ASPECTS OF SYMBOLIC INTEGRATION AND SIMPLIFICATION OF
EXPONENTIAL AND PRIMITIVE FUNCTIONS

by

MICHAEL ROTHSTEIN

A thesis submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

[Computer Sciences]

at the

UNIVERSITY OF WISCONSIN-MADISON

1976
To my optimistic wife Mary, who still doesn't believe it is done,
and to Bob, who has helped me so much.
ACKNOWLEDGMENTS

Many persons and entities have cooperated to bring this thesis to a fruitful conclusion, but I would like to give my special thanks to:

The Fulbright Commission in Bogotá, Colombia, the U.S. Department of State and the Institute of International Education for bringing me to the United States and sponsoring my stay here.

Professor B.F. Caviness now at Rensselaer Polytechnic Institute, for his help and suggestions during the preparation of this thesis.

Professor H.I. Epstein, of Boston College, for his many suggestions in connection with this thesis. In particular, the proof of corollary 3.13 is due to him.

Other persons from whom I thankfully acknowledge their comments regarding this thesis, are the readers (Professors George E. Collins, Richard A. Askey and Edward F. Moore, all of the University of Wisconsin) and Professor M. Rosenlicht of the University of California at Berkeley.

Support for this thesis has been provided in part, by the National Science Foundation through grants GJ 30125x, GJ 32181 and MCS76-15035 A01.

I also acknowledge the interest and support of the members of the Symbolic Manipulation Group at the University of Wisconsin, and of the Utah Computational Physics Group at the University of Utah.

Last but not least, I would like to thank Ms. Mary Ann Woolf, of Salt Lake City, Utah, for her heroics with the typewriter, and her patience with all the last minute additions, corrections, deletions, etc. which have happened to this thesis.
In this thesis we cover some aspects of the theory necessary to obtain a canonical form for functions obtained by integration and exponentiation from the set of rational functions.

These aspects include a new algorithm for symbolic integration of functions involving logarithms and exponentials which avoids factorization of polynomials in those cases where algebraic extension of the constant field is not required, avoids partial fraction decompositions, and only solves linear systems with a small number of unknowns.

We have also found a theorem which states, roughly speaking, that if integrals which can be represented as logarithms are represented as such, the only algebraic dependence that a new exponential or logarithm can satisfy is given by the law of exponents or the law of logarithms.
TABLE OF CONTENTS

Chapter 1 - Introduction ........................................... 1

Chapter 2 - An Elementary Integration Algorithm for Exponential and Logarithmic Functions

Section 1 - Introduction ........................................... 11
Section 2 - Some Necessary Concepts ..................................... 11
Section 3 - An Overview of the Algorithm .................................. 12
Section 4 - An Algorithm to Integrate Elements of $F_n$ (Exponential Case) ........... 14
Section 5 - An Algorithm for a Very Special Case of the Differential Equation $X' + vX = T$ ......... 16
Section 6 - An Algorithm to Integrate Elements of $S_n$ (Logarithmic Case) ............. 45
Section 7 - Integration of Rational Elements of $F_n$, or Rational Integration Revisited ............. 49
Section 8 - Integration of Normal Elements of $F_n$ ........................................ 53
Section 9 - Examples .................................................. 65
Section 10 - Computing Time Analysis for the Rational Function Case .................. 75

Chapter 3 - A Structure Theorem for Exponential and Primitive Functions

Section 1 - Introduction ........................................... 78
Section 2 - Some Basic Lemmas ......................................... 82
Section 3 - Simple-logarithmic Elements and Log-explicit Extensions ..................... 86
Section 4 - Liouville Fields and Extensions .................................. 97
<table>
<thead>
<tr>
<th>Section/Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 5 - Examples</td>
<td>101</td>
</tr>
<tr>
<td>Chapter 4 - Conclusions and Future Work</td>
<td>103</td>
</tr>
<tr>
<td>Appendix A</td>
<td>105</td>
</tr>
<tr>
<td>Appendix B - Finding Rational Roots of Polynomials</td>
<td>110</td>
</tr>
<tr>
<td>Bibliography</td>
<td>115</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

This thesis deals with some aspects of the theory necessary to obtain a canonical form of functions obtained by integration and exponentiation from the set of rational functions in one or several variables. In addition, a new algorithm is given for symbolic integration of functions obtained from the rational functions by composing the logarithm and exponential functions.

The main problems addressed and solved here are:

(a) Can we find an improved algorithm for symbolic integration of functions involving logarithms and exponentials? An algorithm has been found for elementary integration of these functions, which avoids factorization of polynomials in those cases where algebraic extension of the constant field is not required, avoids partial fraction decompositions, and only solves linear systems with a small number of unknowns.

(b) What possible algebraic relationships are there between functions built-up from the rational functions using integration and exponentiation?

A theorem has been found which states, roughly speaking, that if integrals which can be represented by using logarithms are represented as such, the only algebraic dependence that a new exponential or logarithm can satisfy is given by the law of exponents or the law of logarithms.
The search for algorithms for elementary integration of elementary functions has been the subject of many efforts in the past. The pioneer of this subject, Joseph Liouville based his work [LIO 33, LIO 33A, LIO 35, LIO 37, LIO 39, LIO 40 and LIO 41] on the fact that the derivative of an elementary function is again an elementary function. He also found the form an integral could take. C. Hermite was the first person to discover an algorithm for integration of rational functions. His work appeared post-humously in 1912 [HER 12]. The first algorithm sketch for integration in finite terms, was written by D.D. Mordoukhay-Boltovsky [MOR 13]. Still, this sketch had some ambiguities. G.H. Hardy [HAR 28] and J.F. Ritt [RIT 48], both published somewhat more detailed sketches of their algorithms. Ritt, and after him, E.R. Kolchin, also were the primary contributors to differential algebra, which provides the framework for most of the work done today in the area. Further details of their work can be found in [RIT 50] and [KOL 73]. Also of interest during this period were A. Ostrowski's papers [OST 46] generalizing a result due to Liouville about the form of an integral and [OST 46A] giving his integration algorithm for primitive functions.

Then came the electronic computer, and the realization that this was a formidable tool for automatic algebra, and perhaps also integration. Thus Slagle [SLA 61] implemented a heuristic integrator program he called SAINT. J. Moses [MOS 67] implemented an improved heuristic integrator he called SIN, and which, (with one addition) is still being used today as part of the symbolic manipulation system MACSYMA. Also at this time, R.G. Tobey [TOB 67] and E. Horowitz [HOR 69] found better algorithms for integration of rational functions. M. Rosenlicht also contributed to the field with a modern rendering of Liouville's Theorem
on integration in elementary terms [ROS 68] and some additional theorems which are useful for determining algebraic dependence and independence of elementary functions [ROS 69], [ROS 75].

The problems of function representation and simplification which are crucial for symbolic integration, also received attention by D. Richardson [RIC 68], who showed that the lack of a normal form implied the impossibility of having an integration algorithm. W. S. Brown [BRO 69] and B.F. Caviness [CAV 70] attacked the problem by reducing it to finding algebraic relationships between constants. It was also during this period that the first complete, correct, algorithm for integration in finite terms, was published by R.H. Risch [RIS 69], [RIS 70]. His work also included the discovery of a simple method to discover all possible algebraic relationships between the logarithm and exponential functions, provided that certain problems with constants are avoided (Risch Structure Theorem) [RIS 69a].

Risch's 1969 paper spurred a whole lot of activity: J. Moses implemented Risch's algorithm and incorporated it within his SIN program. Comments, further explanations and minor modifications were given by J. Moses [MOS 71], C. Mack [McC 76] and D. Mack [McD 75]. B.M. Trager [TRA 76] found a method to obtain the minimal algebraic extension necessary to express the integral. Based on a result discovered by R.H. Risch in 1976 about the exact form of the algebraic part of the integral, R.H. Risch and A.C. Norman discovered and implemented a new integration algorithm in SCRATCHPAD. Also worthy of mention is a result of Moses [MOS 69] about using the Risch integration algorithm for some primitive functions (implemented in MACSYMA for error functions) a result which has been partially implemented for Spence functions (di-logarithms) by
J.A. Fox and A.C. Hearn in REDUCE [FOH 74].

Insofar as recognizing zero is concerned, there is still much interest in the area, as is shown by the work of J. Ax [AX 71], S.C. Johnson [JOH 71], R. Fateman [FAT 72], J. Fitch [FIT 73], M. Singer [SIN 74], [SIR 75] and H.I. Epstein [EPS 75], who implemented the Risch Structure Theorem in SAC-1.

As mentioned before, research on elementary integration carries with it research on the problem of simplification of elementary functions or the structure of fields of functions, since Risch's integration algorithm (and also the one described herein) require a method for verifying when two expressions differ by at most a constant. The relationship becomes mutual when we note that if we are considering expressions built-up from the rational functions using exponentiation and integration, one known way to distinguish algebraically independent functions requires an operation which we shall call simple-logarithmic integration, similar to elementary integration, but not allowing expansion of the constant field. By Liouville's theorem, this is the same as verifying whether an integral can be expressed by using logarithms but with no algebraic extensions allowed.

Further details and a proof of the above will be presented in Chapter 3 and Appendix A.

The algorithms required for our integration algorithm are the usual polynomial and rational function algorithms, i.e. addition, substraction, multiplication, greatest common divisors, resultants, etc., including an extended Euclidean algorithm (or its equivalent) and an algorithm for finding roots of univariate polynomials over the constant field. One subalgorithm not directly required by our algorithm is polynomial
factorization. In this way, we will obtain better computing time bounds than has been possible, so far, for the special case where no algebraic extensions are necessary.

Let me finish this introduction with some definitions and lemmas from differential algebra. For proofs, see [EpC 74].

**Differential Fields**

By a differential field we will mean a field $F$ of characteristic 0 with operators $D_1, \ldots, D_k$, $k \geq 1$, defined on $F$ and satisfying for all integers $i, j$, $(1 \leq i, j \leq k)$ and $x, y \in R$.

1. $D_i(x + y) = D_i x + D_i y$.
2. $D_i(x y) = x D_i y + y D_i x$.
3. $D_i D_j x = D_j D_i x$.

$F$ will be called **partial** if $k > 1$, if $k = 1$, $F$ will be called **ordinary**. $D_1, \ldots, D_k$ will be called the **derivation operators** of $F$.

We can now prove for any derivation operator $D$ of $F$:

(a) $D(1) = D(1 \cdot 1) = D(1) + D(1)$ so that $D(1) = 0$.

(b) If $a \neq 0$, $0 = D(1) = D(a a^{-1}) = a D(a^{-1}) + a^{-1} D(a)$, so that $D(a^{-1}) = -D(a)/a^2$. This implies that $D(a/b) = b Da - a Db / b^2$.

(c) If $n$ is a positive integer, we can prove (by induction) that $D(a^n) = n a^{n-1} Da$. From (b), we can conclude that this property holds for all integers $n$.

Now let $F$ and $G$ be differential fields. $G$ will be called a **differential extension field of $F$** if $G$ is an extension of $F$ on the algebraic sense, both $F$ and $G$ have the same number of differential operators, say $(D_1, \ldots, D_k)$ and $(\overline{D}_1, \ldots, \overline{D}_k)$ respectively, and for each $x$ in $F$ and integer $i$ between 1 and $k$, we have $D_i(x) = \overline{D}_i(x)$. 
Let $F$ and $G$ be two differential fields with differential operations $D_1, \ldots, D_k$, and $\overline{D}_1, \ldots, \overline{D}_k$ respectively. Let $f$ be a function from $F$ into $G$. Then $f$ will be called a differential homomorphism if $f$ is a homomorphism in the algebraic sense, $k = \overline{k}$, and for all $i$ between 1 and $k$, all $x$ in $F$, we have $f(D_i(x)) = \overline{D}_i(f(x))$.

The concepts of differential isomorphism and differential embedding can be defined similarly.

Let $F$ be a differential field, $U$ a differential extension field of $F$ having the property that every finitely generated differential extension field $G$ of $F$ can be embedded in $U$, the embedding leaving $F$ fixed; under these conditions, $U$ will be called a universal extension field of $F$. It is proven in [KOL 73, page 92] that every differential field has a universal extension. Note that Kolchin calls these semi-universal extensions.

Let $F$ be a differential field with differentiation operators $D_1, \ldots, D_k$. Let $K$ be a set of elements of $F$ which are mapped to 0 by all the $D_i$, $1 \leq i \leq k$. Then $K$ is a subfield of $F$ called the field of constants of $F$, or the constant field of $F$.

In this thesis, we shall denote by $\mathcal{C}$, the constant field of a universal extension field of the differential field under consideration. We shall also denote the constant field of $F$ by $C^F$ so that $\mathcal{C} = C^U$ and $C^F = \mathcal{C} \cap F$.

Now let $F$ be a partial differential field, $G$ a differential extension field of $F$. Let $\emptyset \in G$, $\emptyset \neq 0$. Then

(a) $\emptyset$ will be called primitive over $F$ if for all derivations $D$ of $G$, $D \emptyset \in F$. 
(b) $\Theta$ will be **exponential over $F$** if for all derivations $D$ of $F$, 
$\exists u \in F$ such that $Du = D\Theta/\Theta$. In this case we say $\Theta = \exp u$.

(c) $F(\Theta)$ will denote the smallest partial differential field containing both $F$ and $\Theta$.

(d) $\Theta$ will be called a **regular monomial over $F$** if:

i) $\Theta$ is either primitive or exponential over $F$

ii) $\Theta$ is transcendental over $F$

iii) $F$ and $F(\Theta)$ have the same constant field.

(e) If for some $u$ in $F$ and all derivations $D$ of $F$, $D\Theta = \frac{Du}{u}$, we will write $\Theta = \log u$.

Notice that if for some $\Theta$ and $\overline{\Theta}$, and some $u \in G$, we have for all derivations $D$ of $F$, $D\Theta = Du/u = D\overline{\Theta}$, then $\Theta - \overline{\Theta} \in C$. In a similar vein, if $D\Theta = \Theta Du$ and $D\overline{\Theta} = \overline{\Theta} Du$, for all derivations $D$, then $\Theta/\overline{\Theta} \in C$.

Henceforth, when confronted with this situation, we shall select one particular $\Theta$, call it $\log u$, and all others will be represented by $\log u + c$. $\exp u$ will be treated similarly. Notice also that for $\Theta, \phi$ in $F$, $\Theta = \log \phi + c$ for some $c \in C^F$, if, and only if $\phi = k \exp \Theta$ for another $k$ in $C^F$.

If we let $F$ and $G$ be as before, and $\Theta_1, ..., \Theta_n$ in $G$, then:

(a) $F(\Theta_1, ..., \Theta_n)$ will denote the smallest differential field containing $F$ and the $\Theta_i$.

(b) $F(\Theta_1, ..., \Theta_n)$ will be called a **regular Liouville extension of $F = F_0$** if each $\Theta_i$ is a regular monomial over $F_{i-1} = F(\Theta_1, ..., \Theta_{i-1})$.

(c) $F(\Theta_1, ..., \Theta_n)$ will be called a **generalized Liouville extension** of $F = F_0$ if each $\Theta_i$ is either a regular monomial or algebraic over $F_{i-1} = F(\Theta_1, ..., \Theta_{i-1})$ and $F(\Theta_1, ..., \Theta_n)$ has the same field of constants as $F$. 
In the following lemmas, which have been taken from [EpC 74], let $F$ be a differential field with derivation operators $D_1, \ldots, D_k$, let $U$ be a universal extension of $F$, and let $\theta$ in $U$ be a regular monomial over $F$.

1. For $f, g$ in $U$, exponential over $F$, and $m$ a non-zero integer, $f/g^m$ is a constant (in $U$) if and only if for all $i$, $1 \leq i \leq k$,

$$\frac{1}{m} f D_i f = \frac{1}{g} D_i g .$$

Also, in this case, $f$ is not a regular monomial over $F(g)$.

2. Let

$$P(\theta) = \sum_{i=0}^{n} p_i \theta^i$$

in $F[\theta]$, $p_n \neq 0$. Then $\max(\deg D_i P) = n$ unless $\theta$ is primitive over $F$, $1 \leq i \leq k$ and $p_n$ is a constant in which case

$$\max(\deg D_i P) = n - 1 .$$

3. Let

$$P(\theta) = \sum_{i=0}^{n} p_i \theta^i$$

in $F[\theta]$, $p_n \neq 0$. If, for each $j$, $1 \leq j \leq k$,

$$P(\theta) | D_j P(\theta) ,$$

then $\theta$ is an exponential monomial and $P(\theta) = p_n \theta^n$.

4. If $P(\theta)$ in $F[\theta]$ is square-free, then the only (nontrivial) common factor that $P(\theta), D_1 P(\theta), \ldots, D_k P(\theta)$ can have is $\theta$, and that only if $\theta$ is exponential over $F$. (This is a corollary of 3.)
5. If \( P, Q \) are relatively prime elements of \( F(\Theta) \), \( \text{deg} \ Q > 0 \), and for each \( i, 1 \leq i \leq k \), \( D_i(P/Q) = A_i/B_i \), where \( A_i, B_i \) are relatively prime in \( F(\Theta) \), then each \( B_i \) is square-free if and only if \( Q = q \Theta \), for some \( q \) in \( F \).

6. Let \( R(\Theta) \in F(\Theta) \), \( R(\Theta) \neq 0 \). Then, for any \( i, 1 \leq i \leq k \), if \( D_iR(\Theta)/R(\Theta) = A/B \), where \( A, B \) are relatively prime, then \( B \) is square-free.

7. Let \( S \in U \), and for each \( i, 1 \leq i \leq k \), \( D_iS = P_i/Q_i \), with \( P_i, Q_i \) relatively prime elements of \( F(\Theta) \), \( Q_i \) square-free, \( \text{deg} \ P_i < \text{deg} \ Q_i \).

Assume also that

\[
D_iS = \sum_{j=1}^{m} c_j D_i R_j/R_j + D_i T
\]

where the \( c_j \in C^F \), \( R_j, T \in F(\Theta) \). Then, if \( \Theta \) is primitive over \( F \),

\[
D_i T = \sum_{j=1}^{\ell} k_j D_i f_j/f_j + c
\]

for some integer \( \ell \geq 0 \), \( f_j \in F \), constants \( k_j \), \( c \). If \( \Theta \) is exponential over \( F \), then \( \Theta T \) is in \( F[\Theta] \), and if \( T \in F[\Theta] \), then \( T \in F \).

8. Let \( \psi \) in \( U \), and suppose \( \psi \) is not a regular monomial over \( F \).

Then

(a) If \( \psi \) is primitive over \( F \), \( \psi = g + c_1 \) for some \( g \) in \( F \), \( c_1 \) in \( C^U \).

(b) If \( \psi \) is exponential over \( F \), then \( \psi^m = c_2 h \) for some non-zero integer \( m \), \( h \in F \) and \( c_2 \) in \( C^U \).

9. Suppose \( F^* \) is a differential field and \( \sigma \) is a differential isomorphism from \( F^* \) to \( F \). Let \( A \in F^* \). Then \( \log A \) is a regular monomial over \( F^* \) if and only if \( \log \sigma A \) is a regular monomial over \( F \), and \( \exp A \)
is a regular monomial over $F^*$ if and only if $\exp \sigma A$ is a regular monomial over $F$.

Furthermore, in this case, $\sigma$ can be extended to map $F^*(\log A)$ ($F^*(\exp A)$) into $F(\log \sigma A)$ ($F(\exp \sigma A)$) isomorphically.

10. (Ostrowski) Let $\Theta_1, \ldots, \Theta_m$ be primitive over $F$, and assume $C_F = C^{F(\Theta_1, \ldots, \Theta_m)}$. Then $\Theta_1, \ldots, \Theta_m$ are algebraically dependent over $F$ if, and only if, there are constants $c_1, \ldots, c_m$ in $C_F$, not all zero, such that

$$\sum_{j=1}^{m} c_j \Theta_j$$

is in $F$.

We will also use some results which are corollaries of Risch's algorithm described in [RIS 69], and refer to the form that the integral of an expression $f$ in $F(\Theta)$ can have. The results we use, though, and don't prove, can be proven quite easily using the properties above.
Chapter 2

AN ELEMENTARY INTEGRATION ALGORITHM FOR EXPONENTIAL AND LOGARITHMIC FUNCTIONS

1. Introduction.

In this chapter, we shall present a new integration algorithm for exponential and logarithmic functions, whose main novelties are:

a. This algorithm determines the minimal algebraic extension field in which to express the integral. If no algebraic extension is necessary, this algorithm does not invoke factorization thereby yielding a faster algorithm for this particular case.

b. No systems of linear equations are set-up, except for those solved in the algebraic independence operations. Instead, univariate polynomial equations (U.P.E.'s) of the form \( AX + BY = C \) are solved for \( X \) and \( Y \), where \( A, B, C, X \) and \( Y \) are univariate polynomials with rational function coefficients, \( \gcd(A, B) = 1 \), and \( \deg Y < \deg A \). In [YUN 76a] a p-adic algorithm is suggested for some similar problems which can be applied here.

2. Some Necessary Concepts.

Our algorithm will work on differential fields of the form \( F = \mathbb{F}_n = K(z, \Theta_1, \ldots, \Theta_n) \), where \( K \) is the constant field of \( F \), \( z' = 1 \) (\( z \) is the integration variable) and each \( \Theta_i \) is a regular monomial (either logarithmic or exponential) over \( \mathbb{F}_{n-1} = K(z, \Theta_1, \ldots, \Theta_{i-1}) \), \( F_0 = K(z) \).

We will require the existence of algorithms to perform arithmetic.

11
in $K$, and also, algorithms for the usual arithmetic operation for the
domains $S_i = F_{i-1}[\Theta_i]$ and $E_{k} = \{P/\Theta_i^k, P \in S_i, n \in \mathbb{Z}\}$ though most of these
operations can be replaced by a similar operation in $P_i = R[Z, \Theta_1, \ldots, \Theta_i]$ (where $R$ is a subring of $K$ whose fraction field is $K$).

We will also require solution of U.P.E.'s for elements of $S_i$ and
resultants of elements of $S_i$. We will denote these as resultant $(A,B,\Theta_i)$, where we assume that $A,B \in S_i$, though, on occasion, we will not specify the
third argument when there is no risk of confusion.

Let us also define $D_n = E_n$ if $\Theta_n = \exp u, u \in F_{n-1}$, otherwise

$$D_n = S_n.$$ 

3. An Overview of the Algorithm.

In this section, we decompose the problem of integrating an elementary function, to subproblems which will be solved in the remaining sections of this chapter. But first, we will need to give precise definitions of some well-known concepts. These are:

a. Given two arbitrary non-zero elements of $S_n$, their gcd is always monic.

b. Given a non-zero element $f$ of $F_n$, there exist unique $P$ and $Q$ in
$S_n$ such that $P/Q = f$, $\gcd(P, Q) = 1$, and $Q$ is monic. We shall call $P$
the numerator (denoted by $\text{num } f$) and $Q$ the denominator (denoted by
$\text{den } f$) of $f$.

c. We can also consider $z = \Theta_0$, and $F_{-1} = \mathbb{K}$. Thus all definitions
will also apply to $F_0 = \mathbb{K}(z)$.

d. For $f \in F_n$, we will say that $f$ is a proper element of $F_n$ if
$f = 0$, or $\deg(\text{num } f) < \deg(\text{den } f)$, and also, if $\Theta_n$ is exponential over
$F_{n-1}$ then $\Theta_n$ does not divide $\text{den } f$. This implies that all irreducible
factors q of \( \text{den } f \) satisfy \( \gcd(q, q') = 1 \).

We will also say that \( f \) is a normal element of \( F_n \) (or that \( f \) is normal), if \( f \) is proper and \( \gcd(\text{den } f, (\text{den } f)') = 1 \). The significance of this concept lies in the fact that if \( f \) is normal in \( F_n \) and elementarily integrable, then \( \int f \) can be expressed as

\[
\int f = k u + \sum c_i \log v_i, \quad k, c_i \in K
\]

where the \( v_i \) are monic in \( S_n \). Then \( k \neq 0 \) if and only if \( \Theta_n = \exp u \).

We are now ready to state the outline of our integration algorithm. Starting with an element \( f(z) \) of \( F_n \), we first break it up as

\[
F(z) = P(z) + g(z),
\]

where \( P(z) \) is in \( D_n \) and \( g(z) \) is a proper element of \( F_n \). In order to obtain this decomposition, we first let \( \text{num } f(z) = Q(z), \text{den } f(z) = R(z) \), and find \( P_1(z), S(z) \in S_n \) such that

\[
Q(z) = P_1(z)R(z) + S(z), \quad \deg S(z) < \deg R(z)
\]

so that

\[
f(z) = \frac{Q(z)}{R(z)} = P_1(z) + \frac{S(z)}{R(z)}.
\]

If \( D_n = S_n \) (\( \Theta_n \) is logarithmic over \( F_{n-1} \) or \( n = 0 \)), we are done, otherwise, let \( R(z) = \Theta_n^k T(z) \), where \( \Theta_n \) does not divide \( T(z) \), and solve the U.P.E.

\[
S(z) = \Theta_n^k S_1(z) + P_2(z) T(z),
\]

\[
\deg S_1(z) < T(z), \quad \deg P_2(z) < k
\]
for $S_1$ and $P_2$. Then

$$\frac{S(z)}{R(z)} = \frac{S_1(z)}{T(z)} + \frac{P_2(z)}{\theta n}$$

and we can let

$$P(z) = P_1(z) + \frac{P_2(z)}{\theta n}$$

and

$$g(z) = \frac{S_1(z)}{T(z)}$$

which is proper in $F_n$.

We integrate $P(z)$ in a term-by-term fashion (as described in sections 4, 5 and 6) and decompose

$$\int g(z) = r(z) + \int h(z)$$

where $h(z)$ is a normal element of $F_n$. An algorithm that performs this decomposition, designed by D. Mack (see [McD 75]) will be described in section 7. In section 8, we will describe a new algorithm that finds the necessary logarithmic terms. Section 9 will be devoted to some examples, and section 10 will give a computing time analysis of a subcase of the rational function case.

4. An Algorithm to Integrate Elements of $E_n$ (Exponential Case).

The method used by our algorithm to integrate
\[ P(z) = \frac{p}{\sum_{i=-m}^{n} a_i \Theta_n^i} \]

(m, p \geq 0, a_i \in F_{n-1}, \Theta_n = \exp u, u \in F_{n-1}) works as follows: We have to (recursively) integrate \( a_0 \). Then, for \( i \neq 0 \), \( \int a_i \Theta_n^i = x_i \Theta_n^i \) for some \( s_i \) in \( F_{n-1} \), as was shown in [RIS 69]. If we now differentiate both sides of this equation and divide by \( \Theta_n^i \), we obtain the differential equations:

\[ x'_i + i u' x_i = a_i \]

or

\[ x'_i + v'_i x_i = a_i \]

where \( v_i = i u \) for \(-m \leq i < 0\), or \( 0 < i \leq p \) and \( \exp(v_i) \) is a regular monomial over \( F_{n-1} \). This implies that equations of the form

\[ x' + v'_i x = T \]

(for any \( T \) in \( F_{n-1} \)) can have at most one solution in \( F_{n-1} \).

Notice also, that, if \( X, T, V_i \) are in \( F_{n-1} \) then since the general solution of the equation above is

\[ X = \frac{\int T e^{V_i}}{e^{V_i}} \]

the statements \( x' + v'_i x = T \) does not have a solution in \( F_{n-1} \) and \( \int T e^{V_i} \) is not elementary, are equivalent, since \( \int T e^{V_i} \), if elementary, must be of the form \( X e^{V_i} \), with \( X \) satisfying
\[ X' + v I X = T \]

where \( X \in F_{n-1} \).

So far, the contents of this section are very similar to the method used in [RIS 69]. Our method to solve the differential equation above, though, is very different, as will be seen in the next section.

5. An Algorithm for a Very Special Case of the Differential Equation
\[ X' + v X = T. \]

We want to solve the differential equation \( X' + u' X = T(u', T \in F_n) \) knowing that \( \exp(u) \) is a regular monomial over \( F_n(u) \). The algorithm we shall discuss in this section, operates recursively on \( n \), thus requiring us not to assume that \( u \) is in \( F_n \). This also implies that applying one of the subsections below, will mandate applying another one, or even the same one in most cases.

Actually, the method we use is dependent on the form \( u' \) and \( T \), and thus, we have a total of eight different cases:

i) The simplest case is when \( T/u' \) is a constant. In that case, \( X = T/u' \) is the only solution possible.

ii) Next, we want to solve the case where \( u' \) is a constant, and \( T \in K[z] \) (where \( z \) is the integration variable).

In that case, let \( T \) be of the form

\[ T = \sum_{i=0}^{m} a_i z^i. \]

Then, \( X \) must of the form
\[ X = \sum_{i=0}^{m} x_i z^i \]

and if we substitute these values in the equation above, we obtain:

\[ \sum_{i=0}^{m-1} (i+1)x_i+1 z^i + \sum_{i=0}^{m} x_i u' z^i = \sum_{i=0}^{m} a_i z^i \]

so that

i) \[ x_m = \frac{a_m}{u'} \]

ii) For \( 0 \leq i < m \),

\[ x_i = \frac{a_i - (i+1)x_{i+1}}{u'} \]

As a corollary, we point out that

\[ \int P(z) e^{az+b} \, dz, \]

where \( a, b \in K, P(z) \in K[z] \) is elementary, for any \( a, b, P \).

iii) If \( \Theta_n = \exp v, v \in F_{n-1}, u' \in F_{n-1} \) and \( T \in E_n \), let

\[ T = \sum_{i=-m}^{p} f_i \Theta_n^i, f_i \in F_{n-1} \]

Then \( X \) must be of the form

\[ X = \sum_{i=-m}^{p} x_i \Theta_n^i \]

and substituting, we obtain:
\[ \sum_{i=-m}^{p} (x_i^i \otimes_i n + i x_i v' \otimes_n i + x_i u' \otimes_n i) = \sum_{i=-m}^{p} f_i \otimes_n i \]

and we reach the equations

\[ x_i^i + x_i (iv + u)' = f_i \]

for \(-m \leq i \leq p\) with \(x_i\), \((iv + u)'\), \(f_i \in F_{n-1}\).

Notice that \(\exp(iv + u) = \otimes_n^i \exp(u)\) is a regular monomial over \(F_n(iv + u)\), and a fortiori, over \(F_{n-1}(iv + u)\), so that we simply solve these equations inductively.

iv) If \(\otimes_n^i = \log v, v \in F_{n-1}, u' \in F_{n-1}\), and \(T \in S_n\), let

\[ T = \sum_{i=0}^{m} f_i \otimes_n i \]

Again, \(X\) must be of the form

\[ X = \sum_{i=0}^{m} x_i \otimes_n i, \]

so that substituting, we obtain:

\[ \sum_{i=0}^{m} x_i^i \otimes_n i + \sum_{i=0}^{m-1} (i+1) x_{i+1} v' \otimes_n i + \sum_{i=0}^{m} x_i u' \otimes_n i = \sum_{i=0}^{m} f_i \otimes_n i \]

and noting that \(\exp(u)\) is still a regular monomial over \(F_{n-1}(u)\), we can solve the following differential equations inductively:

a. \(x_m^i + u_x = f_m\)

b. \(x_i^i + u_x = f_i - (i+1) x_{i+1} v' (0 \leq i < m)\) in the order of decreasing \(i\).
v) If we reach a differential equation of the form

$$AX' + BX = C,$$

where $A, B, C, X$ are in $S_n$, we apply algorithm SPDE which will be described after we justify it.

Assume we are given $A, B, C$, in $S_n$, with $\text{gcd}(A, B) = 1$ and $\deg A > 0$, and we want to find a $Y$ in $S_n$ such that

$$AY' + BY = C.$$

Assume, for the moment, that $Y$ were known, and that $Y = QA + R$. If we substitute this value, we obtain:

$$A(Q' A + QA' + R') + B(QA + R)$$

$$= A(Q' A + QA' + B + R') + BR = C$$

We have thus proven the following theorem.

**Theorem:**

Let $A, B, C$ be elements of $S_n$ such that $\text{gcd}(A, B) = 1$ and $\deg A > 0$.

Then, there exists a $Y$ in $S_n$ such that

$$AY' + BY = C$$

if and only if there exist $Q, Z, R$ in $S_n$ such that $\deg R < \deg A$,

$$AZ + BR = C,$$

and

$$A(Q' + (A' + B)Q = Z - R')$$
and then

$$Y = QA + R.$$ 

If they exist, Z and R are unique. (We have really only proven one half of this theorem, the other half can be proven by computing

$$A(QA + R)' + B(QA + R)$$

and applying the two equations above.) This theorem then, proves that the following algorithm will correctly solve a big proportion of the problem without requiring recursion.

So, assume we are given A, B, C as described above, and also n, a bound on the degree of the solution Y. Assume also that such a Y, if it exists, is unique (that will always be the case here).

The algorithm looks like this:

Algorithm $Y \leftarrow SPDE(A, B, C, \ell)$

If $\ell < 0$ then no solution can exist, otherwise, let $G = \gcd(A, B)$. If $G$ does not divide $C$, then no solution can exist either, otherwise, let $\overline{A} = A/G$, $\overline{B} = B/G$, $\overline{C} = C/G$. If $\deg \overline{A} = 0$, we have to apply one of the other cases recursively, otherwise, since $\gcd(\overline{A}, \overline{B}) = 1$, we can find $Z, R$ such that

$$\overline{A}Z + \overline{B}R = \overline{C} \quad \deg R < \deg \overline{A}$$

If $Z = R'$, then the solution to our problem is $R$, otherwise, we return

$$\overline{A} \leftarrow SPDE(\overline{A}, \overline{B} + \overline{A}'$, $Z - R'$, $\ell - \deg \overline{A}) + R$$

assuming no failure occurred within the recursive call of $SPDE$. 
Notes:  a. Notice that this algorithm works, because $Z$ and $R$ are unique. This is also the reason why $\bar{A}$, $\bar{B}$ are used throughout.

b. This algorithm will always terminate, since $\ell$ decreases each time SPDE is called directly.

c. There is a possibility of $\bar{A}$ being 1 in this algorithm. We call this phenomenon "coefficient degradation," and will show that the algorithm will still apply to the resulting equation when it arises.

vi) If $\theta_n = \exp v$, $v \in F_{n-1}$, and $u'$, $T$ in $E_n$, $u'$ not in $F_{n-1}$, we must have $X$ in $E_n$, i.e. $X = \bar{X}/\theta_n^Y$ for some $\bar{X}$ in $S_n$. Thus, we now want to find a similar equation for $\bar{X}$, which we find as follows.

Let $u' = \bar{u}/\theta_n^W$, $T = \bar{T}/\theta_n^S$, $\bar{u}$, $\bar{T} \in S_n$, $\gcd(\theta_n, \bar{u}) = \gcd(\theta_n, \bar{T}) = 1$.

If we replace these values in our differential equation, we obtain:

$$X' + u' X = T = \frac{\bar{X}' - y v'}{\theta_n^Y} + \frac{u \bar{X}}{\theta_n^{y+W}} = \frac{\bar{T}}{\theta_n^S}$$

and, to find $y$, we get the following possibilities:

a. If $w > 0$, then $y + w = s$, so $y = s - w$ and thus, our equation becomes:

$$\bar{X}' \theta_n^W + \bar{X}(-y v' \theta_n^W + \bar{u}) = \bar{T}.$$

This equation can be solved by using SPDE with the following degree bound.

Let

$$\bar{X} = \sum_{i=0}^{t} x_i \theta_n^i, \quad \bar{u} = \sum_{i=0}^{t} u_i \theta_n^i.$$
We then have three cases.

If \( t > w \), then \( r = \deg \bar{T} - t \).

If \( t < w \), then \( r \leq \max(\deg \bar{T} - w, y) \), since if \( r > \deg \bar{T} - w \), and \( r \neq y \), then, looking at the leading coefficient, we see:

\[
x_r x_r' + x_r' + x_r(-y v') = 0
\]

so that

\[
\frac{x_r'}{x_r} + (r - y) v' = 0
\]

so that

\[
x_r \otimes_{n}^{r-y}
\]

is a constant, which is impossible because \( x_r \in F_{n-1} \).

Finally, if \( t = w \), we determine \( r \) by noting that if \( r > \deg \bar{T} - w \), then

\[
x_r' + r x_r v' + x_r(-y v' + u_t) = 0
\]

thus

\[
\frac{x_r'}{x_r} + (r - y) v' + u_t = 0
\]

which says that

\[
f - u_t = (r - y) v + \log x_r + k
\]
In order to find \( r \), we first invoke the integration algorithm recursively with \( y v' - u_t \) obtaining either failure, (so that \( r < \deg \bar{T} - w \)) or

\[
\int (y v' - u t) = j + \sum_{i=1}^{m} c_i \log v_i = r v + \log x_r + k
\]

By repeatedly applying the Risch Structure Theorem (see Corollary 3.13 for a generalization thereof, and also [EPS 75] for an implementation) we may assume that each \( \log v_i \) is a regular monomial over \( \mathbb{F}_{n-1}(\log v_1, \ldots, \log v_{i-1}) \).

We may further assume that the \( c_i \) are linearly independent over the rationals since if (say) \( c_1 = \sum_{i=2}^{m} \frac{p_i}{q} c_i \), where \( q \) and \( p_i \) are rational integers, we can write

\[
\int (y v' - u t) = j + \sum_{i=2}^{m} \left( \frac{p_i}{q} c_i \log v_i + c_i \frac{p_i}{q} \log v_i \right)
\]

\[= j + \sum_{i=2}^{m} c_i \left( \frac{p_i}{q} \log(v_1 v_i) \right) + \text{a constant}
\]

(the first step follows from our assumed value for \( c_1 \) and the collection of like \( c_i \)).

We claim that, if \( r > \deg \bar{T} - w \), then \( m = 1 \), and \( c_1 \) is a rational, since

\[
\log x_r = j + \sum_{i=1}^{m} c_i \log v_i - r v - k
\]

so that
\[ x_r = \exp \left( j + \sum_{i=1}^{m} c_i \log v_i - r v - k \right) \]

is not a regular monomial over \( F_{n-1}(\log v_1, \ldots, \log v_m) \), so that applying the Structure Theorem (in the form of theorem 3.12), we obtain that all the \( c_i \) are rationals, since \( j - r v - k \in F_{n-1} \).

This reasoning also implies that \( \exp(j - r v) \) is not a regular monomial over \( F_{n-1} \) and we obtain the linear equation

\[ j - r v = \sum_{i \in L} q_i \Theta_i + \sum_{i \in E} q_i a_i + \bar{k} \]

where

\[ L = \{ i : 1 \leq i \leq n - 1, \Theta_i = \log a_i, a_i \in F_{i-1} \} \]

\[ E = \{ i : 1 \leq i \leq n - 1, \Theta_i = \exp a_i, a_i \in F_{i-1} \} \]

\( \bar{k} \) is some constant, and the \( q_i \) are rational numbers.

We solve this equation by differentiating it, obtaining

\[ j' - r v' = \sum_{i \in L} q_i \frac{a_i'}{a_i} + \sum_{i \in E} q_i a'_i \]

then, finding \( a_i, \beta, j_o, v_o \) in \( K[z, \Theta_1, \ldots, \Theta_{n-1}] \) such that

\[ j' = \frac{j_o}{\beta}, \quad v' = \frac{v_o}{\beta}, \quad \frac{a_i'}{a_i} = \frac{\alpha_i}{\beta} \]

if \( i \in L \), and
\[ a'_i = \frac{\alpha_i}{\beta} \]

if \( i \in E \), obtaining the equation

\[ j_o - r\nu_o = \frac{\sum_{i=1}^{n-1} q_i \alpha_i}{\nu_o} \]

which can be converted into a system of linear equations with coefficients in \( K \), by equating like powers in \( z \) and the \( \Theta_i \).

If there exists an integer \( r \) satisfying this linear equation, and this value is greater than \( \text{deg } \overline{T} - w \), we use the value as our bound, otherwise, we use \( \text{deg } \overline{T} - w \).

We point out that this equation determines \( r \) uniquely, since otherwise \( \Theta_n = \exp v \) would not be a regular monomial over \( F_{n-1} \).

Finally, let us point out that coefficient degradation cannot happen in our application of interest, which was the equation

\[ \overline{x}' \Theta_n^w + \overline{x}(-y v' \Theta_n^w + u) = \overline{T} \]

with \( \Theta_n = \exp v \), \( v \in F_{n-1} \) since the appearance of coefficient degradation would imply that \( \Theta_n^{|u} \) contradicting our former hypothesis.

b. If \( w = 0 \), then \( \overline{u} = u' \), and since \( u' \) is not in \( F_{n-1} \), \( \text{deg } \overline{u} > 0 \). On the other hand, we might have that \( y > s \). This can happen if, and only if, the trailing coefficient of

\[ \overline{x}' - y v' \overline{x} + \overline{u} \overline{x} \]

is zero.

If we call \( x_o \) the trailing coefficient of \( \overline{x} \), and \( u_o \) the trailing coefficient of \( \overline{u} \), this translates into:
\[ x'_o - y' v' x'_o + u'_o x'_o = 0 \]

which implies

\[ u'_o = y' v' - \frac{x'_o}{x'_o} \]

thus

\[ f u'_o = y' v - \log x'_o \]

We find \( y \) in the same way we found \( r \) in (a) and multiplying through by \( \Theta^n_y \), we obtain:

\[ \overline{x}' + \overline{x}(u' - y' v') = \overline{T} \Theta^n_{y-s} \]

which can be solved using the techniques of case vii).

\[ \text{c. If } w < 0, \text{ we must have } y = s, \text{ and multiplying through by } \Theta^n_y, \]

we obtain

\[ \overline{x}' + \overline{x}(u \Theta^n_{-w} - y' v') = \overline{T} \]

and this equation can be solved as explained in vii) below.

\[ \text{vii) If we arrive at an equation of the form} \]

\[ x' + \overline{u} X = T \]

with \( \overline{u}, T \in S_n \), \( \deg \overline{u} > 0 \), and \( X \) is known to be in \( S_n \), either because \( \Theta^n = \log v, \ v \in \mathbb{F}_{n-1} \), or because vi) was already applied, or when algorithm SPDE was applied, we ran into coefficient degradation and \( \deg \overline{u} > 0 \), we solve this equation as follows.
Let
\[ \bar{u} = \sum_{i=0}^{p} u_i \Theta_i, \quad T = \sum_{i=0}^{q} t_i \Theta_i. \]

In order for this equation to have a solution, we must have that either
\( T = 0 \) (so that \( X = 0 \)) or \( p \leq q \), and then, \( \deg X = q - p = r. \)

Let then
\[ X = \sum_{i=0}^{r} x_i \Theta_i, \]
then \( x_r = \frac{t_r}{u_p} \) (since \( p > 0 \)), and we have to solve the equation
\[ X_1' + \bar{u} X_1 = T - \bar{u} x_r \Theta_r - (x_r \Theta_r)' \]
which we do by repeating this process.

viii) If none of the previous cases applies we then have that at least one of the \( u' \), \( T \) is in \( F_n - D_n \). We must then have \( X \in F_n - D_n \), i.e. \( X = \overline{X} \) with \( \hat{X} \) in \( S_n - F_{n-1} \). We thus have to find both of \( \overline{X} \) (in \( D_n \)) and \( \hat{X} \) (in \( S_n \)).

Before proceeding further, let us define a gcd function for \( F_n \) (and thus for \( D_n \)). If \( a, b \in S_n \), there exist \( p, q \) such that \( \Theta_p \in S_n \), \( \gcd(\Theta_p a, \Theta_q) = 1 \) and similarly for \( \Theta_q b \). Then
\[ \gcd(a, b) = \gcd(\Theta_p a, \Theta_q b). \]

In order to find \( \hat{X} \), let us have
\[ u' = \frac{u_1}{u_2} \quad \text{and} \quad T = \frac{T_1}{T_2}. \]
with \( u_1, T_1 \) in \( D_n \), and \( u_2, T_2 \) in \( S_n \) such that if \( \Theta_n = \exp \nu, \nu \in F_{n-1} \)
then \( \Theta_n \nmid u_1, \Theta_n \nmid T_1 \) (in this case \( D_n = E_n \)). We also require that
\( \gcd(u_1, u_2) = \gcd(T_1, T_2) = 1 \) and that \( u_2 \) and \( T_2 \) are monic.

Let us further separate

\[
\hat{u}_1 = \bar{u} \hat{u} \quad \text{and} \quad \hat{T}_1 = \bar{T} \hat{T}
\]

in such a way that \( \hat{u}, \hat{T} \in S_n \) are monic, \( \gcd(\bar{u}, \hat{T}) = 1 \), \( \bar{u}, \bar{T} \in D_n \), and every square-free factor of \( \hat{u} \) (respectively \( \hat{T} \)) divides \( T_2 \) (respectively \( u_2 \)). Notice that these conditions imply that
\( \gcd(\bar{u}, \hat{u}) = \gcd(\bar{T}, \hat{T}) = 1 \).

Now, let \( p_1, \ldots, p_k \) be a square-free basis for \( \hat{u}, \hat{T}, u_2, T_2 \).
Assume each \( p_i \) is monic. We then have

\[
\hat{u}' = \frac{\bar{u}}{k} \prod_{i=1}^{\pi p_i} b_i, \quad \hat{T} = \frac{\bar{T}}{k} \prod_{i=1}^{\pi p_i} c_i
\]

where at least one of \( b_i, c_i \neq 0 \). Notice also that for each \( i \),
\( \gcd(p_i, \bar{p}_i') = 1 \).

In [RIS 69] a result is proven which implies that

\[
\bar{X} = \frac{\bar{X}}{k} \prod_{i=1}^{\pi p_i} x_i, \quad \bar{X} \in D_n
\]

where no \( p_i \) divides \( \bar{X} \) (even though we may have \( \gcd(\bar{X}, p_i) \neq 1 \)).

Before continuing, let me show that, given \( p_1, \ldots, p_k \), this
representation of \( X \) is unique. To do this, assume
\[
\frac{k}{\prod_{i=1}^{\pi} p_i} x_i = \frac{k}{\prod_{i=1}^{\pi} p_i} x_i,
\]

and assume that no \( p_i \) divides neither of \( x \) nor \( Y \). Assume also (without loss of generality) that \( x_1 > y_1 \). But then

\[
\frac{k}{\prod_{i=1}^{\pi} p_i} \bar{x} = \frac{k}{\prod_{i=1}^{\pi} p_i} \bar{y},
\]

so that

\[
\frac{k}{\prod_{i=1}^{\pi} p_i} \bar{x} = \frac{k}{\prod_{i=1}^{\pi} p_i} \bar{y}.
\]

But this says that \( p_1 \) divides \( \frac{k}{\prod_{i=1}^{\pi} p_i} \bar{x} \) and since \( \gcd(p_i, p_j) = 1 \) if \( i \neq j \), we must have \( p_1 | \bar{x} \), contradicting our hypothesis.

Now let us substitute these values (of \( u', T \) and \( X \)) into our differential equation, to obtain:

\[
\frac{\bar{X'}}{\prod_{i=1}^{\pi} p_i} - \bar{X} = \sum_{i=1}^{\pi} p_i \frac{(x_i p_i' \prod_{j=1}^{\pi} p_j)}{\pi} + \frac{\bar{u} \bar{X}}{\prod_{i=1}^{\pi} p_i} = \frac{\bar{T}}{\prod_{i=1}^{\pi} p_i}
\]

We are now interested in finding bounds for the \( x_i \). Considering each \( i_0 \) separately, we obtain two cases:

a. If \( b_{i_0} \neq 1 \), then

\[
\max(x_{i_0} + 1, x_{i_0} + b_{i_0}) = c_{i_0}, \text{ so, } x_{i_0} = c_{i_0} - \max(b_{i_0}, 1).
\]
b. If $b_i = 1$, we may have $x_{i_0} > c_{i_0} - 1$ (because the numerator on the left hand side of our equation can be divisible by $p_{i_0}$). If we call $a_i = \max(b_i, 1)$ we then obtain in our equation:

$$\frac{\prod_{i=1}^{k} a_i}{\prod_{i=1}^{\bar{X}} p_i} + \frac{\bar{X}}{\prod_{i=1}^{k} p_i} - \sum_{i=1}^{k} x_{i_0} p_i^{a_i-1} p_j^{a_j}$$

$$\frac{\prod_{i=1}^{k} x_{i_0} + a_i}{\prod_{i=1}^{\bar{X}} p_i}$$

$$= \frac{\bar{X}}{\prod_{i=1}^{k} c_i}.$$

If $x_{i_0} > c_{i_0} - 1$, we must have that $p_{i_0}$ divides the numerator of this expression, that is

$$p_i | \prod_{i=1}^{k} a_i - b_i - \sum_{i=1}^{k} x_{i_0} p_i^{a_i-1} p_j^{a_j}.$$}

Since $p_{i_0}$ does not divide $\bar{X}$, we must have

$$\gcd(p_{i_0}, \prod_{i=1}^{k} a_i - b_i - \sum_{i=1}^{k} x_{i_0} p_i^{a_i-1} p_j^{a_j})$$

$$= \gcd(p_{i_0}, \prod_{i=1}^{\bar{X}} a_i - b_i - \sum_{i=1}^{\bar{X}} x_{i_0} p_i^{a_i-1} p_j^{a_j})$$

$$\neq 1.$$
(According to the definition of \( \gcd \) for \( E_n \), if \( D_n = S_n \), we simply take \( \alpha = 0 \).) But this \( \gcd \) is the same as

\[
\gcd(p_i, u \Theta_n \prod_{i=1}^k a_i - b_i, p_i) - x_i \rem(\Theta_n \prod_{i=1}^k a_i, p_i) = \gcd(p_i, \rem(u \Theta_n \prod_{i=1}^k a_i - b_i, p_i), p_i)
\]

where \( \rem(\xi, p_i) \) is the remainder obtained when dividing \( \xi \) by \( p_i \).

Thus, to obtain \( x_i \), we only need to look for integers \( x_i \) such that

\[
\text{resultant}(p_i, \rem(u \Theta_n \prod_{i=1}^k a_i - b_i, p_i), p_i) - x_i \rem(\Theta_n \prod_{i=1}^k a_i, p_i) = 0
\]

If there is such an integer value of \( x_i > c_i - 1 \), we use that, otherwise, we use \( x_i = c_i - 1 \).

Although the biggest such solution for \( x_i \) is enough for our purposes, some further analysis will actually provide us with a smaller problem in most cases.
Notice that if $b_i \neq 1$, the analysis under (a) above proves that $\bar{x}$ and $p_i$ have no common factors; equivalently, if $\bar{x}$ and $p_i$ have a common factor, then $b_i = 1$.

Let us now decompose such $p_i$ into

$$p_i = \prod_{j=1}^{m} q_{ij}$$

in such a way that

$$\hat{x} = \frac{\bar{x}}{\prod_{i=1}^{\pi} p_i} = \frac{\bar{x}}{\prod_{i=1}^{\pi} p_i} \cdot \frac{x_i}{x_i} \cdot \frac{x_i}{x_i} = \frac{\bar{x}}{\prod_{i=1}^{\pi} p_i} \cdot \frac{x_i}{x_i}$$

where the $q_i$, $q_{ij} \in S_n - F_{n-1}$, $\hat{x}$ and the $p_i$, $q_{ij}$ have no common factors, for $j \neq \ell$, $x_{ij} \neq x_{i\ell}$, and $x_i = \max_{1 \leq j \leq k_i}(x_{ij})$.

If we redo our previous analysis, but with this new expression, we obtain that, if $x_{i_0 j_0} > c_{i_0} - 1$, then

$$q_{i_0 j_0} \mid (u \prod_{i=1}^{k_i} p_i - x_{i_0 j_0}) = q_{i_0 j_0}^{\pi} q_{i_0 j_0}^{j=1} q_{i_0 j_0}^{j\neq j_0} q_{i_0 j_0}^{i\neq i_0}$$

since $q_{i_0 j_0}$ and $\hat{x}$ are relatively prime) and since

$$q_{i_0 j_0} \mid \sum_{j=1}^{\pi} q'_{i_0 j} - q_{i_0 i} = p_{i_0}^{j=1} q_{i_0 j}^{i\neq j} q_{i_0 j}^{j\neq j_0}$$
we have that

\[ q_{i_0 j_0} \mid \left( \prod_{i=1}^{k} (x_{i_0} - p_{i_0}^{a_i-b_i}) \right) \prod_{i=1}^{k} p_{i_0}^{a_i} \]

Since, for \( j \neq i \), \( x_{i_j} \neq x_{i_k} \), and \( p_{i_j} \) is square-free, we obtain that

\[ q_{i_0 j_0} = \gcd(p_{i_0}, \prod_{i=1}^{k} (x_{i_0} - p_{i_0}^{a_i-b_i}) \prod_{i=1}^{k} p_{i_0}^{a_i}) \]

\[ = \gcd(p_{i_0}, \prod_{i=1}^{k} (x_{i_0} - p_{i_0}^{a_i-b_i}) \prod_{i=1}^{k} p_{i_0}^{a_i}) \]

This means we have already computed all the \( x_{i_j} \) (as roots of the resultant above) and can easily find the \( q_{ij} \).

This step will decrease the degree of the polynomial we will be seeking, and also factor out some of the \( q_{ij} \).

Still, in order to simplify the verification, of other steps of this algorithm, we will assume (for the moment) that this last simplification was not done.

Since in any case \( x_{i_j} + a_i \geq c_i \), this leaves us with the equation

\[ \prod_{i=1}^{k} (x_{i_j} + a_i - c_i) = \prod_{i=1}^{k} \frac{a_i}{p_i} \prod_{i=1}^{k} p_i \]

\[ \prod_{i=1}^{k} \frac{a_i}{p_i} + \prod_{i=1}^{k} \frac{a_i-b_i}{p_i} \]

\[ = \prod_{i=1}^{k} \frac{x_{i_j} + a_i - c_i}{p_i} \]
which is of the form

$$\bar{X}' A + \bar{X} B = C$$

with $A$, $B$, $C$, $\bar{X} \in D_n$.

If $\Theta_n = \exp v$, $v \in F_{n-1}$ we have one more step to do, otherwise we can directly seek the bound we need for algorithm SPDE.

The extra step we have to do now is similar to what was done in case vi) and consists of the following. We know $A = \pi \sum_{i=1}^{k} p_i \in S_n$, and $\Theta_n \uparrow A$. Now let

$$B = \frac{\hat{B}}{\Theta_n b}, \quad C = \frac{\hat{C}}{\Theta_n c}, \quad \bar{X} = \frac{\hat{X}}{\Theta_n x}$$

with $\Theta_n$ dividing neither of $\hat{B}$, $\hat{C}$ or $\hat{X}$. Substituting, we obtain:

$$\frac{\hat{X}' - x v' \hat{X} A}{\Theta_n x} + \frac{\hat{B} \hat{X}}{\Theta_n x + b} = \frac{\hat{C}}{\Theta_n c}$$

If $b \neq 0$, then $x = c - \max(b, 0)$ otherwise, we can have $x > c$ which can happen if and only if

$$x_o a_o - x v' x_o a_o + b_o x_o = 0$$

where $x_o$, $a_o$, and $b_o$ are the trailing coefficients of $\hat{X}$, $\hat{A}$, and $\hat{B}$ respectively. If we divide this equation by $x_o a_o$ and integrate, we obtain:

$$\int \frac{b_o}{a_o} = x v - \log x_o + k$$
and using the techniques developed in case vi), we either obtain a satisfactory bound on \( x \), or use \( x = c \). We can now clear denominators to obtain a similar equation, which is:

\[
\hat{X}' A \Theta_n \overline{b} + \hat{X}(B \Theta_n \overline{b} - x \nu' A \Theta_n \overline{b}) = \hat{C} \Theta_n x^+ \overline{b} - c
\]

where \( \overline{b} = \max(b, 0) \).

Before closing this section we need two things: we need a bound on the degree of \( \hat{X} \), and we need to show that whenever coefficient degradation turns our equation into

\[
Y' + \beta Y = D
\]

with \( \beta, D \in S_n \), we have deg \( \beta > 0 \), so that case vii) applies, or \( \beta = w' \in F_{n-1} \), where \( \exp(w) \) is a regular monomial over \( F_{n-1}(w) \).

To find our bound, let our differential equation be

\[
A X' + B X = C
\]

with

\[
A = \sum_{i=0}^{a} a_i \Theta_n^i, \quad B = \sum_{i=0}^{b} b_i \Theta_n^i
\]

\( c = \deg C \) and \( X = \sum_{i=0}^{x} x_i \Theta_n^i \)

Notice that \( A \) is monic, because it is a product of monic polynomials.

We must place a bound on \( x \) under three possible cases:

a. If \( \Theta_n = \exp \nu, \nu \in F_{n-1} \) then deg \( X' = \deg X = x \), so if \( a \neq b \) then \( x + \max(a, b) = c \), so \( x = \min(c - a, c - b) \). If \( a = b \), then
\[ x > c - a \text{ if and only if} \]

\[ (x'_x + x x_x v') + b_b x_x = 0, \]

which implies

\[ \int b_b = -x v - \log x_x \]

and we handle this equation (and our result) in the same manner as previously.

b. If \( \Theta_n = z(n = 0) \) then \( \deg X' - \deg X = 1 \), so that if \( a - 1 \neq b \), we must have

\[ \max(a + x - 1, b + x) = c, \]

and thus,

\[ x = c - \max(a - 1, b). \]

If \( a - 1 = b \), then \( x > c - b \) if

\[ x x_x + b_b x_x = 0, \]

so that

\[ x = -b_b \]

and therefore \( x = \max(c - b, -b_b \text{ if this is an integer}) \).

c. If \( \Theta_n = \log v, v \in F_{n-1} \) then \( x - 1 \leq \deg X' \leq x = \deg X \) so that if

\[ a \neq b, \quad a \neq b + 1, \]
then if $x_\ell$ is a constant, then

$$\max(a + x - 1, b + x) = c,$$

otherwise

$$\max(a + x, b + x) = c,$$

so that

$$x \leq \max(c - \max(a - 1, b), c - \max(a, b) = c - \max(a - 1, b)).$$

If $a = b$, we have two choices.

i) $x \leq c - \max(a - 1, b) + 1 = c - b + 1$. This bound of $x$ should be used if the analysis below fails or yields a smaller bound.

ii) $x > c - b + 1$. This implies that the two leading coefficients of $A X' + B X$ are 0, so that (since $a = \deg A = b = \deg B$) $\deg X' = \deg X$, which, in turn, implies that $x_\ell$ (the leading coefficient of $X$) is not a constant.

Now, looking at the two leading coefficients of $A X' + B X$, we notice that

$$x'_x + b_- b \cdot x_x = 0,$$

(so that $x'_x = -b_- b \cdot x_x$) and

$$x \cdot x_v v' + x'_x x_{x-1} + x'_x a_{a-1} + b_- b \cdot x_{x-1} + b_- b \cdot x_x = 0.$$

If we let (in this last equation) $x_{x-1} = w_x x_x$, replace $x'_x = -b_- b \cdot x_x$, and divide by $x_x$, we obtain
\[ x \frac{v'}{v} + w' - b_b a_{a-1} + b_{b-1} = 0 \]

so that

\[ f(b_b a_{a-1} - b_{b-1}) = w + x \log v . \]

We can then find \( x \) by invoking the integration algorithm with \( b_b a_{a-1} - b_{b-1} \) as an argument, obtaining

\[ I = f(b_b a_{b-1} - b_{b-1}) = j + \sum_{i=1}^{k} d_i \log s_i = w + x \log v \]

where the \( d_i \) are constants in \( F_{n-1} \), (since no algebraic expansion of the constant field should be necessary, so none should be done; thus if the constant field was extended, our computation here failed and \( x \leq c - b + 1 \)) and the \( s_i \), \( j \) are in \( F_{n-1} \).

We then test each \( s_i \) to see whether it is regular monomial over \( F_n(\log s_1, \ldots, \log s_{i-1}) \) by using the structure theorem (see theorem 3.12 and corollary 3.13) and perform the necessary substitutions. We should then be able to obtain \( w \) and \( x \) from the resulting expression which may not involve any of the \( \log s_i \), since otherwise \( I \) is not of the desired form.

We then use this value of \( x \) if it is an integer greater than \( c - b + 1 \), otherwise we set \( x = c - b + 1 \).

Finally, if \( a = b + 1 \), we again have two possibilities:

i) If \( x \leq c - \max(a - 1, b) = c - b \), we have our bound on \( x \), otherwise,

ii) \( x < c - b \), so that \( \deg X' = \deg X - 1 \), \( x \) is a constant, and
\[ x x \frac{v'}{v} + x'_{x-1} + b_b \cdot x_x = 0 \]

determining \( x \) in this manner uniquely in the same way as before.

One question remains: What happens if we encounter coefficient degradation? We want to prove that, if we run into coefficient degradation, so that we have the equation

\[ \hat{Y} + \hat{u} \cdot Y = \hat{T} \]

with \( \hat{u} \in F_{n-1} \), \( \hat{T} \in S_n \), then \( \exp(\hat{u}) \) is a regular monomial over \( F_{n-1}(\hat{u}) \).

We actually prove that, if this happens, then

\[ u' = \sum_{i=1}^{k} y_{\hat{q}_i} \cdot q'_i + w \]

where \( w \in F_{n-1} \), the \( y_{\hat{q}_i} \) are integers, the \( q'_i \) are monic elements of \( S_n \), and either \( \hat{u} = w \), or \( \hat{u} = w - x \cdot v' \) and \( \Theta_n = \exp v \), \( x \) is an integer, and our statement then, follows.

So, let us assume we have coefficient degradation. We then have two cases to consider:

If \( \Theta_n = \exp v \), \( v \in F_{n-1} \), we are trying to solve the differential equation

\[ A \cdot x' + B \cdot x = C \]

with
\[
A = \theta_n \prod_{i=1}^{k} p_i^{a_i}
\]

\[
B = \bar{u} \prod_{i=1}^{b-b} p_i^{a_i-b_i} - \bar{b} \prod_{i=1}^{k} (x_i p_i^j p_j^{a_j})
\]

\[
- x v' \prod_{i=1}^{b} p_i^{a_i}
\]

\[
C = \bar{c} \prod_{i=1}^{x+b-c} p_i^{x_i+a_i-c_i}
\]

where \(a_i = \max(b_i, 1), \bar{b} = \max(b, 0)\)

\[
u' = \frac{u}{\theta_n \prod_{i=1}^{k} p_i^{b_i}}
\]

\[
T = \frac{\bar{T}}{\theta_n \prod_{i=1}^{c_i}}
\]

\[
\hat{X} = \frac{X}{\theta_n \prod_{i=1}^{k} x_i}
\]

and our original equation was

\[
\hat{X}' + u' \hat{X} = T
\]

If coefficient degradation is a problem (that is, if SPDE reduces this equation to something like

\[
Y' + \hat{u} Y = \hat{T}
\]

where \(u \in \mathbb{F}_{n-1}\) we must have that: \(\bar{b} - b = \bar{b}\) or \(b = 0\). This is only
one condition, since $\overline{b} - b = \overline{b} i f, and only if b = \overline{b} = 0 since
$\overline{b} = \max(b, 0)$.

We also have that either $a_i - b_i = a_i - 1$ or $b_i = 0$, that is,
either $b_i = 0$ or $b_i = 1$.

Let us then assume (lifting the restriction that $q_i \neq 1, p_i \neq 1$) that

$$u' = \frac{u}{\overline{\pi} q_i} \quad \text{and} \quad T = \frac{T}{\overline{\pi} \,(p_i q_i)^i}$$

where either $p_i \neq 1$, or $q_i \neq 1$, and, in any case, each $p_i$, $q_i$ is monic.

We can then also assume that (as done before)

$$q_i = \prod_{j=1}^{i-1} q_{ij}$$

in such a way that

$$\hat{X} = \frac{X}{\overline{\pi} \prod_{i=1}^{n} p_i \prod_{j=1}^{i-1} q_{ij}}$$

where $X$ and the $p_i$, $q_{ij}$ have no common factors, and for $j \neq m, x_{ij} \neq x_{im}$
for all $i$. (As before $X, p_i, q_{ij}$ are in $S_n$.) Notice that $x_{ij} > i - 1$.
We can assume $x_{i1} = i - 1$, so that if $j > 1, x_{ij} > i - 1$.

We can then obtain

$$A = \prod_{i=1}^{\ell} p_i \prod_{j=1}^{\ell} q_{ij}$$
\[
B = \sum_{i=1}^{\ell} p_i \pi_{i} - x v' \sum_{i=1}^{\ell} p_i \pi_{i} q_{ij} \\
- \sum_{i=1}^{\ell} (i - 1) p_i q_i \pi_{i} p_j q_j \\
- \sum_{i=1}^{\ell} \left( \sum_{j=1}^{\ell} x_{ij} q_{ij} \pi_{m} q_{im} \right) \pi_{m} q_{m} \pi_{m} p_{m} \\
\]

When we trace through the way SPDE operates on these inputs, we notice that

(a) \( p_i \pi_{i} q_{ij} \) is factored out the first time through.

(b) For \( i > 1 \), \( p_i \) divides the factor cancelled on the \( i \)-th recursive call.

(c) If coefficient degradation is to occur we must have that for each \( i \), there exists a \( k_i \) such that \( q_{ii} \) divides the factor cancelled on the \( k_i \)-th recursive call. But this last fact implies that \( q_{ii} \) divides

\[
B + k_i A' = B + k_i \left[ \sum_{i=1}^{\ell} p_i q_i \pi_{i} p_i q_i \right] \\
+ \sum_{i=1}^{\ell} p_i \sum_{i=1}^{\ell} q_{ij} \pi_{i} q_{im} \pi_{m} \]

since the extra summands are all divisible by \( q_{ii} \) and the extra factors are relatively prime to \( q_{ii} \).
Thus, we have found integers \( y_{rs} \) such that for all \( r, 1 \leq r \leq \ell \), all \( s, 1 \leq s \leq \ell \) we have that

\[
q_{rs} \text{ divides } S = \prod_{i=1}^{\ell} \prod_{j=1}^{\ell} p_i - x' \prod_{i=1}^{\ell} \prod_{j=1}^{\ell} q_{ij} - \sum_{i=1}^{\ell} t_i \prod_{j=1}^{\ell} p_j q_j
\]

where the \( t_i \) are any integers.

Since the \( q_{rs} \) are relatively prime, this implies that the product of all the \( q_{rs} \) divides \( S \), and this in turn implies that

\[
\prod_{i=1}^{\ell} \prod_{j=1}^{\ell} q_{ij} | u \prod_{i=1}^{\ell} p_i - \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} y_{ij} q_{ij} q_{im} \prod_{m=1}^{\ell} q_{im} \prod_{m\neq j}^{\ell} p_{m} \prod_{m\neq i}^{\ell} p_{i}
\]

or, equivalently (since the \( p_i, q_{ij} \) are relatively prime)

\[
\prod_{i=1}^{\ell} \prod_{j=1}^{\ell} q_{ij} | u - \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} y_{ij} q_{ij} q_{im} \prod_{m=1}^{\ell} q_{im} \prod_{m\neq j}^{\ell} p_{m} \prod_{m\neq i}^{\ell} p_{i}
\]

Thus, for some \( \tilde{u} \) in \( S_n \), we have

\[
\tilde{u} \prod_{i=1}^{\ell} \prod_{j=1}^{\ell} q_{ij} = \tilde{u} - \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} y_{ij} q_{ij} q_{im} \prod_{m=1}^{\ell} q_{im} \prod_{m\neq j}^{\ell} p_{m} \prod_{m\neq i}^{\ell} p_{i}
\]
and, if we add that sum on the right to both sides and divide them by

\[ \pi^{l} \] 
\[ \pi^{i} \] 
\[ q_{ij} = \pi_{i} q_{i}, \]

we obtain

\[ \bar{u} = u' = u + \sum_{i=1}^{l} \sum_{j=1}^{l} q_{ij} q_{ij}. \]

\[ \pi^{i} \] 
\[ q_{i} \]

On the other hand, since this proof also implies that algorithm SPDE will arrive at the equation

\[ X'_{p} + \hat{u} X_{p} = \hat{T}_{p}, \]

where \( \hat{u} = \tilde{u} - x \nu' \) for some \( \hat{T}_{p} \) in \( S_{n} \), and we are assuming that coefficient degradation occurs and that case (vii) is not applicable, we obtain that \( \hat{u} \) is in \( F_{n-1} \), and, since \( u \) is elementary over \( F_{n} \), \( \tilde{u} = w \), must also be elementary over \( F_{n-1} \).

Furthermore, since \( \exp(u) \) is a regular monomial over \( F_{n} \), we obtain that \( \exp(w) \) is a regular monomial over \( F_{n-1}(w) \).

This concludes our proof for the exponential case. The non-exponential case, i.e., when either \( n = 0 \) or \( \theta_{n} \) is logarithmic over \( F_{n-1} \) can be proved in exactly the same way, taking into account that \( x \) does not appear \( (x = 0) \), since \( \theta_{n} \), if at all, divides one of the \( p_{i} \).

Thus ends the longest section in this chapter. Although the basic idea is simple to explain, some of the proofs have turned out to be rather long, and the large number of special cases in both the algorithms and the proofs caused this blow-up.
6. An Algorithm to Integrate Elements of $\mathbb{S}_n$ (Logarithmic Case).

In this section, we discuss the integration of elements of $\mathbb{S}_n$ for $\Theta_n = \log u$, $u \in \mathbb{F}_{n-1}$. The auxiliary equations are found by Risch's method [RIS 69] but their solution method is new. Let $P(z) = \sum_{i=0}^{m} a_i \Theta_n^i$, with $\Theta_n = \log u$, $u \in \mathbb{F}_{n-1}$. Then

$$\int \frac{P(z)}{z} = \sum_{i=0}^{m+1} b_i \Theta_n^i + \sum_{i=0}^{m} d_i \log v_i = \int \sum_{i=0}^{m} a_i \Theta_n^i$$

where $b_i \in \mathbb{F}_{n-1}$, $v_i \in \mathbb{K}_{n-1}$, $d_i \in \mathbb{K}$ (algebraic closure of $\mathbb{K}$).

If we differentiate both sides and equate like powers of $\Theta_n$, we obtain:

i) $b_{m+1}$ is a constant.

ii) $a_m = (m+1) b_{m+1} \Theta_n^m + b_m = (m+1) b_m \frac{u'}{u} + b_m$

which leads us to determining whether the integral

$$\int a_m = c_{m+1} \log u + b_m$$

where $c_{m+1} = (m+1) b_{m+1}$ is elementary, and of this same form.

Notice that $b_m$ is only solved modulo a constant, so call this solution $\overline{b}_m$. Then $b_m = \overline{b}_m + k_m$ for some constant $k_m$, which we will find below:

iii) If $1 \leq i \leq m$, we obtain (as in (ii)):

$$a_i = (i+1) b_{i+1} \frac{u'}{u} + b_i = (i+1) (\overline{b}_{i+1} + k_{i+1}) \frac{u'}{u} + b_i$$

or

$$\int (a_i - (i+1) \overline{b}_{i+1} \frac{u'}{u}) = b_i + c_{i+1} \log u$$
where \( c_{i+1} = (i + 1) k_{i+1} \) which is an integral similar to the one we had in (ii). As before, we find \( c_{i+1} \) and \( b_i = b_i - k_i \), where \( c_{i+1} \), \( k_i \) are constants.

iv) Finally, we reach:

\[
a_o = b_1 \frac{u'}{u} + b_o + \sum d_i \frac{v_i'}{v_i}
\]

and thus

\[
a_o - \bar{b}_1 \frac{u'}{u} = b_o + k_1 \frac{u'}{u} + \sum d_i \frac{v_i'}{v_i}
\]

so we integrate (recursively)

\[
a_o - \bar{b}_1 \frac{u'}{u}
\]

Since we occasionally required finding \( x \) in \( F_{n-1} \) and \( c \) in \( K \) such that

\[
\int u = x + c \log v
\]

where we knew \( u \) and \( v \), in \( F_{n-1} \); let me add that there are two ways to find \( x \) and \( c \): One is entering the integration algorithm recursively with \( u \) as an argument, simplifying the resulting logarithmic terms as explained before. This involves some algebraic independence determinations.

The second method implies looking at this integral as the differential equation

\[
x' = u - c \frac{v'}{v}
\]
which can be solved by going through the integration algorithm with \( u \) and \( \frac{v'}{v} \) together but independently (i.e. in parallel). This approach leads to solving the differential equation

\[
x' = u - \sum_{i=1}^{\xi} \frac{c_i}{v_i} v_i'
\]

with \( x, u, v_i (1 \leq i \leq \xi) \) all in \( F_n \), the \( c_i \) in \( K \) and unknowns \( x \) and the \( c_i \), with \( c_i \) being given uniquely (if at all) by this equation.

By proceeding further with the integration algorithm, we can assume that \( u \) is a regular element of \( F_m \), and all the \( v_i \) are in \( S_m \), monic, non-constant. (We might have to add some \( c_i \) and equations on the \( c_i \) to achieve this.)

If we let

\[
x = \frac{\bar{X}}{k \prod_{i=1}^{\pi} p_i}
\]

and replace in our equation above, we obtain

\[
\bar{X} \prod_{i=1}^{k} p_i - \bar{X} \sum_{i=1}^{k} \prod_{j=1}^{\pi} p_j = u - \sum_{i=1}^{\xi} \frac{c_i}{v_i} v_i'.
\]

Since we know \( u \) and the \( v_i \) \((1 \leq i \leq \xi)\), we can find \( \bar{u}, \bar{v}_i \) \((1 \leq i \leq \xi)\), \( k > 0, q_1, \ldots, q_{k+1} \) in \( S_m \) such that
\[ u - \sum_{i=1}^{\ell} c_i v_i = \frac{\sum_{i=1}^{\ell} c_i \overline{v_i}}{k+1} \prod_{j=1}^{\pi} q_j \]

Comparison of the last two equations yields that:

a. \( q_1 \) divides \( \overline{u} - \sum_{i=1}^{\ell} c_i \overline{v_i} \),

so that

\[ \text{rem}(\overline{u}, q_1) = \sum_{i=1}^{\ell} c_i \text{rem}(\overline{v_i}, q_1) \]

and

b. \( p_i = q_{i+1} \) for \( 1 \leq i \leq k \), and so, we can clear denominators to obtain the equation

\[ \overline{X'} - \overline{X} = \sum_{i=1}^{k} i \overline{p_i} - \sum_{i=1}^{k} i \overline{p_i} \prod_{j=1}^{\pi} p_j \]

\[ = \text{quot}(\overline{u}, q_1) - \sum_{i=1}^{\ell} c_i \text{quot}(\overline{v_i}, q_1) \]

where

\[ \text{quot}(\overline{c}, q_1) = \frac{\text{rem}(\overline{c}, q_1)}{q_1} \]
This last equation can be solved by a version of SPDE which handles \( \text{quot}(\bar{u}, q_1) \) and the \( \text{quot}(v_i, q_1) \) in parallel, and sets up new linear equations instead of giving up, in a similar way to what we did above. Notice that coefficient degradation is no problem here because of the tight degree bounds we have.

Solving the above differential equation can also lead to seeking solutions to the equation

\[
X' + u' X = T + \sum_{i=1}^{\ell} c_i T_i
\]

for \( X \) and the \( c_i \). Again similar changes to the procedures of section 5 will solve this equation (notice that SPDE has already been changed).

It is not clear which procedure is faster. Still, this second method might be preferable because the \( F_m - S_{m-1} \) case is handled very differently, resulting in less work for the second method, and also, the other algorithm still has the problem of simplifying the logarithmic terms it generated.

7. Integration of Rational Elements of \( F_n \), or Rational Integration Revisited.

There are three known algorithms for rational integration, two of which generalize easily to integration of rational elements of \( F_n \). These two are Hermite's and D. Mack's. In [YUN 76] it is proven that Hermite's algorithm is asymptotically faster, but this result only applies if asymptotically fast multiplication and division algorithms are used. It is pointed out, though, (in [MOE 76]) that these fast multiplication algorithms are impractical in the multivariate case if
the degrees are small, as is usually the case here, so that leaves us with Hermite's algorithm using classical arithmetic algorithms or Mack's algorithm.

In [HOR 71] it is shown that in the rational function case, an algorithm based on solving a system of linear equations is superior to Hermite's algorithm both asymptotically and in practice. Also, in [McD 75], it is shown that Mack's and Horowitz's algorithms have the same asymptotic computing time bounds. This implies that Mack's algorithm is asymptotically superior to Hermite's if classical arithmetic algorithms are used.

We will thus describe Mack's algorithm here for completeness.

The algorithm can be described as follows. Let $f$ be a proper element of $\mathbb{F}_n$, $p = \text{num } f$, and let $\prod_{i=1}^{l} \pi_i q_i^i$ be a square-free decomposition of $\text{den } f$. We intend to integrate $f$ by finding $D, R$ such that

$$\int \frac{p}{\prod_{i=1}^{l} \pi_i q_i^i} = \frac{r}{\prod_{i=1}^{l} \pi_i q_i^{i-1}} + \int \frac{d}{\prod_{i=1}^{l} \pi_i q_i}$$

where

$$\frac{r}{\prod_{i=1}^{l} \pi_i q_i^{i-1}} \text{ and } \frac{d}{\prod_{i=1}^{l} \pi_i q_i}$$

are proper in $\mathbb{F}_n$. If $\lambda = 0$, there is nothing to do, so assume $\lambda > 1$.

The method of finding $D, R$, consists of first finding $D_1, Q_1, R_1$ such that
\[
\frac{p}{\prod_{i=1}^{N} q_i^i} = \frac{D_1}{\prod_{i=1}^{N} q_i^i} + \frac{R_1}{\prod_{i=1}^{N} q_i^{i-1}} + \frac{Q_1}{\prod_{i=1}^{N} q_i^{i-1}}
\]

with \(\deg D_1 < \deg q_1\), \(\deg Q_1 < \deg \prod_{i=1}^{N} q_i^{i-1}\) and \(\deg R_1 < \deg \prod_{i=1}^{N} q_i^{i-1}\), and then repeating the process with the second integral on the right.

The algorithm finds \(D_1, Q_1, R_1\) as follows. First, find \(D_1\) (and a \(\overline{Q}\)) such that

\[
D_1 \prod_{i=2}^{N} q_i^i + \overline{Q} q_1 = P \quad \text{(remember that } \gcd(q_1, \prod_{i=2}^{N} q_i^i) = 1)\]

with \(\deg D_1 < \deg q_1\) so that \(\deg \overline{Q} < \deg (\prod_{i=2}^{N} q_i^i)\). This leaves us with finding \(R_1\) and \(Q_1\), and we point out that

\[
(a) \quad \frac{1}{\prod_{i=2}^{N} q_i^{i-1}} = \prod_{i=2}^{N} q_i^i \prod_{j=2}^{\ell} q_j^{i-1} \quad \text{for } j \neq i
\]

\[
(b) \quad \gcd(\prod_{i=2}^{N} q_i^i, \prod_{i=2}^{\ell} (i-1) q_i^i \prod_{j=2}^{\ell} q_j^{i-1}) = 1.
\]

Thus we find \(\overline{R_1}, \overline{Q_1}\) such that

\[
\overline{R_1}[\prod_{i=2}^{N} (i-1) q_i^i \prod_{j=2}^{\ell} q_j^i] + \overline{Q_1} \prod_{i=2}^{N} q_i^i = \overline{Q},
\]

with \(\deg \overline{R_1} < \deg \prod_{i=2}^{N} q_i^i\). If we divide this last equation by \(\prod_{i=2}^{N} q_i^i\),
integrate, and perform some simplifications, we obtain:

\[
\int \frac{Q_i}{q_i} = \int \left[ -R_i \left( \frac{1}{q_i} \right)' \right] + \frac{\bar{Q}_i}{q_i} \\
\int \frac{Q_i}{q_i} = \frac{-\bar{R}_i}{q_i} + \int \frac{\bar{Q}_i + \bar{R}_i'}{q_i} \\
\]

(by separation, integration by parts and joining the two integrals again) and this leaves us with

\[
R_1 = -\bar{R}_1, \quad Q_1 = \bar{Q}_1 + \bar{R}_1' \]

Notice that since \( \operatorname{deg} \bar{Q} < \max_{i=2}^{\ell} \left( \frac{\bar{Q}_i}{q_i} \right) \) we obtain that

\[
\operatorname{deg} \bar{Q}_1 \leq \max(\operatorname{deg} \bar{Q} - \max_{i=2}^{\ell} \left( \frac{\bar{Q}_i}{q_i} \right), \deg[((\frac{\bar{Q}_i}{q_i})')] - 1) \\
< \max_{i=2}^{\ell} \left( \frac{\bar{Q}_i}{q_i} \right), \deg[((\frac{\bar{Q}_i}{q_i})')] \\
= \deg_{i=2}^{\ell} \frac{\bar{Q}_i}{q_i} \]

Since

\[
\operatorname{deg} \bar{R}_1 < \max_{i=2}^{\ell} \frac{\bar{Q}_i}{q_i} \]

this says that
is also proper in \( F_n \).

8. Integration of Normal Elements of \( F_n \).

To complete our algorithm for integrating proper elements of \( F_n \), we need to integrate normal elements of \( F_n \).

Before we embark on our algorithm for this problem let us first do some theory. Assume that \( f \) is a non-zero normal element of \( F_n \), with \( P = \text{num} f \), \( Q = \text{den} f \). Let us also handle first the simpler case, i.e. assume that if any \( q \) in \( S_n \) is monic, then \( \deg q' < \deg q \). This implies that \( \Theta_n = \log v \), \( v \in F_{n-1}^n \) or that \( n = 0 \), \( \Theta_n = \Theta_0 = z \).

Let us assume that \( K \) has been extended to \( \bar{K} \), the smallest degree extension necessary to express \( \int f = \sum_{i=1}^{k} c_i \log v_i + g \), where we may assume, without loss of generality, that the \( v_i \) are in \( S_n \) (augmented by \( \bar{K} \), but, since there is no danger of confusion, we will still follow the notation used in this chapter), that they are square-free, pairwise relatively prime, and that if \( i \neq j \), then \( c_i \neq c_j \). We can also assume that the \( v_i \) are monic, since \( \text{den} f = \prod_{i=1}^{k} v_i \) is also monic. This last equality can be proven by differentiation of the equation above and the assumption that \( g \in F_{n-1}^n \). A similar reasoning proves that \( g' = 0 \), so assume \( g = 0 \).

Notice that these conditions, though very restrictive, do not really restrict our generality. They do, though, uniquely determine the \( c_i, v_i \), as can be proven (after an integration step) using the Risch structure theorem, a generalization of which will be proved in
Chapter 3.

Thus we have

\[ \frac{p}{Q} = \sum_{i=1}^{k} c_i \frac{v_i'}{v_i} \]

with the $v_i$ suitably restricted. This expression implies that

\[ Q = \prod_{i=1}^{k} v_i. \]

Now, call

\[ u_i = \prod_{j=1}^{k} v_j = \frac{Q}{v_i}. \]

This means that

\[ Q' = \sum_{i=1}^{k} u_i \frac{v_i'}{v_i} \]

and that

\[ \frac{v_i'}{v_i} = \frac{v_i' u_i}{v_i u_i} = \frac{v_i' u_i}{Q} . \]

Thus

\[ \frac{p}{Q} = \sum_{i=1}^{k} c_i \frac{v_i'}{v_i} = \sum_{i=1}^{k} c_i \frac{v_i' u_i}{Q} = \sum_{i=1}^{k} c_i v_i' u_i \]

\[ = \frac{k}{Q}. \]
This implies that

\[ P = \sum_{i=1}^{k} c_i v_i' u_i \]

so that, if \( 1 \leq i_0 \leq k \),

\[ \gcd(P - \sum_{i=1}^{k} c_i v_i' u_i, v_{i_0}) = v_{i_0} \]

But

\[ \gcd(P - \sum_{i=1}^{k} c_i v_i' u_i, v_{i_0}) = \gcd(P - c_{i_0} v_{i_0}' u_{i_0}, v_{i_0}) \]

\[ = \gcd(P - c_{i_0} \sum_{i=1}^{k} v_i' u_i, v_{i_0}) = \gcd(P - c_{i_0} Q', v_{i_0}) \]

Now, let \( 1 \leq j \leq k \), \( j \neq i_0 \). Then

\[ \gcd(P - c_{i_0} Q', v_j) = \gcd(\sum_{i=1}^{k} c_i v_i' u_i - c_{i_0} v_j' u_j, v_j) \]

\[ = \gcd(c_j v_j' u_j - c_{i_0} v_j' u_j, v_j) \]

\[ = \gcd((c_j - c_{i_0}) v_j' u_j, v_j) = 1 \]

and this leads to

\[ \gcd(P - c_{i_0} Q', Q) = v_{i_0} \]
We have thus proven the non-exponential cases of the following theorem.

Theorem 1

Let \( f \) be normal in \( \mathbb{F}_n \), \( f \neq 0 \), num \( f = P \), den \( f = Q \). Then, if \( f \) is elementary over \( \mathbb{F}_n \), all the roots for \( \alpha \) in resultant\((P - \alpha Q', Q)\) (with respect to \( \Theta_n \)) are constants (algebraic over \( \mathbb{F}_{n-1} \)) and then, for some \( g \) in \( \mathbb{F}_{n-1} \) (the smallest field containing \( \mathbb{F}_{n-1} \) and all constants algebraic over \( \mathbb{F}_{n-1} \)),

\[
f - g' = \sum_{i=1}^{k} c_i \frac{v_i'}{v_i}
\]

where \( c_1, \ldots, c_k \) are the roots of the resultant above, and the \( v_i \) are given by

\[
v_i = \gcd(P - c_i Q', Q).
\]

If \( \Theta_n = \exp w, w \in \mathbb{F}_{n-1} \), then

\[
g' = \sum_{i=1}^{k} c_i n_i w',
\]

when \( n_i = \deg v_i \) and the proof is very similar to the one for the other cases. The question now arises: Is this condition we found also sufficient? The answer is yes, as shown by Theorem 2.

Theorem 2

Let \( f \) be normal in \( \mathbb{F}_n \), \( f \neq 0 \), P = num \( f \), Q = den \( f \). Then, if all the roots for \( \alpha \) in the algebraic closure of \( \mathbb{F}_n \) of resultant\((P - \alpha Q', Q)\) (with respect to \( \Theta_n \)) are constants, then \( f \) is elementary over \( \mathbb{F}_n \).
Proof:

Let $c_1, \ldots, c_k$ be the (distinct) roots for $a$ in resultant$(P - a Q', Q)$, and let $v_i = \gcd(P - c_i Q', Q)$. Notice that if $i \neq j$, then $\gcd(v_i, v_j) = 1$, because otherwise

$$\gcd(P - c_i Q', Q, P - c_j Q') = \gcd(Q', Q) \neq 1$$

Thus,

$$r = \prod_{i=1}^{k} v_i$$

divides $Q$, so that $rs = Q$ for some $s$ in $S_n$. Notice that $\gcd(r, s) = 1$ because $Q$ is square-free.

If $s \neq 1$, then $\deg s > 0$, so that resultant$(P - a Q', s)$ is a polynomial of positive degree in $a$, and therefore has a root $a_0$. This says that

$$1 \neq \gcd(P - a_0 Q', s) \mid \gcd(P - a_0 Q', Q)$$

so that resultant$(P - a_0 Q', Q) = 0$ and so $a_0$ is one of our former roots (say) $c_1$. This says that $\gcd(s, v_1) \neq 1$ and thus $\gcd(s, r) \neq 1$. This contradiction implies that $s = 1$ and $r = Q$.

Now, let

$$\overline{P} = \sum_{i=1}^{k} c_i v_i' \prod_{j=1}^{\pi} v_j$$

$$j \neq i$$

Since
\[ Q' = \sum_{i=1}^{k} \sum_{j=1}^{k} v_i^j \]

we have that, for \( 1 \leq i \leq k \),

\[ \bar{P} - c_{i_0} Q' = \sum_{i=1}^{k} (c_i - c_{i_0}) v_i \pi v_i = \sum_{j=1}^{k} (c_i - c_{i_0}) v_i^j \pi v_i^j \]

so that \( v_{i_0} | \bar{P} - c_{i_0} Q' \). Since also \( v_{i_0} | P - c_{i_0} Q' \) this says that \( v_{i_0} | \bar{P} - P \). But also, \( \gcd(v_i, v_j) = 1 \) if \( i \neq j \), so that

\[ \sum_{i=1}^{k} v_i = Q|\bar{P} - P . \]

Now we will separate cases.

In the non-exponential case, since all the \( v_i \) are monic, we have that \( \deg v_i' < \deg v_i \), so that \( \deg \bar{P} < \deg Q \). Since also \( \deg P < \deg Q \), this implies that \( P = \bar{P} \) and

\[ \frac{p}{Q} = \sum_{i=1}^{k} \frac{v_i'}{c_i} \]

so that \( \forall \) is elementary in this case.

In the exponential case, that is, when \( \Theta_n = \exp w, w \in F^{n-1} \), we obtain that \( \text{ldcf}(v_i') = n_i w' \), where \( n_i = \deg v_i \). Thus

\[ \bar{P} - (\sum_{i=1}^{k} c_i n_i w')Q = \sum_{j=1}^{k} (v_i' - n_i w' v_i) \pi v_j \]

\[ j \neq i \]
has degree smaller than \( \deg Q \). In other words,

\[
\begin{align*}
Q \big| \overline{P} - \left( \sum_{i=1}^{k} c_i n_i w' \right) Q - P,
\end{align*}
\]

and

\[
\deg(\overline{P} - \left( \sum_{i=1}^{k} c_i n_i w' \right) Q - P) < \deg Q.
\]

This implies that

\[
\overline{P} - \left( \sum_{i=1}^{k} c_i n_i w' \right) Q = P
\]

and

\[
P = \frac{k}{Q} \sum_{i=1}^{k} \frac{v_i'}{v_i} - \left( \sum_{i=1}^{k} c_i n_i w' \right)
\]

and \( \forall f \) is also elementary in this case.

Actually, we can do a lot more with \( r(\alpha) = \text{resultant}(P - \alpha Q', Q) \). It turns out that we can determine whether \( \forall f \) is elementary without finding the roots of \( r(\alpha) \), as evidenced by the following.

**Theorem 3**

Let \( f \) be a normal element of \( F_n' \), \( P = \text{num } f, Q = \text{den } f, r(\alpha) = \text{resultant}(P - \alpha Q', Q, \theta_n') \). Let \( R \) be a unique factorization domain whose fraction field is \( K \), and let \( P_i = R[z, \theta_1, \ldots, \theta_i] \), for \( 1 \leq i \leq n, P_0 = R[z], P_{-1} = R \). Then \( \forall f \) is elementary over \( F_n \) if, and only if
\[ r(\alpha) = s(\alpha) \frac{t_1}{t_2} \]

where \( s(\alpha) \in R[\alpha], \ t_1, \ t_2 \in P_{n-1} \).

For a proof of this theorem, we will need some lemmas.

**Lemma 1.**

Let \( t \) be algebraic over \( F_n \), and assume that \( t \) is a constant. Then \( t \) is algebraic over \( K \).

**Proof:** Though this lemma is a special case of lemma 5.1.2 in [KAP 57], we include here a proof for completeness. Since \( t \) is algebraic over \( F_n \), there exists an irreducible polynomial in \( F_n[X] \)

\[ P(X) = \sum_{i=0}^{m} a_i X^i, \]

\[ a_m = 1, \ a_i \in F_n \text{ for } 0 \leq i \leq m \text{ such that} \]

\[ P(t) = \sum_{i=0}^{m} a_i t^i = 0. \]

But then,

\[ [P(t)]' = \left( \sum_{i=0}^{m} a_i \ t^{i-1} \right)' + \sum_{i=0}^{m} a_i t^i = 0 \]

and since \( t' = 0 \), we have that

\[ \sum_{i=0}^{m} a'_i t^i = \left( \sum_{i=0}^{m-1} a'_i t^i \right)' = 0. \]
Since \(1, t, t^2, \ldots, t^{m-1}\) form a basis for \(F_n(t)\) over \(F_n\), this says that all the \(a_i^1 = 0\), for \(0 \leq i \leq m - 1\), and thus \(a_0, a_1, \ldots, a_{m-1} \in K\).

Since \(a_m = 1 \in K\), we obtain \(P(X) \in K[X]\) and \(t\) is algebraic over \(K\).

Lemma 2.

Let \(R\) be a unique factorization domain with fraction field \(K\), let \(\Theta\) be transcendental over \(K\), and \(X\) an indeterminate over \(K(\Theta)\). Let \(r\) be a polynomial in \(X\) over \(R[\Theta]\), and assume that all its roots are algebraic over \(K\). Then there exist \(s, t\) such that

\[
r = s \cdot t
\]

for \(s \in R[X], t \in R[\Theta]\).

Proof: Let \(\alpha_1, \ldots, \alpha_n\) be the roots of \(r\), with multiplicities \(\ell_1, \ldots, \ell_n\) respectively. Since the \(\alpha_i\) are algebraic over \(K\), we obtain that

\[
A = \prod_{i=1}^{h} (X - \alpha_i)^{\ell_i}
\]

is a polynomial over \(\overline{K}\) the algebraic closure of \(K\), i.e. \(A \in \overline{K}[X]\). But also \(A \in K(\Theta)[X]\), so that

\[
A \in \overline{K}[X] \cap K(\Theta)[X] = K[X],
\]

since \(\Theta\) is transcendental over \(X\).

Now, we can find \(s \in R[X], B \in R\) such that \(s\) and \(B\) have no common factors, and \(A = \frac{s}{B}\). We claim that, if \(u \in R\) and \(u\) divides \(s\), then \(u\) is a unit of \(R\). The reason is that, since the leading coefficient of \(A\) is 1, the leading coefficient of \(s\) is \(B\), so that if \(u\) divides \(s\), \(u\) divides \(B\), and thus divides the greatest common factor of \(s\) and \(B\),
which happens to be 1.

We do know, also, that $\deg_X r = \Sigma l_i = \deg_X A = \deg_X s$, and that for some $t$ in $K(\Theta)[X]$, $r = s t$. We want to prove that $t \in R[\Theta]$. But

(a) $\deg_X t = \deg_X r - \deg_X s = 0$, so that $t \in K(\Theta)$.

(b) Since $r \in R[\Theta, X]$, $s \in R[X]$, $t = \frac{r}{s} \in K[\Theta]$.

(c) Let, now, $t = \frac{t_1}{u}$, $t_1 \in R[\Theta]$, $u \in R$, $\gcd(t_1, u) = 1$. We will be done if we prove that $u$ is a unit of $R$. But, since $r = s \frac{t_1}{u}$, we have that $u r = s t_1$, so that $u$ divides $s t_1$. Since $u$ and $t_1$ are relatively prime, we have that $u$ divides $s$. But, we proved above that then, $u$ is a unit.

We are now ready for a proof of the theorem.

If we assume that $r(\alpha) = s(\alpha) \frac{t_1}{t_2}$, where $s(\alpha) \in R[\alpha]$, $t_1, t_2 \in P_{n-1}$ then all the roots of $r(\alpha)$ are constants, so that by theorem 2 if $f f$ is elementary over $F_n$.

Now, assume that $f f$ is elementary over $F_n$. By theorem 1 of this section, we must have that all the roots of $r(\alpha)$ are constants, and by lemma 1, this implies that all the roots of $r(\alpha)$ are algebraic over $K$.

Since $R(\alpha) \in F_{n-1}[\alpha]$, we can find $\overline{r}(\alpha) \in P_{n-1}[\alpha]$, $t_2 \in P_{n-1}$, $\overline{r}(\alpha)$, $t_2$ relatively prime, such that $r(\alpha) = \frac{\overline{r}(\alpha)}{t_2}$. Since $r(\alpha)$, $\overline{r}(\alpha)$ have the same roots, we can apply an inductive generalization of lemma 2, to obtain

$$\overline{r}(\alpha) = s(\alpha) t_1,$$

with $s(\alpha) \in R[\alpha]$, $t_1 \in P_{n-1}$ so that

$$r(\alpha) = \frac{\overline{r}(\alpha)}{t_2} = s(\alpha) \frac{t_1}{t_2}.$$
This theorem has two interesting corollaries.

Corollary 1:

Using the same notation as in the theorem 3, if we assume that \( \alpha \) is the main variable of \( s(\alpha) \), \( t_1 = \overline{r}(\alpha) \), then \( \mathcal{f} \) is elementary if and only if the primitive part of \( \overline{r}(\alpha) \) (as polynomial in \( \alpha \) with coefficients in \( P_{n-1}^0 \)) is such that all its coefficients are in \( R \).

**Proof:** If the coefficients of the primitive part of \( \overline{r}(\alpha) \) are all in \( R \), we have then found \( s(\alpha) \) and \( t_1 \) in the theorem (as primitive part and content of \( \overline{r}(\alpha) \) respectively), and the conclusion follows.

So, assume now that some coefficient of the primitive part of \( \overline{r}(\alpha) \) involves either \( z \) or one of the \( \theta_i \). Since \( P_{n-1}^0[\alpha] \) is a unique factorization domain, we must have that \( s(\alpha) \) and \( t_1 \) in the theorem do not exist, so that \( \mathcal{f} \) is not elementary over \( F_n \).

Our second corollary says exactly the same, but seen from the other end.

Corollary 2:

Again, let us use the same notation as above, but now, let \( \overline{r}(\alpha) \in R[\alpha, z, \theta_1, \ldots, \theta_{n-1}] \). Let \( r_n(\alpha) = \overline{r}(\alpha), r_i(\alpha) = \text{content}(r_{i+1}(\alpha)) \), \( 0 \leq i \leq n-1 \), and let \( s(\alpha) = r_0(\alpha) \in R[\alpha] \). Then \( \mathcal{f} \) is elementary over \( F_n \) if, and only if, \( \deg_\alpha s(\alpha) = \deg_\alpha \overline{r}(\alpha) \).

**Proof:** Very similar to the proof of corollary 1.

Let us then, take a look at what this stage of our integration algorithm is like. We enter the algorithm with \( \mathcal{f} \), normal in \( F_n \), and find \( P = \text{num} \mathcal{f}, Q = \text{den} \mathcal{f} \).

We then compute \( r(\alpha) = \text{resultant}(P - \alpha Q', Q, \theta_n) \) and apply either corollary 1 or corollary 2 to it. If \( r(\alpha) \) fails our test, we can return failure, otherwise we find the roots of \( s(\alpha) \) (in the corollary used),
which we will call $c_1, \ldots, c_k$, compute $v_i = \gcd(P - c_i Q', Q_i)$ where if our gcd algorithm returns the cofactors at no extra cost, we can let $Q_i = Q$, and $Q_{i+1} = Q_i/v_i$, otherwise, let $Q_i = Q$ for all $i$.

In the first case, there is actually no need to compute the last gcd, since we must have $Q_k = 1$ so that $Q_{k-1} = v_k$. Then, if $\Theta_n = \exp v$, we return

$$ff = \sum_{i=1}^{k} c_i \log v_i - \sum_{i=1}^{k} c_i n_i v'$$

where $n_i = \deg v_i$, otherwise, we return

$$ff = \sum_{i=1}^{k} c_i \log v_i$$

We are not yet done in this section. In [TRA 76] it is shown how to obtain the least degree extension, of the constant field, necessary to express the integral of $f$. We claim that our algorithm also produces this least degree extension. The reason for this is that we required our expansion of the integral to be in this smallest extension in the first place; our assumptions about the exact form of the $v_i$ and $c_i$ do not increase the degree of this extension, and our polynomial $r(a)$ defines exactly the $c_i$ we need. Thus, this polynomial $r(a)$ actually defines the smallest degree extension required.

One final note: This algorithm assumes that $\deg Q > 1$ in order to compute the resultant. If $\deg Q = 1$, of course, no factoring can be done and we only have to check whether $f/(Q'/Q)$ (respectively $(f - u')/(Q'/Q)$) is a constant.
9. **Examples.**

In this section, we present some examples illustrating the new algorithms found.

**Example 1:** Illustrating SPDE. Find

$$I = \int \frac{3z^3 + 9z^2 + z - 6}{z} \, e^{1/z} \, dz$$

We know (by section 4) that I must be of the form

$$I = X \, e^{1/z}$$

for some rational function $X$, and substituting, yields:

$$X' - \frac{1}{z^2} X = \frac{3z^3 + 9z^2 + z - 6}{z}$$

We now know that $X$ must be of the form

$$X = \frac{\bar{X}}{z^n}$$

with $n$ and $\bar{X}$ to be determined. If we substitute this value of $x$, we obtain

$$\frac{X' \, z - n \, \bar{X}}{z^{n+1}} - \frac{X}{z^{n+2}} = \frac{3z^3 + 9z^2 + z - 6}{z}$$

so that $n = -1$. Multiplying both sides by $z$, we obtain

$$\bar{X}' \, z^2 + \bar{X}(z - 1) = 3z^3 + 9z^2 + z - 6$$
We obtain that $\deg \overline{X} \leq 2$, and we enter algorithm SPDE.

We have to find $P_1, R_1, \deg R_1 < 2$, such that

$$P_1 z^2 + R_1(z - 1) = 3 z^3 + 9 z^2 + z - 6$$

We obtain:

$$P_1 = 3 z + 4, \quad R_1 = 5 z + 6$$

Since $(z - 1) + (z^2)' = 3 z - 1$ and $P_1 - R_1' = 3 z - 1$, we now have to solve the differential equation

$$Q' z^2 + Q(3 z - 1) = 3z - 1$$

so we have to find $P_2, R_2, \deg R_2 < 2$ such that

$$P_2 z^2 + R_2(3 z - 1) = 3 z - 1$$

We obtain

$$P_2 = 0, \quad R_2 = 1$$

and since $P_2 = R_1'$, we obtain:

$$Q = 1, \quad \overline{X} = z^2 + 5 z + 6, \quad X = z^3 + 5 z^2 + 6 z$$

and

$$I = \int \frac{3 z^3 + 9 z^2 + z - 6}{z} e^{1/z} \, dz = (z^3 + 5 z^2 + 6 z) e^{1/z}$$

Notice we still have to include an integration constant, so that
\[ I = (z^3 + 5z^2 + 6z)e^{1/z} + c \]

for some constant c.

**Example 2:** This example will illustrate some of the self-adapting features of our algorithm. Find

\[ I = \int e^{\log z + 2 \log(z + 1) + z} \, dz \]

Notice that the easy way out would be to simplify our integrand to

\[ \int z(z + 1)^2 e^z \, dz \]

and reaching the differential equation

\[ X' + X = z(z + 1)^2 \]

Our algorithm will reach this differential equation directly. If I is elementary, then I is of the form:

\[ I = Ye^{\log z + z \log(z + 1) + z} \]

so that Y satisfies the differential equation:

\[ Y' + \left( \frac{1}{z} + \frac{2}{z + 1} + 1 \right) Y = 1 \]

so that

\[ Y' + Y \frac{z^2 + 4z + 1}{z^2 + z} = 1 \]

This means that we can write
\[ Y = \frac{\bar{Y}}{(z^2 + z)^m} \]

and that implies

\[ \frac{\bar{Y}' (z^2 + z) + \bar{Y}(z^2 + 4z + 1 - m(2z + 1))}{(z^2 + z)^{m+1}} = 1 \]

Thus, we either have that \( m = -1 \), or that \((z^2 + z)|\bar{X}'(z^2 + z) + \bar{X}(z^2 + 4z + 1 - m(2z + 1))\). This is true if

\[ \gcd(z^2 + (4 - 2m)z + (1 - m), z^2 + z) \neq 1 \]

or

\[ \text{resultant}(z^2 + (4 - 2m)z + (1 - m), z^2 + z, z) = (1 - m)(m - 2) \]

\[ = 0 \]

so that (choose the largest root!) \( m = 2 \). We could now obtain that

\[ Y = \frac{\hat{Y}}{z(z + 1)^2} \]

and replace. That way, we obtain the differential equation (after cancelling common factors):

\[ \hat{Y}' + \hat{Y} = z^3 + 2z^2 + z = z(z + 1)^2 \]

We prefer, though, to use another route. We replace \( m = 2 \), and continue to obtain:
\[
\frac{\bar{Y}'(z^2 + z) + \bar{Y}(z^2 - 1)}{z^3(z + 1)^3} = 1
\]

or

\[
\frac{\bar{Y}' z + \bar{Y}(z - 1)}{z^3(z + 1)^2} = 1.
\]

This implies

\[
\bar{Y}' z + \bar{Y}(z - 1) = z^3(z + 1)^2
\]

and we enter algorithm SPDE.

We now have to find Q, R such that

\[
Qz + R(z - 1) = z^3(z + 1)^2 \quad \text{deg } R < 1
\]

Clearly \(Q = z^2(z + 1)^2, R = 0\) is our answer, and we have to solve

\[
x' z + x z = z^2(z + 1)^2
\]

and dividing by \(z\), we obtain

\[
x' + x = z(z + 1)^2 = z^3 + 2z^2 + z
\]

that is, we reached the same differential equation as before.

We now follow the algorithm (but now, we use case (ii) in section 5) and obtain

\[
x = \sum_{i=0}^{3} a_i z^i
\]

where
\[ a_3 = \frac{1}{1} = 1 \ , \]
\[ a_2 = 2 - 3(1) = -1 \ , \]
\[ a_1 = 1 - 2(-1) = 3 \ , \]
\[ a_0 = 0 - 3 = -3 \]

so that

\[ X = z^3 - z^2 + 3z - 3 \]

Replacing, we obtain

\[ I = Y \ e^{\log z + 2 \log(z + 1) + z + k} \]
\[ = \frac{Y}{z^2(z + 1)^2} \ e^{\log z + 2 \log(z + 1) + z + k} \]
\[ = \frac{z X}{z^2(z + 1)^2} \ e^{\log z + 2 \log(z + 1) + z + k} \]
\[ = \frac{z^3 - z^2 + 3z - 3}{z(z + 1)^2} \ e^{\log z + 2 \log(z + 1) + z + k} \]

where \( k \) is the integration constant.

**Example 3:** The classical non-elementary example. Find

\[ I = \int e^{x^2} \ dx \]

The \( I = Y \ e^{x^2} \) and this leads to the differential equation:
\[ Y' + 2xY = 1 \]

Since \( Y \) must be a polynomial in \( x \) (no denominators), \( Y \) cannot exist because \( \text{deg}(2x) = 1 \), but \( \text{deg}(1) = 0 \).

Thus, \( I \) is not elementary.

**Example 4:** This example illustrates the way the algorithm in section 8 works. Find

\[
I = \int \frac{x + 2}{x^2 - 1} \, dx
\]

We notice that \( f = \frac{x + 2}{x^2 - 1} \, dx \) is normal in \( R(X) \), (where \( R \) is the rational number field) so let:

\[
P = x + 2, \quad Q = x^2 - 1.
\]

Then \( Q' = 2x \), \( P - \alpha Q' = x + 2 - 2\alpha x = (1 - 2\alpha)x + 2 \). Then \( r(\alpha) = \text{resultant}(P - \alpha Q', Q, x) = 4\alpha^2 - 4\alpha - 3 \).

If we want \( r(\alpha) = 0 \), then either

\[
\alpha = \frac{3}{2} \quad \text{or} \quad \alpha = -\frac{1}{2}
\]

so let

\[
c_1 = \frac{3}{2}, \quad c_2 = -\frac{1}{2}.
\]

Let

\[
v_1 = \gcd(P - c_1 Q', Q) = \gcd(x + 2 - \frac{3}{2} (2x), Q)
\]

\[
= \gcd(x + 2 - 3x, x^2 - 1) = \gcd(-2x + 2, x^2 - 1) = x - 1
\]
and let
\[ v_2 = \gcd(P - c_2 Q', Q) = \gcd(2x + 2, x^2 - 1) = x + 1. \]

Thus,
\[ I = c_1 \log v_1 + c_2 \log v_2 = \frac{3}{2} \log(x - 1) - \frac{1}{2} \log(x + 1) + k \]

where \( k \) is the integration constant.

**Example 5:** This example illustrates that our algorithm doesn't necessarily completely factor the denominator of normal rational functions. Find
\[ I = \int \frac{1}{x^3 - x} \, dx \]

Then \( f = \frac{1}{x^3 - x} \) is a normal rational function so let \( P = 1, Q = x^3 - x \). Then
\[ Q' = 3x^2 - 1, P - \alpha Q' = 1 - \alpha(3x^2 - 1) = -3\alpha x^2 + 1 - \alpha \]

and
\[ \text{resultant}(P - \alpha Q', Q, x) = (\alpha + 1)(2\alpha - 1)^2 \]

so let \( c_1 = -1, c_2 = 1/2 \). Then
\[ v_1 = \gcd(P - c_1 Q', Q) = \gcd(1 + 3x^2 - 1, x^3 - x) \]
\[ = \gcd(3x^2, x^3 - x) = x \]
\[ v_2 = \gcd(P - c_2 Q', Q) = \gcd(1 - \frac{1}{2} (5x^2 - 1), x^3 - x) \]

\[ = \gcd(- \frac{3}{2} x^2 + \frac{3}{2}, x^3 - x) = x^2 - 1 \]

and

\[ I = -\log x + \frac{1}{2} \log(x^2 - 1) + k \]

**Example 6:** This example illustrates how the minimum degree extension can differ from the splitting field. Find

\[ I = \int \frac{1}{x^3 + x} \, dx \]

Let \( P = 1, Q = x^3 + x, Q' = 3x^2 + 1, P - \alpha Q' = -3 \alpha x^2 + (1 - \alpha), \) and \( \text{resultant}(-3 \alpha x^2 + (1 - \alpha), x^3 + x) = -(\alpha - 1)(2 \alpha + 1)^2. \) So let \( c_1 = 1, c_2 = -1/2. \) Then

\[ v_1 = \gcd(P - Q', Q) = \gcd(1 - 3x^2 - 1, x^3 + x) = x \]

\[ v_2 = \gcd(P + \frac{1}{2} Q', Q) = \gcd(1 + \frac{3}{2} x^2 + \frac{1}{2}, x^3 + x) = x^2 + 1 \]

and

\[ I = \log x - \frac{1}{2} \log(x^2 + 1) + k \]

**Example 7:** This example illustrates a non-elementary integral detected when trying to obtain the necessary logarithmic terms to express the integral. Find
\[ I = \int \frac{1}{(\log x)^2 - x^2} \, dx \]

Again, let \( P = 1, \ Q = (\log x)^2 - x^2 \) so that

\[ Q' = 2 \frac{\log x}{x} - 2x \]

\[ P - \alpha Q' = -\frac{2\alpha \log x}{x} + (2\alpha x + 1) \]

and

\[ \text{resultant}(P - \alpha Q', Q, \log x) = \alpha^2(4x^2 - 4) + 4\alpha x + 1. \]

Since the content of this last expression is 1, we obtain that \( I \) is not elementary.

**Example 8:** This last example illustrates corollaries 8.1 and 8.2.

Find

\[ I = \int \frac{\log x - 1}{(\log x)^2 - x^2} \, dx. \]

Now

\[ f = \frac{\log x - 1}{(\log x)^2 - x^2} \]

is a normal element of \( R(x, \log x) \) (where \( R \) is the rational number field), so let \( P = \log x - 1, \ Q = (\log x)^2 - x^2 \), so that

\[ Q' = 2 \frac{\log x}{x} - 2x, \]

\[ P - \alpha Q' = \log x(1 - \frac{2\alpha}{x}) + (2\alpha x - 1) \]
resultant(P - α Q', Q, log x) = (x^2 - 1)(4 α^2 - 1).

Since (4 α^2 - 1) does not involve x, I is elementary, and we proceed.

Let \( c_1 = 1/2, c_2 = -1/2 \). Then

\[ v_1 = \gcd(P - c_1 Q', Q) = \gcd(\log x (1 - \frac{1}{x}) + (x - 1), \]

\[ (\log x)^2 - x^2 = \log x + x \]

\[ v_2 = \gcd(P - c_2 Q', Q) = \gcd(\log x (1 + \frac{1}{x}) - x - 1, \]

\[ (\log x)^2 - x^2 = \log x - x, \]

and

\[ I = \frac{1}{2} \log(\log x + x) - \frac{1}{2} \log(\log x - x) + k \]


In this section, we will present a computing time analysis of this algorithm for the rational function case. First, recall that, if \( P \) is a polynomial with integer coefficients, \( P = \sum_{i=0}^{n} a_i x^i \), we define

\[ \text{norm } P = |P| = \sum_{i=0}^{n} |a_i| \]

Now, we define \( F(m, n, d) \) as the class of functions \( P/Q \), with \( P, Q \) relatively prime univariate polynomials over the integers,
\[
\max(|P|, |Q|) \leq d, \ \deg P \leq m, \ \deg Q \leq n.
\]

We shall use the definitions and notation for dominance and co-
dominance used, for example, in [COL 71].

Then, we have the following theorem. For \( f \in F(m, n, d) \), the time
required by the algorithm described herein is given by

\[
T_{INTO}(m, n, d) \leq n^8 L^2(dn) + n^6 L^3(dn)
\]

\[+ \max(m + 1 - n, 0) n L^2(d) + 1\]

if we assume that no algebraic extensions are required, and that the
norm of any of the partial results except the resultant, is also
bounded by \( d \), where \( L(d) = \log_2(d) + 1 \).

Proof: We have two cases to consider.

(a) \( m > n \), and

(b) \( m < n \).

If \( m < n \), we do a quotient-remainder operation, and then we
continue with sections 7 and 8. We then have the following computing
times.

Section 3 (quotient-remainder operation) requires constant (1) time.
Section 7 requires time \( n^5 L(nd)^2 \) as proven in [McD 75].
Section 8 requires:

\( n L(d) \) to compute \( Q' \)

\( n L(d) \) to compute \( \overline{P} - \alpha Q' \) (deg \( \overline{P} < n) \)

\( n^3 L(d) \) to compute \( R = \text{resultant}(P - \alpha Q', Q) \).

(We point out that \( \deg R \leq n \), and it's norm is bounded by \( (2n)! d^{2n} \leq 2n^{2n} d^{2n} = (2dn)^{2n} \) and thus \( L(\text{norm } R) \leq n L(dn) \)).
\[ n^8 + n^6 L^2(\text{norm } R) + n^3 L^3(\text{norm } R) \leq n^8 L^2(\text{dn}) + n^6 L^3(\text{dn}) \]

to compute the roots of \( R \) (from Appendix B, assuming number of roots = \( n \)).

\[ n(n^2 L(d) + n L^2(d)) \]
to compute \( \gcd(P - c_i Q', Q) \) for \( 1 \leq i \leq n \)
(assuming there are \( n \) distinct roots of \( R \)).

Adding these times, it is clear that the time to compute the roots of \( R \) dominates all other computing times, and we obtain

\[ n^8 L^2(\text{dn}) + n^6 L^3(\text{dn}) \]

for section 8.

Finally, if \( m > n \), the time to compute the quotient-remainder is given by \( (m + 1 - n)n L^2(d) \) and the time to compute the integral of the polynomial part (by the classical method) is given by \( (m + 1 - n) L^2(d) \).

If we add all these computing times we obtain the result we quoted at the beginning.

Note: The bounds on the time to compute the resultant and the norm of \( R \) were obtained from [COI 71].
Chapter 3

A STRUCTURE THEOREM FOR EXPONENTIAL AND PRIMITIVE FUNCTIONS

1. Introduction.

In this chapter, the structure of fields of functions that are obtained from the rational functions by the use of algebraic operations, integration and exponentiation are studied. Such results are interesting since they provide a basis for algorithms that discover algebraic relations among various classes of functions and lead to automatic tests for equality of expressions.

The main results of this chapter are theorem 12 and corollary 13, which give explicitly the form of any possible algebraic relationship between exponential and primitive functions, i.e. functions definable by integrals, like erf(s), etc., provided that some care is taken in their construction. By the latter we mean that if an integral can be represented by using logarithms, then this should be done for the theorem to apply.

Before proving our structure theorem, we will need some purely algebraic results:

Lemma 1.

Let $F$ be a field, $w$ algebraic over $F$, $\theta$ transcendental over $F$. Then $w$ is algebraic over $F(\theta)$ and $[F(w) : F] = [F(\theta, w) : F(\theta)]$.

Proof: By the first theorem (unnumbered) in [VdW 70] (§73) $w$ is not in $F(\theta)$ if $w$ is not in $F$. 

78
Let $n = [F(w): F]$, and since $[F(\Theta, w): F(\Theta)] \leq n$, assume $[F(\Theta, w): F(\Theta)] = m < n$. Then, there exist $(r_i)_{0 \leq i \leq m}$ such that each $r_i \in F(\Theta)$, such that

$$\sum_{i=1}^{m} r_i w^i = 0 \tag{1}$$

Since $F[\Theta]$ is a unique factorization domain with fraction field $F(\Theta)$, let $r_i = p_i/q_i$, where $p_i$, $q_i$ are relatively prime elements of $F[\Theta]$. Multiplying equation (1) by $\prod_{i=1}^{m} q_i$, we obtain

$$s_i = p_i \prod_{j=1}^{m} q_j \prod_{j \neq i} q_j$$

such that $\sum_{i=0}^{m} s_i w^i = 0$, and $s_i \in F[\Theta]$.

Let $k = \max(\deg(s_i))$. Notice that $k$ must be positive. Thus for $0 \leq i \leq m$ each $i$,

$$s_i = \sum_{j=0}^{k} a_{ij} \Theta^j$$

where $a_{ij} \in F$, for all $i$, $j$ so that

$$0 = \sum_{i=0}^{m} \sum_{j=0}^{k} a_{ij} \Theta^j w^j = \sum_{j=0}^{k} \sum_{i=0}^{m} a_{ij} w^i \Theta^j.$$
Since \( m < n \), and \( k > 0 \), this equation implies that \( \Theta \) is algebraic over \( F(w) \), and thus, \( \Theta \) is algebraic over \( F \), in contradiction to the hypothesis.  \\

Corollary 2.

Let \( F \) be a field, \( \Theta \) transcendental over \( F \), \( w \) separably algebraic over \( F(\Theta) \), \( v \) in \( F(\Theta, w) \), algebraic over \( F \). Then \( v \) is separably algebraic over \( F \).

Proof: Since \( v \in F(\Theta, w) \), \( v \) is separably algebraic over \( F(\Theta) \), therefore, its irreducible polynomial over \( F(\Theta) \) (and by theorem 1, over \( F \)) cannot have any multiple roots. But that is the definition of separably algebraic.  \\

Lemma 3.

Let \( F \) be a field, \( \Theta \) transcendental over \( F \), \( w \) algebraic over \( F(\Theta) \). Then if \( F \) is perfect or \( w \) is separably algebraic over \( F(\Theta) \), there exists \( u \) in \( F(\Theta, w) \), algebraic over \( F \) such that if \( t \in F(\Theta, w) \) is algebraic over \( F(u) \), then \( t \in F(u) \) (i.e. \( F(u) \) is algebraically closed in \( F(\Theta, w) = F(u, \Theta, w) \)).

Proof: If \( F \) is algebraically closed in \( F(\Theta, w) \) then the result follows with \( u = 1 \), otherwise there is a \( t \in F(\Theta, w) \) algebraic over \( F \) but not in \( F \). Thus, we can define a sequence \( t_i \) and \( F_i \) by: \( t_1 = t \), \( F_1 = F(t) \), and for \( n \geq 1 \), let \( s_n \) be an element of \( F(\Theta, w) \) algebraic over \( F_n \) but not in \( F_n \), let \( t_{n+1} \) be such that \( F_{n+1} = F(t_{n+1}) = F(t_n, s_n) \). Notice that \( t_{n+1} \) exists if \( s_n \) exists by the theorem of the primitive element, since \( s_n \) (and also \( t_n \)) are separable over \( F \). However, \( s_n \) does not exist for all \( n \), since \( [F_{n+1}: F_n] > 1 \) for all \( n \) such that \( s_n \) exists, so that \( [F_n: F] \) is a strictly increasing sequence of integers. Since \( [F_n: F] \leq [F(\Theta, w): F(\Theta)] \) (by lemma 1), there is an \( n_0 \) such that \( s_{n_0} = 1 \).
exists, but \( s_{n_0} \) does not. Let \( u = t_{n_0} \), and then \( F(t_{n_0}) = F(u) \) is algebraically closed in \( F(0, w) \).

Let \( F \) be a field, \( G = F(t_1, \ldots, t_n) \) a finitely generated extension of \( F \). We shall denote by \( F_i \), the fields: \( F_i = F(t_1, t_2, \ldots, t_i) \) for \( 1 \leq i \leq n \) and \( F_0 = F \).

Let \( F \) be a field, \( G = F(t_1, \ldots, t_n) \) a finitely generated extension of \( F \). We shall say that \( G \) satisfies the algebraic closure property with respect to \( F \), if there does not exist an \( i \) such that \( F_i \) is an algebraic extension of \( F_{i-1} \) and \( F_{i+1} \) is an algebraic extension of \( F_i \). Furthermore, for each \( i > 1 \) such that \( F_i \) is an algebraic extension of \( F_{i-1} \), \( F_{i-2} \) must be algebraically closed in \( F_i \).

The following lemma could have been proven in a more general setting, but for simplicity, we shall assume our field has characteristic 0, which is our only case of interest.

**Lemma 4.**

Let \( F \) be a field of characteristic 0, \( G = F(t_1, \ldots, t_n) \) be a finitely generated extension of \( F \). Then there exists an integer \( m \) and \( s_1, \ldots, s_m \) in \( G \) such that \( G_m = F(s_1, \ldots, s_m) \) satisfies the algebraic closure property with respect to \( F \), and an integer-valued, non-decreasing function of \( f \), mapping the integers \([1, n]\) onto \([1, m]\) that if \( t_i \) is transcendental over \( F_{i-1} = F(t_1, \ldots, t_{i-1}) \) then \( t_i = s_f(i) \) is transcendental over \( F_{f(i)-1} = F(s_1, \ldots, s_{f(i)-1}) \).

**Proof:** The proof is by induction on \( n \). If \( n = 1 \), let \( G = F(t_1) \) and \( f(1) = 1 \). Now assume \( n > 1 \) and that the theorem is true for \( k < n \). We then have three cases to consider:

i) \( t_n \) is transcendental over \( F_{n-1} \). We then apply the induction hypothesis to \( F_{n-1} \) to obtain \( G_{m-1} \). Thus \( G_m = G_{m-1}(t_n) \) and \( f(n) = m \).
ii) \( t_n \) is algebraic over \( F_{n-1} \), and \( t_{n-1} \) is also algebraic over \( F_{n-2} \). By the theorem of the primitive element, there exists a \( u \) such that \( F_{n-2}(u) = F_{n-2}(t_{n-1}, t_n) \). Now apply the induction hypothesis to \( f_{n-2}(u) \) to obtain \( G_m \). In this case \( f(n-1) = f(n) = m \).

iii) \( t_n \) is algebraic over \( F_{n-1} \) and \( t_{n-1} \) is transcendental over \( F_{n-2} \). By lemma 3, there exists a \( u \) in \( F_n \) algebraic over \( F_{n-2} \) such that \( F_{n-2}(u) \) is algebraically closed in \( F_n \). By the induction hypothesis there exists an extension \( G_{\mu} \) which satisfies the algebraic closure property with respect to \( F \), and such that \( G_{\mu} = F_{n-2}(u) \). Then \( G_{\mu+1} = G_{\mu}(t_{n-1}) \) also satisfies the algebraic closure property. If \( t_n \in G_{\mu+1} \), then \( G_m = G_{\mu+1} \) and \( f(n-1) = f(n) = \mu+1 \), otherwise \( G_m = G_{\mu+1}(t_n) \) and \( f(n-1) = \mu+1, f(n) = \mu+2 \).

2. Some Basic Lemmas

Lemma 5.

Let \( F \) be a differential field, \( \Theta \) a regular monomial over \( F \). Let \( G = F(\Theta) \) or \( G = F(\Theta, t) \) where \( t \) is algebraic over \( F(\Theta) \) and \( F \) is algebraically closed in \( G \).

Let \( c_i, 1 \leq i \leq n \), be constants of \( F \) linearly independent over the rationals, and let \( v, u_i, 1 \leq i \leq n \), be non-zero elements of \( G \) such that for all derivations \( D \) of \( G \), we have

\[
\sum_{i=1}^{n} c_i \frac{D u_i}{u_i} = Dv .
\]

Then, if \( \Theta \) is primitive over \( F \), each \( u_i \) must be in \( F \) and \( v \) must be of the form \( c \Theta + w \), for some \( w \in F \) and \( c \in C^F \). If \( \Theta \) is exponential over \( F \), then \( v \) must be in \( F \) and there exist integers \( m, n_i, 1 \leq i \leq n \),
with \( m \neq 0 \) and \( \hat{u}_i \in F \) such that \( u_i^m = \Theta^n \hat{u}_i \), with \( \overline{u}_i = u_i^m \) satisfying all derivations \( D \) of \( G \),

\[
\sum_{i=1}^{n} c_i \frac{D \overline{u}_i}{\overline{u}_i} = D(mv) .
\]

If \( G = F(\Theta) \), \( m = 1 \) is sufficient in the above equations.

**Proof:** This lemma is the main result of [ROS 69], but a simple proof of a generalization thereof can be found in [ROS 75]. The sufficiency of \( m = 1 \) for the case \( G = F(\Theta) \) can be seen by the fact that since \( F[\Theta] \) is a unique factorization domain, \( m \) must divide \( n_i \) for all \( i \). Thus

\[
\hat{u}_i = u_i^m \Theta^n
\]

as an element of \( F \), has an \( m \)-th root in \( F(\Theta) = G \). Therefore, by the first theorem in §73 of [VdW 70], \( \hat{u}_i \) has an \( m \)-th root in \( F \).  

**Corollary 6.**

Let \( F, G, \Theta, t \) be as in lemma 5 and let \( v \in G \). Then \( \exp v \) is not a regular monomial over \( G \) if and only if \( v \in F(\Theta) = H \) and \( \exp v \) is not a regular monomial over \( H \).

**Proof:** Let \( \phi = \exp(v) \). Then if \( \phi \) is not a regular monomial over \( G \) there exists a constant \( c \) and an integer \( n \) such that \( u = c \phi^n \in G \). Thus for any differentiation operator \( D \) of \( G \), we have

\[
\frac{Du}{u} = nDv
\]
By lemma 5, $v \in H$. Furthermore, if $\Theta$ is primitive over $F$, $u \in F \subseteq H$ otherwise $\Theta$ is exponential over $F$ and $u$ is algebraic over $H$. \[ \]

**Lemma 7.**

Let $v \in F(\Theta)$ with $\Theta$ a primitive regular monomial over the differential field $F$, and assume $\phi = \exp(v)$ is not a regular monomial over $F(\Theta)$.

Then there exist $u, w \in F$, constants $c_1, c_3 \in C, c_2 \in C^F$ and a non-zero integer $m$, such that

$$u = c_3 \phi^m,$$

and

$$v = \frac{1}{m} \log u + c_1 = c_2 \Theta + w.$$

**Proof:** Since $\phi$ is not a regular monomial over $F(\Theta)$ there exist a constant $c_3$ and a non-zero integer $m$, such that $c_3 \phi^m \in F(\Theta)$. Let $c_3 \phi^m = u$. Then, for each derivation operator $D$ of $F$,

$$\frac{Du}{u} = \frac{m c_3 \phi^m Dv}{c_3 \phi^m} = m Dv$$

so that by lemma 5, $u \in F$, $v = c_2 \Theta + w$ for some constant $c_2$, $w \in F$, and by the definition of logarithm, $v = \frac{1}{m} \log u + c_1$ as required. \[ \]

**Note:** From now on, $K(z_1, \ldots, z_n)$ will always denote the differential field with constant field $K$ and $n$ derivation operations $D_1, \ldots, D_n$ where
\[ D_i z_j = \begin{cases} 1 \text{ if } i=j \\ 0 \text{ if } i \neq j \end{cases} \]

Corollary 8.

Let \( v \in F = K(z_1, \ldots, z_n) \), and assume \( \phi = \exp(v) \) is not a regular monomial over \( F \). Then \( v \in K \).

Proof: Though this proposition is a special case of the Risch Structure Theorem as stated in [EpC 74], theorem 5.7, we feel the following proof is instructive.

Since \( z_n \) is primitive over \( F_{n-1} = K(z_1, \ldots, z_{n-1}) \), we can find constants \( c_1, c_3 \in C, c_2 \in K \), and elements \( u, w \in F_{n-1} \) such that \( v = c_2 z_n + w \), and \( v = c_3 \phi^m \). Thus \( \phi \) is algebraic over \( F_{n-1}(c_3) \) so that \( D_n(\phi) = 0 \). But \( 0 = D_n(\phi) = \phi D_n(v) = \phi c_2 \), and since \( \phi \neq 0 \), we must have \( c_2 = 0 \) and \( v = w \in F_{n-1} \). Proceeding by induction on \( n \), we obtain \( v \in K \).

Corollary 9.

Let \( \Theta_1 \) and \( \Theta_2 \) be regular monomials over the differential field \( F \), with \( \Theta_1 \) primitive and \( \Theta_2 \) exponential over \( F \). Then \( \Theta_1 \) is a regular monomial over \( F(\Theta_2) \) and \( \Theta_2 \) is a regular monomial over \( F(\Theta_1) \).

Proof: If \( \Theta_2 \) is not a regular monomial over \( F(\Theta_1) \), then, by lemma 7 we would have \( u = c \Theta_2^m \in F \) for some constant \( c \) and integer \( m \).

Our proof concludes by noting that if \( \Theta_1 \) is not a regular monomial over \( F(\Theta_2) \), then \( \Theta_2 \) is not a regular monomial over \( F(\Theta_1) \).

Lemma 10.

Let \( F \) be a differential field, \( t \in F \), \( \Theta = \exp(t) \) a regular monomial over \( F \). Let \( v \in F(\Theta) \) and assume that \( \phi = \exp(v) \) is not a regular monomial over \( F(\Theta) \). Then there exist \( w \in F \) and a rational number \( q \)
such that exp(w) is not a regular monomial over F, and \( v = w + q \, t \).

**Proof:** Since \( \phi \) is not a regular monomial over \( F(\Theta) \), there exist a \( c \in C \) and an integer \( n \) such that \( u = c \, \phi^n \in F(\Theta) \). Then for all derivation operators \( D \) of \( F \)

\[
\frac{Du}{u} = n \, Dv
\]

so that by lemma 5, \( v \in F \), and \( u = \Theta^k \, s \) for some \( s \in F \), and integer \( k \). Thus,

\[
\frac{Du}{u} = \frac{Ds}{s} + k \, Dt = n \, Dv
\]

and

\[
\frac{Ds}{s} = n \, Dv - k \, Dt.
\]

If we now let \( w = v - \frac{k}{n} \, t \), \( q = \frac{k}{n} \) then exp \( w \) is not a regular monomial over \( F \), because for some constant \( c_1 \), \( c_1 \, s = \exp w \) and \( s \in F \), and also \( v = w + q \, t \) as required. \( \Box \)

3. **Simple-logarithmic Elements and Log-explicit Extensions**

We have by now, accomplished the bulk of an induction proof for one of our main goals i.e. theorem 12. The main obstacle remaining is to show that \( c_2 \) in lemma 7 is rational. For this to hold, however, we need some additional hypothesis.

Let \( F \) be a differential field, \( u \) primitive over \( F \). We shall say that \( u \) is **simple-logarithmic** over \( F \) if there exist \( v_1, \ldots, v_k \in F \) such that for some constant \( c \), \( u + c \in F(\log v_1, \ldots, \log v_k) \). We shall also say that \( u \) is **non-simple** if it is not simple-logarithmic over \( F \).
Let $F_j = F_0(\theta_1, \ldots, \theta_j)$, $F = F_n$ be a regular (generalized) Liouville extension of $F_0$. We shall say that $F$ is a regular (respectively generalized) log-explicit extension of $F_0$ if for each $\theta_j$ one of the following conditions hold:

a. $\theta_j = \exp(a_j)$, $a_j \in F_{j-1}$

b. $\theta_j$ is primitive and non-simple over $F_{j-1}$

c. $\theta_j = \log(a_j)$, $a_j \in F_{j-1}$, or

d. $\theta_j$ is algebraic over $F_{j-1}$.

That is, if $\theta_j$ is simple-logarithmic over $F_{j-1}$, then $\theta_j = \log(a_j)$, $a_j \in F_{j-1}$.

A regular (generalized) Liouville field is a regular (generalized) Liouville extension of $K(z_1, \ldots, z_m)$.

A regular (generalized) log-explicit field is a regular (generalized) log-explicit extension of $K(z_1, \ldots, z_m)$.

Lemma 11.

Let $G = K(z_1, \ldots, z_m, \theta_1, \ldots, \theta_n)$ be a log-explicit field satisfying the algebraic closure property with respect to $K(z_1, \ldots, z_m)$, let $f, v, u_1, \ldots, u_n \in G$, and assume that $G(\log u_1, \ldots, \log u_n)$ is a regular Liouville extension of $G$, with $n \geq 1$.

Let $c_1, \ldots, c_n$ be constants of $G$ such that for all derivation operators $D$ of $G$:

$$Dv + \frac{Df}{f} = \sum_{i=1}^{n} c_i \frac{Du_i}{u_i}$$

Then the set $\{1, c_1, \ldots, c_n\}$ is linearly dependent over the rationals.
Proof: If the set \( \{1, c_1, \ldots, c_n\} \) were not linearly dependent, then we could rewrite our equation above as

\[
Dv = \sum_{i=0}^{n} c_i \frac{Du_i}{u_i},
\]

(2)

where \( u_0 = f \), \( c_0 = -1 \), so that lemma 5 applies, and if we let \( G = F(\Theta, w) \), we obtain two cases:

i) If \( \Theta \) is primitive over \( F \), then each \( u_i \in F \) and \( v = c \Theta + v_1 \) for some constant \( c \) and \( v_1 \in F \). Here we have 3 possibilities.

(a) \( c = 0 \). Then we have proven that \( v \) and the \( u_i \) are in \( F \).

(b) \( c \neq 0 \) and the set \( \{1, c_1, \ldots, c_n, c\} \) is linearly independent over the rationals. Then by (2) we obtain that \( \Theta \) is simple-logarithmic over \( F \), so that \( \Theta = \log(w) \) for some \( w \in F \). If we rename \( c \) and \( w \) as \( -c_{n+1} \) and \( u_{n+1} \) respectively, we obtain an equivalent problem, but with the \( u_i \), \( v \) in \( F \), namely:

\[
Dv_1 = \sum_{i=0}^{n+1} c_i \frac{Du_i}{u_i}.
\]

Notice that \( F(\log u_1, \log u_2, \ldots, \log u_{n+1}) \) is a regular Liouville extension of \( F \).

(c) \( c \neq 0 \) and the set \( \{1, c_1, \ldots, c_n, c\} \) is linearly dependent over the rationals. Thus, for there are rationals \( r_0, \ldots, r_{n+1} \) not all 0, such that

\[
r_0 + \sum_{i=1}^{n} r_i c_i + r_{n+1} c = 0
\]
We might as well assume, without loss of generality that the $r_i$ are integers and notice also that $r_{n+1}$ cannot be 0, because of our linear independence assumption. Thus we obtain a value for $c$ in terms of the $r_i$ and $c_i$. Also, just as in (b), we get $\Theta = \log(u_{n+1})$ (after renaming). If we now replace these values in equation (2) above, we obtain

$$D(r_{n+1} v_1) = \sum_{i=0}^{n} c_i \frac{D(u_i u_{n+1}^{-r_i})}{u_i u_{n+1}^{-r_i}}$$

where

$$F(\log \frac{u_1}{r_1}, \ldots, \frac{u_n}{r_n})$$

$$u_{n+1}$$

is a regular Liouville extension of $F$.

ii) If $\Theta$ is exponential over $F$, then $v \in F$ and there exist integers $m, n_i$, and $u_i \in F$, such that $u_i^m = \Theta^i u_i$, and

$$\sum_{i=1}^{n} c_i \frac{D(u_i^{m})}{u_i^{m}} = D(m v)$$

If we let $\Theta = \exp(t)$, we get

$$D(m v) = \sum_{i=0}^{n} c_i \frac{D(u_i^{m})}{u_i^{m}} = D(\sum_{i=0}^{n} (c_i n_i t) + \sum_{i=0}^{n} \frac{\Delta u_i}{u_i})$$
Thus we obtain (again) a similar equation to our original one:

$$D(v_1) = \sum_{i=0}^{n} c_i \frac{\hat{u}_i}{u_i}$$

where

$$v_1 = m \cdot v - \sum_{i=0}^{n} (c_i \cdot n_i \cdot t).$$

Notice that since \(\log \hat{u}_i = m \cdot \log u_i - n_i \cdot t + k_i\) for some constant \(k_i \in F\), we get that \(F(\log \hat{u}_1, \log \hat{u}_2, \ldots, \log \hat{u}_n)\) is a regular Liouville extension of \(F\).

Thus we have obtained in each case a \(v_1, \hat{u}_0, \ldots, \hat{u}_n \in F\) satisfying the same assumptions as our original \(v\) and \(u_i\), which were in \(G\). Proceeding by induction, we will therefore find constants \(v_p, \tilde{u}_0, \ldots, \tilde{u}_n\) such that \(K(\log \tilde{u}_1, \ldots, \log \tilde{u}_n)\) is a regular Liouville extension of \(K\), the constant field of \(G\) (and \(F\)). But this is clearly impossible (the \(\tilde{u}_i\) being constant) and thus, the \(c_i\) must be linearly dependent over the rationals. \(\blacksquare\)

The following notation will be useful in the discussion to follow. Let \(F_n = F_0(\Theta_1, \ldots, \Theta_n)\) be a regular (generalized) Liouville extension of \(F_0\). Let \(F_i = F_0(\Theta_1, \ldots, \Theta_i), 1 \leq i \leq n\), and define the sets: For \(1 \leq j \leq n\),

\(E_j = \{i | \Theta_i = \exp(u_i), u_i \in F_{i-1}, 1 \leq i \leq j\}\)

\(P_j = \{i | \Theta_i \in F_{i-1}, 1 \leq i \leq j\}\)
\[ L_j = \{ i | \theta_i = \log(u_i), u_i \in F_{i-1}, 1 \leq i \leq j \} \]

Not that in all cases \( L_j \) is a subset of \( P_j \).

**Theorem 12.**

Let \( F \) be a generalized log-explicit field, \( F = F_n = K(z_1, \ldots, z_m, \theta_1, \ldots, \theta_n) \). Let \( u \in F_n \), and assume \( \exp(u) \) is not a regular monomial over \( F_n \). Then there exist rational numbers \( p_i \) and a constant \( c \in C^F \) such that

\[ u = \sum_{i \in L_n} p_i \theta_i + \sum_{i \in E_n} p_i a_i + c \]

where \( \theta_i = \exp(a_i) \) for \( i \in E_n \).

**Proof:** The proof is by induction on \( n \), but by lemma 4, we can assume that \( F \) satisfies the algebraic closure property with respect to \( K(z_1, \ldots, z_m) \). Furthermore, since this transformation preserves the monomials, the statement of the theorem remains the same. For \( n = 0 \), the result follows from corollary 8, so assume the theorem is true for \( K < n \). Let \( \phi = \exp(u) \).

We have 3 cases:

Case 1. \( \theta_n \) is primitive over \( F_{n-1} \). Then by lemma 7, there are \( r, w \in F_{n-1}, c_1, c_3 \in C, c_2 \in C^F \) and an integer \( m \) such that \( r = c_3 \phi^m \) and

\[ u = \frac{1}{m} \log r + c_1 = c_2 \theta_n + w \quad (3) \]

We then have two subcases:

(a) \( c_2 = 0 \). Then \( w = u \), and \( \exp(w) \) is not a regular monomial over \( F_{n-1} \) (by corollary 9). Applying the induction hypothesis we obtain our result for this case.
(b) $c_2 \neq 0$. Then $\Theta_n$ is simple-logarithmic over $F_{n-1}$, so $\Theta_n = \log v, \; v \in F_{n-1}$. Substituting and applying any differentiation operator $D$ of $F$ in equation (3) above, we obtain:

$$\frac{1}{m} \frac{Dr}{r} = c_2 \frac{Dv}{v} + Dw.$$ 

Then by lemma 11, $c_2$ is a rational number, say $p/q$, where $p, q$ are relatively prime integers. Using this result in equation (3), we obtain:

$$u = \frac{p}{q} \log v + w.$$ 

If we now take exponentials and raise to the $q$-th power, we obtain

$$(\exp(u))^q = \phi^q = c v^{p} (\exp(w))^q,$$

where $c$ is a constant in $C$.

Now if we raise this equation to the $m$-th power, and multiply by $c_3^q$, we obtain

$$c_3^q \phi^{mq} = (c_3^q \phi^m)^q = r^q = c^m c_3^q v^{pm} [\exp(w)]^{qm}$$

so that

$$c^m c_3^q [\exp(w)]^{qm} = \frac{r^q}{v^{pm}} \in F_{n-1}$$

and $\exp(w)$ is not a regular monomial over $F_{n-1}$. Therefore, we can apply our induction hypothesis to $w$ to obtain:
\[ u = \frac{p}{q} \log v + w = \frac{p}{q} \log v + \sum_{i \in L_{n-1}} p_i \theta_i + \sum_{i \in E_{n-1}} p_i a_i + c \]

\[ = \sum_{i \in L_n} p_i \theta_i + \sum_{i \in E_n} p_i a_i + c \]

where \( p_n = \frac{p}{q} \).

Case 2. \( \theta_n = \exp(v), v \in F_{n-1} \). Then, by lemma 10, there exist \( w \in F_{n-1} \), and rational numbers \( p, q \) such that

i) \( \exp(w) \) is not a regular monomial over \( F_{n-1} \)

ii) \( u = w + p v \)

Since \( \exp(w) \) is not a regular monomial over \( F_{n-1} \), we can apply the induction hypothesis to obtain

\[ u = w + p v = \sum_{i \in L_{n-1}} p_i \theta_i + \sum_{i \in E_{n-1}} p_i a_i + c + p v \]

\[ = \sum_{i \in L_n} p_i \theta_i + \sum_{i \in E_n} p_i a_i + c \]

where \( p_n = p, a_n = v \).

Case 3. \( \theta_n \) is algebraic over \( F_{n-1} \). Then, by corollary 6, \( u \in F_{n-1} \) and \( \exp(u) \) is not a regular monomial over \( F_{n-1} \). Therefore, we can apply our induction hypothesis to prove the lemma for this case.

Corollary 13.

Let \( F \) be a generalized, log-explicit field, \( F = F_n = K(z_1, \ldots, z_m, \theta_1, \ldots, \theta_n) \). Let \( F_i = K(z_1, \ldots, z_m, \theta_1, \ldots, \theta_i), 0 \leq i \leq n. \)
Let \( u \in F \), and assume \( \log u \) is not a regular monomial over \( F \). Then, there exist rational integers \( p_i, q \) and a constant \( k \) in \( C^F \) such that

\[
  u^q = k \left( \prod_{i \in L_n} a_i^{p_i} \right) \left( \prod_{i \in E_n} \theta_i^{p_i} \right)
\]

where \( \theta_i = \log a_i \) for \( i \in L_n \).

**Proof:** Let for \( i \in E_n \), \( \theta_i = \exp a_i \). Since \( \log u \) is not a regular monomial over \( F \), there exists a \( w \in F \) such that for any differentiation operator \( D(w) = \frac{Du}{u} \), thus \( \exp(w) = c_1 u \) for some \( c_1 \in C \). This says that \( \exp(w) \) is not a regular monomial over \( F \), so that by theorem 12 we can find rational integers \( p_i, q \) and a constant \( c \) such that

\[
  q w = \sum_{i \in L_n} p_i \theta_i + \sum_{i \in E_n} p_i a_i + c.
\]

If we now take exponentials of this equation, we obtain

\[
  (c_1 u)^q = (\exp(w))^q = \left( \prod_{i \in L_n} a_i^{p_i} \right) \left( \prod_{i \in E_n} \theta_i^{p_i} \right) c_2
\]

for some \( c_2 \in C \), or

\[
  u^q = \frac{c_2}{c_1^q} \left( \prod_{i \in L_n} q_i^{p_i} \right) \left( \prod_{i \in E_n} \theta_i^{p_i} \right)
\]

Notice that \( \frac{c_2}{c_1^q} = k \) must be in \( F \), since every other term in this expression is also in \( F \).
In summary we have reached the following algorithmic procedure for canonical forms for logarithms, exponentials and integrals so long as certain problems with constants are avoided. We have a canonical form for what is (essentially) a field of rational functions in the variables $z_1, \ldots, z_n$.

Assume inductively that we have a canonical form for a generalized log-explicit field $F$ of primitive and exponential functions, and we are given a $v$ to put in canonical form. We have 4 cases.

(a) If $v$ is algebraic over $F$, let $p(x)$ be the monic irreducible polynomial over $F$ s.t. $p(v) = 0$. We then obtain a canonical form for $F(v) = F[v]$ by performing all arithmetic modulo $p(v)$.

(b) If $v = \exp(u)$, $u \in F$, we apply theorem 12 and either obtain that $v$ is a regular monomial over $F$, so we can simply add it as another indeterminate, or we obtain (using the notation in theorem 12)

$$u = \sum_{i \in \mathbb{E}_n} p_i a_i + \sum_{i \in \mathbb{L}_n} p_i \theta_i + c,$$

so that

$$v = \prod_{i \in \mathbb{E}_n} \theta_i^{p_i} \prod_{i \in \mathbb{L}_n} a_i^{p_i} e^c$$

is algebraic over $F(e^c)$. If it can be determined whether $e^c$ is transcendental over $F$, we can adjoin $e^c$ to the constant field of $F$ (as a new variable, either independent or algebraic) and then use case (a).

We do have a problem here, though, because no general procedure is known at present to determine whether $e^c$ is algebraic or transcendental over $F$. In this situation, an ad-hoc procedure must be used, since no
In [EPS 75] algorithms are given which apply similar results to fields built-up from the rational functions by repeatedly applying the logarithm and exponential functions. However, his algorithms do not apply to functions defined by general integration.

4. Liouville Fields and Extensions.

Our main purpose here is to develop necessary and sufficient conditions for algebraic relationships to hold between integrals and exponentials. Though we have achieved our goal, as we shall see, the resulting condition seems almost useless for automatic implementation in the Liouville extension case. The equivalent condition for log-explicit extensions, though, is much easier to apply.

Thus, we will reach our goal if we prove that every generalized Liouville extension of a given field can be embedded in a generalized log-explicit extension of the same field. That is our next lemma.

**Lemma 14.**

Let $F_0$ be a differential field. Then for every generalized Liouville extension $F$ of $F_0$, there exist $v_1, \ldots, v_m \in F$ and a generalized log-explicit extension $G$ of $F_0$ such that:

i) $F(\log v_1, \ldots, \log v_m) = \overline{F}$ is a regular Liouville extension of $F$.

ii) $\overline{F}$ and $G$ are differentially isomorphic, and the isomorphism holds $F_0$ fixed.

**Proof:** Let $F_j = F_0(\Theta_1, \ldots, \Theta_j)$, $(1 \leq j \leq n)$, $F = F_n$. The proof will be in induction on $n$, and since $F_0$ is a log-explicit extension of $F_0$ we only have to do the induction step.

Assume then, that we are given $v_1, \ldots, v_m \in F_{n-1}$ and a log-explicit extension $G_{n-1}$ such that
i) \( F_{n-1}(\log v_1, \ldots, \log v_m) = \overline{F}_{n-1} \) is a regular Liouville extension of \( F \).

ii) \( \overline{F}_{n-1} \) and \( G_{n-1} \) are differentially isomorphic. Let \( e \) from \( \overline{F}_{n-1} \) onto \( G_{n-1} \) be this isomorphism. Depending on \( \Theta_n \), we have 5 cases:

(a) \( \Theta_n \) is algebraic over \( \overline{F}_{n-1} \). Let

\[
P(x) = \sum_{i=0}^{p} a_i x^i
\]

be the irreducible polynomial for \( \Theta_n \) over \( \overline{F}_{n-1} \).

Then, by lemma 1, \( \Theta_n \) is algebraic over \( \overline{F}_{n-1} \) of degree \( p \), so that

\[
e(P(x)) = \sum_{i=0}^{p} e(a_i) x^i = Q(x)
\]

is an irreducible polynomial over \( G_{n-1} \), and since \( \overline{F}_{n-1} = F_n(\log v_1, \ldots, \log v_m) \) is a regular Liouville extension of \( F_n \), \( G_n = G_{n-1}(\phi_n) \) is isomorphic to \( \overline{F}_{n-1} \), where \( \phi_n \) is a root of \( Q(x) \).

(b) \( \Theta_n = \exp(u) \), \( u \in F_{n-1} \). Then, by an inductive extension of corollary 9, \( \Theta_n \) is a regular monomial over \( \overline{F}_{n-1} \), and thus \( \overline{F}_{n-1} = F_n(\log v_1, \ldots, \log v_m) \) is a regular Liouville extension of \( F_n \). Finally, \( \phi_n = \exp(e(u)) \) is a regular monomial over \( G_{n-1} \), and \( G_n = G_{n-1}(\phi_n) \) is isomorphic to \( \overline{F}_{n-1} \).

(c) \( \phi_n \) is primitive, non-simple, \( D \Theta_n = u \in F_{n-1} \) for any derivation \( D \) of \( F \). Then \( \Theta_n \) cannot be simple over \( \overline{F}_{n-1} \), so it is still a regular monomial over \( \overline{F}_{n-1} \), thus \( \overline{F}_{n-1} = F_n(\log v_1, \ldots, \log v_m) \) is a regular
Liouville extension of \( F_n \), and so, if for any derivation \( D \) of \( F \), \( D(\phi_n) = e(u_D) \), we can define \( G_n = G_{n-1}(\phi_n) \) as a generalized log-explicit extension of \( F_0 \) and isomorphic to \( F_n \).

(d) \( \Theta_n \) is primitive, simple-logarithmic over \( F_{n-1} \), but not a regular monomial over \( F_{n-1} \). Then, we obtain

\[
\Theta_n = w + \sum_{i=1}^{m} c_i \log v_i + k.
\]

Let \( \ell \) be the smallest \( i \) such that \( c_\ell \neq 0 \). Then \( \overline{F}_n = F_n(\log(v_1, \ldots, \log(v_\ell-1), \log(v_{\ell+1}), \ldots, \log(v_m)) \) is a regular Liouville extension of \( F_n \), and is isomorphic to \( \overline{F}_{n-1} \) and thus isomorphic to \( G_n = G_{n-1} \).

(e) \( \Theta_n \) is simple-logarithmic over \( F_{n-1} \) and a regular monomial over \( F_{n-1} \). Then by the weak Liouville theorem [RIS 69], we can find \( w, u_1, \ldots, u_\ell \in F_{n-1} \) such that for any derivation \( D \) of \( F \)

\[
D \Theta_n = Dw + \sum_{i=1}^{\ell} c_n \frac{Du_i}{u_i}
\]

where the \( c_i \in C_{n-1} \).

Now let \( s_0 = \{v_1, \ldots, v_m\} \), and define (inductively) for \( i = 1, 2, \ldots, \ell \):

\[
s_i = \begin{cases} 
  s_{i-1} & \text{if } \log u_i \text{ is not a regular monomial over } F_{n-1}(s_{i-1}) \\
  s_{i-1} \cup \{u_i\} & \text{otherwise}
\end{cases}
\]
let \( s_{p1} = \{v_1, \ldots, v_p\} \). Then \( \hat{F}_{j-1} = F_{n-1} I \log v_1, \ldots, \log v_p \) is a regular Liouville extension of \( F_{j-1} \) and is differentially isomorphic to

\[
\hat{G}_{n-1} = G_{n-1}(\log(e(v_{m+1})), \ldots, \log(e(v_p)))
\]

which is log-explicit. Since \( \theta_n \) is no longer a regular monomial over \( \hat{F}_{n-1} \), we can proceed as in case (d).

Finally, the fact that the isomorphism between \( \hat{F} \) and \( G \) holds \( F_0 \) fixed, follows from the way the isomorphism is constructed in the lemma.

Something more can be said about the way the isomorphism \( \sigma \) between \( F \) and \( G \) (in lemma 15) operates.

Let \( G = F_0(\theta_1, \ldots, \theta_p) \), \( F = F_0(\theta_1, \ldots, \theta_q) \), \( G_i = F_0(\phi_1, \ldots, \phi_i) \), \( F_i = F_0(\theta_1, \ldots, \theta_i) \). If \( \theta_i = \exp(u_i) \), \( u_i \in F_{i-1} \), then \( \sigma(\theta_i) = \exp(\sigma(u_i)) = \phi_j \) with \( \sigma(u_i) \in G_{j-1} \). If \( \theta_i \) is primitive over \( F_{i-1} \),

\[
\sigma(\theta_i) = \sum_{j=1}^{k} c_j \phi_j + v
\]

where \( c_j \in C^F \), and \( v \in G_{k-1}' \).

As promised, theorem 12 has an analogue, which is:

**Theorem 15.**

Let \( F \) be a generalized Liouville field \( F = F_n = K(z_1, \ldots, z_m, \theta_1, \ldots, \theta_n) \). Let \( u \in F \), and assume \( \exp(u) \) is not a regular monomial over \( F \). Let \( \theta_i = \exp(a_i) \) for \( i \in E_n \). Then there exist constants \( c_i \), with \( c_i \) rational if \( i \in E_n \), otherwise \( c_i \in C^F \), such that

\[
u = \sum_{i \in p_n} c_i \theta_i + \sum_{i \in E_n} c_i a_i + v\]
with

\[ \Sigma_{i \in \mathbb{P}_n} c_i \Theta_i + v = \log(t) \text{ for some } t, v \in F. \]

The proof is an immediate consequence of lemma 14, theorem 12 and the preceding note.

Unfortunately, nothing more can be said about the form of \( u \) above, since if \( \log f \) is a regular monomial over the (univariate) differential field \( F \), then for any \( g \in F \), \( h = \int (\frac{f'}{f} + g') \) is a regular monomial over \( F \), but \( \exp(h - g) \) is not a regular monomial over \( F(h) \).

5. Examples.

We conclude this chapter with some examples; to fix notation, let \( Q \) be the set of rational numbers.

Example 1: What is the transcendence degree of the smallest generalized Liouville field containing \( Q \), the variable \( z \), and the functions

\[ e^z^2, \int_0^e t^2 \text{ dt, } e^0 \text{ over } Q? \]

Answer: Let \( \Theta_1 = e^z^2 \), \( \Theta_2 = \int_0^e t^2 \text{ dt, } \Theta_3 = e^0 \).

Then \( \Theta_1 \) is a regular monomial over \( Q(z) \), and it is well-known that \( \Theta_2 \) is not elementary over \( Q(z, \Theta_1) \) (see Example 2.9.3, for a proof), so that \( \Theta_2 \) cannot be simple-logarithmic over \( Q(z, \Theta_1) \). By theorem 12, \( \Theta_3 \) must be a regular monomial over \( Q(z, \Theta_1, \Theta_2) \). Thus, the answer is 4 and no simplification is possible.
Example 2: Investigate the transcendence degree of the smallest generalized Liouville field containing \( Q \), the variable \( z \) and the functions

\[
\theta_1 = \int \frac{z}{t^2 - 2} \, dt \,, \quad \theta_2 = \exp(\theta_1) \text{ over } Q.
\]

In this case, \( \theta_1 \) is non-simple over \( Q(z) \). (otherwise \( \sqrt{2} \) would have to be rational, since by section 2.8, we need the splitting field of

\[
R = \text{resultant}(1 - \alpha(2t), t^2 - 2, t)
\]

to express the integral, and

\[
R = (1 - 8 \alpha^2),
\]

and thus, by theorem 12, \( \theta_2 \) is a regular monomial over \( Q(z, \theta_1) \).

Therefore, the answer is 3.

Example 3: Investigate the transcendence degree of the smallest generalized Liouville field containing \( Q, \sqrt{2} \), the variable \( z \), and the functions

\[
\theta_1 = \int \frac{z}{t^2 - 2} \, dt \,, \quad \theta_2 = \exp(\theta_1) \text{ over } Q.
\]

This time

\[
\theta_1 = \frac{\sqrt{2}}{4} \log \frac{z - \sqrt{2}}{z + \sqrt{2}}
\]

is simple over \( Q(\sqrt{2}, z) \). Still, \( \theta_2 \) is a regular monomial over

\[
Q(\sqrt{2}, z, \log \frac{z - \sqrt{2}}{z + \sqrt{2}})
\]

since otherwise, by theorem 12, \( \frac{\sqrt{2}}{4} \) would have to be rational.
Chapter 4

CONCLUSIONS AND FUTURE WORK

We have obtained a collection of algorithms which find a canonical form for expressions built-up from the rational functions using exponentiation, integration and arithmetic operations.

In the process, we have also discovered a new integration algorithm which might be faster than hitherto known algorithms. In particular, for the rational function case, if no algebraic extension is required to express the integral, no factorization is required, thereby obtaining a theoretical computing time bound which is a polynomial in the degree and the logarithm of the sum of the coefficients of the input rational function. This result is in sharp contrast to previous algorithms which required a factoring algorithm which has a worst case computing time bound of \( \min(2^r, r^\mu) \cdot (\text{other terms}) \) where \( r \) is less than or equal to the number of factors we are seeking and \( \mu \) is proportional to the degree of the largest factor. (See [MUS 71] for a more precise statement of this computing time bound.)

It remains to be seen for which kinds of inputs the new algorithm is empirically faster, and whether their frequency is sufficiently large to justify large scale development.

Other future projects lie in the direction of providing a general integration algorithm for exponential and primitive functions in terms of a class of primitive functions obtainable from an arbitrary finite set by evaluation. For example if
\[ Ei(t) = \int_{1}^{t} \frac{e^{t}}{t} \, dt , \]

then

\[ \int_{e}^{t} e^{x} \, dx = Ei(e^{t}) . \]

The question we are addressing here is one which would require the system to detect that

\[ \int_{e}^{t} e^{x} \, dx = Ei(e^{t}) \]

from the definition of \( Ei(t) \) directly. So far, no Liouville-type result exists for these functions, and it is not even known what conditions are sufficient for one integral not to be expressible in terms of another, and whether, either one would be capable of expressing a third, even if neither one can be expressed in terms of the other.
APPENDIX A

In this appendix we give the necessary modifications to convert an algorithm for elementary integration into an algorithm for simple logarithmic integration.

By simple logarithmic integration we mean testing whether, given an \( f \) in a regular log-explicit field \( F_n \), there exist \( v_1, \ldots, v_n \) in \( F_n \) and a constant \( c \) such that

\[
\int f + c \in F_n (\log v_1, \ldots, \log v_m)
\]

i.e. to test whether \( \int f \) is simple-logarithmic over \( F \). (Ideally, one would like to have such an algorithm for generalized log-explicit fields, but, even though such an algorithm exists, it is not known at present, whether it is practical [MOS 72].)

Clearly, this problem is strongly related to the problem of elementary integration, so we take the approach of modifying an existing algorithm (the one described in Chapter 2) rather than describing a brand new one.

Notice that most of the description of the algorithm is still valid if we write \( \Theta_n \) instead of \( \log u \), and \( \Theta' \) (or \( u \)) instead of \( u'/u \). The main inconsistencies arise when setting up the problem

\[
\int f = g + k \log h
\]

where \( f \) and \( h \) are known, and we are seeking either \( g \) or \( k \) (or both).
Then translating this problem into our present situation, we obtain

\[ \int f = g + k \Theta_n \]

or

\[ \int (f - k \Theta_n^* ) = \int (f - k u) = g \]

where \( g \) is logarithmic.

The problem, in this form, can be solved as illustrated in the second part of Section 2.6 with the necessary modifications shown above.

The other inconsistency arises when we want (recursively) to integrate

\[ a_0 - b_1 u \]

since, in general, this integral will not be simple logarithmic over \( F_{n-1} \) (and probably not even elementary!).

The solution here, is similar to the one used above. We integrate

\[ a_0 - b_1 u - k_1 u \]

with the constant \( k_1 \) still to be determined.

This leads to the problem of determining, given \( f, g_1, \ldots, g_k \), whether there exist constants \( c_1, \ldots, c_k \) such that

\[ \int (f + \sum_{i=1}^{k} c_i g_i) \]

is simple-logarithmic over the differential field \( F_n \).
As in section 2.3, we first decompose \( f = g_0 \) and the \( g_i \) \((1 \leq i \leq k)\) so that 
\[
g_i = p_i + r_i \quad (0 < i < k)
\]
with \( p_i \in D_n \), \( r_i \) proper in \( F_n \), and then, recalling that for elements 
of \( D_n \), if \( f t \) is elementary, then \( f t \in D_n \), we note that the modifi-
cations mentioned for sections 2.4, 2.5 and 2.6 are also applicable
for this problem. We therefore only have to modify the algorithms
in sections 2.7 and 2.8.

Section 2.7 can be generalized in a similar way to the other steps
of the algorithm. Apply it to each of the \( r_i \) \((0 < i < k)\). The generaliz-
ation of Section 2.8 is somewhat harder. A variety of procedures
suggest themselves; one that seems to be preferable consists of
applying corollary 2.8.1 in the following way.

Given \( f, g_1, \ldots, g_k \), normal in \( F_n \), we let

\[
f + \sum_{i=1}^{k} c_i g_i = \frac{p + \sum c_i p_i}{q}
\]

for suitable \( p, p_i, q \) in \( S_n \). Notice that for any \( c_1, \ldots, c_k \),

\[
\frac{p + \sum c_i p_i}{q}
\]

is normal in \( F_n \). We first test whether there exist constants \( c_1, \ldots, c \),
such that

\[
q'|(p + \sum_{i=1}^{k} c_i p_i)
\]
or

\[ \bar{P} + \sum_{i=1}^{k} c_i \bar{p}_i = 0 \]

where \( \bar{P} = \text{rem}(P, Q') \),

\[ \bar{p}_i = \text{rem}(p_i, Q'). \]

If there do exist such \( c_1, \ldots, c_k \), and

\[ P + \sum_{i=1}^{k} c_i p_i \]

\[ \frac{Q'}{Q'} \]

is a constant, we are done, otherwise, we compute

\[ \text{resultant}(P + \sum_{i=1}^{k} c_i p_i - aQ', Q, \theta_n) \]

obtaining

\[ \sum_{i=1}^{\ell} a_i(c_1, \ldots, c_k) \alpha^i \]

where \( \ell = \deg_{\theta_n} Q \), and

\[ a_i(c_1, \ldots, c_k) \in F_{n-1}(c_1, \ldots, c_k), \quad 0 \leq i \leq \ell. \]
Notice that \( a_2(c_1, \ldots, c_k) \) cannot be 0 for any constant values of \( c_1, \ldots, c_k \), since we would then have
\[
P + \sum_{i=1}^{k} c_i P_i - \alpha Q' = 0
\]
for constants \( c_1, \ldots, c_k, \alpha \), a possibility which we ruled out above.

Let
\[
b_i(c_1, \ldots, c_k) = \frac{a_{i-1}}{a_{\lambda}}, \quad 1 \leq i \leq \lambda.
\]

By corollary 2.8.1, we must have that the \( b_i(c_1, \ldots, c_k) \) are constants so that \( b_i(c_1, \ldots, c_k) = 0 \), for \( 1 \leq i \leq \lambda \).

Using multiple derivatives (if necessary) of the \( b_i \), the equations obtained during previous steps of the algorithm, and equating like powers of \( z \) and the \( \Theta_i \), \( 1 \leq i \leq n \), we can obtain sufficient equations to solve for all the \( c_i \), especially since we know that if there is a solution, it must be unique. Notice that, since we are only interested in solutions in \( K \), the constant field of \( F_n \), we can stop the computations if we find out that there are no solutions in \( K \).

Once we know all the \( c_i \), we can proceed in the same way we did before, to obtain our final answer.
APPENDIX B

FINDING RATIONAL ROOTS OF POLYNOMIALS EXACTLY

In this appendix, we discuss a simple modification due to Rubald [RUB 74] of a procedure due to Heindel [HEI 71] (which computes intervals containing roots of a polynomial over the integers which are smaller than a predetermined error bound) to compute the rational roots of a polynomial exactly. The modification is based on the following well-known result:

**Lemma.**

Let

\[ P(x) = \sum_{i=0}^{n} a_i x^i \]

be a primitive polynomial in \( x \) over the integers, and assume that \( P\left( \frac{a}{b} \right) = 0 \) for some rational integers \( a, b \) which have no common factors. Then \( \frac{a}{b} \) is an integer.

**Proof:** As can be seen from the proof of the main theorem in [VdW 70, §30], since \( X - \frac{a}{b} \) divides \( P(x) \) as polynomials with rational coefficients, \( bX - a \) divides \( P(X) \) as integer polynomials, and our conclusion follows.

We apply this lemma in the following manner. Let

\[ P(X) = \sum_{i=0}^{n} a_i X^i \]
be a primitive integral polynomial whose rational roots are to be determined. We can assume (without loss of generality) that $a_n > 0$, and also (by a construction used in Heindel's algorithm) that $P(X)$ is square-free, so it has only simple roots. We then apply the Collins-Akritas algorithm [CoA 76], followed by Heindel's root refining procedure, if necessary, to obtain $c_1, \ldots, c_m, d_1, \ldots, d_m$ ($m \leq n$) such that

$$c_i < d_i, \quad d_i - c_i < \frac{1}{a_n}$$

and if $\alpha$ is any root of $p$, there exists an $i_{\alpha}$ such that

$$c_{i_{\alpha}} < \alpha \leq d_{i_{\alpha}}.$$

Let $\alpha_i = \frac{p_i}{q_i}$ is a rational number, with $q_i > 0$, $\gcd(p_i, q_i) = 1$, and let

$$r_i = \frac{a_n}{q_i},$$

which is an integer by the lemma.

Then $a_n \alpha_i = r_i p_i$ is an integer satisfying $a_n c_i < r_i p_i \leq a_n d_i$, and since $a_n d_i - a_n c_i < 1$, there can be at most one integer between $a_n c_i$ and $a_n d_i$ so the modification now amounts to checking, for each pair $(c_i, d_i)$ whether there is an integer $m_i$ such that $a_n c_i < m_i \leq a_n d_i$ and, if so, checking whether $\frac{m_i}{a_n}$ is a root of $P(X)$.

We proceed to give a computing time bound due to Collins [COL 76], for this algorithm. Assume that our input is a polynomial $A$ with

$$\deg A \leq m,$$

and assume that $A$ has $n$ distinct roots, and that all the roots of $A$ are real. We first compute $B_1, \ldots, B_d$ such that the $B_i$ are pair-wise relatively prime, each $B_i$ is square-free, and
\[
\prod_{i=1}^{\pi} B_i = A
\]

Then, the computing time to find the \( B_i \) is given by

\[
t_{SQFREE} < m n^2 \log^2(m) + \frac{2}{n} \log(n - n + 1)^2 \log^2 d
\]

(from [HEI 71, theorem 3.2]). If we now let \( n_i = \deg B_i, \overline{d}_i \) be the norm \( B_i \) then, from [MIG 74] we deduce \( \overline{d}_i \leq 2^{n_i} d \).

We now use the algorithm described in [CoA 76] to isolate the roots of \( B \) in time \( n_i^6 \log^2(\overline{d}_i) \).

We point out that since all the intervals are the result of repeatedly bisecting an interval of length \( 2^k \) for some \( k \), this means that all our intervals will of the form \( \left[ \frac{a}{2^i}, \frac{b}{2^j} \right] \) where \( a, b, i, j \) are integers, \( i, j \) are non-negative. Also, if \( i > 0, j > 0 \), then

\[
\frac{b}{2^j} - \frac{a}{2^i} < 2^{-\max(i, j)}
\]

Thus, if we obtain an interval, one of whose end point has a denominator greater than \( 2^k \), there is no need to refine it, and so, we may assume that whenever we refine an interval

\[
\left( \frac{a}{2^i}, \frac{b}{2^j} \right), \quad \max(|a|, |b|, 2^i, 2^j) \leq \overline{d}_i^2
\]

(since \( |b/2^i| \leq \overline{d}_i \) and \( 2^i \leq \overline{d}_i \)). We also know that the maximum length of any interval is \( 2^k \leq \overline{d}_i \) and that \( 1/\varepsilon \leq \frac{1}{|a_n|} \leq \overline{d}_i \), so that, applying theorem 7.1 in [HEI 71] we obtain

\[
n_i^2 \log^3(\overline{d}_i)
\]
as the time to refine any interval. Since there can be as many as \( n_i \) intervals to refine, we obtain a computing time bound for the refining stage of the algorithm, as

\[ n_i^3 L^3(\overline{d}_i) . \]

Finally, the time to do the necessary evaluations is bounded by

\[ n_i^3 L^2(\overline{d}_i) . \]

If we add these computing times together, we obtain

\[ t_i \leq n_i^6 L^2(\overline{d}_i) + n_i^3 L^3(\overline{d}_i) \]

for the time to isolate and refine the roots, plus also the time to determine whether each isolated root is rational. Since \( \overline{d}_i \leq 2^{n_i} L(d_i) \), we obtain

\[ t_i \leq n_i^8 + n_i^6 L^2(d) + n_i^3 L^3(d) . \]

If we add these computing times over \( i \), we obtain a total computing time of

\[ t \leq t_{\text{SQFREE}} + n^8 + n^6 L^2(d) + n^3 L^3(d) \]

(since, for any \( k \),

\[ \sum_{i=1}^{L} n_i^k \leq n^k \]

because

...
Thus, we obtain a total computing time of

\[ t \leq m n^2 L^2(m \cdot d) + n^2(m - n + 1)^2 L^2(d) + n^8 + n^6 L^2(d) \]

\[ + n^3 L^3(d) \leq m^8 + m^6 L^2(d) + m^3 L^3(d) \]

since \( n \leq m \).


COL 76] __________. Private communication.


[LIQ 33a] Liouville, J. Mémoire sur les transcendants elliptiques de première et de seconde espèce, considérées comme fonctions de leur amplitude, Journal de l'École polytechnique, XIV (1833), Section 24, pp. 57-83.


[TRA 76] Trager, Barry M. Algebraic Factoring and Rational Function Integration, in [SYMSAC 76], pp. 219-226.
