APPENDIX: 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 (partially 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 $\mathcal{A}_1(F)$ 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

(i) The map $\mathcal{A}_1(F) \to \mathbb{R}^n$ which sends ε to the n-tuple of values of $\log|\varepsilon^\sigma|$, σ ranging over the embeddings of F into $\mathbb{C}$, maps $\mathcal{A}_1(F)$ to a lattice in an (r_1+r_2-1)-dimensional subspace of $\mathbb{R}^n$; and

(ii) 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 $\mathcal{A}_m(F)$ such that (i) the mth polylogarithm function defines a map from $\mathcal{A}_m(F)$ to a lattice in $\mathbb{R}^n$, 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 (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 in fact this latter statement, not the existence of relations among polylogarithm values, which was the original motivation for the conjectures in [Z].)

The definition of $\mathcal{A}_m(F)$ 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 $F = \mathbb{Q}$ that there are enough elements of $\mathcal{A}_m(F)$ to produce non-trivial relations, and in particular that there are non-trivial relations among values of Li_m with rational arguments for arbitrary 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—in some sense trivial; of course, this is a very non-trivial use of the work "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 $\mathcal{A}_m(F)$. 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 $\mathbb{R}^n$, as sketched in the introductory paragraphs and explained in more detail below, so the group $\mathcal{A}_2(F)$ consists of all good combinations and we automatically get non-trivial 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 $\mathcal{A}_m(F)$, defined as the set of all combinations of elements of F which satisfy this property, then is mapped by the mth polylogarithm map into a certain lattice of finite rank, so we automatically get relations among polylogarithm values if the rank of $\mathcal{A}_m(F)$ (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 $n_i \in \mathbb{Q}$.) For convenience we also allow $[x_i]$ to be 0 or ∞ but set $[0] = [\infty] = 0$. The set of all combinations ξ forms a group $\mathcal{F}_F$, the free abelian group on $F^\times$. An element $\xi \in \mathcal{F}_F$ is *good at level 2* if it satisfies the relation

$$(*) \qquad \sum_i n_i\,[x_i] \wedge [1 - x_i] = 0 \qquad \text{in } \ \Lambda^2(F^\times) \otimes_{\mathbb{Z}} \mathbb{Q}\,,$$

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 ∞.) Explicitly, $(*)$ means that if we pick a basis $p_1, \ldots, p_s$ 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 ζ_i, $\zeta_i' \in F^\times$ are roots of unity and a_{ij}, a'_{ij} belong to $\mathbb{Z}$, then

$$\sum_i n_i \left(a_{ij}\, a'_{ik} - a_{ik}\, a'_{ij}\right) = 0 \qquad (1 \le j < k \le s).$$

In particular, if for some finite subset $S = \{p_1, \ldots, p_s\}$ of $F^\times$ the set of $x \in F$ for which both x and $1 - x$ belong to

$$\langle S\rangle = \left\{\, \zeta\, p_1^{a_1} \ldots p_s^{a_s} \;\middle|\; \zeta \in F^\times \text{ a root of unity, } a_1, \ldots, a_s \in \mathbb{Z}\right\}$$

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 $\xi \in \mathcal{F}_F$ is *good at level m* if it satisfies the relation

$$(*_m) \qquad \sum_i n_i\, \underbrace{[x_i] \otimes \ldots \otimes [x_i]}_{m-1} \otimes [1 - x_i] \ \in\ \mathrm{Sym}^m(F^\times) \otimes_{\mathbb{Z}} \mathbb{Q}\,,$$

where $\mathrm{Sym}^m(A)$ for any abelian group A is the $\mathfrak{S}_m$-invariant subspace of the mth tensor power $A^{\otimes m}$. We can write the $(m-1)$-fold tensor product $[x]\otimes\ldots\otimes[x]$ 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 $\mathcal{G}_m(F)$ of good combinations at level m is the kernel of the linear map

$$\beta_m:\,\mathcal{F}_F\;\rightarrow\;\big((\mathrm{Sym}^{m-1}(F^\times)\otimes F^\times)/\mathrm{Sym}^m(F^\times)\big)\otimes_{\mathbb{Z}}\mathbb{Q}$$

defined on generators by

$$\beta_m\big([x]\big)\;=\;\begin{cases}[x]^{m-1}\otimes[1-x]\;\;(\mathrm{mod}\;\mathrm{Sym}^m(F^\times)) & \text{if } x\neq 0,\,1,\,\infty\,,\\ \qquad 0 & \text{if } x=0,\,1,\,\infty\,.\end{cases}$$

We can also write $\mathcal{G}_m(F)$ in a way more analogous to the definition for $m=2$ as the kernel of the map $[x]\mapsto[x]^{m-2}\otimes\big([x]\wedge[1-x]\big)\in\mathrm{Sym}^{m-1}(F^\times)\otimes\Lambda^2(F^\times)$, 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 $\mathcal{G}_m(F)$ and the polylogarithm function is that functional equations of 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 Li_m cannot be extended to a continuous one-valued function on the whole complex plane, we will work instead with the modified function

$$P_m(x)=\Re_m\Bigg(\sum_{j=0}^{m-1}\frac{2^j\,B_j}{j!}\,\big(\log|x|\big)^j\;\mathrm{Li}_{m-j}(x)\Bigg)\,;$$

here $\Re_m$ denotes $\Im$ or $\Re$ according as m is even or odd and B_j the jth Bernoulli number ($B_0=1,\;B_1=-\frac12,\;B_2=\frac16,\;B_3=0,\;\ldots$). This function is continuous and real-valued on $\mathbb{P}^1(\mathbb{C})=\mathbb{C}\cup\{\infty\}$ and is real-analytic except at 0, 1 and ∞. (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 Li_m also has the advantage that the lower order terms in all functional equations cancel, so that $\sum_i P_m\big(x_i(t)\big)$ =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

$$\beta_m\Big(\frac1x\Big)=\big(-[x]\big)^{m-1}\otimes\big([-1]+[1-x]-[x]\big)=(-1)^{m-1}\beta_m(x)$$

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

$$\beta_m(x^2)=\big(2\,[x]\big)^{m-1}\otimes\big([1-x]+[1+x]\big)=2^{m-1}\,\big(\beta_m(x)+\beta_m(-x)\big)$$

or more generally $\beta_m(x^N)=\sum_{\zeta^N=1}\beta_m(\zeta x)$ 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 $\mathcal{A}_m(F)$. If ϕ is any homomorphism from $F^\times$ to $\mathbb{Z}$ (e.g. $\phi = v_p$ for $F = \mathbb{Q}$, 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 $\iota_\phi : \mathcal{F}_F \to \mathcal{F}_F$ sending $[x]$ to $\phi(x)\,[x]$ sends $\mathcal{G}_m(F)$ to $\mathcal{G}_{m-1}(F)$ (because $\beta_{m-1} \circ \iota_\phi = (\phi \otimes \mathrm{id}^{\otimes(m-1)}) \circ \beta_m$ under the natural identification of $\mathbb{Z} \otimes (F^\times)^{\otimes(m-1)}$ with $(F^\times)^{\otimes(m-1)}$). If we have r homomorphisms ϕ_i ($1 \le i \le r$), then for $m \ge r+2$ the composite map $\iota_{\phi_1} \circ \ldots \circ \iota_{\phi_r} : \mathcal{G}_m(F) \to \mathcal{G}_{m-r}(F)$ is independent of the order of the ϕ_i and will be denoted simply $\iota_{\phi_1} \ldots \iota_{\phi_r}$. The elements $\iota_{\phi_1} \ldots \iota_{\phi_r}(\xi) \in \mathcal{G}_{m-r}(F)$ ($1 \le r \le m-2$) for an element ξ of $\mathcal{G}_m(F)$ will be called the good combinations *associated to* ξ. Up to linear combinations there are only finitely many of them, since only the restrictions of the ϕ_i to the subgroup of $F^\times$ generated by the elements occurring in ξ are important and we can let ϕ_i run over a basis of the group of homomorphisms from this group to $\mathbb{Z}$.

The definition of $\mathcal{A}_m(F)$ is now as follows. We start with $\mathcal{A}_2(F) = \mathcal{G}_2(F)$. For $\xi = \sum n_i[x_i] \in \mathcal{A}_2(F)$ and each embedding $\sigma : F \hookrightarrow \mathbb{C}$, we consider $P_2(\xi^\sigma)$, where ξ^σ denotes $\sum n_i[x_i^\sigma]$ and $P_2(\xi^\sigma)$ denotes $\sum n_i P_2(x_i^\sigma)$ (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 $P_2(\overline{x}) = -P_2(x)$ for $x \in \mathbb{C}$ 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 non-real embeddings, we get a map $\vec{P}_2$ from $\mathcal{A}_2(F)$ to $\mathbb{R}^{r_2}$ by sending ξ to $\{P_2(\xi^\sigma)\}_\sigma$. The image $\vec{P}_2(\mathcal{A}_2(F))$ is contained in a *lattice* (= discrete subgroup of maximal rank) $\mathcal{R}_2 = \mathcal{R}_2(F) \subset \mathbb{R}^{r_2}$. Therefore, if $\mathcal{A}_2(F)$ has rank $> r_2$ (and we will see later that its rank is actually infinite), then the kernel $\mathcal{C}_2(F)$ of $\vec{P}_2$ contains non-zero elements and we get non-trivial relations among dilogarithms of complex algebraic arguments. Furthermore, if $\xi \in \mathcal{C}_2(F)$ and $\sigma : F \hookrightarrow \mathbb{R}$ is a real embedding of F, then $L_2(\xi^\sigma)$, where $L_2(x) = \mathrm{Li}_2(x) + \frac{1}{2}\log|x|\log|1-x|$ ($x \in \mathbb{R}$) is the Rogers dilogarithm function (cf. 1.2.3), is a rational multiple of π^2, so we also get relations among dilogarithms of real algebraic arguments. For $m = 3$ we define $\mathcal{A}_3(F)$ as the subset of $\mathcal{G}_3(F)$ consisting of all good combinations ξ for which $\iota_\phi(\xi) \in \mathcal{A}_2(F)$ belongs to $\mathcal{C}_2(F)$ for every homomorphism $\phi : F^\times \to \mathbb{Z}$. This condition can be checked by numerical computation since the element $P_2(\iota_\phi(\xi))$ belongs to the discrete group $\mathcal{R}_2$ and hence is recognizably zero or non-zero. For $\xi \in \mathcal{A}_3(F)$ 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) σ now ranges over the real and half the complex embeddings of F. This defines a map $\vec{P}_3 : \mathcal{A}_3(F) \to \mathbb{R}^{r_1+r_2}$ whose image is again contained in a lattice $\mathcal{R}_3 \subset \mathbb{R}^{r_1+r_2}$, so again we get a non-trivial kernel $\mathcal{C}_3(F)$ and non-trivial relations among polylogarithm values. Similarly, for $m > 2$ we consider the map $\vec{P}_m : \mathcal{F}_F \to \mathbb{R}^{n_\mp}$, 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 σ ranging over half the complex embeddings (m even) or all the real embeddings and half the complex embeddings (m odd). The subgroup $\mathcal{A}_m(F) \subseteq \mathcal{G}_m(F) \subset \mathcal{F}_F$ will be defined inductively in such a way that its image under $\vec{P}_m$ is contained in a certain lattice $\mathcal{R}_m = \mathcal{R}_m(F) \subset \mathbb{R}^{n_\mp}$, in which case the kernel

$$\mathcal{C}_m(F) = \{\xi \in \mathcal{A}_m(F) \mid P_m(\xi^\sigma) = 0 \quad \text{for all } \sigma : F \hookrightarrow \mathbb{C}\}$$

is a subgroup of corank at most $n_\mp$. Assuming that $\mathcal{A}_{m'}(F)$ and $\mathcal{C}_{m'}(F)$ have been defined for $m' < m$, we define $\mathcal{A}_m(F)$ by

$$\mathcal{A}_m(F) = \{\xi \in \mathcal{G}_m(F) \mid \iota_\phi(\xi) \in \mathcal{C}_{m-1}(F) \quad \text{for all homomorphisms } \phi : F^\times \to \mathbb{Z}\}.$$

Explicitly, this means that starting with $\xi \in \mathcal{G}_m(F)$, we verify $\xi \in \mathcal{A}_m(F)$ by checking first that all associated elements $\iota_{\phi_1} \ldots \iota_{\phi_{m-2}}(\xi) \in \mathcal{A}_2(F)$ map to 0 under $\vec{P}_2$, then that all associated elements $\iota_{\phi_1} \ldots \iota_{\phi_{m-3}}(\xi)$ (which are then known to belong to $\mathcal{A}_3$) map to 0 under $\vec{P}_3$, and so on in succession for all associated elements $\iota_{\phi_1} \ldots \iota_{\phi_r}(\xi)$ for $r = m-2$, $m-3$, ... , 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 $\mathcal{R}_m(F)$ 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 $\mathcal{R}_m \subset \mathbb{R}^{n_\mp}$ is a rational multiple of $|D_F|^{1/2}\,\zeta_F(m)/\pi^{mn_\pm}$. Hence if the map $\vec{P}_m : \mathcal{A}_m(F) \to \mathcal{R}_m$ 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 $\mathcal{R}_2$ was found by Bloch (for this and later developments concerning K_3 and dilogarithms, see Suslin's ICM talk [S]). In this case the map $\vec{P}_2 : \mathcal{A}_2 \to \mathcal{R}_2$ is surjective (after tensoring with $\mathbb{Q}$) and its kernel is the subgroup of $\mathcal{A}_2(F)$ 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 $\vec{P}_3 : \mathcal{A}_3(F) \to \mathbb{R}^{n_+}$ maps not only into, but onto the regulator lattice $\mathcal{R}_3$ (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 $\mathcal{A}_m(F)$ 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 $\mathcal{C}_m(F)$ can be described completely in terms of the functional equations of the mth 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 1.5.1 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 $\xi \in \mathcal{A}_m(F)$ (m odd) and real embeddings σ 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 $\mathcal{A}_m$), 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 $\xi \in \mathcal{A}_m(F)$ for which $P_m(\xi^\sigma) = 0$ for all complex embeddings σ, 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 π^m.

2. Examples of dilogarithm relations. For $m = 2$ there is no distinction between $\mathcal{G}_m$ and $\mathcal{A}_m$, so we just have to look for good combinations at level 2, i.e., combinations $\xi = \sum n_i[x_i]$ satisfying $(*)$. For such ξ, $P_2(\xi) = D(\xi)$ belongs to the r_2-dimensional lattice $\mathcal{R}_2$. Thus if we find r_2 elements ξ 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 ξ, 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 π^2 for the real embeddings σ 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 ξ belonging to $\mathcal{A}_2(\mathbb{Q})$, for each of which $L_2(\xi)$ is a rational multiple of π^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 non-trivial element of $\mathcal{A}_2(\mathbb{Q})$. In fact, calculating

$$\beta_2(\frac{1}{3}) = [\frac{1}{3}] \wedge [\frac{2}{3}] = [2] \wedge [3], \quad \beta_2(-\frac{1}{6}) = [-\frac{1}{6}] \wedge [\frac{7}{6}] = -[2] \wedge [7] - [3] \wedge [7], \quad \text{etc.},$$

we find that two linearly independent combinations of the 7 elements in question are annihilated by β_2, namely $6\left[\frac{1}{3}\right] - \left[\frac{1}{9}\right]$ and $3\left[-\frac{1}{6}\right] - \left[\frac{1}{8}\right] + \left[\frac{1}{9}\right] + \left[\frac{1}{28}\right]$. The image of each of these under the Rogers dilogarithm must be a rational multiple of π^2, and indeed,

$$6\,L_2(\frac{1}{3}) \;-\; L_2(\frac{1}{9}) \;=\; \frac{\pi^2}{3}\,, \qquad 3\,L_2(-\frac{1}{6}) \;-\; L_2(\frac{1}{8}) \;+\; L_2(\frac{1}{9}) \;+\; L_2(\frac{1}{28}) \;=\; -\frac{\pi^2}{12}\,.$$

Taking instead a field with $r_2 > 0$, we let $F = \mathbb{Q}(\sqrt{-7})$. 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 $\mathcal{A}_2(F)$. Hence $2D(x_1)+D(x_2)$ must be a rational multiple of $\zeta_F(2)/\pi^2\sqrt{7}$, and indeed one finds

$$2D\Big(\frac{1+\sqrt{-7}}{2}\Big)+D\Big(\frac{-1+\sqrt{-7}}{4}\Big)=\frac{21\sqrt{7}}{4\pi^2}\,\zeta_{\mathbb{Q}(\sqrt{-7})}(2).$$

As a third example, we consider the number $u=\sqrt{\rho}$, where $\rho=(\sqrt{5}-1)/2$ (cf. Chapter 5, section 5.2.1). Here the relation (5.4) implies that $\xi_0=[u^6]-4[u^3]+6[u]$ is a good element of $\mathcal{F}_F$, $F=\mathbb{Q}(u)$, 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, ξ_0 does not belong to the subgroup $\mathcal{C}_2(F)$ of $\mathcal{A}_2(F)$, because, denoting by σ the embedding of F into $\mathbb{C}$ 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 $F_1=\mathbb{Q}(\sqrt{5})$ and

$$L_{F_1}(s)=\frac{1}{1+5^{-s}}\prod_{p\equiv 11,\,13,\,17,\,19}\frac{1}{1-p^{-2s}}\prod_{p\equiv 3,\,7}\frac{1}{1+p^{-2s}}\prod_{p\equiv 1,\,9}\Big(\frac{1}{1+\varepsilon_p p^{-s}}\Big)^2$$

(here the congruences on primes p are modulo 20 and $\varepsilon_p=\big(\frac{\rho}{p}\big)=\big(\frac{2i+1}{p}\big)$ for $p\equiv 1$ or 9 mod 20, where ρ and i are solutions of $\rho^2+\rho-1\equiv i^2+i\equiv 0 \pmod p$)). 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 $\mathcal{C}_2(F)$ 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 5 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 $\mathcal{A}_2(F)$. 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 $D\big((u^2v)^\sigma\big)+9D\big(u^\sigma\big)$ equals $D(\xi_0^\sigma)$

but that $D\big((u^2v)^\sigma\big) + D\big((v^2/u^3)^\sigma\big)$ vanishes. Hence $[u^2v]+[v^2/u^3]$ and $[u^2v]+9[u]-\xi_0$ belong to $\mathcal{C}_2(F)$ and should map under the Rogers dilogarithm to rational multiples of π^2, and indeed,we have

$$L_2(u^2v)+L_2(v^2/u^3)=\frac{1}{20}\,\pi^2\,,\qquad L_2(u^2v)+9L_2(u)-L_2(\xi_0)=\frac{41}{60}\,\pi^2\,.$$

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 $F=\mathbb{Q}$ (also both taken from [Z]). The first illustrates the necessity of the extra condition in the definition of $\mathcal{A}_m$ as opposed to $\mathcal{G}_m$, 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 $x\in\mathbb{Q}$ 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, $1/4$ and $1/9$ for which the polylogarithm reduces to simpler values by virtue of the duplication equation. These values are $1/2$, $1/3$, $2/3$, $-1/2$, $3/4$, $-1/3$, $8/9$ and $-1/8$. The image of each of them under β_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 $\mathbb{Q}^\times$ 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 $[1/2]$, $[-1/3]-2[1/3]$, $[-1/8]-162[-1/2]$ and $[8/9]-9[3/4]-36[2/3]-18[-1/2]-6[1/3]$. This gives a 4-dimensional subspace of $\mathcal{G}_5(\mathbb{Q})$. However, to get elements of $\mathcal{A}_5(\mathbb{Q})$, we need three further conditions, namely that the images of ξ in $\mathcal{A}_3(\mathbb{Q})$ 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 $\vec{P}_3:\mathcal{A}_3(\mathbb{Q})\to\mathbb{Q}\zeta(3)$. This cuts down the dimension from four to one, the unique surviving relation being $\xi=[-1/8]-126[1/2]-162[-1/2]$. 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 non-trivial element of $\mathcal{A}_7(\mathbb{Q})$ 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 $x\in\mathbb{Q}$ 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 $\mathbb{Q}^\times\otimes_{\mathbb{Z}}\mathbb{Q}$ 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 $\mathcal{R}_{7-r}$. Since 29 is bigger than 28, we must find a solution of this system of equations, i.e., a combination $\xi \in \mathcal{A}_7(\mathbb{Q})$. This solution turns out to be unique up to multiplication by a constant and is given by $\xi = \sum n_i[x_i] \in \mathcal{A}_7(\mathbb{Q})$, where the n_i (normalized for convenience to be in $\frac{1}{6}\mathbb{Z}$ rather than $\mathbb{Z}$) 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 $-\dfrac{1020149599795}{96}\,\zeta(7)$ to high precision.

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 α. The advantage of making this restriction is that the conditions needed to make ξ belong to $\mathcal{A}_m$ 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 $\mathcal{G}_m$ or $\mathcal{A}_m$ 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 α. Since $[x]$ and $[1/x]$ are essentially equivalent for all polylogarithmic purposes, we can restrict to positive powers. Then the combinations ξ we are looking for can be written as

$$\xi = \sum_{j=1}^{\infty} n_j\,[\alpha^j] \qquad (n_j \in \mathbb{Z}, \quad n_j = 0 \text{ for all but finitely many } j)\,. \tag{1}$$

(From now on we will abbreviate the condition in brackets as $\{n_j\} \in \bigoplus_{j\ge1} \mathbb{Z}$.) If we compute the image of this under the map β_m, we find (since $[\alpha^j] = j[\alpha]$ in $F^\times$)

$$\beta_m(\xi) = \sum_{j\ge1} j^{m-1}\, n_j\,[\alpha]^{m-1} \otimes [1-\alpha^j] = [\alpha]^{m-1} \otimes \Big[\prod_{j\ge1} (1-\alpha^j)^{j^{m-1}n_j}\Big].$$

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 α. In other words, turning things around, given any *cyclotomic relation*

$$\prod_{j=1}^{\infty}\bigl(1-\alpha^j\bigr)^{b_j} \;=\; (\text{root of }1)\times\alpha^N \qquad \bigl(\{b_j\}\in\bigoplus_{j\ge1}\mathbb{Z},\ N\in\mathbb{Z}\bigr) \tag{2}$$

we get good combinations $\sum\limits_{j=1}^{\infty} j^{-m+1}\,b_j\,[\alpha^j]$ 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 $n_j\neq0$.) 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 ϕ from V to $\mathbb{Q}$ (namely, the one sending α^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 $\{b_j^{(1)}\},\ldots,\{b_j^{(d)}\}$ be a multiplicatively independent set of cyclotomic relations (2) for the same number α. Then for each m we get d linearly independent elements $\xi_m^{(\nu)}=\sum j^{-m}\,b_j^{(\nu)}\,[\alpha^j]$ 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 $\xi_2^{(1)},\ldots,\xi_2^{(d-r_2)}$, i.e., for each $\nu\le d-r_2$ we have $\sum j^{-2}b_j^{(\nu)}D\bigl((\alpha^\sigma)^j\bigr)=0$ if α^σ is complex, $\sum j^{-2}b_j^{(\nu)}L_2\bigl((\alpha^\sigma)^j\bigr)\in\mathbb{Q}\,\pi^2$ if α^σ is real, and $\xi_3^{(\nu)}=\sum j^{-3}b_j^{(\nu)}[\alpha^j]\in\mathcal{A}_3(F)$. Now the elements $\xi_3^{(1)},\ldots,\xi_3^{(d-r_2)}$ map under D_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 α^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 $\dfrac{2d}{n}$ 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

$$d(\alpha)\underset{\mathrm{DEF}}{=}\mathrm{rk}_{\mathbb{Z}}\Bigl\{\{b_j\}\in\bigoplus_{j\ge1}\mathbb{Z}\ \Bigm|\ \prod_{j=1}^{\infty}(1-\alpha^j)^{b_j}\in\langle\alpha\rangle\Bigr\}$$

is finite for any algebraic number α which is not a root of unity (of course $d(\alpha)=0$ if α is transcendental).

We now illustrate all of this with two examples. We first take α to be the number ω treated in [AL] and in Chapter 4 of the present book, i.e.,the root of $\omega^3=\omega+1$. Here $d(\alpha)$ is (at least) 12, the corresponding *special* exponents (= exponents j for which a positive power of $1-\alpha^j$ belongs to the subgroup generated by α and $1-\alpha^i$ with $i<j$) 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 mth order polylogarithms being 11, 9, 8, 6, 5, 3, 2 for $m=2, 3, \ldots, 8$, respectively. At the 9th level there are no more relations. However, if we include 1 as well as the positive powers of ω 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 $D_9(\omega^n)$ 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 (section 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 $j=360$. Here $r_2=4$ and $r_1+r_2=6$, so the number of conditions to be satisfied at the even and odd steps is a priori 4 and 6, respectively, but because the two real conjugates of α are inverses of one another and $D_m(x^{-1})=D_m(x)$ for m odd, there are in fact only 5 independent conditions at odd levels. We therefore get successively (at least) 71, 67, 62, 58, ... elements of $\mathcal{R}_2$, $\mathcal{R}_3$, $\mathcal{R}_4$, $\mathcal{R}_5$, ... Thus the ladder reaches up to $m=16$.

By the way, to find the special exponents, one proceeds as follows. First calculate the norm of α^j-1 by multiplying the conjugates numerically for j up to, say, 1000. The norm is 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 α, $1-\alpha$, $1-\alpha^2$, $1-\alpha^3$ and $1-\alpha^5$ are needed to generate a group of units of full rank, but we then get 18 independent multiplicative relations among α 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 combinations which are units. For instance, the norms of α^j-1 are positive powers of 3 for $j=4$, 8, 12, 16, 20, 24, 36, and 40, and eliminating α^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 (in practice it is more efficient to use the cyclotomic polynomials $\Phi_j(\alpha)$ rather than the numbers α^j-1) gives the 71 relations mentioned. By the time we get to $j=1000$, the norms are so huge that it seems clear that α^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 α gives the maximum of $\dfrac{d(\alpha)}{[\mathbb{Q}(\alpha):\mathbb{Q}]}$ (indeed, quite possibly even of $d(\alpha)$) for all algebraic numbers, in which case 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 $\mathcal{A}_m(\mathbb{Q})$ 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 ξ under the various maps $\iota_{\phi_{i_1}}\dots\iota_{\phi_{i_r}}$ to map to 0 in $\mathcal{R}_{m-r}\cong\mathbb{Z}$), 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 $\mathcal{A}_m(\mathbb{Q})$ 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 $p_1\dots p_r$ 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 primes $\le t$) much larger than u equals $\dfrac{\pi(t)^{u+\mathrm{o}(u)}}{u!} = (\dfrac{e\pi(t)}{u})^{u+\mathrm{o}(u)}$. The number of pairs (a,c) with $a,\,c\in A$, $c>a$, is therefore equal to $\binom{|A|}{2} = (\dfrac{e\pi(t)}{u})^{2u+\mathrm{o}(u)}$. 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 $N = \binom{|A|}{2}/t^u = (\dfrac{e\pi(t)}{u\sqrt{t}})^{2u+\mathrm{o}(u)}$ ways. We maximize this by choosing $u=\pi(t)/\sqrt{t}$, giving $N = e^{(2+\mathrm{o}(1))\pi(t)/\sqrt{t}}$. Now take $S=\{b\}\cup\{p\le t,\ p \text{ prime}\}$, with cardinality $s = 1+\pi(t) = (1+\mathrm{o}(1))\dfrac{t}{\log t}$; then each representation $b=c-a$ gives a distinct element $x=\dfrac{b}{c}$ of $X(S)$, so $|X(S)| > N = e^{(2+\mathrm{o}(1))\sqrt{s/\log s}}$.

Corollary. *(1) $\mathcal{A}_m(\mathbb{Q})$ 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 "$\mathcal{A}_m(F)$ for any number field F" instead of $\mathcal{A}_m(\mathbb{Q})$, since $\mathcal{A}_m(F)$ contains $\mathcal{A}_m(\mathbb{Q})$, 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

$$\sum_i n_i P_m(x_i) = 0 \qquad (n_i \in \mathbb{Z},\ x_i \in \overline{\mathbb{Q}})\,, \tag{1}$$

(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 $\vec{P}_m : \mathcal{A}_m(F) \to \mathbb{R}^{n_\mp}$ is the group $\mathcal{C}_m(F)$ 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^σ replaces each value $x_i = \phi_i(t)$ by x_i^σ.) 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

$$\sum_{m=1}^{|N|} \mathrm{Cl}_2\big(\theta + \frac{2m\pi}{N}\big) = \frac{1}{N}\,\mathrm{Cl}_2(N\theta) \qquad (0 \ne N \in \mathbb{Z}).$$

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 mth polylogarithm function for arbitrarily large values of m, but for $m > 6$ do not know whether these functions have any non-trivial 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 known functional equations and to find new ones in an algorithmic way. We will express all functional equations in terms of P_m rather than 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 I, section 1.2), which in our notation says that $D\bigl(R_5(x,y)\bigr) = 0$ for any x and y, where $R_5(x,y)$ denotes the formal linear combination

$$R_5(x,y) = [x] + [y] + \Bigl[\frac{1-x}{1-xy}\Bigr] + [1-xy] + \Bigl[\frac{1-y}{1-xy}\Bigr] \tag{1}$$

of elements in $\mathbb{Q}(x,y)$. (The other forms given by Spence, Hill, Abel and Kummer are equivalent to this by the 1-variable relations of Chapter I, 1.1.) To prove it in our language, we have to check that $R_5(x,y)$ is in the kernel of the map $\beta_2 : [z] \mapsto [z] \wedge [1-z]$. We calculate

$$\begin{aligned}\beta_2\bigl(R_5(x,y)\bigr) &= [x]\wedge[1-x] + [y]\wedge[1-y] + \bigl([1-x]-[1-xy]\bigr)\wedge\bigl([x]+[1-y]-[1-xy]\bigr)\\ &\quad + [1-xy]\wedge\bigl([x]+[y]\bigr) + \bigl([1-y]-[1-xy]\bigr)\wedge\bigl([y]+[1-x]-[1-xy]\bigr) = 0.\end{aligned}$$

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 apparently 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 the section. First of all, we can check that R_5 has a cyclic symmetry, i.e.,

$$R_5\bigl(x,\,y\bigr) = R_5\bigl(y,\,\frac{1-x}{1-xy}\bigr) = R_5\bigl(\frac{1-x}{1-xy},\,1-xy\bigr) = R_5\bigl(1-xy,\,\frac{1-y}{1-xy}\bigr) = R_5\bigl(\frac{1-y}{1-xy},\,x\bigr).$$

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 6-fold symmetry plays a role so often in the following that we introduce the special notation

$$x' \underset{6}{\sim} x \;\Leftrightarrow\; x' \in \{x,\,\frac{1}{x},\,1-\frac{1}{x},\,\frac{x}{x-1},\,\frac{1}{1-x},\,1-x\}\,.$$

A symmetric way to express the $\underset{6}{\sim}$-invariance is to say that the function $D([a,b,c,d])$, where $[a,b,c,d] = \dfrac{a-c}{a-d}\dfrac{b-d}{b-c}$ 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,

$$D\big([x_{\pi(1)},\, x_{\pi(2)},\, x_{\pi(3)},\, x_{\pi(4)}]\big) = \operatorname{sgn}(\pi)\, D\big([x_1, x_2, x_3, x_4]\big) \tag{2}$$

for any $x_i \in \mathbb{P}^1(\mathbb{C})$ $(1 \le i \le 4)$ and any $\pi \in \mathfrak{S}_4$, the symmetric group on 4 letters. To verify this in our terminology we compute

$$\begin{aligned}\beta_2\big([x_1, x_2, x_3, x_4]\big) &= \left[\frac{(x_1 - x_3)(x_2 - x_4)}{(x_1 - x_4)(x_2 - x_3)}\right] \wedge \left[\frac{(x_1 - x_2)(x_3 - x_4)}{(x_1 - x_4)(x_2 - x_3)}\right] \\ &= \frac{1}{2} \sum_{\pi \in \mathfrak{S}_4} \operatorname{sgn}(\pi)\, [x_{\pi(1)} - x_{\pi(2)}] \wedge [x_{\pi(2)} - x_{\pi(3)}]\end{aligned}$$

(this has only 12 terms rather than 24, because the πth summand is invariant under $\pi \mapsto \pi\tau$ where τ 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

$$\frac{1}{24} \sum_{\pi \in \mathfrak{S}_5} \operatorname{sgn}(\pi)\, D\big([x_{\pi(1)},\, x_{\pi(2)},\, x_{\pi(3)},\, x_{\pi(4)}]\big) = 0, \qquad (x_1, \ldots, x_5 \in \mathbb{P}^1(\mathbb{C})) \tag{3}$$

(this has only 5 terms rather than 120 because the πth 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 I, 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 non-obvious 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 3×3 matrix

$$Z = \big(z_{ij}\big)_{i,j=1,2,3} = \begin{pmatrix} x & 1/v & v/x \\ 1/w & y & w/y \\ w/x & v/y & xy/vw \end{pmatrix}. \tag{4}$$

Then the functional equation becomes simply $\sum_{i,j} D(z_{ij}) = 0$. The constraints on the z_{ij} are

$$\prod_j z_{ij} = 1 \quad (\forall i), \qquad \prod_i z_{ij} = 1 \quad (\forall j) \tag{5}$$

and (taking the indices i and j modulo 3)

$$(1 - z_{i,j})(1 - z_{i+1,j+1}) = (1 - z_{i,j+1}^{-1})(1 - z_{i+1,j}^{-1}) \qquad (\forall i,\, j). \tag{6}$$

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 8 values of the indices (i, j). Thus 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} $(i,\, j \in \mathbb{Z}/3\mathbb{Z})$ 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 9 equations (6) is equivalent to the condition $\det(T) = 0$, where $T = (t_{ij})$. (The set of 3×3 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 λ_i and μ_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

$$\sum_{i,\,j} D\big([u_i,\, u_{i+1},\, v_j,\, v_{j+1}]\big) = 0 \qquad \text{for any 6 points } u_i,\, v_j \in \mathbb{P}^1(\mathbb{C})\ (i, j \in \mathbb{Z}/3\mathbb{Z}).$$

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

$$\sum_{j} D\big([u_i,\, u_{i+1},\, v_j,\, v_{j+1}]\big) = D\big([u_i, v_1, v_2, v_3]\big) - D\big([u_{i+1}, v_1, v_2, v_3]\big)$$

for each i, and the sum of this over i (mod 3) vanishes. Of course, we can now go back and give a non-motivated and non-symmetric proof of the equation in its original form by adding the three 5-term relations $R_5\big(x,\, 1/v\big)$, $R_5\big(1/w,\, y\big)$ and $R_5\big(w/x,\, v/y\big)$ 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 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 2, 3.2.2.3) has the form $P_3\big(R_9(x, y)\big) = 2\zeta(3)$, where $R_9(x, y)$ denotes the element

$$\begin{aligned} R_9(x,y) =& 2\big[x\big] + 2\big[y\big] + 2\Big[\frac{x(1-y)}{x-1}\Big] + 2\Big[\frac{y(1-x)}{y-1}\Big] + 2\Big[\frac{1-x}{1-y}\Big] + 2\Big[\frac{x(1-y)}{y(1-x)}\Big] \\ & - \big[xy\big] - \Big[\frac{x}{y}\Big] - \Big[\frac{x(1-y)^2}{y(1-x)^2}\Big] \end{aligned} \tag{7}$$

of $\mathbb{Q}(x, y)$. 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 $\mathbb{P}^3(\mathbb{C})$ 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 9 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 sixfold 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 τ 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 9 of these, 3 of the form $[a, \tau(a); c, \tau(c)]$ and 6 of the form $[a, \tau(a); c, d]$ with $d \neq \tau(c)$. Then R_9 is just twice the sum of the latter 6 minus the sum of the former 3. (Take the p_ν to be 0, ∞, 1, xy, x and y with the involution $\tau : t \mapsto xy/t$.) The equivalence of this form of the functional equation to the one just given is seen by taking for $\{p_\nu\}$ the 6 numbers $a_i + a_j$ $(1 \leq i < j \leq 4)$ and for τ the involution $t \mapsto -t$. The symmetry group appears now not as $\mathfrak{S}_4$, but as the semi-direct product of $\mathfrak{S}_3$ (permute the three orbits of τ 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 τ 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 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

$$
\begin{aligned}
&2\left[P_4\big(z(1-z)\big)+P_4\big(-\frac{z}{(z-1)^2}\big)+P_4\big(\frac{z-1}{z^2}\big)\right]\\
&-3\left[P_4\big(\frac{1}{1-z+z^2}\big)+P_4\big(\frac{(1-z)^2}{1-z+z^2}\big)+P_4\big(\frac{z^2}{1-z+z^2}\big)\right]\\
&-6\left[P_4\big(\frac{1-z+z^2}{z(z-1)}\big)+P_4\big(\frac{1-z+z^2}{z}\big)+P_4\big(\frac{1-z+z^2}{1-z}\big)\right] \;=\; 0\,.
\end{aligned}
\tag{8}
$$

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

$$
\begin{aligned}
&\big[P_4\big(-\frac{bd}{a}\big)+P_4\big(-\frac{ac}{d}\big)+P_4\big(-\frac{ab^3}{c^2d}\big)+P_4\big(-\frac{c^3d}{ab^2}\big)+P_4\big(-\frac{d}{ab^2}\big)+P_4\big(-\frac{a}{c^2d}\big)\big]\\
&+\big[P_4\big(-\frac{ad^2}{b}\big)+P_4\big(-\frac{a^2d}{c}\big)\big]+2\big[P_4\big(\frac{b}{c^2}\big)+P_4\big(\frac{c}{b^2}\big)+P_4\big(bc\big)\big]\\
&+3\big[P_4\big(-\frac{b}{ad}\big)+P_4\big(-\frac{c}{ad}\big)+P_4\big(-\frac{1}{ad}\big)\big]\\
&+3\big[P_4\big(-cd\big)+P_4\big(-ab\big)+P_4\big(-\frac{d}{c}\big)+P_4\big(-\frac{a}{b}\big)+P_4\big(-\frac{ab}{c^2}\big)+P_4\big(-\frac{cd}{b^2}\big)\big]\\
&+6\left[P_4\big(\frac{c}{a}\big)+P_4\big(\frac{b}{d}\big)+P_4\big(\frac{1}{d}\big)+P_4\big(\frac{1}{a}\big)+P_4\big(\frac{b}{cd}\big)+P_4\big(\frac{c}{ab}\big)\right] \;=\; 0\,,
\end{aligned}
$$

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 $P:(a,b,c,d)\mapsto(d,c,b,a)$, corresponding to $(y,\,z)\mapsto(1-y,\,1-z)$, and a less obvious one under the involution $Q:(a,b,c,d)\mapsto(a,b^{-1},cb^{-1},db^{-1})$, 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 6 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

$$X \;=\; \Bigl\{(t_1:t_2:t_3:t_4:t_5)\in\mathbb{P}^4 \;\Bigm|\; \sum_{i=1}^{5} t_i=0,\quad \sum_{i=1}^{3} t_i^{-1}=0\Bigr\}\,,$$

with an obvious symmetry of the group $G=\mathfrak{S}_3\times\mathfrak{S}_2\subset\mathfrak{S}_5$. We can parametrize X by

$$\bigl(t_1:t_2:t_3:t_4:t_5\bigr)=\bigl(-y:y-1:y(1-y):y-y(1-y)z:(y-1)^2+y(1-y)z\bigr)\,,$$

(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 β_4.

Example 5. A 2-variable functional equation for the hexalogarithm. Gangl found one-variable functional 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 $z'\underset{6}{\sim}z$. The 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 α, β, γ with even sum we have set

$$T^{\pm}_{\{\alpha,\beta,\gamma\}}=\sum_{\{a,b,c\}=\{\alpha,\beta,\gamma\}}P_6\biggl(\pm\frac{(-1)^a z^b(1-z)^c}{(1-z+z^2)^{(a+b+c)/2}}\biggr)$$

(the sum is over all permutations and has 6 terms in general but only 3 if two of α, β and γ coincide). Later he found a second functional equation with 89 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 $(1 \le i \le 6)$ and w_j $(1 \le j \le 3)$ the elements $y' \underset{6}{\sim} y$ and $z'(1-z')$, $z' \underset{6}{\sim} z$. Then*

$$\sum_{a,b} n_{a,b} \frac{1}{\mu_{a,b}} \sum_{i=1}^{6} \sum_{j=1}^{3} P_6\Big(\big(\frac{w_j}{y_i}\big)^a \big(\frac{1-w_j}{1-y_i}\big)^b\Big) \tag{9}$$

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 96 terms forming 8 orbits if we include the "constant term" as given in the previous remark.
3. Gangl's original 1-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 6 monomials X^kY^{5-k} and there are 7 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 $(1-y_i)/(1-w_j)$ $(1 \le i \le 6,\ 1 \le j \le 3)$, 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 β_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 β_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. Sc. Norm. Sup. Pisa **4** (1977), 613–636.

[BD]. A.A. Beilinson and P. Deligne, in preparation.

[EST]. P. Erdös, C. Stewart and R. Tijdeman, Some diophantine equations with many solutions, Comp. Math. **66** (1988), 36–56.

[G]. A.B. Goncharov, Geometry of configurations, polylogarithms and motivic cohomology, preprint, 1990.

[L]. L. Lewin, *Polylogarithms and Associated Functions*, North-Holland, New York, 1981.

[M]. J. Milnor, Hyperbolic geometry: the first 150 years, Bull AMS **6** (1982), 9–24.

[S]. A.A. Suslin, Algebraic K-theory of fields, in *Proceedings of the International Congress of Mathematicians 1986*, AMS, 1987, pp. 222–244.

[Z]. D. Zagier, Polylogarithms, Dedekind zeta functions, and the algebraic K-theory of fields, in *Arithmetic Algebraic Geometry* (G. v.d. Geer, F. Oort, J. Steenbrink, eds.), Prog. in Math. **89**, Birkhäuser, Boston, 1991, 391–430.

Remark. 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 2π and hence gives $Li_m(z)$ for all z with $.005 < |z| < 230$.