventy-one cyclotomic equations. Because of the arbitrary way in which these cyclotomic equations can be combined, a construction was selected so that, in the vast majority of cases, the powers involved were 1, 2, 3, 4, 5 and $N$. In a few cases, when $N$ is even, a power of $N/2$ needs to be included; and in a very few other cases some other powers are also used. The cyclotomic equation corresponding to (16.65) in fact is a combination of nine of these more "basic" equations.

It is not intended to give here the derivation for each case, but some representative examples are shown. Mostly, they are easy to derive.

16.3.3. If the negative terms in (16.64) are taken to the right-hand side of the equation and summed, the result is

$$x^{10}+x^9+x+1=(1+x)(1+x^9)=x^3(1-x^5)(1-x)^{-1}, \tag{16.67}$$

which can be put in the form

$$1-x^{18}=(1-x^9)(1-x^5)(1-x^2)^{-1}x^3, \tag{16.68}$$

a cyclotomic equation of index 18. Dividing by the factor $(1-x^9)$, (16.68) is $(1-x^2)(1+x^9)=x^3(1-x^5)$ which can be multiplied out and rearranged as either

$$1-x^{11}=(1-x^6)(1-x^2)(1-x)^{-1}x^2, \tag{16.69}$$

or

$$1-x^{16}=(1-x^8)(1-x^7)(1-x^3)^{-1}x^2. \tag{16.70}$$

Many combinations of factors can be deduced in this way. A somewhat different sequence comes by multiplying (16.64) by $x^2-x+1$. Some terms cancel and the resulting formula is

$$1+x^{12}=x^5(1+x+x^2)=x^5(1-x^3)(1-x)^{-1}. \tag{16.71}$$

Hence

$$1-x^{24}=(1-x^{12})(1-x^3)(1-x)^{-1}x^5. \tag{16.72}$$

Equation (16.71) can also be written as $1-x^6+x^{12}=x^5(1+x^2)$, whence

$$1+x^{18}=(1+x^6)(1+x^2)x^5, \tag{16.73}$$

which can be expressed as an index-36 cyclotomic equation:

$$1-x^{36}=(1-x^{18})(1-x^{12})(1-x^6)^{-1}(1-x^4)(1-x^2)^{-1}x^5. \tag{16.74}$$

Equation (16.73) can also be rearranged and factorized as

$$(1-x^7)(1-x^{11})=(1-x^{16})(1-x^8)^{-1}x^5. \tag{16.75}$$

Combining with (16.70) this gives

$$(1-x^{11})(1-x^3)=x^7, \tag{16.76}$$

and this can be combined with (16.69) to give an index-6 result. Equation (16.71) can also be rearranged as

$$(1-x^7)(1-x^5)=x^6. \tag{16.77}$$

16.3.4. In the course of generating these and many other similar equations it was noticed that when apparently disparate indices, for example, 11 and 16 in (16.75), occurred in one formula, they also occurred in a different combination in another, enabling them to be separated out. The resulting equations then exhibited the straightforward type of factorization on which the $r|N$ ladder property was first conjectured. (However, any or all of the indices 1, 2, 3, 4 and 5 may also be involved.) The only apparent exception encountered was a formula combining indices 63 and 132:

$$(1-x^{132})=(1-x^{66})(1-x^{63})(1-x^{24})(1-x^{21})^{-1}(1-x^6)^{-1}x^3. \tag{16.78}$$

No further equations with these two indices could be found, but in the belief that the two should indeed be separable the following direct factorization of $(1-x^{63})$ was generated. Write

$$1-x^{63}=(1-x^{18})+x^{18}(1-x^{45}). \tag{16.79}$$

Using the already developed factorization of $(1-x^{18})$ and $(1-x^{45})$ (see §16.3.7), this reduces to

$$1-x^{63}=[1-x^2+x^{25}-x^{40}](1-x^3)^{-1}(1-x)^{-1}x^{10}, \tag{16.80}$$

with

$$\begin{aligned} 1-x^2+x^{25}-x^{40}&=(1-x^{40})-x^2(1-x^{23})\\ &=[x^5+x-1-x^9](1-x^6)^2(1-x)^{-2}x^8, \end{aligned} \tag{16.81}$$

again on using formulas for $(1-x^{40})$ and $(1-x^{23})$.

Now $x^5+x-1-x^9=-(1-x^5)+x(1-x^8)$, and with the equation for $(1-x^8)$ this reduces to

(16.82)
$$x^5+x-1-x^9=(1-x^5)(1-x)^{-1}x^{-3}[(1-x^5)(1-x^4)(1-x^3)-x^3(1-x)].$$



This result is multiplied by $1-x^7$ and use made of (16.77) to give

$$x^5+x-1-x^9=(1-x^5)(1-x^7)^{-1}(1-x)^{-1}[-1+x+x^3-x^6-x^8+x^{10}]. \tag{16.83}$$

From (16.71), multiplied by $x$, $x-x^6-x^8=x^7-x^{13}$; and from (16.76), $-1+x^3=-x^7-x^{11}-x^{14}$. Substituting in (16.83) gives

$$-1+x+x^3-x^6-x^8+x^{10}=x^{10}-x^{11}-x^{13}+x^{14}=x^{10}(1-x)(1-x^3). \tag{16.84}$$

Insertion back into (16.72) and (16.71) finally yields

$$1-x^{63}=(1-x^7)^{-2}(1-x^6)^2(1-x)^{-3}x^{34}. \tag{16.85}$$

This was by far the most difficult index to track down.

16.3.5. If a cyclotomic equation is written in the form

$$\prod_m^N (1-x^m)^{A_m}=x^n, \tag{16.86}$$

then, since this has to be an algebraic derivation from the symmetric base equation (16.64), the substitution of $1/x$ for $x$ must leave the formula unaltered. This has, as a direct consequence,

$$\begin{cases} \text{(a)} & \sum_m^N A_m \text{ is even},\\ \text{(b)} & n=\frac{1}{2}\sum_m^N mA_m. \end{cases} \tag{16.87}$$

The second part of this equation was found useful for checking the results in §16.3.7.

16.3.6. Corresponding to (16.86) is a dilogarithmic component-ladder

$$L_2(N,x)=\sum_m^N \frac{A_m}{m}\mathrm{Li}_2(x^m)+\frac{n}{2}\log^2(x). \tag{16.88}$$

The leading term has a factor $1/N$ and it is desirable to clear this by defining a *modified component-ladder* by

$$M_2(N,x)=NL_2(N,x). \tag{16.89}$$

In terms of it, (16.65) can be written

(16.90)
$$\begin{aligned} &\tfrac{1}{2}M_2(18,x)-M_2(17,x)+\tfrac{1}{2}M_2(16,x)-M_2(13,x)+2M_2(11,x)\\ &\quad+M_2(10,x)-M_2(9,x)+\tfrac{1}{2}M_2(8,x)+M_2(7,x)=0, \end{aligned}$$

where the component-ladders are defined as in the next section. The terms for $N<6$ are all absorbed, but the remaining terms do not quite follow the coefficients in (16.65) because, as here defined, $L_2(18,x)$ and $L_2(16,x)$ contain, respectively, terms in $\mathrm{Li}_2(x^9)$ and $\mathrm{Li}_2(x^8)$. (A minor redefinition would amend this, which is simply a consequence of the arbitrariness inherent in the ladder definition.)

With a total of seventy-one cyclotomic equations it might be expected that there would exist many more relations of the form (16.90). D. Zagier has predicted that these component-ladders should combine in fives, plus $\zeta(2)$, making a total of sixty-seven valid ladders at $n = 2$.

The next section shows, in Table I, the fifty cyclotomic equations first found; and in Table II Zagier's additional twenty-one results. Tables III and IV show the corresponding valid ladders, found by numerical computation and search.*

16.3.7. As a shorthand in this section we use the notation

$$(m) \equiv (1 - x^m) \tag{16.91}$$

TABLE I. 50 Cyclotomic Equations

| Index $N$ | Cyclotomic Equation |
|---|---|
| 6. | $(6) = (3)^{-1}(2)^{-1}(1)x^5$ |
| 7. | $(7) = (5)^{-1}x^6$ |
| 8. | $(8) = (5)^2(4)(3)(1)^{-1}x^{-4}$ |
| 9. | $(9) = (5)^{-1}(3)^{-1}(2)^2(1)^{-1}x^7$ |
| 10. | $(10) = (5)^{-1}(3)(2)^{-1}x^7$ |
| 11. | $(11) = (3)^{-1}x^7$ |
| 12. | $(12) = (6)(5)(4)^2(1)^{-1}x^{-3}$ |
| 13. | $(13) = (5)(3)(1)^{-1}x^3$ |
| 16. | $(16) = (8)(5)^{-1}(3)^{-1}x^8$ |
| 17. | $(17) = (5)^{-1}(3)^{-1}(2)^{-2}(1)x^{14}$ |
| 18. | $(18) = (9)(5)(2)^{-1}x^3$ |
| 20. | $(20) = (5)^{-1}(4)(1)^{-1}x^{11}$ |
| 21. | $(21) = (7)^{-3}(3)^2(1)^{-2}x^{19}$ |
| 23. | $(23) = (6)^2(1)^{-1}x^6$ |
| 24. | $(24) = (12)(3)(1)^{-1}x^5$ |
| 27. | $(27) = (9)(5)^{-1}(3)(2)^{-2}x^{12}$ |
| 28. | $(28) = (14)(5)(4)(3)^{-1}(2)^{-1}x^5$ |
| 29. | $(29) = (5)(3)^{-1}(1)^{-1}x^{14}$ |

* Subsequent studies at the Max-Planck-Institut für Mathematik have extended these results to at least the thirteenth order, where the combining coefficients become extremely large. Extension to the 16th order is predicted.



TABLE I. 50 Cyclotomic Equations (continued)

| Index $N$ | Cyclotomic Equation |
|---|---|
| 30. | $(30) = (15)(5)^{-1}(2)^{-1}x^{11}$ |
| 34. | $(34) = (17)(5)^2(3)(2)(1)^{-2}x^2$ |
| 36. | $(36) = (18)(12)(6)^{-1}(4)(2)^{-1}x^5$ |
| 37. | $(37) = (5)^{-2}(3)^2(2)^{-2}(1)^{-1}x^{23}$ |
| 38. | $(38) = (19)(9)(2)^{-1}x^6$ |
| 40. | $(40) = (6)^2(4)(1)^{-2}x^{13}$ |
| 42. | $(42) = (21)(14)(3)^{-1}x^5$ |
| 44. | $(44) = (22)(4)(3)(2)^{-1}(1)^{-1}x^9$ |
| 45. | $(45) = (15)(3)^{-1}(1)^{-1}x^{17}$ |
| 47. | $(47) = (5)^{-1}(1)^{-2}x^{27}$ |
| 50. | $(50) = (25)(5)(3)(2)^{-1}(1)^{-1}x^{10}$ |
| 52. | $(52) = (26)(4)(3)^{-2}(2)^{-1}x^{15}$ |
| 56. | $(56) = (28)(7)(3)^{-1}(2)^2(1)^{-2}x^{11}$ |
| 60. | $(60) = (30)(5)^2(4)^2(1)^{-2}x^7$ |
| 62. | $(62) = (31)^2(6)x^{-3}$ |
| 63. | $(63) = (7)^{-2}(6)^2(1)^{-3}x^{34}$ |
| 64. | $(64) = (32)(16)(8)^{-1}(5)^{-1}(3)(2)^{-2}x^{15}$ |
| 65. | $(65) = (5)^{-2}(3)^{-3}(2)^4(1)^{-4}x^{40}$ |
| 66. | $(66) = (33)(22)(11)^{-1}(5)(1)^{-1}x^9$ |
| 70. | $(70) = (35)(14)(1)^{-1}x^{11}$ |
| 74. | $(74) = (37)(5)^{-1}(3)^{-3}(2)(1)^{-1}x^{25}$ |
| 76. | $(76) = (38)(5)^3(4)(3)(2)^{-1}(1)^{-2}x^{10}$ |
| 78. | $(78) = (39)(26)(13)^{-1}(5)^{-1}(1)^{-1}x^{16}$ |
| 84. | $(84) = (42)(7)(4)^2(2)^{-2}(1)^{-1}x^{16}$ |
| 86. | $(86) = (43)(6)^2(5)(3)^{-1}(2)^{-1}(1)^{-1}x^{16}$ |
| 92. | $(92) = (46)(5)(4)(2)(1)^{-3}x^{19}$ |
| 96. | $(96) = (48)(32)(16)^{-1}(3)^{-1}(1)^{-1}x^{18}$ |
| 98. | $(98) = (49)(14)(3)^{-1}(2)^{-2}x^{21}$ |
| 110. | $(110) = (55)(22)(5)^{-1}(3)^{-1}(1)^{-1}x^{21}$ |
| 118. | $(118) = (59)(5)^2(3)^2(2)^{-1}(1)^{-3}x^{24}$ |
| 124. | $(124) = (62)(31)(4)(2)^{-1}(1)^{-1}x^{15}$ |
| 132. | $(132) = (33)(22)(12)(6)(1)^{-3}x^{31}$ |

The additional indices recently found by D. Zagier are

$$N = 57, 75, 105, 108, 130, 138, 144, 154, 160, 165, 175, 182, 186, 195, 204, 212, 240, 246, 270, 286, 360.$$

The corresponding cyclotomic equations are given in Table II.

TABLE II. 21 Additional Cyclotomic Equations

| Index $N$ | Cyclotomic Equation |
|---|---|
| 57. | $(57) = (19)(5)^{-1}(3)^{-1}(2)^{-2}x^{25}$ |
| 75. | $(75) = (25)(15)(3)^{-1}(2)^{-2}x^{21}$ |
| 105. | $(105) = (35)(15)(5)(3)(1)^{-3}x^{25}$ |
| 108. | $(108) = (54)(5)^{3}(4)^{4}(3)^{-1}(2)^{-2}(1)^{-2}x^{16}$ |
| 130. | $(130) = (65)(26)^{2}(13)^{-2}(2)^{-1}(1)^{-1}x^{21}$ |
| 138. | $(138) = (69)(46)(23)^{-1}(5)^{-2}(3)(2)^{-2}(1)^{-1}x^{29}$ |
| 144. | $(144) = (72)(48)(24)^{-1}(5)(2)^{-2}(1)^{-1}x^{24}$ |
| 154. | $(154) = (77)(22)(14)(11)^{-1}(7)^{-1}(3)^{-2}(2)^{-1}(1)^{-1}x^{34}$ |
| 160. | $(160) = (80)(32)(16)^{-1}(5)^{4}(3)^{2}(1)^{-4}x^{21}$ |
| 165. | $(165) = (55)(33)(15)(5)^{-2}(3)^{-1}(2)^{-2}(1)^{-1}x^{40}$ |
| 175. | $(175) = (35)(25)^{2}(5)^{-2}(3)(2)^{-2}(1)^{-3}x^{52}$ |
| 182. | $(182) = (91)(26)(14)^{2}(13)^{-1}(7)^{-2}(5)^{-1}(3)^{-1}(1)^{-2}x^{37}$ |
| 186. | $(186) = (93)(31)(3)^{-1}(2)(1)^{-3}x^{33}$ |
| 195. | $(195) = (39)(15)(5)^{-4}(3)^{-5}(1)^{-4}x^{90}$ |
| 204. | $(204) = (102)(68)(34)^{-1}(5)^{-1}(4)(3)^{-2}(2)(1)^{-3}x^{38}$ |
| 212. | $(212) = (106)(5)^{-1}(4)(3)^{-2}(2)^{-5}(1)^{-1}x^{62}$ |
| 240. | $(240) = (120)(80)(48)(40)^{-1}(24)^{-1}(5)^{-1}(3)^{-1}(1)^{-2}x^{33}$ |
| 246. | $(246) = (123)(82)^{2}(41)^{-2}(5)^{-2}(3)^{-2}(2)^{-3}(1)x^{31}$ |
| 270. | $(270) = (135)(90)(54)(45)^{-1}(27)^{-1}(5)^{2}(3)^{-1}(2)^{-1}(1)^{-2}x^{30}$ |
| 286. | $(286) = (143)(26)(22)(13)^{-1}(11)^{-1}(5)^{-1}(3)^{2}(2)^{-3}(1)^{-4}x^{64}$ |
| 360. | $(360) = (180)(120)(72)(60)^{-1}(36)^{-1}(5)^{-3}(3)^{2}(2)^{-2}(1)^{-3}x^{50}$ |

The following valid dilogarithmic ladders were found using numerical methods to thirty-three decimal places, as calculated by M. Abouzahra, and a multicomponent search algorithm by G. Szekeres; to both of whom I am grateful for their contributions to these computations. We use the shorthand notation

$$[N] \equiv M_2(N, x), \tag{16.92}$$

where $M$ is the modified component-ladder of (16.89). The results read



across the table. Thus, the first ladder is

$$3[6] = 20\zeta(2) + 8[7] + 0[8] - 4[9] - 6[10]. \tag{16.93}$$

All the formulas are expressed in the somewhat arbitrary choice of component-ladders of indices 7, 8, 9, and 10.

TABLE III. 46 Dilogarithmic Valid Ladders

| | $\zeta(2)$ | [7] | [8] | [9] | [10] |
|---|---|---|---|---|---|
| 3[6] | 20 | 8 | 0 | −4 | −6 |
| 3[11] | 17 | 2 | 0 | −1 | −3 |
| [12] | 9 | 9 | 2 | −2 | −4 |
| [13] | −2 | −3 | 1 | 1 | 2 |
| 3[16] | 35 | 17 | 0 | −4 | −9 |
| [17] | 23 | 9 | 0 | −4 | −5 |
| 3[18] | 23 | 5 | 3 | −4 | −3 |
| 3[20] | 44 | 26 | 6 | −4 | −9 |
| 3[21] | −38 | −14 | 9 | 13 | 24 |
| 9[23] | −28 | −13 | 9 | 20 | 30 |
| 3[24] | −26 | −26 | 3 | 10 | 18 |
| 3[27] | 32 | 5 | 3 | −4 | 6 |
| 3[28] | 125 | 59 | 9 | −22 | −36 |
| 3[29] | 89 | 14 | 9 | −4 | −12 |
| 3[30] | 61 | 25 | 3 | −8 | −9 |
| 3[34] | −46 | −46 | 12 | 20 | 27 |
| 3[36] | 122 | 56 | 9 | −22 | −33 |
| 3[37] | 35 | −7 | 9 | 5 | 27 |
| 9[38] | 123 | 15 | 9 | −14 | 0 |
| 9[40] | 181 | 85 | 54 | 16 | 6 |
| [42] | 31 | 11 | 1 | −4 | −8 |
| 3[44] | 80 | 44 | 18 | −10 | −9 |
| 3[45] | 139 | 31 | 12 | −5 | −18 |
| 3[47] | 113 | 5 | 18 | 11 | 12 |
| 3[50] | 49 | −23 | 15 | 4 | 18 |
| 3[52] | 328 | 154 | 18 | −50 | −90 |
| 3[56] | 53 | −31 | 12 | 26 | 9 |
| 3[60] | 232 | 112 | 45 | −26 | −57 |
| 9[62] | −125 | −14 | −9 | 16 | 24 |

TABLE III. 46 Dilogarithmic Valid Ladders (continued)

| | $\zeta(2)$ | [7] | [8] | [9] | [10] |
|---|---|---|---|---|---|
| 3[63] | 104 | 8 | 39 | 32 | 42 |
| 3[64] | 86 | 8 | 9 | −10 | 15 |
| 3[65] | 224 | 20 | 30 | 50 | −3 |
| [66] | 17 | −11 | 5 | 4 | 5 |
| 3[70] | 97 | 7 | 15 | 4 | 3 |
| 3[74] | 379 | 121 | 15 | −20 | −78 |
| 3[76] | 227 | 41 | 57 | −16 | −18 |
| 3[78] | 107 | 11 | 15 | 8 | 9 |
| 3[84] | 541 | 247 | 54 | −80 | −120 |
| 9[86] | 743 | 137 | 81 | −28 | −42 |
| 3[92] | 229 | 37 | 63 | 22 | −3 |
| [96] | 104 | 22 | 8 | −4 | −13 |
| 3[98] | 535 | 157 | 27 | −68 | −90 |
| 3[110] | 355 | 109 | 27 | −14 | −45 |
| [118] | 48 | −54 | 28 | 20 | 42 |
| 3[124] | 491 | 212 | 57 | −58 | −93 |
| [132] | 101 | −19 | 25 | 22 | 28 |

TABLE IV. 21 Additional Dilogarithmic Valid Ladders

| | $\zeta(2)$ | [7] | [8] | [9] | [10] |
|---|---|---|---|---|---|
| 3[57] | 298 | 94 | 12 | −35 | −45 |
| [75] | 135 | 40 | 7 | −17 | −23 |
| 3[105] | 167 | −88 | 66 | 53 | 78 |
| [108] | 397 | 205 | 48 | −64 | −111 |
| 3[130] | 440 | 80 | 48 | −28 | −30 |
| 3[138] | 404 | 92 | 42 | −16 | 27 |
| 3[144] | 653 | 95 | 69 | −52 | −45 |
| 3[154] | 986 | 260 | 66 | −76 | −144 |
| 3[160] | 122 | −322 | 144 | 122 | 198 |
| 3[165] | 1100 | 323 | 69 | −88 | −126 |
| [175] | 297 | 15 | 40 | 10 | 32 |
| [182] | 285 | 91 | 33 | −8 | −26 |
| 3[186] | 685 | 22 | 105 | 64 | 12 |



TABLE IV. 21 Additional Dilogarithmic Valid Ladders (continued)

| | $\zeta(2)$ | [7] | [8] | [9] | [10] |
|---|---|---|---|---|---|
| 3[195] | 2801 | 746 | 183 | −106 | −354 |
| 3[204] | 1315 | 427 | 144 | −32 | −192 |
| 3[212] | 3067 | 1027 | 201 | −368 | −507 |
| 3[240] | 1007 | 143 | 102 | 14 | −39 |
| [246] | 640 | 254 | 12 | −96 | −132 |
| [270] | 512 | 54 | 58 | −20 | −41 |
| 3[286] | 1711 | −65 | 291 | 76 | 267 |
| 3[360] | 1237 | 1 | 228 | 124 | 342 |

**16.4. New functional equations.**

16.4.1. H. Gangl, using extensive machine calculations, has come up with new, single- and double-variable functional equations for $n \leq 6$; these include the first nontrivial results for $n = 6$ in the 150 years since Kummer's studies for $n \leq 5$.

The single-variable results are based on powers of $z$, $(1-z)$, $(1-z+z^2)$, and of $z$, $(1-z)$, $(1-z+z^2)$, $(1+z)$. They are discussed further in Appendix A. As an example of Gangl's findings we quote the following, using the notation

$$[\pm, p, q, r, s] \equiv L_n[\pm z^p(1-z)^q(1-z+z^2)^r(1+z)^s]. \tag{16.94}$$

Then, at $n = 6$ it is found that

(16.95)

$$\begin{aligned}
&-3[-,5,2,-3,-2]-3[+,5,2,-3,-1]-3[+,0,2,-3,3]\\
&-3[-,1,2,-3,-2]-3[+,0,2,-3,-1]-3[+,1,2,-3,3]\\
&+4[+,-3,4,-1,1]+4[+,0,4,-1,1]+4[-,-3,-1,-1,1]\\
&+4[+,5,-1,-1,1]+4[+,2,-1,-1,1]+4[-,0,-1,-1,1]\\
&+4[+,0,-1,-1,-2]+4[-,5,-1,-1,-2]+4[-,0,4,-1,-2]\\
&+5[+,4,-2,1,-1]+5[+,0,2,1,-1]+5[+,-3,-2,1,-1]\\
&+5[+,-3,2,1,-1]+5[+,0,-2,1,-1]+5[+,1,-2,1,-1]\\
&+5[-,-3,-2,1,2]+5[-,-3,2,1,2]+5[-,1,-2,1,2]\\
&+20[-,4,1,-2,-1]+20[-,1,1,-2,-1]+20[-,3,2,-2,-1]\\
&+20[+,3,1,-2,-1]+20[+,0,1,-2,-1]+20[-,0,2,-2,-1]\\
&+20[-,0,1,-2,2]+20[+,1,1,-2,2]+20[+,0,2,-2,2]\\
&+60[+,2,1,0,-1]-60[-,3,-1,0,-1]-60[-,0,2,0,-1]\\
&+60[+,0,1,0,1]+60[-,2,-2,0,-1]-60[-,1,-1,0,1]\\
&-60[-,2,-1,0,-1]+60[+,1,1,0,1]+60[-,1,-2,0,-1]\\
&-60[-,2,-1,0,1]+60[+,1,1,0,-1]+60[-,1,-2,0,1]\\
&+60[-,2,1,0,0]-60[+,3,-1,0,0]-60[+,0,2,0,0]\\
&+60[-,0,1,0,0]+60[+,2,-2,0,0]-60[+,1,-1,0,0]\\
&-90[+,3,0,-1,-1]+90[+,1,0,1,-1]+90[+,2,0,-1,-1]
\end{aligned}$$

(continues)

$$
\begin{aligned}
&+90[+,1,0,-1,-1]-90[+,2,0,-1,1]+90[+,1,0,-1,1]\\
&-90[+,0,2,-1,-1]-90[+,1,2,-1,-1]-90[+,0,2,-1,1]\\
&+90[+,1,-2,1,-1]+90[+,0,0,1,1]-90[+,0,0,1,-1]\\
&+90[-,1,0,1,0]-90[-,3,0,-1,0]+90[-,1,-2,1,0]\\
&-90[-,1,2,-1,0]\\
&-120[+,1,1,0,0]+120[-,2,-1,0,0]+120[-,-1,2,0,0]\\
&+180[-,2,1,-1,-1]+180[+,0,1,-1,-1]+180[-,1,1,-1,-1]\\
&+180[+,1,1,-1,-1]+180[+,-1,1,-1,1]+180[-,1,1,-1,1]\\
&+180[-,0,1,-1,1]+180[+,0,1,-1,1]+180[-,0,0,-1,0]\\
&+180[+,0,0,-1,0]+180[+,2,0,-1,0]+180[-,2,0,-1,0]\\
&+180[+,2,1,-1,0]+180[-,-1,1,-1,0]+180[+,0,2,-1,0]\\
&-180[-,1,0,-1,0]-180[-,1,1,-1,0]-180[+,0,1,-1,0]\\
&-360[+,1,0,-1,0]\\
&=0.
\end{aligned}
$$

This result is obtainable by putting $y=-z$ in the two-variable functional equation of Appendix A, where the highly symmetrical nature of these formulas is discussed.

16.4.2. With the same notation as (16.94), but without the factor $(1+z)^s$, Gangl has found the following functional equation at $n=7$.

(16.96)

$$
\begin{aligned}
&-3[-,0,5,-5]-3[-,5,0,-5]-3[+,5,5,-5]\\
&+4[+,5,-5,0]+4[-,0,5,0]+4[-,5,0,0]\\
&+25[-,3,0,1]+25[-,0,3,1]+25[-,0,5,-1]\\
&+25[-,5,0,-1]+25[+,3,-5,1]+25[+,5,-3,-1]\\
&+25[-,3,3,-1]+25[+,4,-3,1]+25[+,3,-4,-1]\\
&+25[+,0,1,-3]+25[+,0,5,-3]+25[+,1,0,-3]\\
&+25[+,5,0,-3]+25[-,5,1,-3]+25[-,1,5,-3]\\
&+25[+,1,-3,3]+25[+,3,-1,-3]+25[-,1,-4,3]\\
&+25[-,3,4,-3]+25[-,4,3,-3]+25[-,4,-1,-3]\\
&-125[-,0,0,2]-125[-,0,4,-2]-125[-,4,0,-2]\\
&-125[+,0,1,1]-125[+,1,0,1]-125[+,3,0,-1]\\
&-125[+,0,3,-1]-125[-,3,-1,-1]-125[-,1,-3,1]\\
&-150[+,1,3,-1]-150[+,3,1,-1]-150[-,2,-1,1]\\
&-150[-,1,-2,-1]-150[-,3,-2,-1]-150[-,2,-3,1]\\
&+400[+,0,0,1]+400[+,0,2,-1]+400[+,2,0,-1]\\
&-500[+,0,1,-2]-500[+,0,3,-2]-500[-,1,3,-2]\\
&-500[+,1,0,-2]-500[+,3,0,-2]-500[-,3,1,-2]\\
&-500[-,3,-1,0]-500[+,3,-2,0]-500[-,1,-3,0]\\
&-500[+,2,-3,0]-500[+,2,1,0]-500[+,1,2,0]\\
&+900[+,1,1,-2]+900[-,1,2,-2]+900[-,2,1,-2]\\
&-2250[+,1,2,-1]-2250[+,2,1,-1]-2250[+,1,-1,1]\\
&-2250[+,1,-1,-1]-2250[+,2,-1,-1]-2250[+,1,-2,1]
\end{aligned}
$$

(continues)



$$
\begin{aligned}
&-2250[-,2,1,-1]-2250[-,1,2,-1]-2250[-,1,-1,1]\\
&-2250[-,2,-1,-1]-2250[-,1,-2,1]-2250[-,1,-1,-1]\\
&-2750[-,0,2,0]-2750[-,2,0,0]-2750[-,2,-2,0]\\
&+3000[+,1,1,0]+3000[-,2,-1,0]+3000[-,1,-2,0]\\
&+26625[-,0,1,-1]+26625[-,1,0,-1]+26625[+,1,1,-1]\\
&+27000[+,0,1,-1]+27000[+,1,0,-1]+27000[-,1,1,-1]\\
&-62500[-,0,1,0]-62500[-,1,0,0]-62500[+,1,-1,0]\\
&-68000[+,0,1,0]-68000[+,1,0,0]-68000[-,1,-1,0]\\
&=-(25461/8)\zeta(7)
\end{aligned}
$$

16.4.3. Using concepts from the algebraic $K$-theory of fields, Goncharov [7] has produced a new three-variable functional equation for the tri-logarithm. It can be put in the form

(16.97)

$$
\begin{aligned}
\sum_{i=1}^{3}\{&L_3(a_i)+L_3(b_i)-L_3(-b_i/a_{i-1})+L_3(b_i/a_{i-1}a_i)+L_3(a_ib_{i-1}/b_{i+1})\\
&+L_3(-b_i/a_ib_{i-1})-L_3(a_ia_{i-1}b_{i+1}/b_i)\}+L_3(-a_1a_2a_3)=3\zeta(3),
\end{aligned}
$$

where

$$b_i=1-a_i(1-a_{i-1});\quad \text{(indices taken modulo 3).}$$

Kummer's form can be recovered by taking, for example, $a_1, a_2, a_3 = 1, x, (1-y)/(1-x)$, but it is not known, at the time of writing, whether Goncharov's equation can be deduced conversely from Kummer's two-variable formula.

## References


1. P. J. de Doelder, *On the Claussen integral* $\mathrm{Cl}_2(\theta)$ *and a related integral*, J. Comput. Appl. Math. **11** (1984), 325–330.
2. ——, *On the Claussen integral* $\mathrm{Cl}_2(\theta)$ *and a related integral*, Nieuw Tydschrift Wisk., Sept. 1987, 10–18.
3. H. E. Fettis, *Adv. problems*, No. 6448, Amer. Math. Monthly, Jan. 1984.
4. V. S. Adamchick and K. S. Kölbig, *A definite integral of a product of two polylogarithms*, SIAM J. Math. Anal. **19** (1988), 926–938.
5. K. S. Kölbig, *On a definite integral of the product of two polylogarithms*, presented at the Fourth International Conference on Computer Algebra in Physical Research, Dubna, USSR, May 22–26, 1990.
6. G. A. Ray, *Multivariable polylogarithmic identities*, see Chapter 7 of this volume.
7. A. B. Goncharov, *The classical trilogarithm, algebraic K-theory of fields and Dedikind zeta functions*, Dept. Math. Report, MIT, Cambridge, Mass, 1990.

APPENDIX A

# Special Values and Functional Equations of Polylogarithms

DON ZAGIER

**0. Introduction.** In this appendix we want to give a unified explanation of the many linear relations which exist among values of polylogarithms at algebraic arguments, as well as a framework for understanding the functional equations of these functions. We will content ourselves with describing the (partly conjectural, partly proved) results, without explaining the reasons which motivate the form the answer takes. This motivation, as well as many more details of the theoretical setup, can be found in [Z], of which the present appendix (except for §7, which contains some new results) is essentially just a summary.

The classical theorem of number theory in which the "monologarithm" function $\mathrm{Li}_1(x) = -\log(1-x)$ appears is Dirichlet's theorem, which in a weakened form says: Let $\mathscr{A}_1(F)$ denote the group of units of an algebraic number field $F$ of degree $n = r_1 + 2r_2$ in the usual notation, $D_F$ the discriminant of $F$. Then

(1) the map $\mathscr{A}_1(F) \to \mathbf{R}^n$ which sends $\varepsilon$ to the $n$-tuple of values of $\log|\varepsilon^\sigma|$, $\sigma$ ranging over the embeddings of $F$ into $\mathbf{C}$, maps $\mathscr{A}_1(F)$ to a lattice in an $(r_1 + r_2 - 1)$-dimensional subspace of $\mathbf{R}^n$; and

(2) the covolume of (= volume of a fundamental parallelogram for) this lattice is a rational multiple of $|D_F|^{1/2}\pi^{-r_2}$ times the residue at $s = 1$ of $\zeta_F(s)$, the Dedekind zeta-function of $F$.

Our goal is to define similarly, for each $m > 1$, groups $\mathscr{A}_m(F)$ such that (i) the $m$th polylogarithm function defines a map from $\mathscr{A}_m(F)$ to a lattice in $\mathbf{R}^n$, and (ii) the covolume of this lattice is up to a simple factor equal to $\zeta_F(m)$. Assertion (i) implies—since the lattice is finite-dimensional—a large number of relations among the polylogarithms with values in $F$, including all known, and conjecturally all, such relations, while assertion (ii) implies that $\zeta_F(m)$ for any number field $F$ and integer $m > 1$ can be expressed in finite terms in terms of values of polylogarithms with values in $F$. (It was this latter statement, not the existence of relations among polylogarithm values, which was in fact our original motivation.)

The definition of $\mathscr{A}_m(F)$ and more details of the above picture are given in §1. Several examples in the case of the dilogarithm are discussed in §2, and for higher polylogarithms in §3. The reader may want to look at these sections before reading §1, which is more theoretical. The relation between the concepts explained here and Lewin's ladders is discussed in §4. The following section proves in the case $F = \mathbf{Q}$ that there are enough elements of $\mathscr{A}_m(F)$ to produce nontrivial relations, and in particular that there are nontrivial relations among values of $\mathrm{Li}_m$ with rational arguments for arbitrarily large values of $m$. In §6 we state the conjecture that all relations among values of polylogarithms at algebraic arguments come from specializing functional equations (i.e., are "analytically derivable" in Lewin's terminology) and hence—since the processes of verifying or specializing any functional equation are mechanical—are in some sense trivial; of course, this is a very nontrivial use of the word "trivial," since *finding* the functional equations needed is not at all an easy matter. Finally, in §7 we discuss how the same setup as is needed to understand relations among special values of polylogarithms also lets one search efficiently for functional equations, and report on some new functional equations found by H. Gangl, including his spectacular discovery of a functional equation for the hexalogarithm, the first progress beyond $m = 5$ in 150 years.

**1. The basic algebraic relation and the definition of $\mathscr{A}_m(F)$.** We are interested in looking for linear dependences of polylogarithm values at algebraic arguments. It turns out that the only combinations which work are those whose arguments satisfy two conditions. The first condition, which is the basic one, is purely algebraic. We will call collections of numbers satisfying this condition *good*, or more precisely (since the condition depends on $m$ and becomes more restrictive as $m$ grows) *good at level* $m$. For the problem of finding functional equations, discussed in §7, only the algebraic condition is needed, so the whole problem is to find good combinations of functions of one or several variables at various levels $m$. The second condition is inductive. At level 2, any "good" combination of arguments is mapped by the dilogarithm into a certain lattice in $\mathbf{R}^n$, as sketched in the introductory paragraphs and explained in more detail below, so the group $\mathscr{A}_2(F)$ consists of all good combinations and we automatically get nontrivial relations among dilogarithm values whenever the rank of this group is bigger than the dimension of the lattice to which it maps. (Actually, its rank is always infinite, as we will see in §5, so it is merely a question of computation to get as many relations as desired.) In higher levels $m$, any good combination of arguments has associated to it other combinations which are good at some smaller level $m'$, $2 \le m' < m$. Starting at the bottom ($m' = 2$), we require that the image of each of these lower-level combinations under the $m'$th polylogarithm map is 0 in the relevant lattice. The group $\mathscr{A}_m(F)$, defined as the set of all combinations of elements of $F$ which satisfy this property, then is mapped by the $m$th polylogarithm map into a certain lattice of finite rank, so we automatically get relations among polylogarithm values if the rank of $\mathscr{A}_m(F)$ (which, again, turns out actually to be infinite) is larger than that of the lattice.



We now give the precise definitions. Let $F$ be any field. By a "combination of elements of $F$" we will mean a formal linear combination $\xi = \sum_i n_i[x_i]$ where $i$ runs over a finite index set, the $x_i$ are elements of $F^\times$, and the coefficients $n_i$ are integers. (Actually, all our considerations are up to torsion only, so we could allow $n_i \in \mathbf{Q}$.) For convenience we also allow $x_i$ to be 0 or $\infty$ but set $[0] = [\infty] = 0$. The set of all combinations $\xi$ forms a group $\mathscr{G}_F$, the free abelian group on $F^\times$. An element $\xi \in \mathscr{G}_F$ is *good at level* 2 if it satisfies the relation

$$(*) \qquad \sum_i n_i[x_i] \wedge [1 - x_i] = 0 \quad \text{in } \Lambda^2(F^\times) \otimes_{\mathbf{Z}} \mathbf{Q},$$

i.e. if the sum $\sum n_i[x_i] \wedge [1 - x_i]$ is a torsion element in the exterior square of $F^\times$. (The element $[x_i] \wedge [1 - x_i]$ is to be interpreted as 0 if $x_i = 0$, 1, or $\infty$.) Explicitly, $(*)$ means that if we pick a basis $p_1, \dots, p_s$ of the subgroup of $F^\times$ generated by all $x_i$ and $1 - x_i$ (modulo torsion) and write

$$x_i = \zeta_i \prod_{j=1}^{s} p_j^{a_{ij}}, \qquad 1 - x_i = \zeta_i' \prod_{j=1}^{s} p_j^{a_{ij}'}$$

where $\zeta_i, \zeta_i' \in F^\times$ are roots of unity and $a_{ij}$, $a_{ij}'$ belong to $\mathbf{Z}$, then

$$\sum_i n_i(a_{ij}a_{ik}' - a_{ik}a_{ij}') = 0 \qquad (1 \le j < k \le s).$$

In particular, if for some finite subset $S = \{p_1, \dots, p_s\}$ of $F^\times$ the set of $x \in F$ for which both $x$ and $1 - x$ belong to

$$\langle S \rangle = \{\zeta p_1^{a_1} \cdots p_s^{a_s} \mid \zeta \in F^\times \text{ a root of unity}, a_1, \dots, a_s \in \mathbf{Z}\}$$

has cardinality bigger than $\binom{s}{2}$, then we can find good combinations of these elements $x$. A key question, both for finding relations among special values and for finding functional equations of polylogarithms, will be to find as large sets as possible of (numbers or rational functions) $x \in F$ for which all $x$ and $1 - x$ belong to a subgroup of $F^\times$ of small rank.

For any vector space, we can identify $\Lambda^2(V)$ with $V \otimes V / \mathrm{Sym}^2(V)$, where $\mathrm{Sym}^2(V)$ is the subspace of $V \otimes V$ spanned by elements $x \otimes y + y \otimes x$ (or by elements $x \otimes x$). Thus $(*)$ can be interpreted as saying that $\sum n_i[x_i] \otimes [1 - x_i]$ belongs to $\mathrm{Sym}^2(F^\times)$ up to torsion. The generalization to higher $m$ is to say that a combination $\xi \in \mathscr{G}_F$ is *good at level* $m$ if it satisfies the relation

$$(*_m) \qquad \sum_i n_i \underbrace{[x_i] \otimes \cdots \otimes [x_i]}_{m-1} \otimes [1 - x_i] \in \mathrm{Sym}^m(F^\times) \otimes_{\mathbf{Z}} \mathbf{Q},$$

where $\mathrm{Sym}^m(A)$ for any abelian group $A$ is the $\mathfrak{S}_m$-invariant subspace of the $m$th tensor power $A^{\otimes m}$. We can write the $(m-1)$-fold tensor product $[x]\otimes\cdots\otimes[x]$ as $[x]^{\otimes(m-1)}$ or more simply, since it automatically belongs to the subspace $\mathrm{Sym}^{m-1}(F^\times)$ of $\bigotimes^{m-1}(F^\times)$, as $[x]^{m-1}$. Thus the set $\mathscr{G}_m(F)$ of good combinations at level $m$ is the kernel of the linear map

$$\beta_m : \mathscr{F}_F \to ((\mathrm{Sym}^{m-1}(F^\times)\otimes F^\times)/\,\mathrm{Sym}^m(F^\times))\otimes_{\mathbf{Z}}\mathbf{Q}$$

defined on generators by

$$\beta_m([x]) = \begin{cases} [x]^{m-1}\otimes[1-x]\ (\mathrm{mod}\ \mathrm{Sym}^m(F^\times)) & \text{if } x\neq 0,1,\infty, \\ 0 & \text{if } x=0,1,\infty. \end{cases}$$

We can also write $\mathscr{G}_m(F)$ in a way more analogous to the definition for $m=2$ as the kernel of the map $[x]\mapsto[x]^{m-2}\otimes([x]\wedge[1-x])\in\mathrm{Sym}^{m-1}(F^\times)\otimes\Lambda^2(F^\times)$, since for any vector space $V$ there is a natural injection

$$(\mathrm{Sym}^{m-1}(V)\otimes V)/\,\mathrm{Sym}^m(V)\to\mathrm{Sym}^{m-2}(V)\otimes\Lambda^2(V).$$

The first relation between $\mathscr{G}_m(F)$ and the polylogarithm function is that functional equations of $\mathrm{Li}_m$ modulo lower order polylogarithms are given by combinations $\xi=\sum n_i[x_i]$ of rational functions $x_i=x_i(t)$ satisfying $(*_m)$. Since $\mathrm{Li}_m$ cannot be extended to a continuous one-valued function on the whole complex plane, we will work instead with the modified function

$$P_m(x)=\mathfrak{R}_m\left(\sum_{j=0}^{m-1}\frac{2^jB_j}{j!}(\log|x|)^j\,\mathrm{Li}_{m-j}(x)\right);$$

here $\mathfrak{R}_m$ denotes $\mathfrak{I}$ or $\mathfrak{R}$ according as $m$ is even or odd and $B_j$ the $j$th Bernoulli number ($B_0=1$, $B_1=-\frac12$, $B_2=\frac16$, $B_3=0,\ldots$). This function is continuous and real-valued on $\mathbf{P}^1(\mathbf{C})=\mathbf{C}\cup\{\infty\}$ and is real-analytic except at $0$, $1$, and $\infty$. (The function $P_2(x)$ is just the Bloch-Wigner function $D(x)$ defined in 1.5.1 of this book.) Using $P_m$ instead of $\mathrm{Li}_m$ also has the advantage that the lower order terms in all functional equations cancel, so that $\sum_i n_iP_m(x_i(t))=$ constant for any good combination $\sum n_i[x_i(t)]$. The proof, which is not difficult, is given in §7 of [Z] and will not be repeated here. As simple examples, we note that

$$\beta_m\left(\frac1x\right)=(-[x])^{m-1}\otimes([-1]+[1-x]-[x])=(-1)^{m-1}\beta_m(x)$$

for any $x\in F$, since $[-1]$ is a torsion element of $F^\times$ and $[x]^{m-1}\otimes[x]$ belongs to $\mathrm{Sym}^m(F^\times)$, and similarly

$$\beta_m(x^2)=(2[x])^{m-1}\otimes([1-x]+[1+x])=2^{m-1}(\beta_m(x)+\beta_m(-x))$$

or more generally $\beta_m(x^N)=N^{m-1}\sum_{\zeta^N=1}\beta_m(\zeta x)$ for any $N\geq1$, corresponding to the inversion and distribution relations

$$P_m(1/x)=(-1)^{m-1}P_m(x),\qquad P_m(x^N)=N^{m-1}\sum_{\zeta^N=1}P_m(\zeta x)$$



of the polylogarithm. Many other examples will be discussed in §7.

We now return to number fields and the definition of $\mathscr{A}_m(F)$. If $\phi$ is any homomorphism from $F^\times$ to $\mathbf{Z}$ (e.g., $\phi=v_p$ for $F=\mathbf{Q}$, where $p$ is a prime number and $v_p(x)$ denotes the power of $p$ contained in a rational number $x$), then for $m>2$ the map $\iota_\phi:\mathscr{F}_F\to\mathscr{F}_F$ sending $[x]$ to $\phi(x)[x]$ sends $\mathscr{G}_m(F)$ to $\mathscr{G}_{m-1}(F)$ (because $\beta_{m-1}\circ\iota_\phi=(\phi\otimes\mathrm{id}^{\otimes(m-1)})\circ\beta_m$ under the natural identification of $\mathbf{Z}\otimes(F^\times)^{\otimes(m-1)}$ with $(F^\times)^{\otimes(m-1)}$). If we have $r$ homomorphisms $\phi_i$ ($1\leq i\leq r$), then for $m\geq r+2$ the composite map $\iota_{\phi_1}\circ\cdots\circ\iota_{\phi_r}:\mathscr{G}_m(F)\to\mathscr{G}_{m-r}(F)$ is independent of the order of the $\phi_i$ and will be denoted simply by $\iota_{\phi_1}\cdots\iota_{\phi_r}$. The elements $\iota_{\phi_1}\cdots\iota_{\phi_r}(\xi)\in\mathscr{G}_{m-r}(F)$ ($1\leq r\leq m-2$) for an element $\xi$ of $\mathscr{G}_m(F)$ will be called the good combinations *associated to* $\xi$. Up to linear combinations there are only finitely many of them, since only the restrictions of the $\phi_i$ to the subgroup of $F^\times$ generated by the elements occurring in $\xi$ are important and we can let $\phi_i$ run over a basis of the group of homomorphisms from this group to $\mathbf{Z}$.

The definition of $\mathscr{A}_m(F)$ is now as follows. We start with $\mathscr{A}_2(F)=\mathscr{G}_2(F)$. For $\xi=\sum n_i[x_i]\in\mathscr{A}_2(F)$ and each embedding $\sigma:F\hookrightarrow\mathbf{C}$, we consider $P_2(\xi^\sigma)$, where $\xi^\sigma$ denotes $\sum n_i[x_i^\sigma]$ and $P_2(\xi^\sigma)$ denotes $\sum n_iP_2(x_i^\sigma)$ (we shall use this abbreviated notation for the value of functions on combinations of arguments from now on). There are only $r_2$ essentially distinct such values, since the relation $P_2(\bar x)=-P_2(x)$ for $x\in\mathbf{C}$ implies that $P_2(\xi)$ vanishes for the real embeddings of $F$ and is the same up to sign for each conjugate pair of complex embeddings. Thus, picking one of each conjugate pair of nonreal embeddings, we get a map $\vec{P}_2$ from $\mathscr{A}_2(F)$ to $\mathbf{R}^{r_2}$ by sending $\xi$ to $\{P_2(\xi^\sigma)\}_\sigma$. The image $\vec{P}_2(\mathscr{A}_2(F))$ is contained in a *lattice* (= discrete subgroup of maximal rank) $\mathscr{R}_2=\mathscr{R}_2(F)\subset\mathbf{R}^{r_2}$. Therefore, if $\mathscr{A}_2(F)$ has rank $>r_2$ (and we will see later that its rank is actually infinite), then the kernel $\mathscr{C}_2(F)$ of $\vec{P}_2$ contains nonzero elements and we get nontrivial relations among dilogarithms of complex algebraic arguments. Furthermore, if $\xi\in\mathscr{C}_2(F)$ and $\sigma:F\hookrightarrow\mathbf{R}$ is a real embedding of $F$, then $L_2(\xi^\sigma)$, where $L_2(x)=\mathrm{Li}_2(x)+\frac12\log|x|\log|1-x|$ ($x\in\mathbf{R}$) is the Rogers dilogarithm function (cf. 1.2.3), is a rational multiple of $\pi^2$, so we also get relations among dilogarithms of real algebraic arguments. For $m=3$ we define $\mathscr{A}_3(F)$ as the subset of $\mathscr{G}_3(F)$ consisting of all good combinations $\xi$ for which $\iota_\phi(\xi)\in\mathscr{A}_2(F)$ belongs to $\mathscr{C}_2(F)$ for every homomorphism $\phi:F^\times\to\mathbf{Z}$. This condition can be checked by numerical computation since the element $P_2(\iota_\phi(\xi))$ belongs to the discrete group $\mathscr{R}_2$ and hence is recognizably zero or nonzero. For $\xi\in\mathscr{A}_3(F)$ we consider the collection of all $P_3(\xi^\sigma)$ where (since $P_3(x)$ is invariant rather than anti-invariant under complex conjugation of $x$) $\sigma$ now ranges over the real and half the complex embeddings of $F$. This defines a map $\vec{P}_3:\mathscr{A}_3(F)\to\mathbf{R}^{r_1+r_2}$ whose image is again contained in a lattice $\mathscr{R}_3\subset\mathbf{R}^{r_1+r_2}$, so again we get a nontrivial kernel $\mathscr{C}_3(F)$ and nontrivial relations among polylogarithm values. Similarly, for $m > 2$ we consider the map $\vec{P}_m : \mathscr{F}_F \to \mathbb{R}^{n_\mp}$, where, $\pm 1 = (-1)^m$, $n_+$ and $n_-$ denote $r_1 + r_2$ and $r_2$ respectively, and $\vec{P}_m(\xi)$ is the collection of $P_m(\xi^\sigma)$ with $\sigma$ ranging over half the complex embeddings ($m$ even) or all the real embeddings and half the complex embeddings ($m$ odd). The subgroup $\mathscr{A}_m(F) \subseteq \mathscr{G}_m(F) \subset \mathscr{F}_F$ will be defined inductively in such a way that its image under $\vec{P}_m$ is contained in a certain lattice $\mathscr{R}_m = \mathscr{R}_m(F) \subset \mathbb{R}^{n_\mp}$, in which case the kernel

$$\mathscr{C}_m(F) = \{\xi \in \mathscr{A}_m(F) \mid P_m(\xi^\sigma) = 0 \ \text{ for all } \sigma : F \hookrightarrow \mathbb{C}\}$$

is a subgroup of corank at most $n_\mp$. Assuming that $\mathscr{A}_{m'}(F)$ and $\mathscr{C}_{m'}(F)$ have been defined for $m' < m$, we define $\mathscr{A}_m(F)$ by

$$\mathscr{A}_m(F) = \{\xi \in \mathscr{G}_m(F) \mid \iota_\phi(\xi) \in \mathscr{C}_{m-1}(F) \text{ for all } \phi \in \operatorname{Hom}(F^\times, \mathbb{Z})\}.$$

Explicitly, this means that starting with $\xi \in \mathscr{G}_m(F)$, we verify $x \in \mathscr{A}_m(F)$ by checking first that all associated elements $\iota_{\phi_1} \cdots \iota_{\phi_{m-2}}(\xi) \in \mathscr{A}_2(F)$ map to 0 under $\vec{P}_2$, then that all associated elements $\iota_{\phi_1} \cdots \iota_{\phi_{m-3}}(\xi)$ (which are then known to belong to $\mathscr{A}_3$) map to 0 under $\vec{P}_3$, and so on in succession for all associated elements $\iota_{\phi_1} \cdots \iota_{\phi_r}(\xi)$ for $r = m-2,\ m-3, \ldots, 2, 1$.

We do not give any further explanation of these definitions here; the examples in the following sections should make it clear how they work in practice, while the background and motivating ideas are explained in more detail in **[Z]**. However, we should say a few words about the relationship to $K$-theory and about what has been proved so far and by whom. The former can be understood without needing to know the definition of algebraic $K$-groups $K_n(F)$. One need only know that there are such groups and that by a result of Borel **[B]**, the even-index groups are finite while the odd-index groups $K_{2m-1}(F)$ $(m > 1)$ are mapped isomorphically (up to torsion) onto a lattice in $\mathbb{R}^{n_\mp}$ by the so-called "regulator mapping." The lattice $\mathscr{R}_m(F)$ we have been speaking about is the image of the regulator mapping (or more precisely, any lattice commensurable with this image, since everything we are saying is known only up to groups of finite order). Borel also showed that the covolume of $\mathscr{R}_m \subset \mathbb{R}^{n_\mp}$ is a rational multiple of $|D_F|^{1/2}\zeta_F(m)/\pi^{mn_\pm}$. Hence if the map $\vec{P}_m : \mathscr{A}_m(F) \to \mathscr{R}_m$ is surjective, then $\zeta_F(m)$ can be expressed in terms of polylogarithms of order $m$ with arguments in' $F$.

The relationship between the dilogarithm and the regulator lattice $\mathscr{R}_2$ was found by Bloch (for this and later developments concerning $K_3$ and dilogarithms, see Suslin's ICM talk **[S]**). Here the map $\vec{P}_2 : \mathscr{A}_2 \to \mathscr{R}_2$ is surjective (after tensoring with $\mathbb{Q}$) and its kernel is the subgroup of $\mathscr{A}_2(F)$ spanned by the 5-term functional equation of the dilogarithm. The above picture for general $m$ was formulated conjecturally in **[Z]**. The case $m = 3$ was proved completely by Goncharov **[G]**, who showed that the map $P_3 : \mathscr{A}_3(F) \to \mathbb{R}^{n_+}$ maps not only into, but onto the regulator lattice $\mathscr{R}_3$ (at least after tensoring with $\mathbb{Q}$) and also gave a complete description of the kernel in terms of a new 22-term functional equation for the trilogarithm. The fact that $\vec{P}_m$ maps $\mathscr{A}_m(F)$ into the Borel regulator lattice for arbitrary $m$ was proved by Deligne and by Beilinson for cyclotomic fields and by Beilinson in the general case **[BD]**. For $m > 3$ the surjectivity is not known except in the cyclotomic case, so that the desired corollary that $\zeta_F(m)$ can be expressed in terms of special values of polylogarithms at arguments in $F$ is still a conjecture in general. We also conjecture that $\mathscr{C}_m(F)$ can be described completely in terms of the functional equations of the $m$th polylogarithm, but this also is known only for $m = 2$ and 3.



Finally, we should mention that the generalized Rogers function $L_m(x)$ $(x \in \mathbb{R})$ defined in 3.3.2 agrees with $P_m$ for $m = 3$ and $x$ real, while for $m > 3$ odd and $x$ real it is equal to $\sum_{r=0}^{m-2} (\log|x|)^r P_{m-r}(x)/(r+1)!$ (only the terms with $r$ even in this sum contribute, since $P_{m'}(x) = 0$ for $m'$ even and $x$ real). Thus for elements $\xi \in \mathscr{A}_m(F)$ ($m$ odd) and real embeddings $\sigma$ of $F$, the elements $P_m(\xi^\sigma)$ and $L_m(\xi^\sigma)$ agree (because each term $\sum n_i (\log|x_i|)^r P_{m-r}(x_i)$, $r > 0$, vanishes by virtue of the inductive definition of $\mathscr{A}_m$), and therefore our relations among values of $P_m(x)$ at real algebraic arguments can be reinterpreted as relations among the same values of $L_m(x)$. For $m$ even, $P_m(x)$ vanishes identically for $x$ real, but if we take an element $\xi \in \mathscr{A}_m(F)$ for which $P_m(\xi^\sigma) = 0$ for all complex embeddings $\sigma$, then it is apparently a consequence of the results in **[BD]** that each $L_m(\xi^\sigma)$ $(\sigma : F \hookrightarrow \mathbb{R})$ is a rational multiple of $\pi^m$.

**2. Examples of dilogarithm relations.** For $m = 2$ there is no distinction between $\mathscr{G}_m$ and $\mathscr{A}_m$, so we just have to look for good combinations at level 2, i.e., combinations $\xi = \sum n_i [x_i]$ satisfying (*). For such $\xi$, $P_2(\xi) = D(\xi)$ belongs to the $r_2$-dimensional lattice $\mathscr{R}_2$. Thus if we find $r_2$ elements $\xi$ with linearly independent images, we obtain $\zeta_F(2)$ as $\pi^{2(r_1+r_2)}/\sqrt{|D_F|}$ times a rational number times an $r_2 \times r_2$ determinant of integral linear combinations of dilogarithm values, and if we have more than $r_2$ good combinations $\xi$, then they will have linearly dependent images in the lattice and we obtain linear relations over $\mathbb{Q}$ among the values $D(x_i)$ (resp. $L_2(x_i^\sigma)$ modulo $\pi^2$ for the real embeddings $\sigma$ of $F$, where $L_2(x)$ is the Rogers dilogarithm).

As an example over $\mathbb{Q}$, take elements $x \in \mathbb{Q}$ such that $x$ and $1 - x$ contain no primes except 2, 3, 5, and 7. There are exactly 375 such $x$, forming 63 orbits under the group generated by $x \mapsto 1/x$, $x \mapsto 1 - x$. For each one, $\beta_2(x) = [x] \wedge [1 - x]$ is a linear combination of the six elements

$$[2]\wedge[3], \quad [2]\wedge[5], \quad [2]\wedge[7], \quad [3]\wedge[5], \quad [3]\wedge[7], \quad [5]\wedge[7],$$

so we get 57 essentially different linearly independent linear combinations $\xi$ belonging to $\mathscr{A}_2(\mathbb{Q})$, for each of which $L_2(\xi)$ is a rational multiple of $\pi^2$. For instance, if we pick at random the seven elements $\frac{1}{3}$, $-\frac{1}{6}$, $\frac{2}{7}$, $\frac{1}{8}$, $\frac{1}{9}$, $\frac{1}{21}$, and $\frac{1}{28}$ (any other seven would do just as well), then since $7 > 6$ we must find at least one nontrivial element of $\mathscr{A}_2(\mathbf{Q})$. In fact, calculating

$$\beta_2(\tfrac{1}{3}) = [\tfrac{1}{3}] \wedge [\tfrac{2}{3}] = [2] \wedge [3], \quad \beta_2(-\tfrac{1}{6}) = [-\tfrac{1}{6}] \wedge [\tfrac{7}{6}] = -[2] \wedge [7] - [3] \wedge [7],$$

etc., we find that two linearly independent combinations of the 7 elements in question are annihilated by $\beta_2$, namely $6[\frac{1}{3}]-[\frac{1}{9}]$ and $3[-\frac{1}{6}]-[\frac{1}{8}]+[\frac{1}{9}]+[\frac{1}{28}]$. The image of each of these under the Rogers dilogarithm must be a rational multiple of $\pi^2$, and indeed,

$$6L_2(\tfrac{1}{3}) - L_2(\tfrac{1}{9}) = \frac{\pi^2}{3}, \qquad 3L_2(-\tfrac{1}{6}) - L_2(\tfrac{1}{8}) + L_2(\tfrac{1}{9}) + L_2(\tfrac{1}{28}) = -\frac{\pi^2}{12}.$$

Taking instead a field with $r_2 > 0$, we let $F = \mathbf{Q}(\sqrt{-7})$. Then for the two elements $x_1 = (1+\sqrt{-7})/2$ and $x_2 = (-1+\sqrt{-7})/4$, both $1 - x_1$ and $1 - x_2$ belong to the group generated by $-1$, $x_1$, and $x_2$, so both $\beta(x_1)$ and $\beta(x_2)$ are multiples of $[x_1] \wedge [x_2]$. The multiples turn out to be 1 and $-2$, respectively, so $2[x_1] + [x_2]$ belongs to $\mathscr{A}_2(F)$. Hence $2D(x_1) + D(x_2)$ must be a rational multiple of $\zeta_F(2)/\pi^2\sqrt{7}$, and indeed one finds

$$2D\left(\frac{1+\sqrt{-7}}{2}\right) + D\left(\frac{-1+\sqrt{-7}}{4}\right) = \frac{21\sqrt{7}}{4\pi^2}\zeta_{\mathbf{Q}(\sqrt{-7})}(2).$$

As a third example, we consider the number $u = \sqrt{\rho}$, where $\rho = (\sqrt{5}-1)/2$ (cf. Chapter 5, §5.2.1). Here the relation (5.4) implies that $\xi_0 = [u^6] - 4[u^3] + 6[u]$ is a good element of $\mathscr{F}_F$, $F = \mathbf{Q}(u)$, since

$$\begin{aligned}\beta_2(\xi_0) &= [u^6] \wedge [1-u^6] - 4[u^3] \wedge [1-u^3] + 6[u] \wedge [1-u] \\ &= 6[u] \wedge [(1-u^6)(1-u^3)^{-2}(1-u)] = 12[u] \wedge [u] = 0.\end{aligned}$$

However, $\xi_0$ does not belong to the subgroup $\mathscr{C}_2(F)$ of $\mathscr{A}_2(F)$, because, denoting by $\sigma$ the embedding of $F$ into $\mathbf{C}$ which sends $u$ to $u^\sigma = i/u$, we have $D(\xi_0^\sigma) = 8.6124152\ldots \neq 0$. Therefore we should not expect $L_2(\xi_0)/\pi^2$, which is the number given by (5.6), to be rational. What we *do* expect is that $D(\xi_0^\sigma)$ is a rational multiple of $\pi^{-6}\zeta_F(2)$, since for the field $F$ we have $r_1 = 2$, $r_2 = 1$, $D_F = 400$. To compute numerically, we decompose $\zeta_F(s)$ as $\zeta_{F_1}(s)L_{F_1}(s)$ where $\zeta_{F_1}(s)$ is the Dedekind zeta function of $F_1 = \mathbf{Q}(\sqrt{5})$ and

$$L_{F_1}(s) = \frac{1}{1+5^{-s}} \prod_{p \equiv 11,13,17,19} \frac{1}{1-p^{-2s}} \prod_{p\equiv 3,7} \frac{1}{1+p^{-2s}} \prod_{p \equiv 1,9} \left(\frac{1}{1+\varepsilon_p p^{-s}}\right)^2$$

(here the congruences on primes $p$ are modulo 20 and $\varepsilon_p = (\rho/p) = ((2i+1)/p)$ for $p \equiv 1$ or 9 mod 20, where $\rho$ and $i$ are solutions of $\rho^2 + \rho - 1 \equiv i^2 + 1 \equiv 0 \pmod{p}$). We have $\zeta_{F_1}(2) = 20\pi^4/75\sqrt{5}$ and (computing numerically with the Euler product, using all primes up to 2500) $20^{3/2}\pi^{-2}L_{F_1}(2) \approx 8.612406$, agreeing with $D(\xi_0^\sigma)$ to the accuracy of the computation. If we want to find elements of $\mathscr{C}_2(F)$ and hence relations among



the Rogers dilogarithms of elements of $F$, then we must look at combinations involving other numbers of $F$ than just powers of $u$, i.e., it is not enough to consider ladders only. For instance, the group of units of $F$ is generated (up to torsion) by $u$ and $v = 1-u$, and we can look at elements $x \in F$ for which both $x$ and $1-x$ belong to this group. Up to equivalence by the group generated by $x \mapsto 1/x$ and $x \mapsto 1-x$, under which $D(x)$ is invariant up to sign, there are five such elements, namely $x = u$, $-u$, $u^2$, $u^2v$, and $v^2/u^3$. For these elements $1-x$ equals $v$, $u^4/v$, $u^4$, $u^7/v$, and $v/u^6$, respectively, so $\beta_2[x] = [x] \wedge [1-x]$ equals $[u] \wedge [v]$ times $1, -1, 0, -9$, and 9, respectively. We therefore get four linearly independent elements $[u] + [-u]$, $[u^2]$, $[u^2v] + 9[u]$, and $[u^2v] + [v^2/u^3]$ belonging to $\mathscr{A}_2(F)$. The first two of these are proportional by the duplication formula and uninteresting because they reduce to relations from the smaller field $F_1$. Computing numerically, we find that $D((u^2v)^\sigma) + 9D(u^\sigma)$ equals $D(\xi_0^\sigma)$ but that $D((u^2v)^\sigma) + D((v^2/u^3)^\sigma)$ vanishes. Hence $[u^2v]+[v^2/u^3]$ and $[u^2v]+9[u]-\xi_0$ belong to $\mathscr{C}_2(F)$ and should map under the Rogers dilogarithm to rational multiples of $\pi^2$, and indeed, we have

$$L_2(u^2v) + L_2(v^2/u^3) = \tfrac{1}{20}\pi^2, \qquad L_2(u^2v) + 9L_2(u) - L_2(\xi_0) = \tfrac{41}{60}\pi^2.$$

**3. Examples for higher-order polylogarithms.** Many examples of relations among trilogarithms of algebraic arguments are given in §§3–5 of [Z] as motivation for the form of the conjectures explained in §1. Here we give instead two examples for the field $F = \mathbf{Q}$ (also both taken from [Z]). The first illustrates the necessity of the extra condition in the definition of $\mathscr{A}_m$ as opposed to $\mathscr{G}_m$, while the second gives an example of a relation for heptalogarithms, beyond the range of known functional equations.

For the first example, we proceed as in the example for $m = 2$ in §3, but using only the primes 2 and 3 (if we used 2, 3, 5, and 7 again, we would get many more examples). There are 21 numbers $x \in \mathbf{Q}$ for which $x$ and $1-x$ contain only the prime factors 2 and 3, but this number is cut down to 11 if we do not take both $x$ and $x^{-1}$ (for which the values of all polylogarithms are the same up to sign), and further cut down to 8 if we eliminate the numbers $-1$, $\frac{1}{4}$, and $\frac{1}{9}$ for which the polylogarithm reduces to simpler values by virtue of the duplication equation. These values are $\frac{1}{2}$, $\frac{1}{3}$, $\frac{2}{3}$, $-\frac{1}{2}$, $\frac{3}{4}$, $-\frac{1}{3}$, $\frac{8}{9}$, and $-\frac{1}{8}$. The image of each of them under $\beta_5$ belongs to a space of dimension 4 (this is the dimension of $\mathrm{Sym}^4(V) \otimes V/\mathrm{Sym}^5(V)$ for $V$ the 2-dimensional subspace of $\mathbf{Q}^\times$ spanned by 2 and 3), so we must have at least four independent elements $\xi \in \mathrm{Ker}(\beta_5)$, and indeed there are exactly four, namely $[\frac{1}{2}]$, $[-\frac{1}{3}]-2[\frac{1}{3}]$, $[-\frac{1}{8}]-162[-\frac{1}{2}]$, and $[\frac{8}{9}]-9[\frac{3}{4}]-36[\frac{2}{3}]-18[-\frac{1}{2}]-6[\frac{1}{3}]$. This gives a 4-dimensional subspace of $\mathscr{G}_5(\mathbf{Q})$. However, to get elements of $\mathscr{A}_5(\mathbf{Q})$, we need three further conditions, namely that the images of $\xi$ in $\mathscr{A}_3(\mathbf{Q})$ under the three maps sending $[x]$ to $v_2(x)^2[x]$, $v_2(x)v_3(x)[x]$, and $v_3(x)^2[x]$ all map to zero under the trilogarithm map $\vec{P}_3\colon \mathscr{A}_3(\mathbf{Q}) \to \mathbf{Q}\zeta(3)$. This cuts down the dimension from four to one, the unique surviving relation being $\xi = [-\frac{1}{8}] - 126[\frac{1}{2}] - 162[-\frac{1}{2}]$. This element should therefore map to a rational multiple of $\zeta(5)$ under $P_5$ or $L_5$, and indeed we find that $P_5(\xi) = L_5(\xi) = \frac{403}{16}\zeta(5)$.

To get an example for the heptalogarithm takes more work. If we consider the set of all $x$ such that both $x$ and $1-x$ contain only the first $s$ primes for some $s$ (as we did for $m=2$ with $s=4$ and for $m=5$ with $s=2$), then the number of conditions we have to satisfy is so large that the first value of $s$ for which there are enough $x$ to give a nontrivial element of $\mathscr{A}_7(\mathbf{Q})$ is 8 (i.e., use all primes less than 20), for which there are 10946 elements $x$ (up to inversion) and "only" 10662 conditions to be satisfied (cf. [Z], §10A). This system of equations is far too large to solve numerically. Instead we consider $x \in \mathbf{Q}$ for which $x \in \langle\{2, 3\}\rangle$ and $1 - x \in \langle\{2, 3, 5, 7\}\rangle$. There are 29 such $x$ (taking only one of each pair $x$, $1/x$ and omitting squares as before). On the other hand, there are 28 conditions to be fulfilled: first 20 to get good combinations at level 7 (if $V$ and $W$ are the 2- and 4-dimensional subspaces of $\mathbf{Q}^\times \otimes_{\mathbf{Z}} \mathbf{Q}$ generated by $\{2, 3\}$ and $\{2, 3, 5, 7\}$, respectively, the number of conditions is $\dim(\mathrm{Sym}^6(V) \otimes W/\mathrm{Sym}^7(V)) = \binom{7}{6}\cdot 4 - \binom{8}{7})$, and then a further 8 to ensure that each of the combinations $\iota_{v_2}^j \iota_{v_3}^{r-j}(\xi)$ ($r=4$ or $2$, $0 \le j \le r$) maps to 0 in the one-dimensional lattice $\mathscr{R}_{7-r}$. Since 29 is bigger than 28, we must find a solution of this system of equations i.e., a combination $\xi \in \mathscr{A}_7(\mathbf{Q})$. This solution turns out to be unique up to multiplication by a constant and is given by $\xi = \sum n_i[x_i] \in \mathscr{A}_7(\mathbf{Q})$, where the $n_i$ (normalized for convenience to be in $\frac{1}{6}\mathbf{Z}$ rather than $\mathbf{Z}$) and $x_i$ are given by the following table:

| $n_i$ | $x_i$ | $n_i$ | $x_i$ | $n_i$ | $x_i$ |
|---|---|---|---|---|---|
| −25111753072 | 1/3 | −284585110 | 1/8 | 1765911 | 3/128 |
| −27461584367 | −1/3 | 470985412/3 | −1/8 | 478706760 | 2/9 |
| −171330250 | −1/9 | 38987641 | −1/24 | −66158750 | 8/9 |
| 57577037 | −1/27 | 17015061/2 | −1/48 | 15912813 | −9/16 |
| −151540388696 | 1/2 | −11528187258 | 2/3 | 23786119 | 2/27 |
| −136446322032 | −1/2 | −6563312469 | −2/3 | −2879429 | −8/27 |
| −2209899405 | 1/6 | 2802854628 | 3/4 | 2585366 | 27/32 |
| −2199243270 | −1/6 | −751304106 | −3/4 | 111363 | 32/81 |
| 43524 | −1/4374 | −785318380 | 3/8 | −2372265 | −2/243 |
| −7089743800 | −1/4 | 11883921 | −3/32 | | |

The theory now predicts that $\sum n_i P_7(x_i)$ should be a rational multiple of $\zeta(7)$, and indeed, computing numerically we find that it equals $-\frac{1020149599795}{96}\zeta(7)$ to high precision.

REMARKS. For computing numerical examples in number fields it is useful to have a simple and rapidly convergent formula for calculating $Li_m(z)$. For $z$ small (say $|z| < 1/2$), one can of course simply use the defining series of the polylogarithm, and for $z$ large the functional equation relating the polylogarithms of $z$ and $1/z$. For $z$ near the unit circle, a convenient (and pretty) formula which was noticed by Henri Cohen and myself is

$$Li_m(e^x) = \sum_{n=0}^{\infty} \zeta(m-n)\frac{x^n}{n!},$$

where the meaningless term $\zeta(1)$ is to be replaced by $1 + \frac{1}{2} + \cdots + \frac{1}{m-1} - \log(-x)$. This formula, easily proved by $m$-fold differentiation, works for all $x$ of absolute value less than $2\pi$ and hence gives $Li_m(z)$ for all $z$ with $.005 < |z| < 230$.



**4. Examples: ladders.** We now come to the subject of Leonard Lewin's ladders (note once again the fascination with the letter "L" which marks this field; cf. [**L**, p. 191]), the source of most of the examples in the book, and show how they fit into the theory sketched so far. Briefly, ladders are the special case of combinations $\xi = \sum n_i[x_i]$ in which all of the $x_i$ are powers of a single number $\alpha$. The advantage of making this restriction is that the conditions needed to make $\xi$ belong to $\mathscr{A}_m$ become much more transparent and easier to check (and to fulfill) than in the general case. The disadvantage, of course, is that it is very difficult and requires great ingenuity to produce examples, whereas providing examples of general elements in $\mathscr{G}_m$ or $\mathscr{A}_m$ is something which can be done in a mechanical manner. In particular, while we can show (§5) that there exist polylogarithmic relations for arbitrarily high orders $m$, it is not at all clear—indeed, rather unlikely—that there are any valid ladders at all with $m$ larger than, say, 20.

Suppose, then, that all of the $x_i$ are powers of a single number $\alpha$. Since $[x]$ and $[1/x]$ are essentially equivalent for all polylogarithmic purposes, we can restrict to positive powers. Then the combinations $\xi$ we are looking for can be written as

$$(1)\qquad \xi = \sum_{j=1}^{\infty} n_j[\alpha^j] \qquad (n_j \in \mathbf{Z},\ n_j = 0 \text{ for all but finitely many } j).$$

(From now on we will abbreviate the condition in brackets as $\{n_j\} \in \bigoplus_{j\ge1}\mathbf{Z}$.) If we compute the image of this under the map $\beta_m$, we find (since $[\alpha^j] = j[\alpha]$ in $F^\times$)

$$\beta_m(\xi) = \sum_{j\ge1} j^{m-1} n_j[\alpha]^{m-1} \otimes [1-\alpha^j] = [a]^{m-1} \otimes \Big[\prod_{j\ge1}(1-\alpha^j)^{j^{m-1}n_j}\Big]$$

This is 0 modulo $\mathrm{Sym}^m(\langle\alpha\rangle)$ if and only if $\prod_j(1-\alpha^j)^{j^{m-1}n_j}$ is (a root of unity times) a power of $\alpha$. In other words, turning things around, given any *cyclotomic relation*

$$(2)\qquad \prod_{j=1}^{\infty}(1-\alpha^j)^{b_j} = (\text{root of }1)\times\alpha^N \qquad \Big(\{b_j\} \in \bigoplus_{j\ge1}\mathbf{Z},\ \cdot N \in \mathbf{Z}\Big)$$

we get good combinations $\sum_{j=1}^{\infty} j^{-m+1} b_j[\alpha^j]$ for *all levels* $m$. (Of course, if we want integral combinations, we must multiply these by $J^{m-1}$, where $J$ is a common multiple of the $j$ with $n_j \ne 0$.) This is the beauty of ladders: they give an immediate construction of elements of $\mathcal{G}_m(F)$ for all $m$ simultaneously. However, to get relations among polylogarithm values, we need combinations in $\mathcal{A}_m$, not $\mathcal{G}_m$. Here, too, the special property that $V = \langle\{x_i\}\rangle$ is 1-dimensional simplifies life considerably, since it means that there is up to scalar factors only one homomorphism $\phi$ from $V$ to $\mathbb{Q}$ (namely, the one sending $\alpha^j$ to $j$) and hence only one way to associate to a good combination $\xi = \sum n_j[\alpha^j]$ of order $m$ good combinations of smaller order $m' = m - r$, namely $\xi' = \iota_\phi^r(\xi) = \sum_j j^r n_j[\alpha^j]$. Thus we get the following inductive picture. Let $\{b_j^{(1)}\}, \dots, \{b_j^{(d)}\}$ be a multiplicatively independent set of cyclotomic relations (2) for the same number $\alpha$. Then for each $m$ we get $d$ linearly independent elements $\xi_m^{(\nu)} = \sum j^{1-m} b_j^{(\nu)}[\alpha^j]$ of $\mathcal{G}_m(F)$, where $F = \mathbb{Q}(\alpha)$. For $m = 2$ each of these elements belongs to $\mathcal{A}_2$ and maps to the $r_2$-dimensional lattice $\mathcal{R}_2$, so we get at least $d - r_2$ dilogarithm relations. Changing our basis for the set of cyclotomic relations we can assume that these are $\xi_2^{(1)}, \dots, \xi_2^{(d-r_2)}$, i.e., for each $\nu \le d - r_2$ we have $\sum j^{-1} b_j^{(\nu)} D((\alpha^\sigma)^j) = 0$ if $\alpha^\sigma$ is complex, $\sum j^{-1} b_j^{(\nu)} L_2((\alpha^\sigma)^j) \in \mathbb{Q}\pi^2$ if $\alpha^\sigma$ is real, and $\xi_3^{(\nu)} = \sum j^{-2} b_j^{(\nu)}[\alpha^j] \in \mathcal{A}_3(F)$. Now the elements $\xi_3^{(1)}, \dots, \xi_3^{(d-r_2)}$ map under $\vec{P}_3$ to the $(r_1 + r_2)$-dimensional lattice $\mathcal{R}_3$, so we get at least $d - r_2 - (r_1 + r_2) = d - n$ $(n = [F : \mathbb{Q}])$ linearly independent relations among the trilogarithms of the $\alpha^j$ and the same number of elements in the next higher group $\mathcal{A}_4(F)$. Continuing in this way, we find that the dimension goes down alternately by $r_2$ and $r_1 + r_2$, so by $n$ every two steps. Hence after approximately $\frac{2d}{n}$ steps the process terminates, unless we are very lucky (i.e., there happen to be more linear relations than is forced by the number of equations to be satisfied, a rather rare occurrence) or $d$ is infinite. In fact the second alternative cannot happen, since one can show that the number

$$d(\alpha) \underset{\text{DEF}}{=} \operatorname{rk}_{\mathbf{Z}} \left\{ \{b_j\} \in \bigoplus_{j \ge 1} \mathbf{Z} \mid \prod_{j=1}^{\infty} (1 - \alpha^j)^{b_j} \in \langle\alpha\rangle \right\}$$

is finite for any algebraic number $\alpha$ which is not a root of unity (of course $d(\alpha) = 0$ if $\alpha$ is transcendental). We call the $d(\alpha)$ numbers $j$ for which a positive power of $1 - \alpha^j$ belongs to the subgroup generated by $\alpha$ and $1 - \alpha^i$ $(i < j)$ the *special exponents* of $\alpha$.

We now illustrate all of this with two examples. The first is the number $\omega$ treated in [AL] and in Chapter 4 of the present book, i.e., the root of $\omega^3 + \omega^2 = 1$. Here $d(\omega)$ is (at least) 12, the corresponding special exponents being 1, 2, 3, 5, 8, 12, 14, 18, 20, 28, 30, and 42. Since $r_1 = r_2 = 1$, we can go up to $m = 8$, the number of linear relations obtained for the $m$th order polylogarithms being 11, 9, 8, 6, 5, 3, 2 for $m = 2, 3, \dots, 8$, respectively. At the 9th level there are no more relations. However, if we include 1 as well as the positive powers of $\omega$ then there is one more relation in odd levels and in particular we get an expression of $\zeta(9)$ as a rational linear combination of



$P_9(\omega^n)$ with $n > 0$. More details of this example can be found in Chapter 4 and in §9C of [Z].

The most spectacular example is the Salem number treated in Chapter 16 (§16.3), namely the solution of $\alpha^{10} + \alpha^9 - \alpha^7 - \alpha^6 - \alpha^5 - \alpha^4 - \alpha^3 + \alpha + 1 = 0$. Here $d(\alpha)$ is (at least, and probably exactly) 71, the special exponents ranging up to $j = 360$. Since $r_2 = 4$ and $r_1 + r_2 = 6$, the number of conditions to be satisfied at an even or odd step is a priori 4 or 6, respectively, but because the two real conjugates of $\alpha$ are inverses of one another and $P_m(x^{-1}) = P_m(x)$ for $m$ odd, there are in fact only 5 independent conditions at odd levels. We therefore get successively (at least) $71, 67, 62, 58, \dots$ elements of $\mathcal{R}_2$, $\mathcal{R}_3, \mathcal{R}_4, \mathcal{R}_5, \dots$. Thus the ladder must reach up to (at least) $m = 16$, with four linearly independent relations $\sum n_j P_{16}(\alpha^j) = 0$ at order 16; that this is indeed true has just been verified numerically by Henri Cohen. The coefficients of the relations found are 70-digit integers, and finding them required 300 digits of precision for the polylogarithm values. We use this same example to explain the procedure for finding the special exponents. The first step is to calculate the norm of $\alpha^j - 1$ ( by multiplying the conjugates numerically) for $j$ up to, say, 1000. It is equal to 1 for the 22 values $j =$ 1, 2, 3, 5, 6, 7, 9, 10, 11, 13, 17, 18, 21, 23, 27, 29, 34, 37, 47, 63, 65, and 74. Since the unit rank is $r_1 + r_2 - 1 = 5$, we lose 4 relations, because $\alpha$, $1 - \alpha$, $1 - \alpha^2$, $1 - \alpha^3$ and $1 - \alpha^5$ are needed to generate a group of units of full rank, but we then get 18 independent muliplicative relations among $\alpha$ and the $1 - \alpha^j$. Now we look at other $j$ for which the norm of $1 - \alpha^j$ is no longer 1 but still factors into small prime factors which have already occurred for previous $j$, and then try to form multiplicative combinations which are units. For instance, the norms of $\alpha^j - 1$ are positive powers of 3 for $j = 4, 8, 12, 16, 20, 24, 36$, and 40; and eliminating $\alpha^4 - 1$, which together with the units generates the group of all elements of $F$ whose norm is a power of 3, we get 7 further relations. Proceeding in this way gives the 71 relations mentioned. (In practice it is more efficient to use the cyclotomic polynomials $\Phi_j(\alpha)$ rather than the numbers $\alpha^j - 1$; for instance, by doing this one gets 66 rather than 22 cases of norm 1.) By the time we get to $j = 1000$, the norms are so huge that it seems clear that $\alpha^j - 1$ will never again be a combination of smaller values, and this could be proved by a finite effort if really required.

It seems very likely that this particular 10th degree Salem number gives the maximum of $d(\alpha)/[\mathbb{Q}(\alpha) : \mathbb{Q}]$, and indeed, quite possibly even of $d(\alpha)$, for all algebraic numbers $\alpha$. (It is the same number which has been conjectured for almost 60 years to give the absolute minimum of the "Mahler measure," defined as the product of the absolute values of the conjugates of an algebraic integer outside the unit circle.) If this is the case, then there are probably no valid ladders at all of order bigger than 16. In any case, the study of the number $d(\alpha)$, motivated by the ladder concept, seems to be a very interesting problem in the field of diophantine approximation.

**5. Existence of relations among polylogarithm values of arbitrarily high order.** Let $S$ be a set of $s$ numbers in $\mathbb{Q}$ and $X(S)$ the set of all $x \in \mathbb{Q}$ such that both $x$ and $1-x$ belong to the multiplicative group generated by the elements of $S$. The number of independent requirements on a combination $\xi = \sum_{x \in X} n_x[x]$ to belong to $\mathscr{A}_m(\mathbb{Q})$ is

$$(m-1)\binom{s+m-2}{m} + \sum_{\substack{0 \le r < m-2 \\ m-r \text{ odd}}} \binom{s+r-1}{r}$$

(the first term is the dimension of $\mathrm{Sym}^{m-1}(V) \otimes V / \mathrm{Sym}^m(V)$ where $V = \langle S \rangle$ is $s$-dimensional; the other terms give the number of successive conditions for the images of $\xi$ under the various maps $\iota_{\phi_{i_1}} \cdots \iota_{\phi_{i_r}}$ to map to 0 in $\mathscr{R}_{m-r} \cong \mathbb{Z}$), which is a polynomial in $s$. Therefore if there are sets $S$ of arbitrarily large cardinality $s$ for which $|X(S)|$ grows more than polynomially with $s$, then it follows that $\mathscr{A}_m(\mathbb{Q})$ has infinite rank and hence that there are infinitely many relations among the values of $L_m(x)$ $(x \in \mathbb{Q})$ for every $m$. The existence of such $S$ is the content of a theorem of Erdös-Stewart-Tijdeman [EST]. A simplified presentation of their proof was given in [Z] and a further simplification (with a slightly weaker bound, but not appealing to any results from analytic number theory) is given here.

THEOREM (ERDÖS-STEWART-TIJDEMAN). *For any $\varepsilon > 0$ there exist sets $S$ of arbitrarily large cardinality $s$ for which $|X(S)| > e^{(2-\varepsilon)\sqrt{s/\log s}}$.*

PROOF. Let $t$ and $u$ be large numbers, to be chosen presently, and let $A$ denote the set of products $p_1 \cdots p_r$ with $r \le u$ and each $p_i$ a prime $\le t$. The cardinality of $A$ is then $\binom{u+\pi(t)}{u}$, which for $\pi(t)$ (number of prime $\le t$) much larger than $u$ equals $\pi(t)^{u+o(u)}/u! = (e\pi(t)/u)^{u+o(u)}$. The number of pairs $(a, c)$ with $a, c \in A$, $c > a$, is therefore equal to $\binom{|A|}{2} = (e\pi(t)/u)^{2u+o(u)}$. For each such pair the difference $b = c - a$ is an integer between 0 and $t^u$, so by the pigeonhole principle there is a number $b$ which is expressible as $c - a$ in at least $N = \binom{|A|}{2}/t^u = (e\pi(t)/u\sqrt{t})^{2u+o(u)}$ ways. We maximize this by choosing $u = \pi(t)/\sqrt{t}$, giving $N = e^{(2+o(1))\pi(t)/\sqrt{t}}$. Now take $S = \{b\} \cup \{p \le t,\ p \text{ prime}\}$, with cardinality $s = 1 + \pi(t) = (1+o(1))t/\log t$; then each representation $b = c - a$ gives a distinct element $x = b/c$ of $X(S)$, so $|X(S)| > N = e^{(2+o(1))\sqrt{s/\log s}}$.

COROLLARY. (1) *$\mathscr{A}_m(\mathbb{Q})$ has infinite rank for all $m$.* (2) *For any $m$, there are infinitely many linearly independent relations over $\mathbb{Q}$ among the values of $L_m(x)$, $x \in \mathbb{Q}$.*

REMARKS. **1.** The original result of Erdös-Stewart-Tijdeman was slightly stronger in that 2 was replaced by 4 in the exponent, and $S$ was required to consist only of primes; for the proof of the stronger statement, see [EST] or [Z].



**2.** In part (1) of the corollary, we could have written "$\mathscr{A}_m(F)$ for any number field $F$" instead of $\mathscr{A}_m(\mathbb{Q})$, since $\mathscr{A}_m(F)$ contains $\mathscr{A}_m(\mathbb{Q})$, but of course the interesting question is whether one gets infinitely many *new* relations on passing from $\mathbb{Q}$ to $F$, and this is not clear without doing more work.

**3.** Of course, part (2) of the corollary is not a corollary of the Erdös-Stewart-Tijdeman theorem alone, but of this theorem together with Beilinson's deep result on the relation between the polylogarithm and regulator mappings.

**6. A conjecture on linear independence.** We conjecture that the only linear relations over $\mathbb{Q}$ among polylogarithm values at algebraic arguments are those which follow from the theory explained in §1. More precisely, this says that if there is a relation

$$(1) \qquad \sum_i n_i P_m(x_i) = 0 \qquad (n_i \in \mathbb{Z},\ x_i \in \overline{\mathbb{Q}}),$$

(respectively $\sum n_i L_m(x_i) = 0$ with $n_i \in \mathbb{Z}$ and $x_i \in \overline{\mathbb{Q}} \cap \mathbb{R}$), then

(i) $\sum n_i [x_i]$ satisfies the algebraic relation $(*_m)$, and

(ii) the conjugate equations $\sum n_i P_m(x_i^\sigma) = 0$ $(\sigma \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}))$, as well as the associated equations $\sum n_i \phi_1(x_i) \cdots \phi_r(x_i) P_{m-r}(x_i) = 0$ and *their* conjugates for all $r \le m-2$ and all homomorphisms $\phi_1, \dots, \phi_r : \overline{\mathbb{Q}}^\times \to \mathbb{Q}$, also hold.

Since we further conjecture that the kernel of the maps $\vec{P}_m : \mathscr{A}_m(F) \to \mathbb{R}^{n_\mp}$ is the group $\mathscr{C}_m(F)$ defined by specializing functional equations, we can state the combined conjecture more concisely by saying that the only relations (1) are specializations of functional equations $\sum n_i P_m(\phi_i(t)) = 0$ $(\phi(t) \in \mathbb{Q}(t))$ to arguments $t \in \overline{\mathbb{Q}}$. (This includes conditions (i) and (ii) since the arguments of functional equations of $P_m$ automatically satisfy $(*_m)$ and since replacing $t \in \overline{\mathbb{Q}}$ by $t^\sigma$ replaces each value $x_i = \phi_i(t)$ by $x_i^\sigma$.) This conjecture, which is discussed in more detail in §10 of [Z], contains as a special case Milnor's conjecture [M] that the only linear relations over $\mathbb{Q}$ of the Clausen function $\mathrm{Cl}_2(\theta) = \sum_{n=1}^\infty \sin(n\theta)/n^2$ at arguments $\theta \in \mathbb{Q}\pi$ are those arising from the distribution relations

$$\sum_{m=1}^{|N|} \mathrm{Cl}_2\left(\theta + \frac{2m\pi}{N}\right) = \frac{1}{N}\mathrm{Cl}_2(N\theta) \qquad (0 \ne N \in \mathbb{Z}).$$

The only evidence in its support is its naturalness and internal consistency and the fact that the many known examples of algebraic relations among polylogarithms all conform to it. On the theoretical level nothing is known; so far as I know, for instance, one cannot prove that there is even a single pair of values $D(x)$, $D(y)$ with $x$ and $y$ algebraic which are linearly independent over $\mathbb{Q}$, or even that there is a single value of $D(x)$ which is not a rational number! The conjecture is also particularly daring because we know from §5 that there are linear dependences among values of the $m$th polylogarithm function for arbitrarily large values of $m$, but for $m > 7$ do not know whether these functions have any nontrivial functional equations at all.

**7. Functional equations.** The basic fact here, already mentioned in §1, is that the only requirement for a functional equation is the algebraic condition $(*_m)$, i.e., that for any combination $\sum n_i[x_i]$ of functions (of one or several variables $t$) $x_i(t)$ which satisfies $(*_m)$, the corresponding sum of polylogarithms $\sum n_i P_m(x_i(t))$ is independent of $t$. In this section we discuss how this criterion can be used to check functional equations and to find new ones in an algorithmic way. We will express all functional equations in terms of $P_m$ rather than $\mathrm{Li}_m$ in order to have well-defined values for complex arguments and to avoid lower-order terms in the equations; for real values of the arguments, the function $L_m$ would do just as well because of the relation between $L_m$ and $P_m$ mentioned at the end of §1. A striking property of many of the functional equations is their high level of symmetry; we will emphasize this aspect in our discussion.

EXAMPLE 1. THE 5-TERM RELATION FOR THE DILOGARITHM. The basic functional equation of the dilogarithm function is the 5-term relation (cf. Chapter 1, §1.2), which in our notation says that $D(R_5(x, y)) = 0$ for any $x$ and $y$, where $R_5(x, y)$ denotes the formal linear combination

$$(1)\qquad R_5(x, y) = [x] + [y] + \left[\frac{1-x}{1-xy}\right] + [1-xy] + \left[\frac{1-y}{1-xy}\right]$$

of elements in $\mathbf{Q}(x, y)$. (The other forms given by Spence, Hill, Abel, and Kummer are equivalent to this by the 1-variable relations of Chapter 1, §1.1.) To prove it in our language, we have to check that $R_5(x, y)$ is in the kernel of the map $\beta_2\colon [z] \mapsto [z] \wedge [1-z]$. We calculate

$$\begin{aligned}\beta_2(R_5(x, y)) &= [x]\wedge[1-x] + [y]\wedge[1-y]\\ &\quad + ([1-x]-[1-xy])\wedge([x]+[1-y]-[1-xy])\\ &\quad + [1-xy]\wedge([x]+[y])\\ &\quad + ([1-y]-[1-xy])\wedge([y]+[1-x]-[1-xy])\\ &= 0.\end{aligned}$$

This is the simplest example of the use of the calculus with wedge products in checking functional equations (and also, as far as the dilogarithm is concerned, the basic example, since it is conjecturally the case that all functional equations of the dilogarithm are consequences of the 5-term relation.)

We use this simple example to give a first illustration of the comments on symmetry made at the beginning of this section. First of all, we can check



that $R_5$ has a cyclic symmetry, i.e.,

$$\begin{aligned}R_5(x, y) &= R_5\left(y, \frac{1-x}{1-xy}\right) = R_5\left(\frac{1-x}{1-xy}, 1-xy\right)\\ &= R_5\left(1-xy, \frac{1-y}{1-xy}\right) = R_5\left(\frac{1-y}{1-xy}, x\right).\end{aligned}$$

However, this is only part of the full symmetry group. The basic invariance property of the dilogarithm is that for any $x$ the six numbers $D(x)$, $D(1/x)$, $D(1-1/x)$, $D(x/(x-1))$, $D(1/(1-x))$, and $D(1-x)$ are equal up to sign. This six-fold symmetry plays a role so often in the following that we introduce the special notation

$$x' \underset{6}{\sim} x \Leftrightarrow x' \in \left\{x, \frac{1}{x}, 1-\frac{1}{x}, \frac{x}{x-1}, \frac{1}{1-x}, 1-x\right\}.$$

A symmetric way to express the $\underset{6}{\sim}$ invariance is to say that the function $D([a, b, c, d])$, where $[a, b, c, d] = (a-c)(b-d)/(a-d)(b-c)$ denotes the cross-ratio of $a, b, c$, and $d$, is up to sign a symmetric function of its four arguments, since changing the order of four numbers replaces their cross-ratio $x$ by a number $x' \underset{6}{\sim} x$. More precisely,

$$(2)\qquad D([x_{\pi(1)}, x_{\pi(2)}, x_{\pi(3)}, x_{\pi(4)}]) = \mathrm{sgn}(\pi) D([x_1, x_2, x_3, x_4])$$

for any $x_i \in \mathbf{P}^1(\mathbf{C})$ $(1 \le i \le 4)$ and any $\pi \in \mathfrak{S}_4$, the symmetric group on four letters. To verify this in our terminology we compute

$$\begin{aligned}\beta_2([x_1, x_2, x_3, x_4]) &= \left[\frac{(x_1-x_3)(x_2-x_4)}{(x_1-x_4)(x_2-x_3)}\right] \wedge \left[\frac{(x_1-x_2)(x_3-x_4)}{(x_1-x_4)(x_2-x_3)}\right]\\ &= \frac{1}{2}\sum_{\pi\in\mathfrak{S}_4} \mathrm{sgn}(\pi)[x_{\pi(1)} - x_{\pi(2)}] \wedge [x_{\pi(2)} - x_{\pi(3)}]\end{aligned}$$

(this has only twelve terms rather than twenty-four, because the $\pi$th summand is invariant under $\pi \mapsto \pi\tau$ where $\tau$ is $1 \leftrightarrow 3$). This has the desired invariance property under the action of $\mathfrak{S}_4$, so $D([x_1, x_2, x_3, x_4])$ also does. Now the 5-term relation becomes simply

$$(3)\qquad \frac{1}{24}\sum_{\pi\in\mathfrak{S}_5} \mathrm{sgn}(\pi) D([x_{\pi(1)}, x_{\pi(2)}, x_{\pi(3)}, x_{\pi(4)}]) = 0, \qquad (x_1, \ldots, x_5 \in \mathbf{P}^1(\mathbf{C}))$$

(this has only 5 terms rather than 120 because the $\pi$th summand is unchanged by $\pi \mapsto \pi\tau$, $\tau \in \mathfrak{S}_4$), with a symmetry group of order 120.

EXAMPLE 2. THE 9-TERM RELATION FOR THE DILOGARITHM. As a second example, we consider the 3-variable functional equation (2.40) of Chapter 2, §2.3.2. This equation has an obvious 8-fold symmetry generated by the involutions $x \leftrightarrow y$, $v \leftrightarrow w$, and $(x, y) \leftrightarrow (v, w)$, but in fact has a symmetry group of order 72, a typical nonobvious symmetry being given by

$(x, y, v, w) \mapsto (y, xy/vw, y/w, y/v)$. To make the symmetries obvious, we first define a $3 \times 3$ matrix

$$(4) \qquad Z = (z_{ij})_{i,j=1,2,3} = \begin{pmatrix} x & 1/v & v/x \\ 1/w & y & w/y \\ w/x & v/y & xy/vw \end{pmatrix}.$$

Then the functional equation becomes simply $\sum_{i,j} D(z_{ij}) = 0$. The constraints on the $z_{ij}$ are

$$(5) \qquad \prod_j z_{ij} = 1, \qquad (\forall i), \qquad \prod_i z_{ij} = 1 \qquad (\forall j),$$

and (taking the indices $i$ and $j$ modulo 3)

$$(6) \qquad (1 - z_{i,j})(1 - z_{i+1,j+1}) = (1 - z_{i,j+1}^{-1})(1 - z_{i+1,j}^{-1}) \qquad (\forall i, j).$$

Indeed, (5) is equivalent to the fact that $Z$ can be expressed by (4) for *some* numbers $x, y, v$, and $w$, and (6) for $i = j = 1$ is just the constraint $(1-x)(1-y) = (1-v)(1-w)$. A calculation shows that this equation implies the validity of (6) for the other eight values of the indices $(i, j)$. Thus we have we have the symmetry group of order 72 given by permuting the indices $i$, permuting the indices $j$, and interchanging the roles of $i$ and $j$. However, this is still not satisfactory because we cannot yet "see" why equation (6) for one pair of indices $(i, j)$ implies (in conjunction with (5)) its truth for all $i, j$. To remedy this, we choose new coordinates $t_{ij}$ $(i, j \in \mathbf{Z}/3\mathbf{Z})$ and set $z_{ij} = t_{i,j} t_{i,j+1}^{-1} t_{i+1,j}^{-1} t_{i+1,j+1}$. Then the equations (5) are true identically and each of the nine equations (6) is equivalent to the condition $\det(T) = 0$, where $T = (t_{ij})$. (The set of $3 \times 3$ matrices $T$ with $\det(T) = 0$ is 8-dimensional, but there are only 3 free parameters because replacing $t_{ij}$ by $\lambda_i \mu_j t_{ij}$ for any $\lambda_i$ and $\mu_j$ leaves $z_{ij}$ unchanged.) Finally, we give a parametric solution of $\det(T) = 0$ by setting $t_{ij} = u_i - v_j$. Then the constraints (5) and (6) are automatically satisfied and $z_{ij}$ is simply the cross-ratio of the four numbers $u_i$, $u_{i+1}$, $v_j$, and $v_{j+1}$. Therefore the functional equation can be written

$$\sum_{i,j} D([u_i, u_{i+1}, v_j, v_{j+1}]) = 0 \quad \text{for any } u_i, v_i \in \mathbf{P}^1(\mathbf{C}) \ (i \in \mathbf{Z}/3\mathbf{Z})$$

Now not only the 72-fold symmetry of the equation is obvious (permute the $u_i$'s or the $v_j$'s or interchange $\mathbf{u}$ and $\mathbf{v}$), but also its proof: applying (3) to the 5-tuple $(u_i, u_{i+1}, v_1, v_2, v_3)$ gives

$$\sum_j D([u_i, u_{i+1}, v_j, v_{j+1}]) = D([u_i, v_1, v_2, v_3]) - D([u_{i+1}, v_1, v_2, v_3])$$



for each $i$, and the sum of this over $i$ (mod 3) vanishes. Of course, we can now go back and give a nonmotivated and nonsymmetric proof of the equation in its original form by adding the three 5-term relations $R_5(x, 1/v)$, $R_5(1/w, y)$, and $R_5(w/x, v/y)$ and using the 1-variable functional equations of the dilogarithm to simplify the result.

EXAMPLE 3. THE 9-TERM RELATION FOR THE TRILOGARITHM. As the first example of the verification of a functional equation for a higher-order polylogarithm we consider Kummer's 2-variable equation for the trilogarithm. Again we use $P_3$ rather than $\mathrm{Li}_3$ in order to eliminate lower-order terms. (The function $L_3$ would do just as well if we restricted our attention to real values of the variables.) The functional equation in question (cf. Chapter 3, 3.2.2.3) has the form $P_3(R_9(x, y)) = 2\zeta(3)$, where $R_9(x, y)$ denotes the element

$$(7) \qquad \begin{aligned} R_9(x, y) = 2[x] + 2[y] + 2\left[\frac{x(1-y)}{x-1}\right] + 2\left[\frac{y(1-x)}{y-1}\right] + 2\left[\frac{1-x}{1-y}\right] \\ + 2\left[\frac{x(1-y)}{y(1-x)}\right] - [xy] - \left[\frac{x}{y}\right] - \left[\frac{x(1-y)^2}{y(1-x)^2}\right] \end{aligned}$$

of $\mathbf{Q}(x, y)$. This equation has an obvious symmetry group of order 8 (invert $x$ or $y$ or interchange $x$ and $y$), but in fact has a symmetry group of order 24. We can write it symmetrically in two ways.

*Symmetric form: first version.* Let $H$ be the hypersurface in $\mathbf{P}^3(\mathbf{C})$ consisting of 4-tuples $(a_1 : a_2 : a_3 : a_4)$ with $\sum_i a_i = 0$, with the obvious action of the group $\mathfrak{S}_4$. Let $V$ be the 3-dimensional space generated multiplicatively by all ratios $a_i/a_j$ and $W \supset V$ the 6-dimensional space generated by all quotients $(a_i + a_j)/a_i$, $j \neq i$ (this space has dimension only 6 because up to sign there are only three distinct quantities $a_i + a_j$). Each of the elements $x = -a_i/a_j$ $(j \neq i)$ and $a_i a_j / a_k a_l$ $(i, j, k, l$ distinct) has the property that $x \in V$ and $1 - x \in W$, and up to inversion there are exactly nine of them. The argument $R_9$ of Kummer's functional equation can then be written symmetrically as

$$2\left(\left[-\frac{a_1}{a_2}\right] + 5 \text{ permutations}\right) - \left(\left[\frac{a_1 a_2}{a_3 a_4}\right] + 2 \text{ permutations}\right).$$

(To see this, normalize the common value of $a_1 + a_2$ and $-a_3 - a_4$ to be 1, in which case $H$ is parametrized by two parameters $x = -\frac{a_1}{a_2}$, $y = -\frac{a_3}{a_4}$ as $\left\{\left(\frac{x}{x-1} : \frac{1}{1-x} : \frac{y}{1-y} : \frac{1}{y-1}\right)\right\}$.) The proof of $P_3(R_9) =$ constant (the value of the constant is then found by specializing to $x = 0$, $y = 1$) can be obtained easily by computing $\beta_3(-a_1/a_2)$ and $\beta_3(a_1 a_2 / a_3 a_4)$ and symmetrizing with respect to the group action.

*Symmetric form: second version.* As with the dilogarithm, we interpret the arguments of the trilogarithms as cross-ratios of 4-tuples of points in $\mathbb{P}^1(\mathbb{C})$. However, we no longer have the 6-fold symmetry of the function $D(x)$, but only the twofold symmetry $P_3(x) = P_3(1/x)$. This means that the argument of $P_3(x)$ must be interpreted as a cross-ratio $[a, b, c, d]$ of four points $a, b, c, d \in \mathbb{P}^1(\mathbb{C})$ where the only symmetries allowed are the ones generated by the interchanges of $a$ and $b$, of $c$ and $d$, or of $(a, b)$ and $(c, d)$, i.e., $P_3([a, b, c, d])$ depends only on the unordered pair of unordered pairs $\{\{a, b\}, \{c, d\}\}$. To emphasize this, we write the cross-ratio $x$ as $[a, b; c, d]$ rather than $[a, b, c, d]$, although the definition is the same as before. Now suppose that we have six points $p_\nu \in \mathbb{P}^1(\mathbb{C})$ and that there is an involution $\tau$ on $\mathbb{P}^1(\mathbb{C})$ which interchanges these points in pairs. We consider all cross-ratios $[a, b; c, d]$ of four distinct points $a, b, c, d$ from the set $\{p_\nu\}$ for which $\{a, b\}$ is disjoint from $\{\tau(c), \tau(d)\}$. Using the symmetries of $[a, b; c, d]$ and the invariance of the cross-ratio under automorphisms of $\mathbb{P}^1(\mathbb{C})$, we find that there are nine of these, three of the form $[a, \tau(a); c, \tau(c)]$ and six of the form $[a, \tau(a); c, d]$ with $d \neq \tau(c)$. Then $R_9$ is just twice the sum of the latter six minus the sum of the former three. (Take the $p_\nu$ to be $0, \infty, 1, xy, x$, and $y$ with the involution $\tau\colon t \mapsto xy/t$.) The equivalence of this form of the functional equation to the one just given is seen by taking for $\{p_\nu\}$ the six numbers $a_i + a_j$ $(1 \le i < j \le 4)$ and for $\tau$ the involution $t \mapsto -t$. The symmetry group appears now not as $\mathfrak{S}_4$, but at the semidirect product of $\mathfrak{S}_3$ (permute the three orbits of $\tau$ on $\{p_\nu\}$) with a Klein 4-group (interchanging the two elements within each orbit is an automorphism of order 2, but the product of these three automorphisms is just $\tau$ and has no effect on the cross-ratio).

The functional equations discussed up to now are classical, but the method of verifying them can also be used to discover new equations. The basic desideratum, as in the case of relations among special values, is to find as many $x$ as possible such that all the $x$ and $1-x$ belong to subspaces $V$ and $W$ of $F^\times$ of small dimension. Typically one first chooses the generators of $V$ (certain irreducible polynomials in one or several variables), preferably with a lot of symmetry to reduce the number of independent conditions which have to be checked later, and then looks for many $x_i$ in $V$ for which the prime factors of the elements $1 - x$ are repeated many times; then one computes $\beta_m(x)$ for these $x$'s and some small $m$ and uses linear algebra to solve if possible the system of equations which expresses that a linear combination of them vanishes. This algorithm, which can be carried out by hand in simple cases, can also be programmed, although not easily. This has been carried out by H. Gangl, who in this way found a large number of new functional equations. In the remainder of the appendix we present a few of these.

EXAMPLE 4. FUNCTIONAL EQUATIONS FOR THE TETRALOGARITHM. The simplest of Gangl's functional equations is the 9-term equation



$$
\begin{aligned}
&2\left[P_4(z(1-z)) + P_4\left(-\frac{z}{(z-1)^2}\right) + P_4\left(\frac{z-1}{z^2}\right)\right] \\
&-3\left[P_4\left(\frac{1}{1-z+z^2}\right) + P_4\left(\frac{(1-z)^2}{1-z+z^2}\right) + P_4\left(\frac{z^2}{1-z+z^2}\right)\right] \\
&-6\left[P_4\left(\frac{1+z+z^2}{z(z-1)}\right) + P_4\left(\frac{1-z+z^2}{z}\right) + P_4\left(\frac{1-z+z^2}{1-z}\right)\right] = 0.
\end{aligned}
\tag{8}
$$

This has a 6-fold symmetry $G_9(z') = G_9(z)$ for $z' \underset{6}{\sim} z$ and up to this symmetry has only three rather than nine terms, as shown by the square brackets.

More interesting is the equation

$$
\begin{aligned}
&\left[P_4\left(-\frac{bd}{a}\right) + P_4\left(-\frac{ac}{d}\right)\right. \\
&\qquad \left. + P_4\left(-\frac{ab^3}{c^2d}\right) + P_4\left(-\frac{c^3d}{ab^2}\right) + P_4\left(-\frac{d}{ab^2}\right) + P_4\left(-\frac{a}{c^2d}\right)\right] \\
&+\left[P_4\left(-\frac{ad^2}{b}\right) + P_4\left(-\frac{a^2d}{c}\right)\right] + 2\left[P_4\left(\frac{b}{c^2}\right) + P_4\left(\frac{c}{b^2}\right) + P_4(bc)\right] \\
&+3\left[P_4\left(-\frac{b}{ad}\right) + P_4\left(-\frac{c}{ad}\right) + P_4\left(-\frac{1}{ad}\right)\right] \\
&+3\left[P_4(-cd) + P_4(-ab) + P_4\left(-\frac{d}{c}\right)\right. \\
&\qquad \left. + P_4\left(-\frac{a}{b}\right) + P_4\left(-\frac{ab}{c^2}\right) + P_4\left(-\frac{cd}{b^2}\right)\right] \\
&+6\left[P_4\left(\frac{c}{a}\right) + P_4\left(\frac{b}{d}\right) + P_4\left(\frac{1}{d}\right) + P_4\left(\frac{1}{a}\right) + P_4\left(\frac{b}{cd}\right) + P_4\left(\frac{c}{ab}\right)\right] = 0,
\end{aligned}
$$

in two variables $y$ and $z$, where we have abbreviated

$$a = 1 - z + yz, \quad b = -y, \quad c = y - 1, \quad d = 1 - y + yz.$$

This equation has an obvious symmetry under the involution $P\colon (a, b, c, d) \mapsto (d, c, b, a)$, corresponding to $(y, z) \mapsto (1-y, 1-z)$, and a less obvious one under the involution $Q\colon (a, b, c, d) \mapsto (a, b^{-1}, cb^{-1}, db^{-1})$, corresponding to $(y, z) \mapsto (y^{-1}, -yz)$. These two involutions generate a group of order 12 under which the 26 terms of the functional equations fall into only six orbits (grouped by square brackets in the formula above). One can introduce new variables which make the symmetries obvious. Let $F$ be the function field (over $\mathbb{Q}$) of the variety

$$X = \left\{(t_1 : t_2 : t_3 : t_4 : t_5) \in \mathbb{P}^4 \;\middle|\; \sum_{i=1}^{5} t_i = 0, \sum_{i=1}^{3} t_i^{-1} = 0\right\},$$

with an obvious symmetry of the group $G = \mathfrak{S}_3 \times \mathfrak{S}_2 \subset \mathfrak{S}_5$. We can parametrize $X$ by

$$\begin{aligned}&(t_1 : t_2 : t_3 : t_4 : t_5)\\ &\quad = (-y : y-1 : y(1-y) : y - y(1-y)z : (y-1)^2 + y(1-y)z),\end{aligned}$$

(set $y = t_3/t_2$, $z = -(t_1+t_4)/t_3$) and under this identification $G$ is identified with the symmetry group $\langle P, Q\rangle$ of the functional equation. The 26 arguments $x$ of $P_4$ satisfy $x \in V$, $1-x \in W$, where $V$ is the 4-dimensional subspace of $F^\times \otimes_{\mathbb{Z}} \mathbb{Q}$ spanned by the quotients $t_i/t_j$ $(i \neq j)$ and $W$ the 11-dimensional space spanned by $V$ and the elements $(t_i+t_j)/t_i$ $(i \neq j,\ \max\{i, j\} \geq 4)$. Using this description and the symmetry one can easily check that the linear combination of the $x$'s specified by the functional equation is in the kernel of $\beta_4$.

EXAMPLE 5. A 2-VARIABLE FUNCTIONAL EQUATION FOR THE HEXALOGARITHM. Gangl found one-variable equations for the penta- and hexalogarithm which had a similar structure to equation (8) above, namely, the arguments of the polylogarithms involve only the irreducible polynomials $z$, $1-z$, and $1-z+z^2$ and the whole equation is unchanged if $z$ is replaced by any $z' \underset{6}{\sim} z$. His functional equation for the hexalogarithm, in particular, has 60 terms forming 13 orbits under the 6-fold symmetry and can be written efficiently as

$$\begin{aligned}&3T^-_{\{0,1,5\}} + 4T^+_{\{-5,0,3\}} + 5T^-_{\{-4,3,3\}} - 20T^-_{\{0,1,3\}} - 20T^-_{\{0,0,4\}}\\ &- 60T^-_{\{-3,1,2\}} + 90T^-_{\{-1,1,2\}} + 95T^-_{\{-1,0,3\}} + 120T^+_{\{-2,1,1\}} - 180T^+_{\{-1,1,2\}}\\ &- 360T^+_{\{0,1,1\}} + 540T^-_{\{0,1,1\}} - 544T^+_{\{0,0,2\}} = 0,\end{aligned}$$

where for a set of three integers $\alpha$, $\beta$, $\gamma$ with even sum we have set

$$T^\pm_{\{\alpha,\beta,\gamma\}} = \sum_{\{a,b,c\}=\{\alpha,\beta,\gamma\}} P_6\left(\frac{\pm(-1)^a z^b (1-z)^c}{(1-z+z^2)^{(a+b+c)/2}}\right)$$

(the sum is over all permutations and has six terms in general but only three if two of $\alpha$, $\beta$, and $\gamma$ coincide). Later he found a second functional equation with eighty-nine terms and where the prime factors of the hexalogarithm arguments involved four prime factors $z$, $1-z$, $1+z$, and $1-z+z^2$. Many of the coefficients in these two equations agreed, suggesting that they might be different specializations of the same two-variable equation. After a considerable amount of trial and error, guided by the symmetry, it was found that this is true and that the two-variable functional equation has an 18-fold symmetry:

THEOREM. *Let $y$ and $z$ be variables and denote by $y_i$ $(1 \leq i \leq 6)$ and $w_j$ $(1 \leq j \leq 3)$ the elements $y' \underset{6}{\sim} y$ and $z'(1-z')$, $z' \underset{6}{\sim} z$. Then*

$$\sum_{a,b} n_{a,b} \frac{1}{\mu_{a,b}} \sum_{i=1}^{6}\sum_{j=1}^{3} P_6\left(\left(\frac{w_j}{y_i}\right)^a \left(\frac{1-w_j}{1-y_i}\right)^b\right) \tag{9}$$

*is independent of $y$, where the coefficients $n_{a,b}$ and $\mu_{a,b}$ are given by the following table:*

| $(a,b)$ | $(-2,3)$ | $(-1,-1)$ | $(-2,1)$ | $(1,-2)$ | $(1,0)$ | $(0,1)$ | $(1,-1)$ |
|---|---|---|---|---|---|---|---|
| $n_{a,b}$ | 3 | 4 | 5 | 20 | 60 | 90 | 180 |
| $\mu_{a,b}$ | 3 | 2 | 2 | 2 | 1 | 1 | 1 |

REMARKS. **1.** The value of (9) can be found by specializing $y$ in any way we want. Taking $y = \infty$ (or 0 or 1, which are equivalent under $\underset{6}{\sim}$), we find that it equals $60P_6(G_9(z))$, where $G_9(z)$ is the combination of arguments occurring in Gangl's tetralogarithmic functional equation (8).

**2.** The coefficient of each inner double sum in (9) has been written as $n_{a,b}/\mu_{a,b}$ rather than simply $n_{a,b}$ because the 18 terms of the double sums occur with multiplicity $\mu_{a,b}$, so that the $(a, b)$th summand in (9) in fact consists of $18/\mu_{a,b}$ terms with coefficient $n_{a,b}$. Thus there are 87 terms altogether, forming 7 orbits under a group of order 18 (or 90 terms forming 8 orbits if we include the "constant term" as given in the previous remark).

**3.** Gangl's original one-variable equations are obtained by specializing to $y = z$ and $y = -z$.

**4.** The coefficients $n_{a,b}$ are determined by the requirement that the homogeneous polynomial $\sum n_{a,b}(aX+bY)^5$ vanish identically (since this polynomial is a combination of six monomials $X^kY^{5-k}$ and there are seven indices $(a, b)$, this has a solution). To see the necessity of this condition, note that the arguments $x$ of the hexalogarithms in (9) belong to the 16-dimensional space generated by the irreducible polynomials $y_i/w_j$, and $(1-y_i)/(1-w_j)$ $(1 \leq i \leq 6,\ 1 \leq j \leq 3)$, while $1-x$ belongs to the 34-dimensional space with the additional generators $p_{i,j} = y_i - w_j$. The contribution of one of the new primes $p = p_{i,j}$ to $\beta_6(x)$, where $x$ is one of the arguments in (9), is $\mathrm{ord}_p(x)[x]^5 \otimes [p] \in \mathrm{Sym}^5(F^\times) \otimes F^\times (F = \mathbb{Q}(y, z))$, so a short consideration shows that the total coefficient of $[p_{i,j}]$ under $\beta_6$ equals $\sum n_{a,b}(aX+bY)^5$ with $X = [y_i/w_j]$, $Y = [(1-y_i)/(1-w_j)] \in F^\times$. Thus the functional equation stated in the theorem can only hold for the values of $n_{a,b}$ given in the table; that it actually does hold depends on luck (at least at our present stage of understanding), since the coefficients are already uniquely determined by requiring that the "new" primes $p_{i,j}$ drop out under the map $\beta_6$, and we have no further freedom to ensure that the "old" primes also give a zero contribution modulo $\mathrm{Sym}^6(F^\times)$.

## REFERENCES


[AL] M. Abouzahra and L. Lewin, *The polylogarithm in algebraic number fields*, J. Number Theory **21** (1985), 214–244.

[B] A. Borel, *Cohomologie de $SL_n$ et valeurs de fonctions zêta aux points entiers*, Ann. Scuola Norm. Sup. Pisa **4** (1977), 613–636.

[BD] A. A. Beilinson and P. Deligne, in preparation.

[EST] P. Erdős, C. Stewart, and R. Tijdeman, *Some diophantine equations with many solutions*, Compositio Math. **66** (1988), 36–56.

[G] A. B. Goncharov, *Geometry of configurations, polylogarithms and motivic cohomology*, Max-Planck-Institut für Mathematik, Preprint, 1990.

[L] L. Lewin, *Polylogarithms and associated functions*, North-Holland, New York, 1981.

[M] J. Milnor, *Hyperbolic geometry: the first 150 years*, Bull Amer. Math. Soc. **6** (1982), 9–24.

[S] A. A. Suslin, *Algebraic K-theory of fields*, Proc. Internat. Congr. Math. (Berkeley, 1986), Amer. Math. Soc., Providence, R.I., 1987, pp. 222–244.

[Z] D. Zagier, *Polylogarithms, Dedekind zeta functions, and the algebraic K-theory of fields*, Arithmetic Algebraic Geometry (G. v.d. Geer, F. Oort, J. Steenbrink, eds.), Prog. Math., Vol. **89**, Birkhäuser, Boston, 1991, pp. 391–430.

APPENDIX B

# Summary of the Informal Polylogarithm Workshop November 17–18, 1990, MIT, Cambridge, Massachusetts

**List of Participants.** The informal Saturday workshop was attended by about 50 mathematicians including the following list of individuals who signed the log book.

| | |
|---|---|
| G. W. Anderson (CIT) | E. Bachmat (MIT) |
| A. A. Beilinson (Moscow, MIT) | S. Bloch (Chicago) |
| P. Deligne (IAS) | J. Dupont (Aarhus) |
| W. Gerdes (Brandeis) | A. B. Goncharov (Moscow, MIT) |
| S. Govindachar (Cornell, Brown) | R. Greenberg (BU) |
| L. Guo (BU) | T. Hagedorn (Harvard) |
| R. Hain (Washington, Duke) | F. Hajir (MIT) |
| M. Hanamura (IAS) | B. Harris (Brown) |
| M. Harris (Brandeis) | Y. Hu (MIT) |
| D. Husemuller (Haverford) | K. Igusa (Brandeis, Brown) |
| S. Kaminski (Brandeis) | M. Kim (MIT) |
| S. Kleiman (MIT) | M. Knorr (Brandeis) |
| M. Levine (Northeastern) | N. C. Leung (MIT) |
| S. Lichtenbaum (Brown) | R. MacPherson (MIT) |
| Y. Manin (Steklov, MIT) | W. Niziot (Princeton) |
| C. H. Sah (Stony Brook) | V. V. Schechtman (IAS) |
| T. Shiota (Brandeis) | V. Tschinkel (MIT) |
| A. N. Varchenko (IAS) | J. Yang (Washington) |

The workshop began at 9 a.m., Saturday, November 17, with the drafting of the program. Nine talks (each lasting between 30 minutes to an hour) were then presented on Saturday. Rather than having evening sessions, two talks were presented beginning at 10 a.m., Sunday, November 18, for the 20 or so remaining "diehards". The speakers and the titles of their talks (in order of presentation) were as follows:

G. W. Anderson, Drinfeld polylogarithms.
P. Deligne, Zagier's conjecture, I.
A. A. Beilinson, Zagier's conjecture, II.
A. B. Goncharov, Zagier's conjecture, III.
S. Bloch, Cycles and polylogarithms.
R. Hain, Extensions, variations, and trilogarithms.
M. Hanamura, Polylogarithms on Grassmann complex.
K. Igusa, Morse theory and relations with polylogarithms.
J. Yang, Trilogarithm and value of zeta function.
J. Dupont, Projective configurations and homology of $\mathrm{PGL}(n+1)$.
A. A. Beilinson, Elliptic polylogarithms.

The organizers would like to thank all the participants for their enthusiastic contributions. All the participants would like to thank Ms. Lisa Court for generous donation of invaluable administrative assistance and to MIT for providing the free space. Without their help the workshop would not have been possible.

Bob MacPherson
Han Sah

**An abbreviated summary.** Given the limited amount of space, time, as well as the ability of the reporter, it was impossible to present details of the proceedings at the workshop. The following article presented an excellent introduction as well as a summary of a modern view on higher logarithms:

R. Hain and R. MacPherson, *Higher logarithms*, Illinois J. Math. **34** (1990), 392–475.

The plenary address in ICM–90–Kyoto of A. Varchenko contained a rapid introduction to a number of the talks presented at the workshop in addition to a number of related topics which could not be covered at the workshop. See the preprint:

A. Varchenko, *Multidimensional hypergeometric functions in conformal field theory, algebraic K-Theory, algebraic geometry*, (Institute for Advanced Study, Princeton, November, 199).

About half of the talks at the workshop were devoted to the recent conjectures of Zagier on the relation between the values of regulator maps on $K_{2n-1}(F)$, $F$ a number field, and rational linear combinations of values of polylogarithm function. These conjectures were given in the following



Max-Planck-Institut preprints:

D. Zagier, *The Bloch–Wigner–Ramakrishnan polylogarithm function*, Max-Planck-Institut, 89–39.

D. Zagier, *Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields*, Max-Planck-Institut, 90-44.

Deligne's lecture was related to his preprint:

P. Deligne, *Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs*, (Institute for Advanced Study, Princeton, November, 1990).

Beilinson's lecture presented additional views. Goncharov's lecture then extended the views from number fields to arbitrary fields. The idea was related to the premise already hinted in the unpublished Irvine Lectures of Bloch:

S. Bloch, *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*, Irvine Lecture Notes, 1978.

Namely, results on the level of function fields could be specialized to give results in number fields. Beilinson's talk on Elliptic Polylogarithms was also in this direction. Bloch's "futuristic" lecture was probably best summarized by his statement:

Mixed Tate Motive = Matrix of Algebraic Cycles satisfying a condition of the form:

$$d\Omega + \Omega \cdot \Omega = 0.$$

Anderson's lecture extended polylogarithms to the realm of fields of positive characteristics. This brought out a mysterious new direction.

The remaining half of the workshop was devoted to some recent results (either in preprint form or under preparation) that are somewhat more down to earth. In particular, the talks of Hain, Hanamura, and Yang were in this direction. Naturally, additional conjectures were discussed. The lecture of Igusa was in the form of "applications". The basic idea was to use di- and tri-logarithms to detect elements that arose in topological investigations. The lecture of Dupont dealt with the role of projective configurations in the investigation of the Eilenberg-MacLane homology of the projective general linear group. It was an incredible sight watching a group of high powered mathematicians on a Sunday morning passionately discussing the secrets hiding behind high school geometric pictures involving only points, lines, and planes.

For an impromptu workshop of this type, the informal discussions between talks were just as important if not more important. Unfortunately, it was not possible to report on all of them. The following could perhaps be viewed as central.

One basic question that was discussed informally was "What should be a 'functional equation' of 'polylogarithm'?". One viewpoint suggested that there would be "many equations". Another viewpoint suggested that "functional equation(s) should be related to the description of the relevant homology group. Namely, one restricts the role of the polylogarithm in terms of the cohomology class it represents and then tries to describe the functional equation in terms of the domain of the cohomology class". This latter could be viewed as more or less the sentiment of the entire workshop. The principal issue was one of "what kind of homology and cohomology theories should be used?" This, of course, depended on the kind of problems one wanted to tackle. It was clear from the discussions at the workshop that there were and will be an enormous range of interesting questions involving polylogarithms.

In the most down to earth terms, one expectation was that a "suitably defined" $n$-logarithm, $n > 1$, should satisfy a functional equation involving $(2n+1)$ groups of terms. For large $n$, each group would involve a "huge number" of terms where each term should be the composition of various functions of $n(n-1)$ variables and a function of 1 variable. This would not resolve the question about "functional equation(s) for the classical $n$-logarithm function". The special case of $n = 1$ is excluded because PGL(1) is trivial.

The following reference book provided much information on volumes in non-Euclidean geometries. The concept of volume had been hiding in the background in many of the discussions:

> J. Böhm and E. Hertel, *Polyedergeometrie in n-dimensionalen Räumen constanter Krümmung*, Birkhäuser, Basel, 1981.

The abbreviated format of the workshop provided no opportunity to explore the connections between polylogarithms and hypergeometric functions as well as physics. In the USSR-US Mathematical Physics Symposium held in Philadelphia, December 1988, I. M. Gelfand delivered a lecture on Hypergeometric Functions. At that lecture, Gelfand revealed a projected 6 volume treatise on hypergeometric functions under preparation with several of his collaborators. At the opposite end of the spectrum, a few copies of a "science fiction" on the level of "UFO sightings" of polylogarithms in recent physics related speculations were privately "peddled" at the workshop. As indicated at the beginning, these topics were summarized in the preprint of Varchenko (with 113 references). It should be clear that one of the oldest topics in mathematics is now at the cutting edge of diverse investigations.

C. H. Sah

# Bibliography

The following works, in alphabetical order, represent recent publications; for the most part published subsequent to the updated bibliography in "Polylogarithms and Associated Functions" (1981).


[AK] Adamchik, V. S., Kölbig, K. S., *A definite integral of a product of two polylogarithms*, SIAM J. Math. Anal. **19** (1988), 926–938.

[AL1] Abouzahra, M., and Lewin, L., *The polylogarithm in algebraic number fields*, J. Number Theory **21** (1985), 214–244.

[AL2] ——, *The polylogarithm in the field of two irreducible quinitcs*, Aequationes Math. **31** (1986), 315–321.

[AL3] ——, *Supernumary polylogarithmic ladders and related functional equations*, Aequationes Math. **39** (1990), 210–253.

[ALX] Abouzahra, M., Lewin, L., and Xiao, H., *Polylogarithms in the field of omega (a root of a given cubic): Functional equations and ladders*, Aequationes Math. **33** (1987) 23–45. Addendum, Aequationes Math. **35** (1988), 304.

[A1] Aomoto, K., *A generalization of Poincaré normal functions on a polarized manifold*, Proc. Japan Acad., **55** (1979), 353–358.

[A2] ——, *Addition theorems of Abel type for hyperlogarithms*, Nagoya Math. J. **88** (1982), 55–71.

[A3] ——, *Configurations and invariant Gauss-Mannin connections of integrals* I, Tokyo J. Math. **5** (1982), 249–287.

[A4] ——, *Special values of hyperlogarithms and linear difference schemes*, Illinois J. Math. **34** (1990), 191–216.

[Ba] Baddoura, J., *A dilogarithmic extension of Liouville's theorem on integration in finite terms*, MIT Press, Cambridge, MA.

[Be1] Beilinson. A., *Higher regulators and values of L-functions*, J. Soviet Math. **30** (1985), 2036–2070. Translated from: Sovr. Probl. Mat. **24**, Mosc. VINITI (1984), 181–238.

[Be2] ——, *Notes on absolute Hodge cohomology*, Applications of Algebraic $K$-Theory to Algebraic Geometry and Number Theory, Contemporary Math., vol. 55, part I, Amer. Math. Soc., Providence, R. I., 1986, pp. 35–68.

[Be3] ——, *Polylogarithm and cyclotomic elements*, 1990 preprint.

[Be4] ——, *Motivic polylogarithm*, preprint.

[BGSV] Beilinson, A. A., Goncharov, A. B., Schechtman, V. V., and Varchenko, A. N., *Aomoto dilogarithms, mixed Hodge structures, and motivic cohomology of pairs of triangles on the plane*, The Grothendieck Festschrift, vol. 1, Birkhauser, Progress in Mathematics Series, vol. 86, Boston, 1990.

[BMS] Beilinson, A., MacPherson, R., and Schechtman, V. V., *Notes on motivic cohomology*, Duke Math. J. **54** (1987), 679–710.

[B1] Bloch, S., *Applications of the dilogarithm function in algebraic K-theory and algebraic geometry*, International Symposium on Algebraic Geometry, Kyoto, 1977, Kinokuniya Book-Store Co. Ltd., Tokyo (1978), pp. 103–114.

[B2] ——, *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*, unpublished manuscript, 1978.

[B3] ——, *The dilogarithm and extensions of Lie algebras*, Algebraic $K$-Theory, (M. Stein and E. Friedlander, eds.), LNM 854, Springer-Verlag, Berlin, Heidelberg, New York, 1984, pp. 1–23.

[Bo] Boldy, M. C., *The Bloch group and its relation to $K$-theory*, Report 8917, Catholic University, Nijmegen, 1989.

[Br1] Browkin, J., *Conjectures on the dilogarithm*, $K$-theory J. 3 (1989), 29–56.

[C] Chern, S. S., *Web Geometry*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), 1–9.

[Co] Coleman, R. F., *Dilogarithms, regulators, and $p$-adic $L$-functions*, Invent. Math. **69** (1982), 171–208.

[Cx1] Coxeter, H. S. M., *Star polytopes and the Schläfi function $f(\alpha, \beta, \gamma)$*, El. Math. **44** (1989), 25–36.

[Cx2] ——, *Trisecting an orthoscheme*, Computers Math. Applic. **17** (1989), 59–71.

[D1] Deligne, P. *Le groupe fondamental de la droite projective moins trois points*, (Y. Ihara, K. Ribet, and J. -P. Serre, eds.) Galois Groups over $\mathbf{Q}$, Proc. of a Workshop held March 23–27, 1987, Springer-Verlag, New York, 1989.

[D2] ——, *Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et regulateurs*, preprint.

[DD1] de Doelder, P. J., *On the Clausen integral $Cl_2(\theta)$ and a related integral*, Comp. and Appl. Math. **11** (1984), 325–330.

[DD2] ——, *On the Clausen integral $Cl_2(\theta)$ and a related integral*, Nieuw Tÿdschrift voor Wiskwande, Sept. 1987, 10–18. (This is a modified version of DD1.)

[DeD] Devoto, A., and Duke, D. W., *Table of integrals and formulae for Feynman diagram calculations*, Department of Physics Report, Florida State University, October 1983.

[Dp1] Dupont, J., *The dilogarithm as a characteristic class for flat bundles*, J. Pure and App. Alg. **44** (1987), 137–164.

[Dp2] ——, *On polylogarithms*, Nagoya Math. J. **114** (1989), 1–20.

[DS] Dupont, J., and Sah, C. H., *Scissors congruences* II, J. Pure and App. Algebra **25** (1982), 159–195.

[FK] Fornberg, B., and Kölbig, K. S., *Complex zeros of the Jonquière or polylogarithm function*, Math. Comp. **29** (1975), 582–599.

[G] Goncharov, A. B., *The classical trilogarithm, algebraic $K$-theory of fields and Dedekind zeta functions*, Department of Mathematics report, MIT, Cambridge, MA, 1990.

[GGL] Gabrielov, A. M., Gelfand, I. M., and Losik, M. H., *Combinatorial calculation of characteristic classes*, Functsional Anal. i Prilozheniya **9** (1975), no. 1, 54–55; no. 2, 12–28; no. 3, 5–26.

[GM] Gelfand, I. M., and MacPherson, R., *Geometry in grassmannians and a generalization of the dilogarithm*, Adv. in Math. **44** (1982) 279–312.

[Gr] Grayson, D., *Dilogarithm computations for $K_3$*, Lecture Notes Math., Springer-Verlag, Berlin, 1981, pp. 168–177.

[HM] Hain, R., and MacPherson, R., *Higher logarithms*, Illinois J. Math. **34** (1990), 392–475.

[K1] Kellerhals, R., *Ueber den Inhalt hyperbolischer Polyeder in den Dimensionen drei und vier*, Inauguraldissertation, Math. Inst. Univ. Basel, 1988.

[K2] ——, *On the volume of hyperbolic polyhedra*, Math. Ann. **285** (1989), 541–569.

[K3] ——, *On Schläfli's reduction formula*, Max-Planck-Institut für Mathematik, Report MPI/89-81.

[Ki] Kirillov, A. N., *On identities for Rogers' dilogarithm function related to simple Lie algebras*, Differential Geometry Lie Groups and Mechanics. IX, (Zap. nauchn. Semin. LOMI, v. 164), L. "Nauka", 1987, 121–133.

[Ko1] Kölbig, K. S., *Nielsen's generalized polylogarithms*, SIAM J. Math. Anal. **17** (1986), 1232–1258.

[Ko2] ——, *On a definite integral of the product of two polylogarithms*, Fourth International Conference on Computer Algebra in Physical Research, Dubna, USSR, May 22–26, 1990.

[L1] Lewin, L., *Polylogarithms and Associated Functions*, Elsevier North-Holland, New York, 1981.





[L2] ——, *The dilogarithm in algebraic fields*, J. Austral. Math. Soc. (A) **33** (1982), 302–330.

[L3] ——, *The inner structure of the dilogarithm in algebraic fields*, J. Number Theory **19** (1984), 345–373.

[L4] ——, *The order independence of the polylogarithmic ladder structure—implications for a new category of functional equations*, Aequationes Math. **30** (1986), 1–20.

[L5] ——, Supernumary polylogarithmic and Clausen-function ladders, Hayashibara Forum '90, International Symposium on Special Functions, Conference Proceedings, 1991. (RIMS, Kyoto University, Japan).

[L6] ——, *Further results on supernumary polylogarithmic ladders*, Max-Planck-Institut für Mathematik, Bonn 1991 preprint.

[LR] Lewin, L., and Rost, E., *Polylogarithmic functional equations: A new category of results developed with the help of computer algebra* (MACSYMA), Aequationes Math. **31** (1986), 223–242.

[Li] Lichtenbaum, S., *The construction of weight-two arithmetic cohomology*, Invent. Math. **88** (1987), 183–215.

[Lo] Loday, J. -L., *Comparaison des homologies du groupe linéaire et de son algèbre de Lie*, Ann. Inst. Fourier 37 (1987), 167–190.

[Lx] Loxton, J. H., *Special values of the dilogarithm function*, Acta Arith. **43** (1984), 155–166.

[Me] Meijer, J. W., *Subscriber line statistics, exponential integrals and polylogarithms*, Report 527 Tm/85, Dr. Neher Laboratories, Netherlands PTT, 1985.

[M1] Milnor, J., *Hyperbolic geometry: the first 150 years*, Bull. Amer. Math. Soc. (N.S.) **6** (1982), 9–24.

[M2] Milnor, J., *On polylogarithms, Hurwitz zeta functions, and the Kubert identities*, L'Enseingnement Math. **29** (1983), 281–322.

[O] Oda, T., *An interpretation of Lewin's ladder relation of polylogarithms*, Hayashibara Forum '90, International Symposium on Special Functions, Conference Proceedings 1991, (RIMS, Kyoto University, Japan).

[R1] Ramakrishnan, D., *A regulator for curves via the Heisenberg group*, Bull. Amer. Math. Soc. (N.S.) **5** (1981), 191–195.

[R2] ——, *On the monodromy of higher logarithms*, Proc. Amer. Math. Soc. **85** (1982), 596–599.

[R3] ——, *Analogs of the Bloch-Wigner function for higher polylogarithms*, Applications of Algebraic $K$-Theory to Algebraic Geometry and Number Theory, Contemporary Math., vol. 55 part I, Amer. Math. Soc., Providence, R.I. 1986, pp. 371–376.

[Rs] Richmond, B., and Szekeres, G., *Some formulas related to dilogarithms, the zeta function and the Rogers-Ramanujan identities*, J. Austral. Math. Soc. (A) **31** (1981), 362–373.

[V] Varchenko, A., *Multidimensional hypergeometric functions in conformal field theory, algebraic $K$-theory, algebraic geometry*, ICM, Kyoto, 1990.

[W1] Wojtkowiak, Z., *A construction of analogs of the Bloch-Wigner function*, Math. Scand. **64** (1989), 140–142.

[W2] ——, *A note on functional equations of polylogarithms*, Max-Planck-Institut für Mathematik, Bonn, 1990 preprint.

[W3] ——, *A note on the monodromy representation of the canonical unipotent connection on $P^1(\mathrm{C}) \setminus \{a_1, \dots, a_n\}$*. Max-Planck-Institut für Mathematik, Bonn, 1990 preprint.

[W4] ——, *A note on functional equations of the $p$-adic polylogarithms*, Institut des Hautes Études Scientifiques, France, 1991 preprint.

[Y] Yang, J., *The Hain-MacPherson trilogarithm, the third Borel regulator and the values of Dedekind zeta functions at* 3, 1990 preprint.

[Z1] Zagier, D., *The remarkable dilogarithm*, J. Math. Phys. Sci. **22** (1988), 131–145.

[Z2] ——, *The dilogarithm function in geometry and number theory*, Number theory and related topics, Ramanujan Colloquium, Bombay, January 1988, 231–249. (This is the same material as in [Z1]).

[Z3] ——, *The Bloch-Wigner-Ramakrishnan polylogarithm function*, Math. Ann. **286** (1990), 613–624.

[Z4] ——, *Polylogarithms, Dedekind zeta functions and the algebraic K-theory of fields*, MPI 90-44, Max-Planck-Institut für Mathematik, Bonn, 1990, 40 pp., preprint.

[Z5] ——, *Polylogarithms in geometry and arithmetics*, Hayashibara Forum '90, International Symposium on Special Functions, Conference Proceedings 1991. (RIMS, Kyoto University, Japan).

[Zu] Zucker, I. J., *On the series* $\sum_{k=1}^{\infty} \binom{2k}{k}^{-1} k^{-n}$ *and related sums*, J. Number Theory, **20** (1985), 92-102.

# Index