Summation in Finite Terms

MICHAEL KARR

Massachusetts Computer Associates, Wakefield, Massachusetts

ABSTRACT. Results which allow either the computation of symbolic solutions to first-order-linear difference equations or the determination that solutions of a certain form do not exist are presented. Starting with a field of constants, larger fields may be constructed by the formal adjunction of symbols which behave like solutions to first-order-linear equations (with a few restrictions) It is in these extension fields that the difference equations may be posed and in which the solutions are requested. The principal application of these results is in finding formulas for a broad class of finite sums or in showing the nonexistence of such formulas.

KEY WORDS AND PHRASES. summation, finite terms, algebraic symbol manipulation, difference algebra

CR CATEGORIES: 5.7

1. *Introduction*

1.1 BACKGROUND. This paper is concerned with the problem of finding formulas for finite sums. The approach taken here is to pose the problem in algebraic terms and then to derive constructive conditions for summability. While the motivation for this research was to produce the algorithms implicit in the proofs, the theory itself is of independent mathematical interest; the emphasis of this paper is on the overall structure of the subject. An attempt has been made to avoid both theory which is irrelevant to the algorithms and algorithmic details which obscure the theory.

Consider some "closed form" solution to a summation,

$$g(n) = \sum_{a \le i < n} f(i), \qquad h(n) = \sum_{a < i \le n} f(i).$$

From the classical calculus of finite differences [1] we know that we may apply the upper (lower) difference operator Δ (∇) and obtain

$$\Delta g(n) \triangleq g(n+1) - g(n) = f(n),$$
$$\nabla h(n) \triangleq h(n) - h(n-1) = f(n).$$

Thus, $\Delta g = f = \nabla h$. Furthermore, for *any* g or h such that $\Delta g = f = \nabla h$,

$$\sum_{a \le i < n} f(i) = g(n) - g(a), \qquad \sum_{a < i \le n} f(i) = h(n) - h(a).$$

Hence, given f, if we could solve the equation $\Delta g = f$ or $\nabla h = f$, we would have a formula for the finite sum.

The obvious analogy between this situation and the fundamental theorems of calculus has not gone unnoticed, and there is an extensive theory, "operator methods"

Author's present address: GSG, Inc., 51 Main Street, Salem, NH 03079

or "symbolic methods," which connects ordinary calculus with the calculus of finite differences [1, 17].

The problem of *integrating* in finite terms has been solved now for several years [14, 15]. Given the power and elegance of operator methods, one might think that the problem of summation in finite terms would also have been solved. However, the difficulty is that "finite terms" are not preserved under the transformations of operator methods, and there appears to be no simple way to adapt the integration work for summation formulas. More will be said about this issue in the conclusion.

Up to this point, I have been discussing "the" problem of finite summation as if there were a widely accepted view as to what the problem is. This is not the case, as a brief review of the literature reveals. In the algebraic symbol manipulation literature, one of the first efforts is [9], which can be used to verify formulas for sums (essentially by applying Δ) but does not consider the problem of finding formulas. The problem of summing rational functions was considered in an earlier version of this paper [10], as well as in [13]. The problem of exponentials and rational functions was also treated in [10]. A decision procedure for a class of summands involving indefinite products of rational functions was obtained in [6, 7]. Mention should also be made of the lookup/transformation approach found in [2, 3].

This paper describes techniques which greatly broaden the scope of what is meant by "finite terms"; consequently, the class of sums which can be meaningfully simplified is considerably enlarged. Examples of sums for which formulas can be produced by methods of this paper (and earlier ones) include

$$\sum_{i=1}^{n} i, \qquad \sum_{i=0}^{n} 2^i, \qquad \sum_{i=1}^{n} i^2 2^i, \qquad \sum_{i=1}^{n} \frac{1}{i^2 + 2i}, \qquad \sum_{i=1}^{n} i \cdot i!.$$

Similarly, these methods will show that the following sums have no formula as a rational function of n:

$$\sum_{i=1}^{n} \frac{1}{i}, \qquad \sum_{i=1}^{n} \frac{1}{i^2}, \qquad \sum_{i=1}^{n} \frac{2^i}{i}, \qquad \sum_{i=1}^{n} i!. \tag{1}$$

A glimpse of the additional power of the techniques presented here may be seen in a small example. Consider the first sum of (1). This is known as the nth harmonic number H_n. This symbol may then occur in other sums, for example,

$$\sum_{i=1}^{n} H_i, \qquad \sum_{i=1}^{n} iH_i, \qquad \sum_{i=1}^{n} \frac{H_i}{i}$$

The algorithms presented below yield formulas, as a rational function of n and the symbol H_n, for the first two of these sums; they also show that the third has no formula of that type.

The techniques here do not cover such formulas as

$$\sum_{i=0}^{n} \frac{1}{c^{2^i}}.$$

Although one can probably extend the techniques of this paper to deal with c^{2^i}, this has not yet been done. It should be noted that the techniques of this paper cannot be extended to deal with a summation problem in which one of the limits is also involved in the summand, for example,

$$\sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

Such summations cannot be solved by inverting $\Delta g = f$.

The algorithms implicit in the proofs of this paper use features available in algebraic manipulation systems, principally factorization and greatest common divisor calculations. Some of the techniques presented here have already been implemented [4, 19].

The reader of this paper is expected to be acquainted with the concepts and notation of modern algebra—a one-year undergraduate course in the subject, using [8, 12, 18] as texts, should easily supply the required background. Given a basic knowledge of algebra, this paper is self-contained, with the exception of the omission of many proofs, particularly those not involving algorithms. These proofs may be supplied by the reader or found in [11].

1.2 DIFFERENCE FIELDS.[1] When one studies integration procedures, the algebraic object of interest is the differential field, that is, a field together with a map which is linear over addition in the field, and obeys the familiar product rule for differentiation. In studying summation, it seems reasonable to begin by studying the operators Δ and ∇. These operators are clearly linear; the product rules, though simple, are not as widely known as those for differentiation:

$$\Delta(f \cdot g) = f \cdot \Delta g + \Delta f \cdot g + \Delta f \cdot \Delta g,$$
$$\nabla(f \cdot g) = f \cdot \nabla g + \nabla f \cdot g - \nabla f \cdot \nabla g.$$

Hence, we might propose the following axioms for an upper (or lower) "difference field," in addition to the usual field axioms: Letting $\delta = \Delta, \nabla$ and $\pm$ be $+, -$, respectively,

DIFF1 for all f, g, $\delta(f + g) = \delta f + \delta g$;
DIFF2 for all f, g, $\delta(f \cdot g) = f \cdot \delta g + \delta f \cdot g \pm \delta f \cdot \delta g$.

These two axioms are not quite enough; the last remaining axiom may be stated in a wide variety of ways.

PROPOSITION 1. *Let F be any field, and let $\delta: F \to F$ be any map satisfying DIFF1 and DIFF2. The following conditions are equivalent (in their usual respective cases):*

DIFF3

(*a*) $\delta 1 = 0$.
(*b*) $\delta 1 \neq \mp 1$.
(*c*) $\delta f \neq \mp f$ *for some* f.
(*d*) $\delta f \neq \mp f$ *for all* $f \neq 0$.
(*e*) *Let* $\rho f \triangleq f \pm \delta f$. *Then* ρ *is an endomorphism of* F. (*We use* σ *and* τ *for the respective cases of* ρ.)

Definition 1. An *upper* (respectively, *lower*) *difference field* is a field F together with a map Δ (respectively, ∇) from F to F, satisfying DIFF1, DIFF2, and any of the equivalent conditions of DIFF3.

The appearance of the endomorphism ρ from the product rule for the abstract difference operators should not be viewed as mysterious. It is the algebraic vestige of the shift operator. In the concrete case

$$\sigma f(x) = f(x + 1), \qquad \tau f(x) = f(x - 1). \tag{2}$$

Note that in the concrete case, σ and τ are inverses of each other. This leads us to the

[1] Throughout this paper all fields are assumed to have characteristic 0.

beautifully simple definition for a field which has properly related upper and lower difference operators.

Definition 2. A *difference field* is a field F together with an automorphism σ of F.

This seems almost too trivial to be useful, but it is the algebraic object which we shall study throughout most of the rest of this paper, and it leads to summation decision procedures.

Example 1. The complex numbers, with their automorphism $\sigma(a + b\sqrt{-1}) = a - b\sqrt{-1}$, form a difference field. One should not artificially exclude this as a difference field, since $\sqrt{-1}$ is algebraically indistinguishable from the function of an integer n, $I(n) \triangleq (-1)^n\sqrt{-1}$; note that $[I(n)]^2 = -1$. Conjugation is then the shift operator applied to $I(n)$, as in (2):

$$\sigma(a + b \cdot I(n)) = a + b \cdot I(n + 1) = a - b \cdot I(n). \qquad \square$$

Conventions. In any difference field, we have

$$\tau \triangleq \sigma^{-1};$$
$$\Delta f \triangleq \sigma f - f, \qquad \nabla f \triangleq f - \tau f;$$
$$K \triangleq \{f \in F \mid \sigma f = f\} \quad \text{(the } \textit{constants} \text{ of } F\text{)}.$$

If $K = F$, we have a *constant* difference field.

Observe that $f \in F$ is a constant $\Leftrightarrow \Delta f = 0 = \nabla f$; also, the set of constants K is actually a subfield of F, called the fixed field of σ.

Example 2. Continuing Example 1, if the complex numbers with conjugation are considered a difference field, the field of constants is the field of real numbers. The complex numbers can also be a field of constants for a (perhaps) larger difference field, by letting σ be the identity on this field. $\square$

The subject of difference algebra is treated in [5], which has a number of references to the mathematical literature. The problem of finding formulas has apparently not received much attention by pure mathematicians.

1.3 Outline of the Problem and Solution. With the concept of a difference field understood, it is possible to describe somewhat more precisely what it means to find a sum in "finite terms." Given f and a difference field F of which f is an element, we look for solutions g of $\Delta g = f$ (or $\nabla g = f$) *only in* F. Thus *the choice of field* F *is the means by which "in finite terms" is given precise meaning.* For example, in the context of integration, "finite terms" usually means starting with the rational functions over some constant field and allowing algebraic, exponential, and logarithmic extensions, nested to arbitrary depth. This paper does not consider algebraic extensions. The extensions which it does consider are analogous to exponential and logarithmic extensions, but in fact are much more general than just these. A precise definition of these extensions is postponed until Section 2.

We now reconsider the equations $\Delta g = f$ and $\nabla g = f$. In terms of σ and τ, we have

$$\sigma g - g = f, \qquad g - \tau g = f.$$

If F, σ is a difference field, so is F, τ. Thus we lose no essential generality in considering only one of these equations; we arbitrarily pick the former.

Ideally, this paper would proceed as follows. We would first observe that $\sigma g - g = f$ is easy to solve in a constant field (a solution exists $\Leftrightarrow f = 0$; if a solution

exists, any element of K is a solution). Then we would prove that if there exists an algorithm for solving $\sigma g - g = f$ in some field F, σ, we could somehow use this procedure to help us obtain an algorithm to solve the same problem in E, σ, where E, σ is some difference field extension of F, σ.

Unfortunately, things are more complicated than this. The solvability of the simple equation $\sigma g - g = f$ seems to be an inconvenient property to lift through the extensions we are considering. In Section 3 we consider a more general problem, which at least has the virtue that its form is preserved in the process of proof by induction.

1.4 CONVENTIONS AND NOTATION. Consider $\sum_{m \le i < n} f(i)$. If $m < n$, this has an obvious meaning. If $m \ge n$, this is summation over the null set, which is customarily defined to be zero. We shall follow this convention when $m = n$, but when $m > n$, we shall say that

Definition 3

$$\sum_{m \le i < n} f(i) \triangleq - \sum_{n \le i < m} f(i), \qquad \text{where} \quad m > n.$$

N.B. This abuse of notation means that "$m \le i < n$," when written under a $\sum$, does not imply that $m < n$.

There is more to this than an attempt to ape the calculus convention; it is genuinely useful in several places. For starters, we observe

PROPOSITION 2

(*a*) $\Delta g = f \Rightarrow \sum_{m \le i < n} f(i) = g(n) - g(m)$, *regardless of the ordering of* m, n.
(*b*) $\sum_{l \le i < n} f(i) = \sum_{l \le i < m} f(i) + \sum_{m \le i < n} f(i)$, *regardless of the ordering of* l, m, n.

The same convention is used with the product symbol $\prod$, with the understanding that reciprocation replaces negation in Definition 1.

It is sometimes convenient to introduce several variables which are determined implicitly by some relationship. For example, given positive integers m and n, division with remainder determines q and r such that $m = q \cdot n + r$, where $0 \le r < n$. We would introduce these implicitly defined variables thus:

$$\text{def } q, r: \qquad m = q \cdot n + r, \quad 0 \le r < n.$$

We use the standard notation for the ring or field formed by adjoining an element t to a ring R or field F: $R[t]$ or $F(t)$.

This paper is frequently concerned with transcendental field extensions, in which case $R[t]$ is a ring of polynomials and $F(t)$ is a field of "rational functions." The numerators and denominators of elements of $F(t)$ are defined thus:

$$\text{def num}(f), \text{den}(f): \qquad f = \frac{\text{num}(f)}{\text{den}(f)},$$

$\text{num}(f), \text{den}(f) \in F[t]$, their gcd is 1, and $\text{den}(f)$ is monic.

A *factor* $\mathscr{f}$ of $f \in F(t)$ is a polynomial which divides $\text{num}(f)$ or $\text{den}(f)$. The *power to which* $\mathscr{f}$ *occurs* in f is defined to be the power to which it occurs in $\text{num}(f)$ or the negative of the power to which it occurs in $\text{den}(f)$. The unique factorization of polynomials extends to the unique factorization of any $f \in F(t)$:

$$f = u \cdot \prod_{\mathscr{f}} \mathscr{f}^{n_{\mathscr{f}}},$$

where

(F1) $\mathscr{f}_i$ is monic irreducible, degree > 0, $u \in F$;
(F2) $i_1 \neq i_2 \Rightarrow \mathscr{f}_1 \neq \mathscr{f}_2$;
(F3) $n_i \neq 0$ (but may be positive or negative).

A tuple may be abbreviated by a boldface version of the same letter used for its components: $\mathbf{f} \triangleq \langle f_1, \ldots, f_k \rangle$; the length of a tuple may be written $|\mathbf{f}|$. Concatenation of tuples is done thus:

$$\mathbf{f} \frown \mathbf{g} \triangleq \langle f_1, \ldots, f_k, g_1, \ldots, g_l \rangle.$$

This notation may be abused when appending a single element:

$$\mathbf{f} \frown g \triangleq \mathbf{f} \frown \langle g \rangle = \langle f_1, \ldots, f_k, g \rangle.$$

If the components of $\mathbf{f}$ are drawn from a set S, we write

$$\mathbf{f} \in S^{\omega}.$$

In other words,

$$S^{\omega} = \{\text{null-tuple}\} \cup S \cup S \times S \cup \cdots.$$

A more standard notation for S^{ω} would be S^*; but since S is often a field F, F^* also brings to mind the multiplicative group of F, which notation we in fact use.

We occasionally work in a module R^n where R is $\mathbb{Z}$ (the integers) or a field. We use the standard inner product when $|\mathbf{f}| = |\mathbf{g}|$:

$$\mathbf{f} \cdot \mathbf{g} = \sum_{i=1}^{|\mathbf{f}|} f_i \cdot g_i. \tag{3}$$

We also need the annihilator of an element:

$$\mathrm{Ann}_R(\mathbf{f}) \triangleq \{\mathbf{g} \mid \mathbf{g} \in R^{\omega}, \mathbf{f} \cdot \mathbf{g} = 0\}.$$

We let $\mathbf{0}_n$ be a tuple consisting of n zeros. For stating and proving certain results it is convenient to allow the null tuple, denoted by $\mathbf{0}_0$ in certain examples, and to let $\mathbf{0}_0 \cdot \mathbf{0}_0 = 0$ (justified by $|\mathbf{0}_0| = 0$ and the summation convention). We use $\mathbf{0}_{mn}$ for an m by n matrix of zeros. Again, for convenience, there are situations in which m is zero (the matrix has no rows) or n is zero (no columns). Certain obvious conventions apply, for example, $\mathbf{0}_{m0}\mathbf{0}_0 = \mathbf{0}_m$ and $\mathbf{0}_{m0} \cdot \mathbf{0}_{0n} = \mathbf{0}_{mn}$.

2. *Extensions*

2.1 Defining Extensions. Suppose that one is faced with a complicated sum,

$$\sum_{1 \leq i < n} \left[\sum_{1 \leq j < i} \frac{j-1}{j^2 + j} + i \cdot \sum_{1 \leq j < i} \frac{1}{j} \right]. \tag{4}$$

It seems unlikely that this will simplify to a rational function of n (and indeed it will not). However, it is reasonable to ask if the sum can be expressed using the terms of the summand with "i" replaced by "n," that is, to ask if the sum can be expressed as a rational function of the following:

$$n, \qquad \sum_{1 \leq j < n} \frac{j-1}{j^2 + j}, \qquad \sum_{1 \leq j < n} \frac{1}{j}. \tag{5}$$

The answer is yes, but that is getting ahead of the story. The purpose of this section

is to completely formalize the question, both mathematically and computationally. Given a constant field, we characterize those extensions in which we are interested. These extensions include more than the adjunction of just formal sums and products, but the extra generality comes at no extra expense—the techniques used appear to be necessary to handle only sums and products. Extensions are to be "declared" one step at a time, and algorithms which verify that the extensions satisfy the mathematical criteria are given. As an example of what can happen, consider the summation in (4). One might naively try a three-level extension, successively including the three terms of (5). But the algorithms will reveal that the last two terms are not independent and that only a two-level extension is required (each of the sums of (5) can be expressed in terms of the other). In order to apply the algorithms of this paper, the summand of (4) must be rewritten so that this relationship is evident.

2.2 Types of Extensions. All of the extensions considered in this paper have the following properties.

Definition 4. Let F, σ be a difference field and E, σ an extension difference field. We call this extension *affine* $\Leftrightarrow$

$$E = F(t), \qquad \text{where} \quad \alpha, \beta \in F \quad \text{with} \quad \sigma t = \alpha \cdot t + \beta.$$

Given a difference field F with computable field operations (including σ), there is a simple way to compute σ for $F(t)$, given α, β, and some knowledge of the representation of elements of $F(t)$. The representation and the extension of the other operations from F to $F(t)$ depend on whether t is transcendental over F. This paper concentrates on extension difference fields which are transcendental and have the following additional property.

Definition 5. Let $F(t), \sigma$ be an extension difference field of F, σ. This extension is *first-order-linear* $\Leftrightarrow$

(a) the extension is affine;
(b) t is transcendental over F;
(c) $K(E) = K(F)$ (i.e., the constant field is not extended).

By condition (a), we may think of t as a solution to the first-order-linear equation $\sigma t - \alpha \cdot t = \beta$. Conditions (b) and (c) have to do with the necessity of going to an extension field for a solution. They are considered in more detail later.

Example 3. Let $F = \mathbb{R}$, the real numbers, and let $E = \mathbb{R}((-1)^n \sqrt{-1})$, the difference field version of $\mathbb{C}$, the complex numbers. The extension is affine, because $\sigma(t) = \sigma((-1)^n \sqrt{-1}) = (-1)^{n+1} \sqrt{-1} = -t$. But it is not first-order-linear, because $t^2 = -1 \in \mathbb{R}$, violating condition b. □

Example 4. Let F be an arbitrary field and σ the identity on F, so that $K = F$. Let $E = K(x)$, where x is transcendental over K, and $\sigma x \triangleq x + 1$. We may view E as the field of rational functions from the integers to F, in which case σ is the classical shift operator. We later show that the constant field is not extended, so that the extension is indeed first-order-linear. □

Example 5. Let $F = K(x)$ as above and $\alpha(x) \in F$ be any nonzero element. We may extend F by the "factorial" of $\alpha(x)$, obtaining another function from the integers to F:

$$t(x) \triangleq \prod_{0 \le i < x} \alpha(i).$$

Thus

$$\sigma t(x) = \prod_{0 \le i < x+1} \alpha(i) = \alpha(x) \cdot \prod_{0 \le i < x} \alpha(i) = \alpha(x) \cdot t(x) \qquad \text{(i.e., } \beta = 0\text{)}.$$

Note that $t(x) = x! = 1 \cdot 2 \cdots x$ is of this type, with $\alpha = x + 1$; similarly $t = c^x$, $c \in K$, is of this type, with $\alpha = c$. When doing algebraic symbol manipulation, it is convenient to think of adjoining a symbol "t" to F and, in imitation of how $\prod$ works, simply defining $\sigma t = \alpha \cdot t$, where $\alpha \in F(x)$ is viewed not as a function, but as a symbolic expression. As long as conditions (b) and (c) can be shown to be true (we consider this problem later), we have an algebraic means of manipulating formal products, namely, by first-order-linear extensions in which $\beta = 0$. □

Example 6. Let $F = K(x)$ and $\beta(x) \in F$ be an arbitrary element. We may extend F by the "indefinite sum" of $\beta(x)$. Let

$$t(x) \triangleq \sum_{0 \le i < x} \beta(i).$$

Thus

$$\sigma t(x) = \sum_{0 \le i < x+1} \beta(i) = \beta(x) + \sum_{0 \le i < x} \beta(i) = t(x) + \beta(x).$$

(Of course, Example 3 is a special case of this, where $x = \sum_{0 \le i < x} 1$.)

By analogy to the remarks made in Example 4, in the symbolic case we may think of adjoining a symbol "t" to F and defining $\sigma t = t + \beta$. Once conditions (b) and (c) are taken care of, we have an algebraic means of handling formal sums. □

The following facts are simple to prove but are used over and over.

LEMMA 1. *Let $F(t)$, σ be a first-order-linear extension of F, σ. Then*

(*a*) *σ is an automorphism of $F[t]$.*
(*b*) *Let $f, g \in F[t]$. Then $\sigma\, gcd(f, g) = gcd(\sigma f, \sigma g)$ (with the understanding that gcd's are unique up to a factor from F).*

We have seen in Example 5 that the case in which $\beta = 0$ is of special importance. It would be natural to call such extensions "homogeneous" because t is the solution to the homogeneous equation $\sigma t - \alpha \cdot t = 0$. Our definition of such extensions will certainly include affine extensions in which $\beta = 0$, but we wish to define the term for all extensions, and in a manner which is intrinsic to the extension.

Definition 6. Let E, σ be a difference field extension of F, σ. We say that $g \in E$ is *homogeneous* over $F \Leftrightarrow g \notin F$ but $\sigma g / g \in F$. We say that the extension is *homogeneous* $\Leftrightarrow$ there exists $g \in E$ which is homogeneous over F.

Observe that homogeneous extensions include all those which extend the constant field.

The following result gives various useful (both computationally and mathematically) criteria for determining whether certain extensions are homogeneous.

THEOREM 1. *Let $F(t)$, σ be a difference field extension of F, σ in which $\sigma t = \alpha \cdot t + \beta$. The following conditions are each equivalent to the extension's being homogeneous:*

(*a*) *There exists $g \in F[t]$, $g \notin F$, with $\sigma g / g \in F$.*
(*b*) *The following equation can be solved for $w \in F$:*

$$\sigma w - \alpha \cdot w = \beta.$$

Condition (b) is important both mathematically, because it is used in many later results, and computationally, because it reduces the question of the homogeneity of an extension to a question for which Section 3 gives an algorithm. This condition also allows us to provide a "change of basis," so that extensions which are homogeneous can be written with $\beta = 0$: Given that $\sigma t - \alpha \cdot t = \sigma w - \alpha \cdot w = \beta$, we let $t' = t - w$; so $\sigma t' = \alpha \cdot t'$. Because of this, we hereafter consider only the following type of homogeneous extension.

Definition 7. We say that $F(t)$, σ is a $\prod$-*extension* of F, $\sigma \Leftrightarrow$

(a) the extension is first-order-linear;
(b) $\sigma t = \alpha \cdot t$ (i.e., $\beta = 0$).

We now turn to the question of which homogeneous extensions are also first-order-linear. In other words, we would like to know when the constant field is not extended, and when t is transcendental over F. Our first step in solving this problem is to define a special subset of a difference field.

Definition 8. Given any difference field F, σ, we have the *homogeneous group*,

$$H(F, \sigma) \triangleq \{\sigma g/g \mid 0 \neq g \in F\}.$$

When the automorphism in question is understood, we simply write $H(F)$; when the field also is understood, we write H.

PROPOSITION 3. *The elements of H, together with multiplication from F, form a group (a subgroup of F^*, the multiplicative group of F). Under the assignment $g \mapsto \sigma g/g$, $H \cong F^*/K^*$ (in particular $H(K) = \{1\}$).*

Using the homogeneous group, we are able to obtain a simple (though not yet computational) criterion for the two "technical" conditions required by first-order-linear homogeneous extensions.

THEOREM 2. *Let $F(t)$, σ be a difference field extension of F, σ, and let $\sigma t = \alpha \cdot t$, $\alpha \in F$. This extension is first-order-linear (i.e., does not extend the constant field and t is transcendental) $\Leftrightarrow \alpha^n \notin H(F)$ for all $n > 0$.*

It seems surprising that the properties of extending the constant field or being algebraic are so closely tied, as the above result indicates; more surprising still is the fact that, roughly speaking, the two properties cannot be cleaved within the confines of difference field theory. To illustrate this point, let $F = \mathbb{Q}(2^x)$ so that $2 \in H(F)$; let $t = (\pi + e) \cdot 2^x$. Then $F(t)$ is surely a homogeneous extension of F and either is algebraic over F (with degree 1 if $\pi + e \in \mathbb{Q}$) or extends the constant field ($\pi + e \notin \mathbb{Q}$). But it is currently an open question whether t is transcendental over F, that is, whether $\pi + e$ is transcendental over $\mathbb{Q}$. Clearly, such a problem is outside the domain of difference field theory.

In spite of the fact that the above result is not quite as precise as one might hope for, it has great utility in proving the algebraic independence of factorials, as will become apparent when we analyze the homogeneous group more closely.

Before studying the homogeneous group, we consider inhomogeneous extensions. Here, the situation is much simpler.

THEOREM 3. *Let $F(t)$, σ be an inhomogeneous extension of F, σ in which $\sigma t = \alpha \cdot t + \beta$. Then the extension is first-order-linear.*

This paper considers a slightly restricted class of such extensions.

Definition 9. We say that $F(t)$, σ is a $\sum$-*extension* of F, $\sigma \Leftrightarrow$

(a) the extension is inhomogeneous;
(b) for $n \neq 0$, $\alpha^n \in H \Rightarrow \alpha \in H$, where $\sigma t = \alpha \cdot t + \beta$.

Note that $\sum$-extensions include all inhomogeneous extensions arising from the adjunction of a sum, since in this case $\sigma t = t + \beta$, that is, $\alpha = 1$. As stated earlier, no essential complications are added by considering $\sum$-extensions as opposed to only those extensions in which $\alpha = 1$, that is, extensions arising from the adjunction of a $\sum$.

In summary, this subsection has studied extensions presented in the form $F(t)$, where $\sigma t = \alpha \cdot t + \beta$. We have discovered that certain extensions are homogeneous, that this property may be computed if we can determine whether a certain first-order-linear difference equation has a solution in F, and that given this solution, we may rewrite the extension in a form in which $\beta = 0$. Whether or not such an extension is a $\prod$-extension has been reduced to a question about the homogeneous group. We have also seen that if an extension of the above type is inhomogeneous, then it is automatically first-order-linear. Whether or not it is a $\sum$-extension is again a question about the homogeneous group. We can finally make a precise statement characterizing the fields studied in this paper.

Definition 10. An extension $F(t)$, σ is called a $\prod\sum$-*extension* of F, $\sigma \Leftrightarrow$ it is a $\prod$-extension or a $\sum$-extension. Given a constant difference field K, σ, we say that F, σ is a $\prod\sum$-*field over* $K \Leftrightarrow$ there is a tower of fields,

$$K = F_0 \subset \cdots \subset F_n = F,$$

in which F_i, σ is a $\prod\sum$-extension of F_{i-1}, σ, for $i = 1, \ldots, n$.

The basic result of this paper is that if we know how to compute answers to questions about the homogeneous group and solutions of first-order-linear difference equations in a $\prod\sum$-field F, then

(1) given α, $\beta \in F$, it can algorithmically be determined whether $F(t)$ with $\sigma t = \alpha \cdot t + \beta$ is a $\prod\sum$-extension of F;
(2) if $F(t)$ described in (1) is indeed a $\prod\sum$-extension of F, then the computations concerning the homogeneous group and difference equations may be lifted to $F(t)$.

The construction starts with a constant field, in which certain computations are possible, and is guided solely by the successive choices of the pairs α, β from ever larger fields.

2.3 An Equivalence Relation. In this section we consider the equivalence relation of a difference field extension. This object plays a central role throughout the rest of this paper.

Definition 11. Let E, σ be a difference field extension of F, σ. For nonzero $f, g \in E$, we say that f is *equivalent* to g, written $f \sim g$, $\Leftrightarrow$

$$\frac{\sigma^k f}{g} \in F \qquad \text{for some} \quad k \in \mathbb{Z}.$$

The word "equivalent" is used throughout this paper for this purpose only.

Proposition 4. *The relation* $\sim$ *is an equivalence relation.*

We are interested in computing whether two elements of E are equivalent, and, if they are, we would like to know the possible values of k. The form of the solution is described by the following definitions and result.

Definition 12. The *period* of a nonzero $f \in E$,

$$\operatorname{per}(f) \triangleq \begin{cases} 0 & \text{if } \sigma^p f/f \notin F, \text{ all } p > 0, \\ \min\{p > 0 \mid \sigma^p f/f \in F\} & \text{otherwise.} \end{cases}$$

Definition 13. If $f \sim g$, a *specification of the equivalence* is an integer k such that

(a) $\sigma^k f/g \in F$;
(b) $\operatorname{per}(f) \neq 0 \Rightarrow 0 \leq k < \operatorname{per}(f)$.

We use the following convention, once we know the specification exists and is unique:

$$\operatorname{spec}(f, g) \triangleq \begin{cases} \text{the specification of the equivalence} & \text{if } f \sim g, \\ * & \text{otherwise.} \end{cases}$$

The "$*$" is an arbitrary symbol meaning that $f \not\sim g$.

PROPOSITION 5. *If $f \sim g$, then there exists a unique specification of the equivalence.*

We next define certain functions which arise when dealing with iterated applications of σ.

Definition 14. For $k \in \mathbb{Z}$. $f \in F$,

$$f_{(k,\sigma)} \triangleq \prod_{0 \leq i < k} \sigma^i f \quad \text{(the factorial function)},$$
$$f_{\{k,\sigma\}} \triangleq \sum_{0 \leq i < k} f_{(i,\sigma)}.$$

We let the summation/product convention (introduced in subsection 1.4) determine the values for $k < 0$. If σ is clear from context, it is omitted.

Definition 15. Observe that if $\sigma t = \alpha \cdot t + \beta$, $\sigma^k t$ will be a polynomial in t of degree at most 1. Then

$$\text{def } \alpha_k, \beta_k: \qquad \alpha_k \cdot t + \beta_k = \sigma^k t.$$

We collect in one place various equations involving the above-defined objects.

Identities. For $k, l \in \mathbb{Z}$ (regardless of sign),

(1) $(f \cdot g)_{(k)} = f_{(k)} \cdot g_{(k)}$, $(\sigma f)_{(k)} = \sigma(f_{(k)})$;
(2) $f_{(k+l)} = \sigma^k f_{(l)} \cdot f_{(k)}$;
(3) $f_{(k \cdot l, \sigma)} = (f_{(l,\sigma)})_{(k,\sigma^l)}$;
(4) $f_{(-k)} = \dfrac{1}{\sigma^{-k} f_{(k)}}$;
(5) $\left(\dfrac{\sigma^k f}{f}\right)_{(l)} = \left(\dfrac{\sigma^l f}{f}\right)_{(k)}$;
(6) $\sigma^k f - a_{(k)} \cdot f = a_{(k)} \cdot \sum_{0 \leq i < k} \sigma^i \dfrac{[\sigma f - a \cdot f]}{a_{(i+1)}}$, $a \neq 0$;
(7) $\alpha_k = \alpha_{(k)}$, $\beta_k = \alpha_{(k)} \cdot \sum_{0 \leq i < k} \dfrac{\sigma^i \beta}{\alpha_{(i+1)}}$.

We now start to work on the problem of computing spec in $\prod\sum$-fields. This is not easy, and the path to this result is necessarily indirect.

Definition 16. We say that a difference field F, σ is $\prod$-*regular* $\Leftrightarrow$ given $f, g \in F$, with f not a root of unity (including $f \neq 1$), there exists at most one k such that

$$f_{(k)} = g.$$

We say that the difference field is *computably* $\prod$-*regular* if there exists an algorithm which, given f and g, determines k or declares its nonexistence.

Definition 17. A difference field F, σ is $\sum$-*regular* $\Leftrightarrow$ given $f, g \in F, f \neq 0$, and $f = 1$ or f not a root of unity, there is at most one k such that

$$f_{(k)} = g.$$

We say that F, σ is *computably* $\sum$-*regular* $\Leftrightarrow$ F, σ is $\sum$-regular and there is an algorithm which, given $f, g \in F$, determines k or declares that no such k exists.

Definition 18. Let F, σ be a difference field. Then $H(F)$ is *torsion-free* $\Leftrightarrow$ for all $k \neq 0$ and $f \in H, f^k = 1 \Rightarrow f = 1$.

Using torsion-freeness and (computable) $\prod$- and $\sum$-regularity, we have the following result, whose proof contains the algorithms for computing spec in $\prod$- or $\sum$-extensions. The result also yields some important facts regarding the periods of elements.

THEOREM 4. *Let $F(t), \sigma$ be an extension of F, σ.*

(*a*) *If the extension is a $\prod$-extension, and if F is $\prod$-regular, then the only elements of the extension with nonzero period are of the form $u \cdot t^n$, $n \in \mathbb{Z}$ (which have period 1). If F is computably $\prod$-regular, then spec is computable.*

(*b*) *If the extension is a $\sum$-extension, if F is $\sum$-regular, and if $H(F)$ is torsion-free, then all the elements of the extension have period 0. If F is computably $\sum$-regular, then spec is computable.*

PROOF. Some of the nonalgorithmic aspects of this proof may be found in the corresponding theorem in [11]; these parts are labeled "Note 1," etc.

The first step in this proof is to reduce the question to one of polynomials, regardless of the type of the extension. It can be shown [11, Note 1] that

$$\frac{\sigma^k f}{g} \in F \Leftrightarrow \sigma^k \frac{\operatorname{num}(f)}{\operatorname{num}(g)} \in F \text{ and } \sigma^k \frac{\operatorname{den}(f)}{\operatorname{den}(g)} \in F.$$

If the period and specification of polynomials can be determined, we have a simple means of determining period and specification in $F(t)$, namely,

$$\operatorname{spec}(f, g) = \begin{cases} \operatorname{spec}(\operatorname{num}(f), \operatorname{num}(g)) & \text{if } \operatorname{per}(\operatorname{den}(f)) = \operatorname{per}(\operatorname{den}(g)) = 1, \\ \operatorname{spec}(\operatorname{den}(f), \operatorname{den}(g)) & \text{if } \operatorname{per}(\operatorname{num}(f)) = \operatorname{per}(\operatorname{num}(g)) = 1, \\ k & \text{(see below)}, \\ * & \text{otherwise.} \end{cases}$$

The condition for the third choice above is

$$\operatorname{per}(h) = 0 \qquad \text{where} \quad h = \operatorname{num}(f), \operatorname{den}(f), \operatorname{num}(g), \operatorname{den}(g)$$

and

$$k \triangleq \operatorname{spec}(\operatorname{num}(f), \operatorname{num}(g)) = \operatorname{spec}(\operatorname{den}(f), \operatorname{den}(g)).$$

We now consider the equivalence and period of polynomials in $\prod$-extensions. Let $f, g \in F[t]$, and write

$$f = \sum_{i=0}^{m} v_i t^i, \qquad g = \sum_{i=0}^{n} w_i t^i, \qquad v_m \neq 0, \qquad w_n \neq 0.$$

Then

$$\frac{\sigma^k f}{g} = u \in F \Rightarrow \sum_{i=0}^{m} \sigma^k v_i (\alpha_{(k)} t)^k = \sum_{i=0}^{n} u \cdot w_i t^i \qquad \text{(Identity 6)}$$

$$\Rightarrow m = n \text{ and } \sigma^k v_i \cdot \alpha_{(k)}^i = u \cdot w_i, \qquad i = 0, \ldots, m \qquad \text{(matching coefficients).}$$

Thus the coefficients of f and g must be zero and nonzero in corresponding positions. If only the leading coefficients are nonzero, then clearly $f \sim g$, and the period of each is 1. Assume that there is some $j < n$ with v_j, w_j nonzero. Then, since v_m, w_n are nonzero also, we conclude ([11, Note 2]) that

$$\left(\alpha^{m-j} \cdot \frac{\sigma v}{v}\right)_{(k)} = \frac{w}{v} \qquad \text{where} \quad w \triangleq \frac{w_m}{w_j} \quad \text{and} \quad v \triangleq \frac{v_m}{v_j}.$$

We would like to use $\prod$-regularity of F, but to do so, we must show that $\alpha^{m-j} \cdot \sigma v / v$ is not a root of unity. This is done in [11, Note 3]. Thus we may indeed invoke the $\prod$-regularity of F (computationally, if we desire to compute k) and try to find l such that

$$\left(\alpha^{m-j} \cdot \frac{\sigma v}{v}\right)_{(l)} = \frac{w}{v}.$$

Note that $\prod$-regularity leaves only one possible choice for l, thus proving that any polynomial with more than one nonzero coefficient of t has period 0. This concludes the proof of part (a).

We next consider the equivalence and period of polynomials in $\sum$-extensions. Again, let $f, g \in F[t]$, and expand them as above. It is immediately clear that if $f \sim g$, their degrees are equal; call this degree m. Define the following:

$$v \triangleq \frac{-v_{m-1}}{m \cdot v_m}, \qquad w \triangleq \frac{-w_{m-1}}{m \cdot w_m}, \qquad z \triangleq \frac{\sigma v - \alpha \cdot v - \beta}{\alpha}.$$

It is shown in [11, Note 4] that $z \neq 0$ and that

$$\frac{\sigma^k f}{g} \in F \Rightarrow \left(\frac{\sigma z / z}{\alpha}\right)_{(k)} = \frac{w - v}{z}.$$

We must show that $\alpha \cdot \sigma z / z$ obeys the necessary condition for $\sum$-regularity, that is, that $(\sigma z / z)/\alpha$ is either equal to 1 or is not a root of unity ([11, Note 5]). Thus we can use (computable) $\sum$-regularity of F to try to find l such that

$$\left(\frac{\sigma z / z}{\alpha}\right)_{(l)} = \frac{w - v}{z}.$$

If no such l exists, then k cannot exist, so $f \not\sim g$. Otherwise

$$\operatorname{spec}(f, g) = \begin{cases} l & \text{if } \dfrac{\sigma^l g}{f} \in F, \\ * & \text{otherwise.} \end{cases}$$

As in part (a), there is at most one value of l, proving that every polynomial has period 0. □

To summarize this subsection, we have introduced an equivalence relation, whose importance will become clear later, and have characterized it (and shown how to compute it) in terms of torsion-freeness and (computable) $\prod$- and $\sum$-regularity. The concepts of regularity, particularly $\sum$-regularity, seem abstruse compared to the innocent-looking equivalence relation, and it might seem that little progress has been made. Realize, however, that regularity is intrinsic to a difference field, whereas the equivalence relation was defined for an extension of difference fields. We shall see that torsion-freeness and (computable) regularity can be lifted through $\prod\sum$-extensions, so that they are properties of all $\prod\sum$-fields.

2.4. $\prod$-REGULARITY. This and the next subsection do most of the work in proving the eventual result that $\prod\sum$-fields are $\prod$- and $\sum$-regular. They are not easy, but because they yield the computability of spec (Theorem 4), and because spec is the key to computing answers to questions about the homogeneous group and to solving difference equations, these sections contain the backbone of the theory (and algorithms) of $\prod\sum$-fields. We start at the constant field, first by introducing a property required for computability.

Definition 19. Let K be a field. We say that K has *recognizable powers* $\Leftrightarrow$ there is an algorithm which when given c, $d \in K$ either produces a k such that $c^k = d$ or declares that there is no such k. If c and d are roots of unity, the algorithm yields the smallest positive k, if one exists.

LEMMA 2. *A constant difference field is $\prod$-regular, and computably so, provided that it has recognizable powers.*

PROOF. In a constant field, $f_{(k)} = f^k$ (negative values too of course). Thus

$$f_{(k)} = g \text{ and } f_{(l)} = g \Rightarrow f^k = f^l \Rightarrow f^{k-l} = 1.$$

If $k \neq l$, then f must be a root of unity, proving $\prod$-regularity. The fact that recognizable powers implies computable $\prod$-regularity is just a tautology, in light of $f_{(k)} = f^k$. □

To obtain the results of these two subsections, it is convenient to extend the notion of degree and leading coefficient from $F[t]$ to $F(t)$.

Definition 20. Let $f \in F(t)$.

$$\deg(f) \triangleq \deg(\text{num}(f)) - \deg(\text{den}(f)); \qquad \deg(0) = -\infty;$$
$$\text{lc}(f) \triangleq \text{lc}(\text{num}(f)) \qquad (\textit{leading coefficient}).$$

There are the following trivial facts regarding these functions.

Facts. For $f, g \neq 0$,

(1) $\deg(f \cdot g) = \deg(f) + \deg(g)$,
(2) $\deg(\sigma^k f) = \deg(f)$,
(3) $\deg(f_{(k)}) = k \cdot \deg(f)$,
(4) $\text{lc}(\sigma^k f) = \sigma^k(\text{lc}(f))\alpha_{(k)}^{\deg(f)}$,
(5) $\deg(f) = 0 \Rightarrow \text{lc}(f_{(k)}) = \text{lc}(f)_{(k)}$,
(6) $\deg(f) = \deg(g)$ and $\text{lc}(f) + \text{lc}(g) \neq 0 \Rightarrow \text{lc}(f+g) = \text{lc}(f) + \text{lc}(g)$,
(7) $\deg(f+g) \leq \max(\deg(f), \deg(g))$.

THEOREM 5. *Let $F(t)$, σ be a $\prod\sum$-extension of F, σ. Suppose that F is (computably) $\prod$-regular (and that polynomials over F can be factored). Then $F(t)$, σ is (computably) $\prod$-regular.*

PROOF. Suppose $f, g \neq 0, f$ is not a root of unity. We desire to determine whether a k exists such that

$$f_{(k)} = g.$$

If k exists, we wish to prove that it is unique. We shall follow the convention that if k does not exist, $k = *$.

If $f \in F$, then $f_{(k)} \in F$. If $g \notin F$, then clearly $k = *$; otherwise the solution is immediate by (computable) $\prod$-regularity of F. Thus we may assume $f \notin F$.

Let $m \triangleq \deg(f)$, $n \triangleq \deg(g)$. Then

$$k \neq * \Rightarrow k \cdot m = n.$$

Case 1. $m \neq 0$. Let

$$l \triangleq \frac{n}{m}.$$

If k exists, then $k = l$. In other words,

$$k = \begin{cases} l & \text{if } l \in \mathbb{Z} \text{ and } f_{(l)} = g, \\ * & \text{otherwise.} \end{cases}$$

Case 2. $m = 0$ and $n \neq 0$. Then $k \cdot m$ cannot equal n, so

$$k = *.$$

Case 3. $m = 0$ and $n = 0$. Suppose the extension is homogeneous, and let r, s be the powers to which t occurs in f, g, respectively. Then $k \cdot r$ is clearly the power to which t occurs in $f_{(k)}$, so

$$k \cdot r = s.$$

Case 3.1. $r \neq 0$. As in Case 1, let

$$l \triangleq \frac{s}{r}.$$

Then k may be defined from l as before.

Case 3.2. $r = 0$ and $s \neq 0$. Then $k \cdot r$ cannot equal s, so

$$k = *.$$

Case 3.3. f has no factors with period 1. Let $\mathfrak{f}$ be an irreducible factor of f. (Under computability hypotheses, an $\mathfrak{f}$ may be calculated.) In the following definition we use the convention that $\min \varnothing = +\infty$ and $\max \varnothing = -\infty$.

$$\lambda(h) \triangleq \min\{\operatorname{spec}(\mathfrak{f}, \mathfrak{h}) \mid \mathfrak{f} \sim \mathfrak{h} \text{ and } \mathfrak{h} \text{ is an irreducible factor of } h\},$$
$$\mu(h) \triangleq \max\{\operatorname{spec}(\mathfrak{f}, \mathfrak{h}) \mid \mathfrak{f} \sim \mathfrak{h} \text{ and } \mathfrak{h} \text{ is an irreducible factor of } h\},$$
$$p \triangleq \delta(f), \qquad q \triangleq \mu(f), \qquad r \triangleq \lambda(g), \qquad s \triangleq \mu(g).$$

(Under computability hypotheses, spec is computable by Theorem 4, so that λ and μ are computable.) Note that $p \leq 0 \leq q$ and that $r = +\infty \Leftrightarrow s = -\infty$; in this situation

we clearly cannot have $f_{(k)} = g$ for any k. Otherwise, both $r, s \in \mathbb{Z}$ and $r \leq s$. It is shown in [11] that the only admissible value for k is

$$l \triangleq \begin{cases} s - q + 1 & \text{if } r = p \text{ and } s \geq q, \\ r - p & \text{if } s = q - 1 \text{ and } r < p, \\ * & \text{otherwise.} \end{cases}$$

(Because of the ordering relations on p and q and on r and s, the first two conditions are mutually exclusive.) We obtain k from this l as before. This completes the proof of $\prod$-regularity, since there has never been more than one allowable value for k. We have also seen that under computability hypotheses of the theorem, $F(t)$, σ is computably $\prod$-regular. □

2.5. $\sum$-REGULARITY. As with $\prod$-regularity, we start with the constant field.

LEMMA 3. *A constant difference field is $\sum$-regular, and computably so, provided that K has recognizable powers and that $\mathbb{Z}$ is a computable subset of K.*

PROOF. The fact that recognizable powers implies (computable) $\sum$-regularity for $f \neq 1$ follows from

$$f_{(k)} = g \Rightarrow \frac{f^k - 1}{f - 1} = g \Rightarrow f^k = g(f - 1) + 1.$$

If more than one value of k satisfies this equation, then clearly f is a kth root of unity. If $f = 1$, $f_{(k)} = k$ and computability comes from being able to ask whether $g \in \mathbb{Z}$, allowed by the assumption that integers are a computable subset of K. □

The next result is the crucial part of the induction proof; it shows that if F, σ is a $\prod\sum$-field, then so is F, σ^k for $k \neq 0$. This fact is used in the lifting of $\sum$-regularity.

LEMMA 4. *Let F, σ be $\prod$- and $\sum$-regular, and suppose $H(F)$ is torsion-free. If $F(t)$, σ is a $\prod$-extension (respectively, $\sum$-extension) of F, σ, then for $k \neq 0$, $F(t)$, σ^k is a $\prod$-extension (respectively, $\sum$-extension) of F, σ^k.*

THEOREM 6. *Let $F(t)$, σ be a $\prod\sum$-extension of F, σ. Suppose that for all $i \neq 0$,*

(a) F, σ^i is (computably) $\sum$-regular;
(b) $H(F, \sigma^i)$ is torsion-free;
(c) (polynomials can be factored over F).

Then $F(t)$, σ is (computably) $\sum$-regular.

PROOF. The proof of this theorem is long, so some guideposts to its overall structure are given. We are concerned with the equation

$$\sum_{0 \leq i < k} c_i \cdot f_{(i)} = g,$$

where $c_i \in K$ and not all c_i are 0. (The c_i have other special properties which are needed in [11].)

To show $\sum$-regularity alone, we assume that $k > 0$ and $g = 0$ and try to show that $f \in F$, so that the desired condition regarding f follows by $\sum$-regularity of F, σ. This proof consists of deriving a contradiction from the assumption that $f \notin F$. Most of the details appear in "Notes" given under this theorem in [11]. This case is referred to as the "$g = 0$" case.

To show computable $\sum$-regularity, we assume that $c_i = 1$ for all i but that g is arbitrary, and try to determine a k or show that none exists. Note that if $g = 0$, then

$k = 0$ is the answer; we hereafter assume $g \neq 0$ when proving computable $\sum$-regularity and call this the "$g \neq 0$" case. Note that

$f \in F$ and $g \in F \Rightarrow k$ may be determined by computable $\sum$-regularity in F,
$f \in F$ and $g \notin F \Rightarrow k = *$ (because $f \in F \Rightarrow f_{(k)} \in F$).

Thus we may assume that throughout this proof $k \neq 0$ in the $g \neq 0$ case and $f \notin F$.

In most parts of the proof, the technique is to narrow the possible values of k to a single value l, as we did with computable $\prod$-regularity. We again use the convention $k = *$, and occasionally $l = *$ or l is not an integer. Then

$$k = \begin{cases} l & \text{if } l \in \mathbb{Z} \text{ and } f_{(l)} = g, \\ * & \text{otherwise.} \end{cases}$$

The proof (and algorithm) is broken into its major parts by consideration of the following degrees:

$$m \triangleq \deg(f), \qquad n \triangleq \deg(g).$$

In Part 1 we consider the case in which $m = 0$. Here, after excluding various impossibilities, we consider leading coefficients and use $\sum$-regularity in F to obtain l. Even this involves a slight twist however. In Part 2 we consider certain "easy" cases for determining l, on the basis of those situations in which $\deg(\sum f_{(i)})$ becomes larger as k becomes larger in absolute value. After Part 2 the sign of k is determined, and certain relationships of m and n may be assumed. Part 3 involves examining factors in $\operatorname{den}(f_{(k)})$ and matching them up with factors in $\operatorname{den}(g)$.

Part 1. $m = 0$. In the $g = 0$ case we derive a contradiction in Part 3. The remainder of this part is concerned only with $g \neq 0$ (recall that $c_i = 1$ in this case). On the basis of the consideration that $\deg(\sum f_{(i)}) \leq 0$ (by facts (3) and (7), above), we see that

$$\deg(g) > 0 \Rightarrow k = *.$$

Thus we may assume that $\deg(g) \leq 0$. Now let

$$v \triangleq \operatorname{lc}(f),$$
$$w \triangleq \begin{cases} \operatorname{lc}(g) & \text{if } \deg(g) = 0, \\ 0 & \text{otherwise (i.e., } \deg(g) < 0). \end{cases}$$

By facts (5) and (6),

$$\sum_{0 \leq i < k} v_{(i)} = w.$$

Now if v is 1 or not a root of unity, we may use $\sum$-regularity in F to determine that the above equation is impossible or to determine the unique l such that $v_{(l)} = w$. This l is then the trial value of k, and we proceed as in the introduction.

Thus the only remaining problem is when v is a root of unity other than 1. Using recognizability of powers in K, we may determine the smallest positive p such that $v^p = 1$. Then we must have

$$r \equiv k \bmod p \Rightarrow v_{(k)} = \frac{v^k - 1}{v - 1} = \frac{v^r - 1}{v - 1} = w.$$

Again using recognizability of powers in K, we can determine r by the equation

$$v^r = w \cdot (v - 1) + 1.$$

Of course,

$$r \text{ does not exist} \Rightarrow k = *.$$

Hence we have at least determined k mod p, and we let $k = p \cdot q + r$ and set about trying to find q. By the $g = 0$ case of Part 3 (using $f \notin F$), we will see that $f_{(p)} \neq 0$. This allows the mysterious definitions,

$$\mathsf{g} \triangleq \frac{\sigma^{-r}(g - \sum_{0 \leq i < r} f_{(i)})/f_{(r)}}{f_{(p)}},$$

$$\mathsf{f} \triangleq \frac{f_{(p)} \cdot \sigma f_{(p)}}{f_{(p)}}.$$

It is shown in [11, Note 1 under this theorem] that

$$f_{(p\ q+r)} = g \Rightarrow \mathsf{f}_{(q,\sigma^p)} = \mathsf{g}.$$

Thus we have converted the problem from one in $F(t)$, σ to one in $F(t)$, σ^p for $p > 1$. This may not seem like a reduction, but actually it is. First, note that we have assumed that F, σ^p is computably $\prod$- and $\sum$-regular and that $H(F)$, σ^p is torsion-free. Lemma 4 then says that $F(t)$, σ^p is a $\sum$-extension of F, σ^p, so that the hypotheses of this theorem remain true with σ^p substituted for p. Second, note that $\deg(\mathsf{f}) = 0$, so that when this proof (or algorithm) is used for f and g, it is still Part 1 that applies, and we need not worry about what happens in the rest of this proof. Third, it is shown in [17, Note 2] that lc(f) is either equal to 1 or is not a root of unity, so that a trial l is obtained earlier in Part 1—the proof (or algorithm) does not pass this way twice.

Thus we can determine q in $F(t)$, σ^p. Then

$$l \triangleq \begin{cases} q \cdot p + r & \text{if } q \text{ exists,} \\ * & \text{otherwise.} \end{cases}$$

This completes the determination of k when $m = 0$.

Part 2. $m \neq 0$. The proof of the $g = 0$ case for this part may be found in [17, Note 3]. Observe that the terms of $\sum f_{(i)}$ have distinct degree. Tabulating the various possibilities, we have

$$\deg\left(\sum_i f_{(i)}\right) = \begin{cases} m \cdot (k-1) & \text{if } m > 0, \quad k > 0, \\ 0 & \text{if } m < 0, \quad k > 0, \\ -m & \text{if } m > 0, \quad k < 0, \\ m \cdot k & \text{if } m < 0, \quad k < 0. \end{cases}$$

Case 1. $m > 0$, $n \geq 0$. If $k < 0$, then together with $m > 0$, $\deg(\sum f_{(i)}) = -m < 0 \leq n$. Thus if k exists, $k > 0$, and in fact,

$$m \cdot (k-1) = n.$$

Accordingly, the only possible choice for k is

$$l \triangleq \frac{n}{m} + 1.$$

Case 2. $m < 0$, $n > 0$. If $k > 0$, then together with $m > 0$, $\deg(\sum f_{(i)}) = 0 < n$. Thus if k exists, $k < 0$, and in fact,

$$m \cdot k = n.$$

The only possible choice for k is

$$l \triangleq \frac{n}{m}.$$

Case 3. $m < 0$, $n \leq 0$. If $k < 0$, then together with $m < 0$, $\deg(\sum f_{(i)}) = m \cdot k > 0 \geq n$. Thus $k > 0$. We shall worry about its exact value in Part 3.

Case 4. $m > 0$, $n < 0$. If $k > 0$, then together with $m > 0$, $\deg(\sum f_{(i)}) = m \cdot (k - 1) \geq 0 > n$. Thus $k < 0$. We may convert this into a case 2 or 3 problem by substituting $\mathsf{f} \triangleq f_{(-1)}$ and $\mathsf{g} \triangleq -\sigma(g/\mathsf{f})$; observe that $\deg(\mathsf{f}) = -\deg(f) < 0$ (by identity (4) of Subsection 2.3). It is shown in [11, Note 4] that

$$\sum_{0 \leq i < k} f_{(i,\sigma)} = g \Leftrightarrow \sum_{0 \leq i < -k} \mathsf{f}_{(i,\sigma^{-1})} = \mathsf{g}.$$

Part 3. $k > 0$ and ($m = 0$ ($g = 0$ case) or $s = 0$, $m < 0$, and $n \leq 0$ ($\mathsf{g} \neq 0$ case)). In this part we inductively consider the more complicated equation,

$$\sum_{0 \leq i < k} c_i \cdot \sigma^i h \cdot f_{(i)} = g, \qquad 0 \neq h \in F[t], \quad \text{all factors of } h \text{ have period } 0.$$

Given any irreducible factor $\mathfrak{p}$ of $\mathrm{den}(f)$, period 0, we give a procedure which has one of the following outcomes:

(1) A contradiction is derived for the continuation of Part 1, or the possible values for k are narrowed down to a single choice l for the continuation of Part 2.
(2) It is shown that h, f, g exist such that the above equation can be satisfied for a given $k \Leftrightarrow$ the same equation can be satisfied with h, f, g substituted in it. Further, the hypothesis regarding h is maintained, $\deg \mathsf{f} = \deg f$, and $g = 0 \Leftrightarrow \mathsf{g} = 0$. A reduction occurs because
   (a) $\deg(\mathrm{den}(\mathsf{f})) < \deg(\mathrm{den}(f))$, or
   (b) $\deg(\mathrm{den}(\mathsf{f})) = \deg(\mathrm{den}(f))$ and $\deg \mathsf{h} < \deg h$.

Note that if this reduction procedure cannot be applied, $f \in F$ or all of the factors of $\mathrm{den}(f)$ have period 1, that is, the extension is homogeneous and $\mathrm{den}(f)$ is a power of t. We worry later about what this means, first presenting the reduction (the first time through, $h = 1$).

Step 1. Eliminate the possibility that $\sigma^{-1}\mathfrak{p} \mid h$. If it does, then let

$$\mathsf{f} \triangleq \mathfrak{p} \cdot \frac{f}{\sigma^{-1}\mathfrak{p}} \qquad \text{(note that } \deg(\mathrm{den}(\mathsf{f})) = \deg(\mathrm{den}(f)),\ \deg \mathsf{f} = \deg f\text{)},$$

$$\mathsf{g} \triangleq \frac{g}{\sigma^{-1}\mathfrak{p}} \qquad \text{(note that } g = 0 \Leftrightarrow \mathsf{g} = 0\text{)},$$

$$\mathsf{h} \triangleq \frac{h}{\sigma^{-1}\mathfrak{p}} \qquad \text{(note that } \mathsf{h} \in F[t],\ \deg \mathsf{h} < \deg h, \text{ so outcome (2b) holds)}.$$

It is shown in [11, Note 5] that the equivalence is maintained. By repeatedly applying this step, we may assume henceforth that

$$\mathfrak{p} \mid \mathrm{den}(f) \Rightarrow \sigma^{-1}\mathfrak{p} \nmid h.$$

Step 2. Suppose $\mathfrak{p} \mid \mathrm{den}(f)$ and $\sigma^j\mathfrak{p} \mid \mathrm{num}(f)$ for some $j > 0$. Then let

$$\mathsf{f} \triangleq \frac{f}{\sigma^k\mathfrak{p} \,/\, \mathfrak{p}} \qquad \text{(}\deg(\mathrm{den}(\mathsf{f})) < \deg(\mathrm{den}(f))\text{, so outcome (2a) holds)},$$

$$\mathsf{g} \triangleq g \cdot \mathfrak{p}_{(j)} \qquad \text{(note that } g = 0 \Leftrightarrow \mathsf{g} = 0\text{)},$$

$$\mathsf{h} \triangleq h \cdot \mathfrak{p}_{(j)}.$$

It is shown in [11, Note 6] that the equivalence is maintained. Again, the hypotheses regarding h, f, and g remain true. Thus, by repeatedly applying this step we may assume that

$$\mathfrak{p} \mid \mathrm{den}(f) \text{ and } \sigma^j \mathfrak{p} \mid \mathrm{num}(f) \Leftrightarrow j < 0.$$

Step 3. We use the definition of μ from the proof of $\prod$-regularity.

$$p \triangleq \mu(\mathrm{den}(f)),$$
$$q \triangleq \mu(\mathrm{den}(g)).$$

We also require

$$r \triangleq \max\{i \mid c_i \neq 0\}.$$

By the assumption that the c_i are not all zero, $r \geq 0$; when $s = 0$ in the $g \neq 0$ case, r is of course $k - 1$. We prove in [11, Note 7] that

$$\mu\left(\sum_{0 \leq i < k} c_i \cdot \sigma^i h \cdot f_{(i)}\right) = p + r - 1.$$

For $g = 0$, the existence of k implies $q \in \mathbb{Z}$ (as opposed to $-\infty$), and

$$p + r - 1 = p + k - 2 = q \quad (\text{because } g \neq 0 \Leftrightarrow r = k - 1).$$

Accordingly,

$$l \triangleq \begin{cases} q - p + 2 & \text{if } q \in \mathbb{Z} \text{ and } q > p - 2, \\ * & \text{otherwise.} \end{cases}$$

As for $g \neq 0$, the fact that $\sum c_i \cdot \sigma^i h \cdot f_{(i)}$ has a denominator contradicts the fact that it is equal to 0. Thus step 1 or 2 must have applied.

End of reduction process. We now consider the impact of this reduction process. As pointed out before, the only way that $\mathrm{den}(f)$ can have a nontrivial denominator is if the extension is homogeneous and the denominator is a power of t. Assume that this power is $p \neq 0$ (different p from above). Then, since $t \nmid h$, t occurs to the power $i \cdot p$ in $\mathrm{den}(\sigma^i h \cdot f_{(i)})$, and to the power $r \cdot p$ in $\mathrm{den}(\sum c_i \cdot \sigma^i h \cdot f_{(i)})$, where, as before, r is the largest value $< k$ such that $c_r \neq 0$. If t occurs to the power q (different q from above) in $\mathrm{den}(g)$, then

$$r \cdot p = q.$$

For Part 2 we have as the only possible value of k,

$$l \triangleq \frac{q}{p} + 1 \qquad (\text{because } r = k - 1 \text{ in this case}).$$

Since $\deg(f) < 0$, we *must* have $p \neq 0$ in Part 2, thereby completing the determination of k (at long last). The final contradiction for Part 1 is shown in [11, Note 8]. □

2.6 The Homogeneous Group. In this subsection we answer various questions which have arisen concerning the homogeneous group. We begin with a necessary result about torsion-freeness.

Lemma 5. *Let $F(t)$, σ be a $\prod\sum$-extension of F, σ. Suppose*

(a) F, σ is $\prod$-regular;
(b) $H(F)$ is torsion-free.

Then $H(F(t))$ is torsion-free.

We turn to computational questions. First, the problem of deciding whether an extension where $\sigma t = \alpha \cdot t$ is an $\prod$-extension requires an answer to

Given $f \in F$, does there exist a nonzero $n \in \mathbb{Z}$ such that $f^n \in H(F)$?

Second, the problem of deciding whether an inhomogeneous extension is a $\sum$-extension requires answers to the above question and to

Is $f \in H(F)$?

The fact that F is an arbitrary $\prod\sum$-field seems to force the consideration of the following problem, the solution to which easily suffices to resolve the original questions:

Given $f_1, \ldots, f_k \in F$, describe the set of $n_1, \ldots, n_k \in \mathbb{Z}$ such that $f_1^{n_1} \cdots f_k^{n_k} \in H(F)$.

Convention. $\mathbf{f}^{\mathbf{n}} \triangleq f_1^{n_1} \cdots f_k^{n_k}$. We do not reserve k for the length of $\mathbf{f}$.

One might worry that there is an infinite set of $\mathbf{n}$ such that $\mathbf{f}^{\mathbf{n}} \in H$; however, the set of these values has a convenient algebraic structure which makes a finite description possible.

Definition 21. $M(\mathbf{f}, F) \triangleq \{\mathbf{n} \mid \mathbf{f}^{\mathbf{n}} \in H\}$.

LEMMA 6. *Let* $\mathbf{f} \in F^\omega$. *Then* $M(\mathbf{f}, F)$ *forms a submodule of* $\mathbb{Z}^{|f|}$ (*the underlying ring is* $\mathbb{Z}$).

Since $M(\mathbf{f}, F)$ is a submodule of a finite-dimensional free module, it itself is free [12, Th. X.12, p. 358] and can be described simply by giving a basis for its elements. Thus we may further refine the above question as follows:

Given $\mathbf{f} \in F^\omega$, compute a basis for $M(\mathbf{f}, F)$.

We shall be able to do this in arbitrary $\prod\sum$-fields.

We first consider the problem in a constant field. This paper does not consider how to go about making computations in K, but it states precisely the requirements for such computations. Since $H(K) = \{1\}$, we have the following.

Definition 22. We say a field K has *recognizable powers* $\Leftrightarrow$ there is an algorithm which, given $\mathbf{c} \in K^\omega$, produces a basis for the set of all $\mathbf{n} \in \mathbb{Z}^\omega$ such that

$$\mathbf{c}^{\mathbf{n}} = 1.$$

We have taken the liberty of reusing the term "recognizable powers" of the previous section, because the ability required in Definition 22 above easily implies the ability required in Definition 19. This problem is not difficult to solve in $\mathbb{Q}$ or in any transcendental extension of a field in which it is already solved. Lifting this ability through arbitrary algebraic extensions is an open problem.

The key to lifting the computability of a basis for M is a certain canonical form for the elements of $F(t)$; this form relies heavily on the equivalence relation $\sim$. We first introduce this form, prove that it is canonical, and show how to compute it. Then we

use it in computing a basis for M. Throughout this section, we assume that $F(t)$, σ is a $\prod$- or $\sum$-extension of F, σ.

Definition 23. Let $\mathbf{f} \in F(t)^{\omega}$, and suppose that

(σF0) $f_i = u_i \cdot t^{e_i} \cdot \prod_{j,k} \sigma^k f_j^{e_{ijk}}$, $u_i \in F$, $f_j \in F[t]$.

We say that $\langle \langle f_j \rangle_j, \langle u_i, e_i, \langle e_{ijk} \rangle_{j,k} \rangle_i \rangle$ is a σ-factorization of $\mathbf{f} \Leftrightarrow$

(σF1) f_j is monic, has degree >0, and is irreducible;
(σF2a) $t \neq f_j$ if the extension is homogeneous;
(σF2b) $e_i = 0$ if the extension is inhomogeneous;
(σF3) $j_1 \neq j_2 \Rightarrow f_{j_1} \not\sim f_{j_2}$;
(σF4) for all j, there exists some i, k such that $e_{ijk} \neq 0$.

The idea behind this definition is that equivalent irreducible factors are grouped together (with a special case made for the one possible factor with nonzero period).

THEOREM 7. *Let F, σ be $\prod$- and $\sum$-regular. Then there exists a σ-factorization of every element of $F(t)^{\omega}$, and it is unique up to the u_i, permutation of the f_j, and translation of the last index of e_{ijk}. The σ-factorization may be computed, provided that there is an algorithm for factoring polynomials over F, and that spec for the extension is a computable function.*

PROOF. Uniqueness is proved in [11]; we examine here the existence (or computation) question. Consider (or calculate) the set of all irreducible factors of the f_i. Any factors of period 1 determine e_i for f_i. All other factors have period 0. Group these into equivalence classes under $\sim$, and pick an arbitrary representative of each equivalence class, calling these representatives f_j (this may be computed if spec is computable). Once the set f_j is fixed, all of the factors of period 0 may be written in the form $\sigma^k f_j$ (perhaps multiplying by an element of F). Then e_{ijk} is simply the power to which $\sigma^k f_j$ occurs in f_i (by convention, e_{ijk} is 0 if $\sigma^k f_j$ is not a factor of f_i). The u_i are determined by dividing f_i by t^{e_i} and $\sigma^k f_j^{e_{ijk}}$ for appropriate j and k; this forces $u_i \in F$, and guarantees σF0. The conditions σF1–σF4 are clear from the construction. □

We now come to the main result of this section, which characterizes $M(\ldots, F(t))$ in terms of $M(\ldots, F)$ and various functions of the components of a σ-factorization.

THEOREM 8. *Let $\mathbf{f} \in F(t)^{\omega}$ have a σ-factorization, denoted as in Definition* 22. *Then*

$$M(\mathbf{f}, F(t)) = M_1 \cap M_2 \cap M_3,$$

where

(*a*) $M_1 \triangleq \begin{cases} M(\mathbf{u}, F) & \text{in the case of a } \sum\text{-extension,} \\ \{\mathbf{n} \mid \mathbf{n} \wedge d \in M(\mathbf{u} \wedge 1/\alpha, F)\} & \text{otherwise,} \end{cases}$

(*b*) $M_2 \triangleq Ann(\mathbf{e})$,

(*c*) $M_3 \triangleq \bigcap_j Ann \left\langle \sum_k e_{ijk} \right\rangle_i$.

Furthermore, M_1, M_2, and M_3 are each submodules of $\mathbb{Z}^{|\mathbf{f}|}$.

We do not discuss in this paper how to obtain the basis for an annihilator, given the elements to be annihilated; nor how to obtain the basis for M_1, given that for

$M(\mathbf{u} \wedge \langle 1/\alpha \rangle, F)$; nor how to compute the basis for the intersection of two modules, given the basis for each. These are problems in standard linear algebra, which do not concern us here. Given techniques for those problems and the ability to compute σ-factorization, it is clear that the computability of a basis for $M(\mathbf{f}, \ldots)$ may be lifted through $\prod\sum$-extensions, which was the goal we set for ourselves in this section.

We now give several examples which illustrate the combined power of the above theorem and Theorem 2.

Example 7. Consider the difference field $\mathbb{Q}(x)$, where $\sigma x = x + 1$; let $E = \mathbb{Q}(x)(p_1^x, p_2^x, \ldots)$, where p_i is the ith prime. We show that $\{x\} \cup \{p_i^x\}_i$ is an algebraically independent set (not a surprising result, of course—what relationships could there possibly be?), in particular, that p_k^x is transcendental over $F_{k-1} \triangleq \mathbb{Q}(x)(p_1^x, \ldots, p_{k-1}^x)$. To begin with, if 2^x is algebraic over $\mathbb{Q}(x)$, then $2^n \in H(\mathbb{Q}(x))$ for some n, by Theorem 2. In Theorem 8, let $\mathbf{f} = \langle 2 \rangle$, $F = \mathbb{Q}$, and $t = x$. In the σ-free factorization of $\mathbf{f}$, j ranges over the null set, and we have $u_1 = 2$; also, $e_1 = 0$. Thus M_2 and M_3 of the theorem are both $\mathbb{Z}$ (considered as a module over itself), and M_1 is simply $M_{\langle 2 \rangle}$ in $\mathbb{Q}$. But

$$\{n \mid 2^n \in H(\mathbb{Q})\} = \{n \mid 2^n = 1\} = \{0\}.$$

Hence $2^n \notin H(\mathbb{Q}(x))$ for $n > 0$, and 2^x is transcendental over $\mathbb{Q}(x)$. (We could arrive at the same conclusion "analytically" because 2^x grows too fast as $x \to \infty$.)

Inductively, assume that p_l^x is transcendental over F_{l-1} for $l < k$. By Theorem 2, p_k^x is algebraic over F_{k-1} only if $p_k^{n_k} \in H(F_{k-1})$ for $n_k > 0$. This leads to a contradiction. Let $\mathbf{p} \triangleq \langle p_{l+1}, \ldots, p_k \rangle$, $\mathbf{n} \triangleq \langle n_{l+1}, \ldots, n_k \rangle$. Starting at $l = k - 1$, we may inductively assume that $\mathbf{p}^{\mathbf{n}} \in H(F_l)$. In Theorem 7 let $f_l = p_l^x$, $F = F_{l-1}$. In the σ-factorization of $\mathbf{f}$, j ranges over the null set, $\mathbf{u} = \mathbf{p}$, and $\mathbf{e} = 0$. Thus $M_2 = M_3 = \mathbb{Z}^{|\mathbf{f}|}$ (i.e., yields no information), and $M_1 = M(\mathbf{p} \wedge \langle 1/p_l \rangle)$. Thus there exists $n_l \in \mathbb{Z}$ such that $p_l^{n_l} \cdot \mathbf{p}^{\mathbf{n}} \in H(F_{l-1})$. Thus we have decreased l by 1 in the induction process. When $l = 0$, we have $\mathbf{p} = \langle p_1, \ldots, p_k \rangle$, $\mathbf{n} = \langle n_1, \ldots n_k \rangle$, with $\mathbf{p}^{\mathbf{n}} \in H(\mathbb{Q}(x))$. Here we use Theorem 4 in the inhomogeneous case and obtain $\mathbf{p}^{\mathbf{n}} \in H(\mathbb{Q})$, that is, $\mathbf{p}^{\mathbf{n}} = 1$. But since the p_i are all primes, this forces $\mathbf{n} = 0$, in particular $n_k = 0$. Thus p_k^x is transcendental over F_{k-1}, so that $\{x, p_1^x, \ldots, p_k^x\}$ is an algebraically independent set for any k. □

Example 8. Again consider $\mathbb{Q}(x)$, but this time extend $\mathbb{Q}(x)$ by $x!$, so $\sigma x! = (x + 1)x!$. This extension is algebraic only if $(x + 1)^n \in H(\mathbb{Q}(x))$, for some n. Apply Theorem 8 with $\mathbf{f} = \langle x + 1 \rangle$, $t = x$, $F = \mathbb{Q}$. In the σ-factorization of $\mathbf{f}$,

$$f_1 = x + 1, \qquad u_1 = 1, \qquad e_1 = 0, \qquad e_{110} = 1.$$

Condition (c) of Theorem 8 requires that $n \cdot e_{110} = 0$. Since $e_{110} \neq 0$, we conclude that $n = 0$, that is, $M(\mathbf{f}, \mathbb{Q}(x)) = M_1 \cap M_2 \cap \{0\} = \{0\}$, without even calculating $M_1 \cap M_2$. Thus $x!$ is transcendental over $\mathbb{Q}(x)$.

Consider *any* extension $F(x)$ which is inhomogeneous. Then using exactly the same proof, we see that $x!$ is transcendental over $F(x)$. Since we showed in Example 7 that $\mathbb{Q}(p_1^x, \ldots)(x)$ is an inhomogeneous extension of $\mathbb{Q}(p_1^x, \ldots)$ (because the set $\{x, p_1^x, p_2^x, \ldots\}$ is algebraically independent), we can conclude that $\{x, x!, p_1^x, \ldots\}$ is an algebraically independent set, as one would expect. □

2.7. Review. The purpose of this section has been to introduce $\prod\sum$-fields and prove basic facts about them. We have seen that a constant field K is vacuously a

$\prod\sum$-field over itself. More important, given any $\prod\sum$-field F, σ and elements $\alpha, \beta \in F$, we have considered an extension $F(t)$, σ, where $\sigma t = \alpha \cdot t + \beta$, and have been able to characterize whether it is a $\prod$-extension (possibly by a "change of basis"), a $\sum$-extension, or neither.

In the process of characterizing $\prod\sum$-fields it has been necessary to study various properties which all such fields have. The crucial work in doing this has been scattered throughout previous subsections; the following result concisely summarizes these results, and its proof [11] ensures that all the pieces fit together as intended. Note that all the results of this chapter are tied together in an inductive way along the tower of a $\prod\sum$-field.

THEOREM 9. *Let F, σ be a $\prod\sum$-field over a constant field K. Then*

(*a*) *F, σ^k is a $\prod\sum$-field over K whenever $k \neq 0$.*
(*b*) *F, σ is $\prod$-regular and $\sum$-regular.*
(*c*) *$H(F, \sigma)$ is torsion-free.*
(*d*) *Let $F(t)$, σ be a $\prod\sum$-extension of F, σ. All elements of $F(t)$, σ have periods of 1 or 0; the only elements of period 1 are either in F or, in the case of a $\prod$-extension, are of the form $u \cdot t^n$, $u \in F$.*

Suppose that K has the following properties:

(*i*) *Polynomials in several variables may be factored over K.*
(*ii*) *K has recognizable powers (by Definition 22).*
(*iii*) *The integers are a computable subset of K.*

Then we can also conclude that

(*e*) *F, σ is computably $\prod$- and $\sum$-regular.*
(*f*) *For $F(t)$, σ a $\prod\sum$-extension of F, σ, spec is a computable function.*
(*g*) *Given $\mathbf{f} \in F^\omega$, a basis for $M(\mathbf{f}, F)$ may be calculated.*

The only question of computation which remains concerning $\prod\sum$-extensions is the ability either to solve the first-order linear equation $\sigma g - a \cdot g = f$ in F, σ or to declare that a solution does not exist. This will be answered in the next section.

3. *Solutions*

3.1. THE GENERAL EQUATION. We mentioned earlier that $\sigma g - g = f$ is not a convenient equation to consider when we try to lift solvability through $\prod\sum$-extensions. There are two ways in which the equation must be generalized. First, we are forced to consider general first-order-linear equations $\sigma g - a \cdot g = f$; the need for solving these arose anyway in the characterization of $\prod\sum$-extensions. Even more generally, it is necessary to consider the following problem:

Given $a \in F$, $\mathbf{f} \in F^\omega$, solve $\sigma g - a \cdot g = \mathbf{c} \cdot \mathbf{f}$ for $g \in F$, $\mathbf{c} \in K^\omega$.

As part of the inductive process, we are interested in "all" solutions to $\sigma g - a \cdot g = \mathbf{c} \cdot \mathbf{f}$. The first thing to notice about "all" these solutions is that they have a familiar algebraic structure.

Definition 24. Let $S \subseteq F$, $a \in F$, $\mathbf{f} \in F^\omega$. The *solution space* for a, $\mathbf{f}$ in S,

$$V(a, \mathbf{f}, S) \triangleq \{\mathbf{c} \frown g \in K^{|\mathbf{f}|} \times S \mid \sigma g - a \cdot g = \mathbf{c} \cdot \mathbf{f}\}.$$

Observe that $\mathbf{f}$ is not necessarily in S^ω.

PROPOSITION 6. *If S is a vector space over K, then for any a, $\mathbf{f}$, $V(a, \mathbf{f}, S)$ is a vector space over K, with dimension $\leq |\mathbf{f}| + 1$.*

Definition 25. Let S be a vector space over K, $a \in F$, $\mathbf{f} \in \mathbf{F}^\omega$. A *solution basis* for a, $\mathbf{f}$ in S is by definition a basis for $V(a, \mathbf{f}, S)$. Elements of this basis are usually denoted $\mathbf{c}_i \wedge g_i$, and the totality of the elements may be written $\mathbf{C} \wedge \mathbf{g}$ where $\mathbf{C}$ is a matrix whose ith row is $\mathbf{c}_i$, and $\mathbf{g}$ a column vector whose ith element is g_i (note abuse of "$\wedge$" notation). If $V(a, \mathbf{f}, S) = \{0\}$, then a basis for it is $\mathbf{0}_{0,|\mathbf{f}|} \wedge \mathbf{0}_0$.

This section has many worked examples in which we need conventions for matrices and vectors. Matrices are written with parentheses in the usual fashion; I_n is the identity matrix of size n; for example,

$$I_1 = (1), \qquad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Vectors are throught of as column vectors (for convenience in multiplying by matrices), but the tuple convention may be used for typographical convenience.

$$\langle 1, 2 \rangle = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

When the vector has a length of one, the notations are interchangeable: $\langle 1 \rangle = (1)$.

In the midst of all this generality, it might be well to point out how knowledge of the solution basis tells us whether f is summable. Suppose we have a basis for $V(1, \langle f \rangle, F)$. This space will *always* include $\langle 0, 1 \rangle$ (because $\sigma 1 - 1 \cdot 1 = 0 \cdot f$) and so will be at least one-dimensional. If and only if the dimension is two, we will have nonzero c and some g such that $\sigma g - 1 \cdot g = c \cdot f$. Since c is nonzero, $\sigma(g/c) - 1 \cdot (g/c) = f$, and f is summable.

Our first result concerning solutions is for the constant field, where the recursion must stop.

THEOREM 10. *Given* $a \in K$, $\mathbf{f} \in K^\omega$,

$$V(a, \mathbf{f}, K) = Ann_K(\mathbf{f} \wedge (a - 1)).$$

Given a vector, there are well-known ways to produce a basis of its annihilator; these will not be discussed here.

It is convenient to introduce at this time a method for transforming the a in the equation $\sigma g - a \cdot g = \mathbf{c} \cdot \mathbf{f}$.

LEMMA 7. *Let* $a \in F$, $\mathbf{f} \in F^\omega$, *and let* $a_0 \in F$ *be nonzero. Then*

$$\mathbf{C} \wedge \mathbf{g} \text{ is a basis for } V(a \cdot (\sigma a_0 / a_0), \sigma a_0 \cdot \mathbf{f}, F)$$
$$\Rightarrow \mathbf{C} \wedge \mathbf{g}/a_0 \text{ is a basis for } V(a, \mathbf{f}, F).$$

In words, we may modify an equation by multiplying a by an element of $H(F)$ and adjusting $\mathbf{f}$ accordingly, and thus obtain the basis for the original equation from that for the modified equation.

3.2. REDUCTION TECHNIQUES. In the language of solution spaces, the problem we are trying to solve is: Given the ability to find a basis for $V(u, \mathbf{v}, F)$, given arbitrary $u, v_i \in F$, devise a method to find a basis for $V(a, \mathbf{f}, F(t))$. This problem breaks down naturally into a series of smaller problems of the following type. Note that there are infinitely many vector spaces U over K lying between $\{0\}$ and $F(t)$. Depending upon a and $\mathbf{f}$, we judiciously pick a finite number of these and consider

the following tower:

$$F(t) = U_n \supseteq U_{n-1} \supseteq \cdots \supseteq U_0 = \{0\}.$$

Then, for $i = n, \ldots, 1$, we reduce the problem of finding a basis for $V(\ldots, \ldots, U_i)$ to that of finding a basis for $V(\ldots, \ldots, U_{i-1})$, perhaps using a solution from $V(\ldots, \ldots, F)$. This subsection considers various general properties of the reduction process, and the rest of Section 3 is concerned with particular U_i.

Our first result describes what happens at the end of the reduction process.

PROPOSITION 7. *Given* $a \in F(t)$, $\mathbf{f} \in F(t)^\omega$,

$$V(a, \mathbf{f}, \{0\}) = Ann_K(\mathbf{f}) \times \{0\}.$$

The techniques of this section rely on some measure of the "complexity" of an element of $F(t)$. There are many such measures, depending upon the particular U_i/U_{i-1} in which we are interested. These measures can be defined so that they all have certain general properties.

Definition 26. Let W be a vector space over a field K, and let $\| \ \|$ be a map from W to the integers. We say that $\| \ \|$ *grades* $W \Leftrightarrow$

(G1) $\|f + g\| \leq \max(\|f\|, \|g\|)$ for $f, g \in W$.
(G2) $\|c \cdot f\| \leq \|f\|$ for $c \in K, f \in W$.

Given such a function, the *grades* of W are

$$W_m \triangleq \{f \in W \mid \|f\| \leq m\}.$$

We use the following notational convention for $\mathbf{f} \in W^\omega$:

$$\|\mathbf{f}\| \triangleq \max_i \|f_i\|.$$

Typically, the range of $\| \ \|$ is the set of all integers greater than some given integer.

PROPOSITION 8. *Let* $\| \ \|$ *grade* W. *Then the grades of* W *form an ascending chain of vector spaces whose limit is* W; *that is,*

$$\cdots \subseteq W_{m-1} \subseteq W_m \subseteq \cdots \to W.$$

Conversely, given such a sequence of vector spaces, let

$$\|f\| \triangleq min\{m \mid f \in W_m\}.$$

Then $\| \ \|$ *grades* W.

We have the following additional properties for any grading of a vector space:

(*a*) $\|\mathbf{c} \cdot \mathbf{f}\| \leq \|\mathbf{f}\|$ *for* $\mathbf{c} \in K^\omega, \mathbf{f} \in W^\omega$;
(*b*) $\|f_1\| > \|f_2\| \Rightarrow \|f_1 + f_2\| = \|f_1\|$ *for* $f_1, f_2 \in W^\omega$.

Our first reduction result is the simplest; it tells when two solution spaces are the same, even though one of them may be relative to a much larger W than the other. This result is used to "bound" $\|g\|$.

THEOREM 11. *Let* W *be a vector space graded by* $\| \ \|$, *let* $f \in W$, *and suppose there is some* m *such that*

$$\sigma g - a \cdot g = f \Rightarrow \|g\| \leq m.$$

Then

$$V(a, \mathbf{f}, W) = V(a, \mathbf{f}, W_m).$$

Let $W_1 \supseteq W_2$ be vector spaces between $F(t)$ and F. There are infinitely many ways of choosing a subspace of W_1 isomorphic to the quotient space of W_1 by W_2. In this section, we use the symbol W_1/W_2 to refer to both the quotient space and to a specific (but arbitrary) subspace, with the distinction made clear by context. We use the fact from linear algebra that

$$W_1 \cong W_2 \oplus W_1/W_2.$$

Once a specific $W_1/W_2 \subseteq W_1$ is fixed, the $\cong$ may be replaced by equality.

In this section, it will always be the case that there exists some $l \in \mathbb{Z}$ such that for any g,

$$\|\sigma g - a \cdot g\| \leq \|g\| + l.$$

This number is known in any given application and is used in defining the object at the heart of the reduction process.

Definition 27. Let $W_m \supseteq W_{m-1}$ be vector spaces over K; let $a \in F(t)$, $\mathbf{f} \in W_{m+l}^{\omega}$. The incremental solution space for $a, \mathbf{f}$ relative to a fixed W_m/W_{m-1} is

$$V(a, \mathbf{f}, W_m/W_{m-1}) \triangleq \{\mathbf{c} \wedge g \mid \mathbf{c} \in K^{\omega}, g \in W_m/W_{m-1}, \exists g_1 \in W_{m-1} \text{ with } \sigma(g + g_1) - a \cdot (g + g_1) - \mathbf{c} \cdot \mathbf{f} \in W_{m+l-1}\}.$$

Observe that this object depends not only upon the choice of $W_m/W_{m-1} \subseteq W_m$, but also upon l, which is too much to bother with notationally.

PROPOSITION 9. *Let $m, l, a, \mathbf{f}$ be as in Definition* 27. *Then $V(a, \mathbf{f}, W_m/W_{m-1})$ is a finite-dimensional vector space over K.*

Since an incremental solution space is a finite-dimensional vector space over K, it can be represented by its basis, just as a solution space may be. The following result shows how to compute a basis for a solution space of a larger space, given a basis for a smaller space and a basis for the incremental solution space between the two spaces.

THEOREM 12. *Let $m, l, a, \mathbf{f}$ be as above. Perform the following construction.*

(i) Let $\mathbf{C} \wedge \mathbf{g}$ be a basis for $V(a, \mathbf{f}, W_m/W_{m-1})$.
(ii) Let $\mathbf{f}_{m-1} \triangleq \mathbf{Cf} - (\sigma\mathbf{g} - a \cdot \mathbf{g})$ (observe that $\mathbf{f}_{m-1} \in W_{m+l-1}^{\omega}$).
(iii) Let $\mathbf{D}_{m-1} \wedge \mathbf{h}_{m-1}$ be a basis for $V(a, \mathbf{f}_{m-1}, W_{m-1})$.
(iv) $\mathbf{D}_m \triangleq \mathbf{D}_{m-1}\mathbf{C}$, $\mathbf{h}_m \triangleq \mathbf{D}_{m-1}\mathbf{g} + \mathbf{h}_{m-1}$.

Then

$$\mathbf{D}_m \wedge \mathbf{h}_m \quad \text{is a basis for} \quad V(a, \mathbf{f}, W_m).$$

In some cases an incremental solution basis may be particularly easy to compute. One such example is the following.

THEOREM 13. *Let $m, l, a, \mathbf{f}$ be as above. Suppose that for $g \in W_m/W_{m-1}$,*

$$\|\sigma g - a \cdot g\| < m + l \Leftrightarrow g = 0.$$

Suppose also that there exists $\mathbf{g} \in (W_m/W_{m-1})^{\omega}$ with

$$\|\mathbf{f} - (\sigma\mathbf{g} - a \cdot \mathbf{g})\| < m + l.$$

Then

$$\mathbf{I} \wedge \mathbf{g} \qquad \textit{is a basis for} \qquad V(a, \mathbf{f}, W_m/W_{m-1}),$$

where I is the identity matrix.

3.3. THE POLYNOMIAL PART. In this subsection we consider a vector space W with the property that

$$F[t] \subseteq W \subseteq F(t).$$

The goal is to "eliminate" the polynomial part of the solution. The reduction process for doing this is based upon the following obvious grading.

Definition 28. Let $f \in F(t)$, and let f_0 be the polynomial part of f. Then

$$\|f\| \triangleq \begin{cases} -1 & \text{if } f_0 = 0, \\ \deg f_0 & \text{otherwise.} \end{cases}$$

The first two results of this section are concerned with bounding $\|g\|$ in the case where $\|a\| = 0$. In most cases, Theorem 11 may be used in bounding $\|g\|$, but here the problem is more subtle, and we must use other means of finding a bound. It is necessary to treat $\prod$- and $\sum$-extensions separately.

THEOREM 14. *Let $F(t)$, σ be a $\sum$-extension of F, σ, and let $a \in F(t)$ with $\|a\| = 0$ (i.e., the polynomial part of a is a nonzero element of F). For any $f \in W$, if there exists $g \in W$ such that $\sigma g - a \cdot g = f$ and $\|g\| > \|f\| + 1$, then*

(*a*) *$a \notin F$, so we may write*

$$a = u \cdot \frac{t^p + u_1 t^{p-1} + \cdots}{t^p + u_2 t^{p-1} + \cdots} \qquad \textit{where} \quad p > 0, \quad u, u_1, u_2 \in F;$$

(*b*) $u_1 \neq u_2$;

(*c*) *there exists a unique $m \in \mathbb{Z}$, $m > \|f\| + 1$, such that for some $w \in F$,*

$$\sigma w - \alpha \cdot w = \alpha(u_1 - u_2) - m \cdot \beta;$$

(*d*) $\|g\| \leq m$.

Furthermore, the existence of the m of part (c), *and its value if it exists, may be calculated provided that*

(*i*) *a basis for $V(\ldots, \ldots, F)$ may be computed; and*

(*ii*) $\mathbb{Z}$ *is a computable subset of K.*

PROOF. Most of the proof is in [11]. We examine here only the issue of calculating m. To do so, first calculate a basis $\mathbf{C} \wedge \mathbf{w}$ for $V(\alpha, \langle \alpha \cdot (u_2 - u_1), -\beta \rangle, F)$. Without loss of generality we may assume that $\mathbf{C} \wedge \mathbf{w}$ is in reduced row echelon form. If the only row of the matrix is of the form $(0\ 0\ w)$, then the equation of part (c) cannot be satisfied for $w \in F$, $m \in K$, let alone $m \in \mathbb{Z}$. If the matrix has a row of the form $(0\ 1\ w)$, then $\sigma(-w) - \alpha \cdot (-w) = \beta$ and the extension is homogeneous, by Theorem 1, contradicting our assumption. We may thus assume that $\mathbf{C} \wedge \mathbf{w}$ has a row of the form $(1\ c\ w)$, $c \in K$, where the other row, if any, has zeros in the first two columns. If $c \in \mathbb{Z}$, part (c) will be satisfied $\Leftrightarrow c > \|f\| + 1$, and of course, $m = c$. If $c \notin \mathbb{Z}$, then no such m exists. □

Realize what a strange thing is going on here. The ability to find a solution basis in a smaller field is being used not to find "part" of the solution (e.g., a basis for

some incremental solution space), but rather to find (or disprove the existence of) the integer m satisfying part (c). Note that what we would ordinarily think of as the solution to the difference equation, w, is not used anywhere else in the result.

Using this result, we can calculate an m such that

$$\sigma g - a \cdot g = f \Rightarrow \|g\| \leq m,$$

namely, m is either determined in part (c) above, or $m \triangleq \|f\| + 1$. Thus, for $\|a\| = 0$ and inhomogeneous extensions, we have reduced the problem of W to W_m, for a finite m.

Example 9. Let $F = \mathbf{Q}$, the rational numbers, and let $F(t) = \mathbf{Q}(x)$, the rational functions of x, so that $\sigma x = x + 1$. Let $a = (x^2 + 4x + 7)/(x^2 + 1)$ and $f = 1$, so we are considering the difference equation,

$$\sigma g - \frac{x^2 + 4x + 7}{x^2 + 1} \cdot g = -13.$$

Applying Theorem 14, observe that

(i) Part (a) is satisfied since $a \notin F$. The definitions require

$$u = 1, \qquad u_1 = 4, \qquad u_2 = 0.$$

(ii) Part (b) is satisfied since $4 \neq 0$.

Using the proof of the result, we want a basis for $V(1, \langle 4, -1\rangle, \mathbf{Q})$. By Theorem 10 we need $\mathrm{Ann}(\langle 4, -1, 1 - 1\rangle)$. Linear algebra algorithms yield the following basis in the required form:

$$\begin{pmatrix} 1 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using the notation from the proof of Theorem 14, $c \triangleq 4$, so $c \in \mathbf{Z}$ and $c > \|f\| + 1 = 1$. Thus we let $m \triangleq 4$, and by part (d) we know that the polynomial part of g has degree ≤ 4. □

This example shows why the generality of Theorem 14 is required (we will eventually see that the difference equation has a polynomial solution of degree 4). When dealing with the difference equation arising directly from a sum, the following result suffices.

COROLLARY 1. *Let $F(t)$, σ and F, σ be as above. If there exists a $g \in W$ such that $\sigma g - g = f$, then $\|g\| \leq \|f\| + 1$.*

PROOF. In this case, $a = 1 \in F$, and we may use Theorem 14(a). □

We come to the corresponding result in $\prod$-extensions.

THEOREM 15. *Let $F(t)$, σ be a $\prod$-extension of F, σ. Let $a \in F(t)$, with $\|a\| = 0$ and $u \in F$ the polynomial part of a. For any $f \in W$, if there exists $g \in W$ such that $\sigma g - a \cdot g = f$ and $\|g\| > \|f\|$, then*

(a) there exists a unique $m \in \mathbf{Z}$, $m > \|f\|$, such that

$$\frac{u}{\alpha^m} \in H(F);$$

(b) $\|g\| \leq m$.

The existence and value, if any, of the m of part (a) may be calculated, provided that a basis for $M(\ldots, F)$ may be calculated.

PROOF. We consider here only the problem of calculating m; the rest of the proof is in [11]. Given a basis for $M(\langle u, 1/\alpha\rangle)$, we may assume that there is at most one row which is nonzero in the first column. If this row has ± 1 in the first column (without loss of generality, say $+1$), and if the value in the second column is $> \|f\|$, then the value in the second column is necessarily m. If there is no such row, m does not exist. □

Again, we have succeeded in reducing the problem from W to W_m, given the ability to perform certain calculations in F, σ.

Example 10. Let $F = \mathbb{Q}$, and think $t = 2^x$; so $\sigma t = 2 \cdot t$. Consider the equation

$$\sigma g - 1048576 \cdot g = 0.$$

The method of Theorem 15 requires a basis for $M(\langle 1048576, 1/2\rangle)$. By the results of Section 2, this is a problem of recognizable powers in $\mathbb{Q}$; the module has the basis $(1, 20)$. Since $20 > \|0\| = -1$, $m = 20$ (and obviously $g = t^{20}$ is a solution in this case). □

COROLLARY 2. *Let $F(t), \sigma$ and F, σ be as above. If there exists a $g \in F(t)$ such that $\sigma g - g = f$, then $\|g\| \leq max(0, \|f\|)$.*

PROOF. Since $u = 1$ in this case, assuming the contrary would force $\alpha^m \in H(F)$, $m > 0$, contradicting Theorem 2 concerning $\prod$-extensions. □

By Theorems 14 and 15, when $\|a\| = 0$ we need consider only the problem of finding a basis for $V(\ldots, \ldots, W_m)$ for $m \geq -1$. The reduction to $F(t)_{-1}$ also handles the case in which $\|a\| = -1$, for which we first find a bound for m.

LEMMA 8. *If $\|a\| = -1$, then $\|\sigma g - a \cdot g\| = \|g\|$.*

In particular, $\sigma g - a \cdot g = f \Rightarrow \|g\| = \|f\|$, so we may choose $m \triangleq \|f\|$ in Theorem 11.

Convention. Let $m \geq 0$. Then

$$W_m / W_{m-1} \triangleq \{vt^m \mid v \in F\}.$$

It is clear that this is a legitimate choice for a quotient space.

THEOREM 16. *Suppose $F(t), \sigma$ is a $\prod\sum$-extension of F, σ. Let $a \in F(t)$ have polynomial part $u \in F$ (so $\|a\| \leq 0$). Note in this case that $\|\sigma g - a \cdot g\| \leq \|g\|$, so $l = 0$. Let $\mathbf{f} \in W_m^\omega$, $m \geq 0$.*

$$\begin{aligned} &\textit{def } \mathbf{v}: \quad \mathbf{v}t^m + \cdots = \mathbf{f}, \quad \mathbf{v} \in F, \quad \|\cdots\| < m, \\ &\mathbf{C} \curvearrowleft \mathbf{w} \triangleq \textit{a basis for } V(u/\alpha^m, \mathbf{v}/\alpha^m, F). \end{aligned}$$

Then

$$\mathbf{C} \curvearrowleft \mathbf{w}t^m \qquad \textit{is a basis for} \qquad V(a, \mathbf{f}, W_m/W_{m-1}).$$

By repeated application of this result, we may henceforth assume that the polynomial part of g is zero when trying to solve $\sigma g - a \cdot g = f$, where $\|a\| \leq 0$. This is one of the most important reduction techniques, so several examples of its use are given. The examples are preceded by a result which is useful in short-cutting the reduction process.

PROPOSITION 10. *Let F, σ be a difference field and $W \subseteq F$ a vector space over K. Then a basis for $V(1, \mathbf{0}_n, W)$ is,*

$$\begin{array}{ll} I_{n+1} & \text{if } W \cap K = K, \\ I_n \frown \mathbf{0}_n & \text{if } W \cap K = \{0\}, \end{array}$$

where I_n is the identity matrix of size n. (These are the only possibilities.)

Example 11. We consider an extremely simple sum, $\sum_{i=1}^{n} 1$. This leads to the difference equation in $\mathbb{Q}(x)$ (as usual, $\sigma x = x + 1$),

$$\sigma g - g = 1.$$

By Corollary 1, $\|g\| \leq \|1\| + 1 = 1$. Thus we are looking for a basis for $V(1, \langle 1\rangle, \mathbb{Q}(x)_1)$. Apply Theorem 12 with $m = 1$. It first requires a solution basis for $V(1, \langle 1\rangle, \mathbb{Q}(x)_1/\mathbb{Q}x)_0)$. For this, we apply Theorem 16 (where $m = 1$, $\mathbf{f} = \langle 1\rangle$, $\mathbf{v} = \langle 0\rangle$), which requests a basis $\mathbf{C}_1 \frown \mathbf{w}_1$ for $V(1, \langle 0\rangle, \mathbb{Q})$. By Proposition 10,

$$\mathbf{C}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{w}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Letting $\mathbf{g}_1 \triangleq \mathbf{w}_1 x$, Theorem 16 yields $\mathbf{C}_1 \frown \mathbf{g}_1$ as a basis for the incremental solution space $V(1, \langle x\rangle, \mathbb{Q}(x)_1/\mathbb{Q}(x)_0)$. Then Theorem 12 requires the computation

$$\mathbf{f}_0 = \mathbf{C}_1\mathbf{f} - (\sigma\mathbf{g}_1 - \mathbf{g}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \left[\begin{pmatrix} 0 \\ x+1 \end{pmatrix} - \begin{pmatrix} 0 \\ x \end{pmatrix} \right] = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The next step is to obtain a basis for $V(1, \mathbf{f}_0, \mathbb{Q}(x)_0)$. This is done by a "recursive" use of Theorem 12, this time with $m = 0$. This use first requires a solution basis for $V(1, \mathbf{f}_0, \mathbb{Q}(x)_0/\mathbb{Q}(x)_{-1})$. For this, we again apply Theorem 16 (where $m = 0$, $\mathbf{f} = \langle 1, -1\rangle$, $\mathbf{v} = \mathbf{f}$), which requests a basis for $V(1, \langle 1, -1\rangle, \mathbb{Q})$, for which Theorem 10 and a little linear algebra yield

$$\mathbf{C}_0 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{w}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Letting $\mathbf{g}_0 = \mathbf{w}_0 x^0 = \mathbf{w}_0$, Theorem 16 yields $\mathbf{C}_0 \frown \mathbf{g}_0$ as a basis for the incremental solution space. Then Theorem 12 requires

$$\mathbf{f}_{-1} = \mathbf{C}_0\mathbf{f}_0 - (\sigma\mathbf{g}_0 - \mathbf{g}_0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

By Proposition 10, a basis for $V(1, \langle 0, 0\rangle, \mathbb{Q}(x)_{-1})$ is $(\mathbb{Q} \cap \mathbb{Q}(x)_{-1} = \{0\})$:

$$\mathbf{D}_{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{h}_{-1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

We may now finish the recursive use which Theorem 12 has made of itself. For $m = 0$,

$$\mathbf{D}_0 = \mathbf{D}_{-1}\mathbf{C}_0 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{h}_0 = \mathbf{D}_{-1}\mathbf{g}_0 + \mathbf{h}_{-1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then, for the original use of Theorem 12, where $m = 1$,

$$\mathbf{D}_1 = \mathbf{D}_0\mathbf{C}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{h}_1 = \mathbf{D}_0\mathbf{g}_1 + \mathbf{h}_0 = \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix}.$$

This is the solution basis for $V(1, \langle 1\rangle, \mathbb{Q}(x))$. Since it is two-dimensional, we know that $\sum_{i=1}^{n} 1$ is a rational function of n:

$$\sum_{i=1}^{n} 1 = \sum_{0\le i<n} 1 = x_{\text{evaluated at } n} - x_{\text{evaluated at } 0} = n - 0 = n. \qquad \square$$

At this point the serious reader will go through the same exercise for the sum $\sum_{i=1}^{n} i$. The computation is essentially the same as in Example 11, except the recursion goes one level deeper (see [11, Exer. 1]).

In Example 11 and the suggested exercise, the results are familiar, as are more efficient techniques for their derivation. The following example, though somewhat lengthy, reveals more subtle aspects of the procedure.

Example 12. Let us continue Example 9. In this case, $u = 1$, $\alpha = 1$, and we have already seen that we may start with $m = 4$. For conciseness, the example is presented in tableau form, with many details suppressed.

Use Theorem 12 with $m = 4$, $\mathbf{f} = \langle -13\rangle$.

(1) Use Theorem 16 with $m = 4$, $\mathbf{f} = \langle -13\rangle$:

$$\mathbf{C}_4 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{w}_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \mathbf{g}_4 = \mathbf{w}x^4 = \begin{pmatrix} 0 \\ x^4 \end{pmatrix}.$$

(2) $$\mathbf{f}_3 \triangleq \begin{pmatrix} 1 \\ 0 \end{pmatrix}(-13) - \left[\sigma\begin{pmatrix} 0 \\ x^4 \end{pmatrix} - a \cdot \begin{pmatrix} 0 \\ x^4 \end{pmatrix}\right] = \begin{pmatrix} -13 \\ \dfrac{-8x^3 - 7x^2 - 4x - 1}{x^2 + 1} \end{pmatrix}.$$

(3) Use Theorem 12 with $m = 3$, $\mathbf{f} = \mathbf{f}_3$.

(3.1) Use Theorem 16 with $m = 3$, $\mathbf{f} = \mathbf{f}_3$, $\mathbf{v} = \langle 0, 0\rangle$:

$$\mathbf{C}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{w}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{g}_3 = \mathbf{w}x^3 = \begin{pmatrix} 0 \\ 0 \\ x^3 \end{pmatrix}.$$

(3.2) $$\mathbf{f}_2 \triangleq \mathbf{C}_3\mathbf{f}_3 - (\sigma\mathbf{g}_3 - a \cdot \mathbf{g}_3) = \begin{pmatrix} \mathbf{f}_3 \\ \dfrac{x^4 + 3x^3 - 4x^2 - 3x - 1}{x^2 + 1} \end{pmatrix}.$$

(3.3) Use Theorem 12 with $m = 2$, $\mathbf{f} = \mathbf{f}_2$.

(3.3.1) Use Theorem 16 with $m = 2$, $\mathbf{f} = \mathbf{f}_2$, $\mathbf{v} = \langle 0, 0, -1\rangle$:

$$\mathbf{C}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{g}_2 = \mathbf{w}x_2^2 = \begin{pmatrix} 0 \\ 0 \\ x^2 \end{pmatrix}.$$

(3.3.2) $$\mathbf{f}_1 \triangleq \mathbf{C}_2\mathbf{f}_2 - (\sigma\mathbf{g}_2 - a \cdot \mathbf{g}_2) = \begin{pmatrix} \mathbf{f}_3 \\ \dfrac{2x^3 + 5x^2 - 2x - 1}{x^2 + 1} \end{pmatrix}.$$

(3.3.3) Use Theorem 12 with $m = 1$, $\mathbf{f} = \mathbf{f}_1$.

(3.3.3.1) Use Theorem 16 with $m = 1$, $\mathbf{f} = \mathbf{f}_1$, $\mathbf{v} = \langle 0, -8, 2\rangle$:

$$\mathbf{C}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{w}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{g}_1 = \mathbf{w}_1 x = \begin{pmatrix} 0 \\ 0 \\ x \end{pmatrix}.$$

(3.3.3.2) $$\mathbf{f}_0 = \mathbf{C}_1\mathbf{f}_1 - (\sigma\mathbf{g} - a\cdot\mathbf{g}) = \begin{pmatrix} -13 \\ \dfrac{13x^2 - 12x - 5}{x^2 + 1} \\ \dfrac{3x^2 + 6x - 1}{x^2 + 1} \end{pmatrix}.$$

(3.3.3.3) Use Theorem 12 with $m = 0$, $\mathbf{f} = \mathbf{f}_0$.

(3.3.3.3.1) Use Theorem 16 with $m = 0$, $\mathbf{f} = \mathbf{f}_0$, $\mathbf{v} = \langle -13, 13, -3\rangle$:

$$\mathbf{C}_0 = \begin{pmatrix} 1 & 0 & \frac{13}{3} \\ 0 & 1 & -\frac{13}{3} \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{w}_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \mathbf{g}_0 = \mathbf{w}x^0 = \mathbf{w}_0.$$

(3.3.3.3.2) $$\mathbf{f}_{-1} = \mathbf{C}_0\mathbf{f}_0 - (\sigma\mathbf{g} - a\cdot\mathbf{g}) = \frac{1}{x^2 + 1}\cdot\begin{pmatrix} 26x - \frac{52}{3} \\ -38x - \frac{2}{3} \\ 4x + 6 \end{pmatrix}.$$

(3.3.3.3.3) We see later that $V(a, \mathbf{f}_1, \mathbb{Q}(x)_{-1}) = V(a, \mathbf{f}_1, \{0\})$. By Proposition 7 we determine a basis for $\mathrm{Ann}_{\mathbb{Q}}(\mathbf{f}_{-1}) \times \{0\}$:

$$\mathbf{C}_{-1} = (1 \quad 1 \quad 3)\mathbf{g}_{-1} = (0).$$

(3.3.3.3.4) $\mathbf{D}_0 = \mathbf{C}_{-1}\mathbf{C}_0 = (1\ 1\ 0)$, $\mathbf{h}_0 = \mathbf{C}_{-1}\mathbf{g}_0 + \mathbf{g}_{-1} = 3$.

(3.3.3.4) $\mathbf{D}_1 = \mathbf{D}_0\mathbf{C}_1 = (1\ 1\ 4)$, $\mathbf{h}_1 = \mathbf{D}_0\mathbf{g}_1 + \mathbf{h}_0 = 3$.

(3.3.4) $\mathbf{D}_2 = \mathbf{D}_1\mathbf{C}_2 = (1\ 1\ 0)$, $\mathbf{h}_2 = \mathbf{D}_1\mathbf{g}_2 + \mathbf{h}_1 = 4x^2 + 3$.

(3.4) $\mathbf{D}_3 = \mathbf{D}_2\mathbf{C}_3 = (1\ 1)$, $\mathbf{h}_3 = \mathbf{D}_2\mathbf{g}_3 + \mathbf{h}_2 = 4x^2 + 3$.

(4) $\mathbf{D}_4 = \mathbf{D}_3\mathbf{C}_4 = (1)$, $\mathbf{h}_4 = \mathbf{D}_3\mathbf{g}_4 + \mathbf{h}_3 = x^4 + 4x^2 + 3$.

Thus a solution basis for $V(a, -13, \mathbb{Q}(x))$ is $\langle 1, x^4 + 4x^2 + 3\rangle$; in other words, $x^4 + 4x^2 + 3$ is a solution to the original difference equation in Example 9. □

In Examples 11 and 12, we have considered the $\sum$-extension $\mathbb{Q}(x)$ of the constant field $\mathbb{Q}$. We give an example of the $\prod$-extension $\mathbb{Q}(x, x!)$ of the nonconstant field $\mathbb{Q}(x)$; recall that $\sigma x! = (x + 1)\cdot x!$.

Example 13. Consider the difference equation arising from $\sum_{i=1}^{n} i\cdot i!$:

$$\sigma g - g = x\cdot x!.$$

In this case $\|f\| = 1$ (the degree as a polynomial in $x!$), so by Corollary 2, $\|g\| \leq 1$.

Use Theorem 12 with $m = 1$, $\mathbf{f} = \langle x\cdot x!\rangle$, $F = \mathbb{Q}(x)$, $t = x!$.

(1) To obtain a basis for $V(1, x\cdot x!, \mathbb{Q}(x)(x!)_1/\mathbb{Q}(x)(x!)_0)$ use Theorem 16 with $m = 1$, $\mathbf{v} = \langle x\rangle$.

(1.1) We require a basis for $V(1/(x+1), x/(x+1), \mathbf{Q}(x))$. To obtain it, note that $\|a\| = -1$; so by Lemma 8 we may use $m = \|f\| = 0$. Thus, use Theorem 12 with $m = 0$, $\mathbf{f} = \langle x/(x+1)\rangle$, $a = 1/(x+1)$, $F = \mathbf{Q}$, $t = x$.

(1.1.1) To obtain a basis for $V(a, \mathbf{f}, \mathbf{Q}(x)_0/\mathbf{Q}(x)_{-1})$, use Theorem 16.

(1.1.1.1) Because $\mathbf{v} = \langle 1\rangle$, we require a basis for $V(0, \langle 1\rangle, \mathbf{Q})$, which by Theorem 10 is

$$\mathbf{C}_0 = (1), \qquad \mathbf{w}_0 = (1).$$

(1.1.1.2) The required basis is thus $\mathbf{C}_0 \wedge \mathbf{w}_0 x^0 = \mathbf{C}_0 \wedge \mathbf{w}_0$, so $\mathbf{g}_0 = \mathbf{w}_0$.

(1.1.2) $\mathbf{f}_{-1} = \mathbf{C}_0\langle x/(x+1)\rangle - (\sigma 1 - (1/(x+1))\cdot 1) = x/(x+1) - x/(x+1) = (0)$.

(1.1.3) A basis is required for $V(a, (0), \mathbf{Q}(x)_{-1})$, which will later be seen to be $V(a, (0), \{0\})$, which by Proposition 7 is $\mathrm{Ann}_{\mathbf{Q}}((0)) \times \{0\}$. Thus

$$\mathbf{D}_{-1} = (1), \qquad \mathbf{h}_{-1} = (0).$$

(1.1.4) $\mathbf{D}_0 = \mathbf{D}_{-1}\mathbf{C}_0 = (1)$, $\mathbf{h}_0 = \mathbf{D}_{-1}\mathbf{g}_0 + \mathbf{h}_{-1} = (1)$.

(1.2) A basis for $V(1, x\cdot x!, \mathbf{Q}(x)(x!)_1/\mathbf{Q}(x)(x!)_0)$ is thus $\mathbf{C}_1 \wedge \mathbf{g}_1$:

$$\mathbf{C}_1 = \mathbf{D}_0 = (1), \qquad \mathbf{g}_1 = \mathbf{h}_0 x!^1 = (x!).$$

(2) $\mathbf{f}_0 = \mathbf{C}_1\mathbf{f} - (\sigma\mathbf{g}_1 - \mathbf{g}_1) = (x\cdot x! - ((x+1)\cdot x! - x!)) = (0)$.

(3) A basis is required for $V(1, (0), \mathbf{Q}(x)(x!)_0)$. By Proposition 10,

$$\mathbf{D}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{h}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad \text{(other } \mathbf{D}_0, \mathbf{h}_0 \text{ now inactive)}.$$

(4) $\mathbf{D}_1 = \mathbf{D}_0\mathbf{C}_1 = \binom{1}{0}$, $\mathbf{h}_1 = \mathbf{D}_0\mathbf{g}_1 + \mathbf{h}_0 = \binom{x!}{0} + \binom{0}{1} = \binom{x!}{1}$. Thus

$$\sum_{i=1}^{n} i\cdot i! = \sum_{1\le i<n+1} i\cdot i! = (n+1)! - 1. \qquad \square$$

Example 14. To show that $\sum_{i=1}^{n} i!$ has no formula in $\mathbf{Q}(n, n!)$, examine

$$\sigma g - g = x!$$

The process begins as in Example 13, but with $\mathbf{f} = \langle x!\rangle$, $\mathbf{v} = \langle 1\rangle$ in line (1).

(1.1) We require a basis for $V(1/(x+1), \langle 1/(x+1)\rangle, \mathbf{Q}(x))$. Note that $\|a\| = -1$, so by Lemma 8 we may use $m = \|f\| = -1$. Thus a basis is required for $V(1/(x+1), \langle 1/(x+1)\rangle, \mathbf{Q}(x)_{-1})$. We will see in Example 23 that this solution space is equal to $\{\langle 0, 0\rangle\}$, so the basis is null:

$$\mathbf{C}_1 = \mathbf{0}_{01}, \qquad \mathbf{w}_1 = \mathbf{0}_0.$$

(1.2) If $\mathbf{w}_1 = \mathbf{0}_0$, so is $\mathbf{g}_1 = \mathbf{w}_1 x!$.

(2) $\mathbf{f}_0 = \mathbf{C}_1\mathbf{f} - (\sigma\mathbf{g}_1 - \mathbf{g}_1) = \mathbf{0}_0$.

(3) A basis is required for $V(1, \mathbf{0}_0, \mathbf{Q}(x)(x!)_0)$. By Proposition 10,

$$\mathbf{D}_0 = \mathbf{0}_{10}, \qquad \mathbf{h}_0 = (1).$$

(4) $\mathbf{D}_1 = \mathbf{D}_0\mathbf{C}_1 = (0)$, $\mathbf{h}_1 = \mathbf{D}_0\cdot\mathbf{g}_1 + \mathbf{h}_0 = (1)$.

Since this basis is one-dimensional, $\sum_{i=1}^{n} i!$ is not summable in $\mathbf{Q}(n, n!)$. $\square$

The reader is invited to use these results to find formulas for $\sum_{i=1}^{n} 2^i$, $\sum_{i=1}^{n} i^2 2^i$ ([11, Exer. 2, 3]). It may be assumed (see Example 15 below) that $V(1/2, \mathbf{0}, \mathbf{Q}(x)_{-1})$ equals $V(1/2, \mathbf{0}, \{0\})$.

We now return to the main point of this subsection, which is to eliminate the polynomial part of the solution. We consider the case when a has a nontrivial polynomial part; that is, $\|a\| > 0$. As in the case where $\|a\| = -1$, it is simple to find a bound on the $\|\ \|$ of a solution.

LEMMA 9. *If* $\|a\| > 0$, *then*

$$\|\sigma g - a \cdot g\| \leq \|g\| + \|a\|, \qquad \textit{with equality when} \quad \|g\| \geq 0.$$

Accordingly, for this case, $l \triangleq \|a\|$. *A direct consequence of this is*

$$\sigma g - a \cdot g = f \Rightarrow \|g\| = \begin{cases} \|f\| - l & \text{if } \|f\| \geq l, \\ -1 & \text{otherwise.} \end{cases}$$

Thus, Theorem 11 may be applied with $m \triangleq \|f\| - \|a\|$ or -1 as appropriate. The only remaining case in which g might have a nonzero polynomial part is when $\|f\| \geq \|a\| > 0$. This is handled by the following result.

THEOREM 17. *Let* a *be as above and* $f \in W_{m+l}$ *for* $m \geq 0$. *Perform the following construction*:

$f_0 \triangleq$ *the polynomial part of* f;
$a_0 \triangleq$ *the polynomial part of* a;
def g, h: $\quad g \cdot a_0 + h = -f_0, \quad g, h \in F[t], \quad \|h\| < \|a_0\|$
(*division with remainder*).

Then

$$\|f - (\sigma g - a \cdot g)\| < m + l.$$

For $0 \neq g \in W_m / W_{m-1}$, $\|g\| \geq 0$, and by Lemma 9, $\|\sigma g - a \cdot g\| = \|g\| + l$, and $\|\sigma 0 - a \cdot 0\| = -1 \neq m + l$ for any $m \geq 0$. This fact, in combination with the above theorem, allows Theorem 13 to be used to reduce $\|\mathbf{f}\|$ until $\|\mathbf{f}\| < l$. Then, by the other part of Lemma 9, the polynomial part of any solution must be zero. Examples of Theorem 17 are not very interesting; readers may construct their own.

In summary, this subsection has reduced the problem of computing a basis for $V(\ldots, \ldots, W)$ to $V(\ldots, \ldots, W_{-1})$. Furthermore, we have the following bound on $\mathbf{f}$:

$$\|\mathbf{f}\| \leq \max(-1, \|a\| - 1).$$

3.4 FACTORS OF PERIOD 1 IN THE FRACTIONAL PART. The analysis of the fractional part proceeds by examining various possible factors of the denominator of a solution. This subsection is concerned with the factor t in $\prod$-extensions; this restriction is implicit throughout the subsection. We are concerned with vector spaces W such that

$$\left\{ \sum_{i \geq 0} \frac{v_i}{t^i} \;\middle|\; v_i \in F \right\} \subseteq W \subseteq F(t).$$

Definition 29. Let $f \in W$. In this subsection,

$$\|f\| \triangleq \begin{cases} -1 & \text{if } t \mid \operatorname{num}(f), \\ m & \text{otherwise, where } t \text{ occurs to the power } m \text{ in } \operatorname{den}(f). \end{cases}$$

As with the polynomial part, we have a case analysis depending upon $\|a\|$.

THEOREM 18. *Let* $a \in F(t)$, $\|a\| = 0$. *Write*

$$a = \frac{u + \cdots}{1 + \cdots},$$

where $u \in F$ *and* "$\cdots$" *consists of terms with degree* >0 (*perhaps none*). (*The fact that a can be written this way uses the fact that* $\|a\| = 0$; *observe that* $u \neq 0$.) *For any* $f \in W$, *if there exists* $g \in W$ *such that* $\sigma g - a \cdot g = f$ *and* $\|g\| > \|f\|$, *then*

(*a*) *there exists a unique* $m \in \mathbb{Z}$, $m > \|f\|$, *such that*

$$u \cdot \alpha^m \in H(F);$$

(*b*) $\|g\| \leq m$.

The existence and value, if any, of the m of part (*a*) *may be calculated provided that a basis for* $M(\ldots, F)$ *may be calculated.*

As usual, this reduces the problem to W_m when $\|a\| = 0$. There is an even easier case when $\|a\| = -1$.

LEMMA 10. *If* $\|a\| = -1$, *then* $\|\sigma g - a \cdot g\| = \|g\|$.

Again, a simple bound for m has been obtained using Theorem 11.

Convention. Let $m > 0$. Then

$$W_m / W_{m-1} \triangleq \{v/t^m \mid v \in F\}.$$

We now have the result which allows us to compute the incremental solution space when $\|a\| \leq 0$.

THEOREM 19. *Let* $a \in F(t)$ *have* $\|a\| \leq 0$, *and write a as in Theorem* 18:

$$a = \frac{u + \cdots}{1 + \cdots}, \qquad \textit{where} \quad u \in F \quad (\textit{perhaps } u = 0).$$

In this case $\|\sigma g - a \cdot g\| \leq \|g\|$, *so* $l = 0$. *Let* $\mathbf{f} \in W_m^{\omega}$, *where* $m > 0$.

$$\textit{def}\ \mathbf{v}: \quad \frac{\mathbf{v}}{t^m} + \cdots = \mathbf{f}, \quad \mathbf{v} \in F, \quad \|\cdots\| < m,$$

$$\mathbf{C} \wedge \mathbf{w} \triangleq \textit{a basis for } V(u \cdot \alpha^m, \mathbf{v} \cdot \alpha^m, F).$$

Then

$$\mathbf{C} \wedge \mathbf{w}/t^m \qquad \textit{is a basis for} \qquad V(a, \mathbf{f}, W_m / W_{m-1}).$$

Repeatedly applying this result, we have reduced the problem to W_0. In other words, we may assume that t is not a factor of the denominator of g when $\|a\| \leq 0$ and the extension is homogeneous. There is a remaining case to consider, when $\|a\| > 0$. First we bound the grade of g.

LEMMA 11. *If* $\|a\| > 0$, *then*

$$\|\sigma g - a \cdot g\| \leq \|g\| + \|a\|, \qquad \textit{with equality when} \quad \|g\| \geq 0.$$

Accordingly, for this case, $l \triangleq \|a\|$. *A direct consequence of this is*

$$\sigma g - a \cdot g = f \Rightarrow \|g\| = \begin{cases} \|f\| - l & \textit{if} \quad \|f\| \geq l; \\ -1 & \textit{otherwise.} \end{cases}$$

Thus the grade of any solution can be bounded. Finally, we show how to get to W_0.

$$\text{def } u: \quad a = \frac{u}{t^l} + \cdots, \quad u \in F, \quad \|\cdots\| < l,$$

$$\text{def } v: \quad f = \frac{v}{t^{m+l}} + \cdots, \quad v \in F, \quad \|\cdots\| < m + l,$$

$$g \triangleq \frac{(-v/u)}{t^m}.$$

Then

$$\|f - (\sigma g - a \cdot g)\| < m + l.$$

Thus we may apply Theorem 13 until $\|\mathbf{f}\| \leq l$, so $\|g\| \leq 0$. This is the final step in showing that henceforth we may assume that t is not a denominator of g in the homogeneous case. Observe that the reduction of this subsection and that of the previous subsection may be applied in any order; only in the second reduction does the zeroth-order term $u \cdot t^0 = u = u/t^0$ actually have to be eliminated.

3.5. Factors of Period 0 in the Fractional Part. As a preliminary to studying nonhomogeneous factors of the fractional part, it is convenient to consider only equations whose a's have the following property.

Definition 30. Let $a \in F(t)$. Then a is *pure* $\Leftrightarrow$ for any $\mathfrak{f}_1, \mathfrak{f}_2$ which are factors of a,

$$\mathfrak{f}_1 \neq \mathfrak{f}_2 \Rightarrow \mathfrak{f}_1 \not\sim \mathfrak{f}_2.$$

For example, if $\sigma x = x + 1$, then $x \cdot (x + 2)$ is impure; $x \cdot (2x + 1)$ is pure.

Lemma 12. *Let $0 \neq a \in F(t)$. Then there exists an $a_0 \in F(t)$ such that $a \cdot (\sigma a_0/a_0)$ is pure. Given the σ-factorization of a, a_0 may be computed.*

Proof. Let $\langle \langle \mathfrak{f}_j \rangle_j, u, e, \langle e_{jk} \rangle_{jk} \rangle$ be the σ-factorization of a (the "i" index may be dropped, because a is only a one-tuple). It is possible to choose $\mathfrak{f}_j$ so that $e_{jk} = 0$ for $k < 0$ and $e_{j0} \neq 0$. Assuming that this is done, let

$$d_{jk} \triangleq -\sum_{l \geq k} e_{jl} \qquad \text{for} \quad k \geq 0,$$

$$a_0 \triangleq \prod_{j,k \geq 0} \sigma^k \mathfrak{f}_j^{d_{j,k+1}}.$$

A simple calculation [11] shows that

$$a \cdot \frac{\sigma a_0}{a_0} = u \cdot \frac{t^e}{\prod_j \mathfrak{f}_j^{d_{j0}}}.$$

Since the $\mathfrak{f}_j$'s of the original set are pairwise inequivalent by definition of σ-factorization, it is clear that all the factors of $a \cdot (\sigma a_0/a_0)$ are pairwise inequivalent. Thus $a \cdot (\sigma a_0/a_0)$ is pure $\square$

By Lemma 7, we lose no generality in assuming that a is pure.

In this subsection we consider an irreducible polynomial $\mathfrak{f}$ with period 0; that is, $\mathfrak{f}$ is not equal to t in a $\prod$-extension. We show that algorithmically, only a finite number of such polynomials need be considered. The vector spaces W in which we are interested have the following property:

$$\left\{ \sum_{i>0, k \in Z} \frac{f_{ik}}{\sigma^k \mathfrak{f}^i} \;\middle|\; \deg f_{ik} < \deg \mathfrak{f} \right\} \subseteq W \subseteq F(t).$$

Several different grades of W may be necessary for the complete reduction; the first of these is

Definition 31. Let $f \in W$. Then

$$\|f\| \triangleq \max\{i \mid \exists k \text{ with } \sigma^k \mathscr{f}^i \text{ dividing } \operatorname{den}(f)\}.$$

N.B. This grade depends upon $\mathscr{f}$, but the notation suppresses this.

LEMMA 13. *Let a be pure and $\|a\| = 0$. Then $\|\sigma g - a \cdot g\| = \|g\|$.*

By Theorem 11, if $\mathscr{f}$ has the property that $\|a\| = 0$, then we know that $\|g\| \leq \|\mathbf{f}\|$; in particular, if $\|\mathbf{f}\| = 0$, then $\sigma^k \mathscr{f}$ cannot be a factor of the denominator of g for any k. This is a crucial result for the construction, because it eliminates all but a finite number of equivalence classes for $\mathscr{f}$.

Example 15. It was claimed in Subsection 3.3 (regarding the summation of $\sum_{i=1}^{n} i^2 2^i$) that $V(1/2, \mathbf{0}, \mathbf{Q}(x)_{-1}) = V(1/2, \mathbf{0}, \{0\})$. This is an immediate consequence of Lemma 13 and the above remarks, since for any $\mathscr{f}$, $\|1/2\| = \|0\| = 0$. Thus g must have a trivial fractional part. □

For those $\mathscr{f}$ for which $\|\mathbf{f}\| > 0$, a reduction may be used. First we define the quotient space for this process.

Convention. For $m > 0$,

$$W_m / W_{m-1} \triangleq \left\{ \sum_k f_k / \sigma^k \mathscr{f}^m \mid f_k \in F[t], \deg f_k < \deg \mathscr{f} \right\}.$$

In the following reduction process we use the fact that if $\sigma^k \mathscr{f} \not\mid \operatorname{den}(a)$, then for any $g_k \in F[t]$, $\deg g_k < \deg \mathscr{f}$, there exists (by the theory of partial fraction decomposition and $\operatorname{per}(\mathscr{f}) = 0$) a unique $h_k \in F[t]$, $\deg h_k < \deg \mathscr{f}$, such that

$$\frac{h_k}{\sigma^k \mathscr{f}^m} + \cdots = a \cdot \frac{g_k}{\sigma^k \mathscr{f}^m}, \qquad \text{where} \quad \text{“}\cdots\text{”} \in W_{m-1}.$$

THEOREM 21. *Let $a \in F(t)$ be pure and $\|a\| = 0$. Let $\mathbf{f} \in W_m^{\omega}$ be written*

$$\sum_{k=\iota}^{J} \frac{\mathbf{f}_k}{\sigma^k \mathscr{f}^m} + \cdots = \mathbf{f}, \qquad \deg \mathbf{f}_k < \deg \mathscr{f}, \quad \text{“}\cdots\text{”} \in W_{m-1}.$$

Choose i so that

$$k < i \Rightarrow \sigma^k \mathscr{f} \not\mid \operatorname{num}(a). \tag{6}$$

Note that $\mathbf{f}_i$ may well be zero. Next, perform the following construction:

$$\left.\begin{aligned} &\mathbf{g}_{J-1} \triangleq \sigma^{-1}\mathbf{f}_J, \\ &\text{def } \mathbf{h}_k\text{:} \qquad \frac{\mathbf{h}_k}{\sigma^k \mathscr{f}^m} + \cdots = a \cdot \frac{\mathbf{g}_k}{\sigma^k \mathscr{f}^m}, \\ &\mathbf{g}_{k-1} \triangleq \sigma^{-1}(\mathbf{f}_k + \mathbf{h}_k), \end{aligned}\right\} \quad i \leq k < j.$$

This is done in the order $j-1, j-2, \ldots, i$ so that $\mathbf{g}_k$ is defined before $\mathbf{h}_k$, and $\mathbf{h}_k$ before $\mathbf{g}_{k-1}$. Let

$$\mathbf{g} \triangleq \sum_{k=\iota}^{J-1} \frac{\mathbf{g}_k}{\sigma^k \mathscr{f}^m},$$

$$\mathbf{C} \qquad \text{be a basis for} \qquad \operatorname{Ann}(\mathbf{g}_{\iota-1}).$$

Then

$$\mathbf{C}(I \wedge \mathbf{g}) \qquad \textit{is a basis for} \qquad V(a, \mathbf{f}, W_m/W_{m-1}).$$

Example 16. One of the simplest applications of this result is the demonstration that H_n (see Subsection 1.1) is not a rational function of n. Consider

$$\sigma g - g = 1/x.$$

To calculate $V(1, \langle 1/x \rangle, \mathbb{Q}(x))$, we begin with Corollary 1, which reduces $\mathbb{Q}(x)$ to $\mathbb{Q}(x)_0$. To apply Theorem 12, we must obtain a basis for $V(1, \langle 1/x \rangle, \mathbb{Q}(x)_0/\mathbb{Q}(x)_{-1})$. Theorems 16 and 10 give this basis as $\mathbf{C} \wedge \mathbf{g}$, where

$$\mathbf{C} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{g}_0 = \mathbf{w} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Theorem 12 also needs the calculation,

$$\mathbf{f}_{-1} = \mathbf{C}_0\mathbf{f} - (\sigma\mathbf{g} - \mathbf{g}) = \begin{pmatrix} 1/x \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/x \\ 0 \end{pmatrix}.$$

Then Theorem 12 requires a basis for $V(1, \mathbf{f}_{-1}, \mathbb{Q}(x)_{-1})$, which is where the methods of this subsection come in. Applying Lemma 13 with any $\mathscr{f} \in \mathbb{Q}[x]$, where $\mathscr{f} \not\sim x$, we further narrow $\mathbb{Q}(x)_{-1}$ to

$$W \triangleq \left\{ \sum_{j>0,k} \frac{c_{jk}}{(x+k)^j} \;\middle|\; c_{jk} \in \mathbb{Q} \right\}.$$

Fixing $\mathscr{f} = x$, Lemma 13 narrows this choice further to

$$W_1 = \left\{ \sum_k \frac{c_k}{x+k} \;\middle|\; c_j \in \mathbb{Q} \right\}.$$

Since $W_0 = \{0\}$, $W_1/W_0 = W_1$. We now apply Theorem 21, with $\mathbf{f} = \langle 1/x, 0 \rangle$. In the statement of that result, we have $i = j = 0$, $\mathbf{f}_0 = \langle 1, 0 \rangle$, and $\mathbf{g}_{-1} = \sigma^{-1}\mathbf{f}_0 = \langle 1, 0 \rangle$. Observe that there are no k with $i \le k$ and $k < 0$. Then

$$\mathbf{g} = \sum_{k=0}^{-1} \frac{\mathbf{g}_k}{x+k} = \langle 0, 0 \rangle.$$

A basis $\mathbf{C}$ for $\mathrm{Ann}(\mathbf{g}_{-1}) = \mathrm{Ann}(\langle 1, 0 \rangle)$ is $(0, 1)$. Thus, the desired basis is

$$\mathbf{C}\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right] = (0 \quad 1) \wedge (0).$$

Therefore, letting $\mathbf{D}_{-1} = (0\ 1)$ and $\mathbf{h}_{-1} = (0)$, we may return to Theorem 12 and calculate a solution basis for $V(1, \langle 1/x \rangle, \mathbb{Q}(x))$:

$$\mathbf{D}_0 = \mathbf{D}_{-1}\mathbf{C}_0 = (0), \qquad \mathbf{h}_0 = \mathbf{D}_{-1}\mathbf{g}_0 + \mathbf{h}_{-1} = (1) + (0) = (1).$$

This is only a one-dimensional solution space, so we have shown that H_n is not a rational function of n. Since $1 \in H(\mathbb{Q})$, we know by Theorem 1 that $\mathbb{Q}(x, H_x)$ is a $\sum$-extension of $\mathbb{Q}(x)$. This fact will be used when H_i occurs in the summand of $\sum_i$. □

The reader may follow a similar path to show that $\sum_{i=1}^{n} 1/i^2$ is not a rational function of n [11, Exer. 1].

Example 17. We now find a simple expression for $\sum_{i=1}^{n} 1/(i^2 + 2i)$; we solve

$$\sigma g - g = \frac{1}{x^2 + 2x}.$$

Because $x^2 + 2x = x(x + 2) = x \cdot \sigma^2 x$, the calculation here is essentially the same as in the previous example, until we begin to apply Theorem 21 with $\mathbf{f} = \langle 1/(x^2 + 2x), 0\rangle$. Then we find that

$$\left\langle \frac{1}{x^2 + 2x}, 0 \right\rangle = \left\langle \frac{1/2}{x} + \frac{0}{x+1} + \frac{-1/2}{x+2}, 0 \right\rangle, \qquad \text{so} \quad i = 0, \quad j = 2,$$

$$\mathbf{f}_2 = \left\langle -\frac{1}{2}, 0 \right\rangle, \qquad \mathbf{f}_1 = \mathbf{0}_2, \qquad \mathbf{f}_0 = \left\langle \frac{1}{2}, 0 \right\rangle.$$

Then compute the $\mathbf{g}$, $\mathbf{h}$ sequence,

$$\mathbf{g}_1 = \sigma^{-1}\mathbf{f}_2 = \left\langle -\frac{1}{2}, 0 \right\rangle,$$

$$\text{def } \mathbf{h}_1\text{:} \quad \frac{\mathbf{h}_1}{x+1} + \cdots = \frac{\langle -1/2, 0\rangle}{x+1}, \quad \text{so} \quad \mathbf{h}_1 = \left\langle -\frac{1}{2}, 0 \right\rangle,$$

$$\mathbf{g}_0 = \sigma^{-1}(\mathbf{f}_1 + \mathbf{h}_1) = \left\langle -\frac{1}{2}, 0 \right\rangle,$$

$$\text{def } \mathbf{h}_0\text{:} \quad \frac{\mathbf{h}_0}{x} + \cdots = \frac{\langle -1/2, 0\rangle}{x}, \quad \text{so} \quad \mathbf{h}_0 = \left\langle -\frac{1}{2}, 0 \right\rangle,$$

$$\mathbf{g}_{-1} = \sigma^{-1}(\mathbf{f}_0 + \mathbf{h}_0) = \langle 0, 0\rangle.$$

Finally,

$$\mathbf{g} = \sum_{k=0}^{1} \frac{\mathbf{g}_k}{\sigma^k x} = \left\langle \frac{-1/2}{x} + \frac{-1/2}{x+1}, 0 \right\rangle = \left\langle \frac{-x - 1/2}{x^2 + x}, 0 \right\rangle,$$

$$\mathbf{C} \text{ a basis for } \operatorname{Ann}(\langle 0, 0\rangle) \Rightarrow \mathbf{C} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Thus the desired basis is

$$\mathbf{D}_{-1} = \mathbf{C}I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{h}_{-1} = \mathbf{C}\mathbf{g} = \begin{pmatrix} \dfrac{-x - 1/2}{x^2 + x} \\ 0 \end{pmatrix}.$$

Returning to Theorem 12, we calculate a solution basis for $V(1, 1/(x^2 + 2x), \mathbb{Q}(x))$:

$$\mathbf{D}_0 = \mathbf{D}_{-1}\mathbf{C}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{h}_0 = \mathbf{D}_{-1}\mathbf{g}_0 + \mathbf{h}_{-1} = \begin{pmatrix} \dfrac{-x - 1/2}{x^2 + x} \\ 0 \end{pmatrix}.$$

Thus,

$$\sum_{i=1}^{n} \frac{1}{i^2 + 2i} = -\frac{n + 1 + 1/2}{(n+1)^2 + (n+1)} - \left(-\frac{3/2}{2}\right) = \frac{3n^2 + 5n}{4n^2 + 24n + 12}. \qquad \square$$

As a consequence of Theorem 21, the only case remaining is when $\|a\| \neq 0$, that is, $\sigma^k \not| \operatorname{den}(a)$. Without loss of generality, we may choose $\not$ so that it itself divides $\operatorname{den}(a)$. We continue to assume that a is pure, so that $\sigma^k \not| \operatorname{den}(a) \Rightarrow k = 0$, as well as the converse.

The elimination of factors of the form $\sigma^k f$ as potential denominators of g is done in two steps. The first of these steps eliminates the possibility that k is negative. For this, we need a new grade.

Definition 32. Let $f \mid \mathrm{den}(a)$. Then for $f \in W$,

$$\|f\| \triangleq \max\{i \mid \exists k < 0 \text{ with } \sigma^k f^i \mid \mathrm{den}(f)\}.$$

N.B. This grade depends upon f, but the notation suppresses this.

LEMMA 14. *Let a be pure, $f \mid den(a)$. Then $\|\sigma g - a \cdot g\| = \|g\|$.*

The remarks made after Lemma 13 may be adapted to this lemma as well, namely, that if $f \mid \mathrm{den}(a)$ and $\|\mathbf{f}\| = 0$, then $\sigma^k f$ cannot be a factor of the denominator of g, for any $k < 0$.

Example 18. We previously encountered in Example 12 (step (3.3.3.3.3)) this space:

$$V(1/(x^2 + 1), \langle 26x - 52/3, -38x - 2/3, 4x + 6\rangle, \mathbb{Q}(x)_{-1}).$$

Using Lemmas 13 and 14, $\mathbb{Q}(x)_{-1}$ may be narrowed to the following space:

$$W = \left\{ \sum_{i>0, k\geq 0} \frac{f_{ik}}{\sigma^k(x^2+1)^i} \;\middle|\; f_{ik} \in \mathbb{Q}[x],\ \deg f_{ik} \leq 1 \right\}. \qquad \square$$

Example 19. We saw in Example 13 (step (1.1.3)) the following space:

$$V(1/(x + 1), \langle 0\rangle, \mathbb{Q}(x)_{-1}).$$

Again, Lemmas 13 and 14 reduce $\mathbb{Q}(x)_{-1}$ to a smaller space,

$$W = \left\{ \sum_{i>0, k\geq 0} \frac{c_{ik}}{\sigma^k(x+1)^i} \;\middle|\; c_{ik} \in K \right\}. \qquad \square$$

In order to perform the desired reduction, it is necessary to introduce a preliminary result concerning partial fractions.

PROPOSITION 11. *Let $a \in F(t)$ be pure, $f \mid den(a)$, $k < 0$. For any $h \in F[t]$, deg $h <$ deg f, there exists a unique $g \in F[t]$, deg $g <$ deg f, with*

$$a \cdot \frac{g}{\sigma^k f^m} + \cdots = \frac{h}{\sigma^k f^m}, \qquad \textit{where} \quad \text{“}\cdots\text{”} \in W_{m-1}.$$

Such a g is introduced using "def:" (see, e.g., Theorems 22 below).

Convention. Under the grade of Definition 32, for $n > 0$,

$$W_m / W_{m-1} \triangleq \left\{ \sum_{k<0} \frac{f_k}{\sigma^k f^m} \;\middle|\; f_k \in F[t],\ \deg f_k < \deg f \right\}.$$

For this reduction we can use Theorem 13. By Lemma 14 we see that for $g \in W_m / W_{m-1}$,

$$\|\sigma g - a \cdot g\| < m \Leftrightarrow g = 0.$$

The other requirement for Theorem 13 is contained in the following result.

THEOREM 22. *Let $a \in F(t)$ be pure, $f \mid den(a)$. Let $f \in W_m$, and write*

$$\sum_{k=i}^{-1} \frac{f_k}{\sigma^k f^m} + \cdots = f, \qquad \textit{where} \quad \text{“}\cdots\text{”} \in W_{m-1}.$$

Without loss of generality, i may be chosen so that $f_i \neq 0$. Construct the sequences h_k, g_k as follows:

$$\left.\begin{aligned} & g_{i-1} \triangleq 0, \\ & h_k \triangleq \sigma g_{k-1} - f_k, \\ & \text{def } g_k: \quad a \cdot \frac{g_k}{\sigma^k \not{f}^m} + \cdots = \frac{h_k}{\sigma^k \not{f}^m}, \end{aligned}\right\} \; i \leq k \leq -1.$$

This is done in the order i, $i+1$, ..., so g_{k-1} is defined before h_k, and h_k before g_k. Proposition 11 allows the construction. Then set

$$g \triangleq \sum_{k=i}^{-1} \frac{g_k}{\sigma^k \not{f}^m}.$$

We have

$$\| f - (\sigma g - a \cdot g) \| < m.$$

Example 20. Consider the difference equation,

$$\sigma g - \frac{1}{x} \cdot g = \frac{1}{(x-1)^2}.$$

It is quickly seen that any solution g is in $\{\sum c_k/\sigma^k x\}$, so we examine here only the case in which $\not{f} = x$; observe that $\| f \| = 2$. Apply Theorem 22, noting that $i = -1$ and $f_{-1} = 1$ and performing the following construction:

$$\begin{aligned} & g_{-2} = 0, \qquad h_{-1} = \sigma 0 - f_{-1} = -1, \\ & \text{def } g_{-1}: \quad \frac{1}{x} \cdot \frac{g_{-1}}{(x-1)^2} + \cdots = \frac{-1}{(x-1)^2} \\ & \qquad = \frac{1}{x} \frac{-x}{(x-1)^2} = \frac{1}{x}\left(\frac{-1}{(x-1)^2} - \frac{1}{x-1}\right) \\ & \qquad \Rightarrow g_{-1} = -1, \\ & g = \frac{-1}{(x-1)^2}, \\ & f - (\sigma g - a \cdot g) = \frac{1}{(x-1)x} + \frac{1}{x^2} = \frac{1}{x-1} - \frac{1}{x} + \frac{1}{x^2}. \end{aligned}$$

We again apply Theorem 22, this time to $f - (\sigma g - a \cdot g)$, whose grade is 1. We still have $i = -1$ and $f_{-1} = 1$, so again $g_{-2} = 0$ and $h_{-1} = -1$. Multiplying both sides of the previous definition of g_{-1} by $x - 1$ provides the proper definition of g_{-1} in this case, so $g_{-1} = -1$ again. Thus g for this case is $-1/(x-1)$, and we can write the original equation as follows:

$$\sigma\left(g + \frac{1}{(x-1)^2} + \frac{1}{x-1}\right) + \frac{1}{x}\left(g + \frac{1}{(x-1)^2} + \frac{1}{x-1}\right) = \frac{1}{x^2} + \frac{1}{x}.$$

In other words, a solution to $\sigma h - a \cdot h = 1/x^2 - 1/x$ allows the recovery of g, and the impossibility of a solution means that g does not exist. □

For the remainder of this subsection we consider a more limited set of vector spaces than before, namely,

$$\left\{ \sum_{i>0, k \geq 0} \frac{f_{ik}}{\sigma^k \not{f}^i} \;\middle|\; \deg f_{ik} < \deg \not{f} \right\} \subseteq W \subseteq F(t).$$

For any such W we introduce yet another grade.

Definition 33. For the rest of this subsection,

$$l \triangleq \text{the power to which } \mathscr{p} \text{ occurs in } \operatorname{den}(a).$$

Then for any $f \in W$,

$$\|f\| \triangleq \max\{i + k \cdot l \mid \sigma^k \mathscr{p}^i \text{ divides } \operatorname{den}(f), i > 0\} \cup \{0\}.$$

LEMMA 15. *For $g \in W$, $\|\sigma g - a \cdot g\| \leq \|g\| + l$, with equality if $\|g\| > 0$.*

By Theorem 11 we can bound the grade of solutions, namely,

$$\sigma g - a \cdot g = f \Rightarrow \|g\| \leq \begin{cases} \|f\| - l & \text{if } \|f\| > l, \\ 0 & \text{otherwise.} \end{cases}$$

Example 21. We continue Example 18. In this case $\mathscr{p} = x^2 + 1$, so $l = 1$. The denominators of $\mathbf{f}$ are all 1, and for any polynomial f, $\sigma^k \mathscr{p}^i$ does not divide $\operatorname{den}(f)$, so $\|\mathbf{f}\| = 0$. Thus the W of Example 18 may further be narrowed to W_0, that is, $\{0\}$. □

Example 22. We continue Example 19. Here $\mathscr{p} = x + 1$, so $l = 1$. The denominator of 0 is 1, and the remarks of the previous example apply to this case. □

Example 23. In Example 14, step (1.1), we claimed that the solution space $V(1/(x+1), \langle 1/(x+1)\rangle, \mathbf{Q}(x)_{-1})$ was $\{\langle 0, 0\rangle\}$. We can now prove this. By Lemmas 13 and 14 we need consider only those g in the space,

$$W = \left\{ \sum_{i>0, k\geq 0} \frac{c_{ik}}{\sigma^k (x+1)^i} \;\middle|\; c_{ik} \in K \right\}.$$

Choose $\mathscr{p} = x + 1$ to define the current grade. Then $l = 1$, and $\|1/(x+1)\| = 1$. Since $\|f\| \leq l$, we conclude that $\|g\| \leq 0$, by the above remark. The only such $g \in W$ is 0, proving the claim. □

Having bounded the grade of the solution, we turn to the reduction process. As usual, we first introduce the quotient space.

Convention. For $m > 0$,

$$W_m / W_{m-1} \triangleq \left\{ \sum_{k=0}^{\lfloor m/l \rfloor - 1} f_k / \sigma^k \mathscr{p}^{m-l\cdot k} \;\middle|\; f_k \in F[t], \deg f_k < \deg \mathscr{p} \right\}.$$

THEOREM 23. *Let a, l be as above. Let $\mathbf{f} \in F(t)^{\omega}_{m+l}$, with $m > 0$, and write*

$$\sum_{k\geq 0} \frac{\mathbf{f}_k}{\sigma^k \mathscr{p}^{m+l-k\cdot l}} + \cdots = \mathbf{f}, \qquad \textit{with} \quad \text{“}\ldots\text{”} \in F(t)_{m+l-1}.$$

Perform the following construction:

$\mathbf{g}_k \triangleq \sigma^{-1}\mathbf{f}_{k+1}, \qquad k \geq 0,$

$\mathbf{g} \triangleq \sum_{k\geq 0} \dfrac{\mathbf{g}_k}{\sigma^k \mathscr{p}^{m-k\cdot l}},$

def $\mathbf{q}, \mathbf{r}$: $\quad \mathbf{q}\cdot\mathscr{p} + \mathbf{r} = \textit{num}(a \cdot \mathscr{p}^l \cdot \sigma^{-1}\mathbf{f}_1 + \mathbf{f}_0)$ *(division with remainder)*,

$\mathbf{C} \triangleq$ *a basis for* $\textit{Ann}(\mathbf{r})$.

Then

$$\mathbf{C}(I \curvearrowright \mathbf{g}) \qquad \textit{is a basis for} \qquad V(a, \mathbf{f}, W_m / W_{m-1}).$$

Example 24. Continuing Example 20, we are interested in $V(1/x, \langle 1/x^2 + 1/x\rangle, W)$, where

$$W = \left\{ \sum_{i>0,k\geq 0} \frac{c_{ik}}{\sigma^k x^i} \,\middle|\, c_{ik} \in K \right\}.$$

Here, $\mathscr{f} = x$, $l = 1$, and we observe that $\|1/x^2 + 1/x\| = 2 > 1 = l$. Thus $\|g\| \leq 2 - l = 1$, and we would like a basis for $V(a, \mathbf{f}, W_1/W_0)$, where $\mathbf{f} = \langle 1/x^2 + 1/x\rangle$. Applying Theorem 23 with $m = 1$,

$$\frac{1}{x^2} + \frac{0}{x+1} + \cdots = \frac{1}{x^2} + \frac{1}{x} \Rightarrow \mathbf{f}_0 = \langle 1\rangle \text{ and } \mathbf{f}_1 = \langle 0\rangle.$$

Following the construction of that result,

$$\mathbf{g}_0 = \sigma^{-1}\mathbf{f}_1 = 0, \qquad \mathbf{g} = \frac{\mathbf{g}_0}{x} = 0,$$

$$\text{def } q, r: \qquad q\cdot x + r = \text{num}\left(\frac{1}{x}\cdot x\cdot \sigma^{-1}\, 0 + 1\right) = 1 \Rightarrow q = 0, r = 1,$$

$$\mathbf{C} = \text{a basis for Ann}(\langle 1\rangle) = \mathbf{0}_{01}.$$

Thus the basis for the solution space is

$$\mathbf{C}((1)\wedge 0) = \mathbf{0}_{01}\wedge \mathbf{0}_0.$$

Since the quotient space is $\{\langle 0, 0\rangle\}$, the desired solution space will also be trivial; in other words, $\sigma g - g/x = 1/x + 1/x^2$ has no solution in $\mathbb{Q}(x)$. □

Reviewing this subsection, we need consider only a finite number of possible $\mathscr{f}$ such that $\sigma^k\mathscr{f}$ is a factor of the denominator of g. By the use of several reduction techniques, we can lower the degree to which any $\sigma^k\mathscr{f}$ might occur, until $\sigma^k\mathscr{f}$ can no longer be in the denominator of g. At the same time, the process "simplifies" f, so that eventually the power of $\mathscr{f}$ in den(f) is no greater than the power of $\mathscr{f}$ in den(a), including the case where this power is zero.

4. *Conclusion*

4.1 A FUNDAMENTAL THEOREM. In Section 2 we saw how to verify whether a tower of affine extensions is a $\prod\sum$-field over some field of constants, and in Section 3 we saw how to solve arbitrary first-order-linear difference equations in any $\prod\sum$-field. The question which this section answers is: How does one choose a particular $\prod\sum$-field in which to look for a solution to an equation?

In the various examples in this paper, we have always attempted to solve a difference equation in the smallest field in which it can be posed. As a case in point, consider $\sum i!$. In Example 14, we tried to find an expression for this in $\mathbb{Q}(n, n!)$ and failed. We will shortly see why looking in larger fields is essentially futile.

Definition 34. A *basis* for the $\prod\sum$-tower $F = F_0 \subseteq \cdots \subseteq F_n = E$ is a sequence $t_1, \ldots, t_n$, where $F_i = F_{i-1}(t_i)$. As usual, we define $\alpha_i, \beta_i \in F_{i-1}$ by $\sigma t_i = \alpha_i t_i + \beta_i$. Associated with a tower's basis and $a \in F$ is

$$S \triangleq \{i \mid 0 \neq \beta_i \in F \text{ and } \alpha_i = a\}.$$

(Recall that F_i is a $\prod$-extension of F_{i-1} if $\beta_i = 0$; otherwise it is a $\sum$-extension.)

Using S, there are "obvious" solutions to $\sigma g - a\cdot g = f$ for certain pairs a, f:

$$f = \sigma v - a\cdot v + \sum_{i\in S} c_i\cdot\beta_i \text{ where } v \in F$$

$$\Rightarrow v + \sum_{i\in S} c_i\cdot t_i \text{ is a solution.}$$

The interesting point is that if $f \neq 0$, *all* solutions in a $\prod\sum$-tower to $\sigma g - a \cdot g = f$ are essentially of the above form. The only technicality is that one must adjust the choice of basis relative to F and $a \in F$.

Definition 35. Given a tower and its basis as in Definition 34, and $a \in F$, we say that the basis is *normalized* wrt (with respect to) $a \Leftrightarrow$ *for* $i = 1, \ldots, n$,

$$\beta_i \neq 0 \text{ and } \frac{\alpha_i}{a} \in H(F_{i-1}) \Rightarrow \alpha_i = a.$$

We say that the basis is *reduced* wrt $F \Leftrightarrow$ for $i = 1, \ldots, n$,

$$\beta_i \neq 0 \text{ and } h \in F_{i-1} \text{ with } \beta_i + \sigma h - \alpha_i \cdot h \in F \Rightarrow \beta_i \in F.$$

We now see how to normalize and reduce a basis.

PROPOSITION 12. *Let a tower of height n be normalized through height* $n - 1$, *and suppose* $\alpha_n/a = \sigma w/w$ *for* $w \in F_{n-1}$. *Then letting* $t'_n = t_n/w$ *produces a normalized basis* $t_1, \ldots, t_{n-1}, t'_n$.

PROPOSITION 13. *Consider a tower of height n which is reduced through height* $n - 1$. *If* $\beta_n \neq 0$ *and* $\exists h \in F_{n-1}$ *with* $\sigma h - \alpha_n \cdot h + \beta_n \in F$, *let* $t'_n = t_n + h$. *Then* $t_1, \ldots, t_{n-1}, t'_n$ *is reduced wrt* F.

The following result is the difference field analog to Liouville's theorem on elementary integrals [16].

THEOREM 24. *Let* $t_1, \ldots, t_n$ *be a basis for the* $\prod\sum$*-tower* $F = F_0 \subseteq \cdots \subseteq F_n = E$. *Given* $a \in F$, *suppose that this basis is normalized wrt a and reduced wrt* F. *Given also a nonzero* $f \in F$, *suppose that* $\sigma g - a \cdot g = f$ *has a solution in* E. *Then*

$$\exists v \in F, \quad c_i \in K \qquad \text{with} \quad f = \sigma v - a \cdot v + \sum_{i \in S} c_i \beta_i.$$

In the summation case, $a = 1$. Loosely speaking, if f is summable in E, then part of it is summable in F, and the rest consists of pieces whose formal sums have been adjoined to F in the construction of E. This makes the construction of extension fields in which f is summable somewhat uninteresting and justifies the tendency to look for sums of $f \in F$ only in F.

4.2 FURTHER RESEARCH. There are a number of directions in which this work can be pursued. One place to look for generalization is in the class of equations which can be solved. The techniques of this paper rely very heavily upon linearity, suggesting that the generalization to nth order (or simultaneous) linear difference equations may not be too difficult. It is not so clear how to get interesting results on even small classes of nonlinear equations; the general case is very likely recursively unsolvable.

Another part of the theory which could stand generalization is the class of extensions considered. One might want to imbed difference fields in more general difference rings, so that $(-1)^n$ can be handled in a general way (note that $1/(1 + (-1)^n)$ is unreasonable, since it is $1/0$ half the time). Even staying within fields, there are perhaps other interesting cases to be examined; picking up an example from Subsection 1.1, one can model c^{2^t} by $\sigma t = t^2$. How difficult is this to analyze?

Even staying within $\prod\sum$-fields and first-order-linear equations, there is analysis of algorithms to be done. For example, there are techniques for summing rational functions or answering questions about $H(\mathbb{Q}(x))$ which do not require complete factorization of polynomials [7, 10]. Can similar techniques be profitably applied in

$\prod\sum$-fields? There are very likely other approaches which would improve the efficiency of the algorithms of this paper.

Finally, at a more mathematical level, one is always drawn back to the similarities of the differential versus difference cases. Particularly in light of Theorem 24, one might hope for a theory which would unify the results of this paper and the corresponding results for at least the transcendental case of differential field theory.

ACKNOWLEDGMENTS. I am especially grateful to Massachusetts Computer Associates, which supported a substantial part of this research, as well as preparation of the manuscript. I am indebted to Jane Cotreau, who patiently typed and retyped many versions of the paper. Thanks also go to Tom Cheatham, who inadvertently got me started.

REFERENCES

1. BOOLE, G. *Calculus of Finite Differences.* Chelsea Publishing Co., New York, 1970 (originally published in 1860).
2. CHEATHAM, T.E., AND TOWNLEY, J. Symbolic evaluation of programs. A look at loop analysis Proc. 1976 Symp. on Symbolic and Algebraic Computation, Yorktown Heights, N.Y, 1976, pp 90–96.
3. COHEN, J., AND KATCOFF, J. Symbolic solutions of finite-difference equations. *ACM Trans. Math. Softw. 3,* 3 (Sept. 1977), 261–271).
4. COHEN, J. Personal communication.
5. COHN, R M. *Difference Algebra.* Interscience, New York, 1965.
6. GOSPER, R.W JR. Indefinite hypergeometric sums in MACSYMA. Proc 1977 MACSYMA User's Conf., Berkeley, Calif, 1977, pp. 237–252.
7. GOSPER, R.W. JR Decision procedure for indefinite hypergeometric summation *Proc. Nat Acad Sci. USA 75,* 1 (1978), 40–42.
8. HERSTEIN, I.N. *Topics in Algebra.* Blaisdell Publishing Co., Waltham, Mass., 1964.
9 JOHNSON, S.C. On the problem of recognizing zero. *J. ACM 18,* 4 (Oct. 1971), 559–565.
10. KARR, M. Summation in finite terms (preliminary version). Tech. Rep. CA-7602-2911, Massachusetts Computer Associates, Inc., Wakefield, Mass., 1976.
11. KARR, M. Unpublished pieces of summation in finite terms. Tech Rep. CA-7904-1211, Massachusetts Computer Associates, Wakefield, Mass., 1979.
12. MACLANE, S., AND BIRKHOFF, G. *Algebra* MacMillan, New York, 1967, Th. X.12, p. 358.
13. MOENCK, R. On computing closed forms for summation. Proc. 1977 MACSYMA User's Conf., Berkeley, Calif., 1977, pp. 225–236.
14. RISCH, R. The problem of integration in finite terms. *Trans. AMS, 139* (1969), 167–189.
15. RISCH, R. The solution to the problem of integration in finite terms, *Bull. AMS 76* (1970), 605–608.
16. ROSENLICHT, M. Liouville's theorem on functions with elementary integrals. *Pacific. J. Math. 24,* 1 (1968), 153–161.
17. ROTA, G.-C., AND MULLIN, R. On the foundations of combinatorial theory. In *Graph Theory and Its Applications,* B. Harris, Ed., Academic Press, New York, 1970, pp. 167–213.
18. VAN DER WAERDEN *Modern Algebra.* Frederick Ungar Publishing Co., New York, 1931.
19. WASHINGTON, D. Personal communication.

RECEIVED SEPTEMBER 1977, REVISED APRIL 1979; ACCEPTED FEBRUARY 1980