4-27-2004

Factorization of polynomials and real analytic function

Radoslaw L. Stefanski
University of Richmond

Math
Ste

Factorization of Polynomials and Real Analytic Functions

Radosław L. Stefański
Honors thesis[1]
Department of Mathematics & Computer Science
University of Richmond

April 27, 2004

[1] Under the direction of Dr. William T. Ross

Abstract

In this project, we address the question: When can a polynomial $p(x,y)$ of two variables be factored as $p(x,y) = f(x)g(y)$, where f and g are polynomials of one variable. We answer this question, using linear algebra, and create a *Mathematica* program which carries out this factorization. For example,

$$3+3x-5x^3+y+xy-\frac{5}{3}x^3y+y^2+xy^2-\frac{5}{3}x^3y^2 = (1+x-\frac{5}{3}x^3)(3+y+y^2).$$

We then generalize this concept and ask: When can $p(x,y)$ be written as

$$p(x,y) = f_1(x)g_2(y) + f_2(x)g_2(y) + \cdots + f_r(x)g_r(y),$$

where f_j, g_j are polynomials? This can certainly be done (for large enough r). What is the minimum such r? Again, we have a *Mathematica* program which carries out this computation. For example,

$$1+2x+x^2+2x^3+2y+2x^2y+7xy^2+7x^3y^2 = (1+x^2)(1+2y)+(x+x^3)(2+7y^2).$$

We generalize this further to larger number of variables (with an appropriate *Mathematica* program to carry out this computation). We then apply this and consider the domains of convergence of certain types of real analytic functions and try to relate the domain of convergence with the rank of the polynomial.

This paper is part of the requirements for honors in mathematics. The signatures below, by the advisor, a departmental reader, and a representative of the departmental honors committee, demonstrate that Radosław L. Stefański has met all the requirements needed to receive honors in mathematics.

(advisor)

(reader)

(honors committee representative)

Contents

1. Introduction

If $\mathbb{R}[x]$ is the set of all polynomials with real coefficients, then the standard factorization of a polynomial $p \in \mathbb{R}[x]$, is derived from a theorem in abstract algebra[1]. This theorem states that $p \in \mathbb{R}[x]$ can be written as,

$$p(x) = v_1(x)^{s_1} \cdots v_n(x)^{s_n},$$

where $s_i \in \mathbb{N}$ and $v_i(x), \cdots, v_n(x)$ are irreducible (i.e., they cannot be non-trivially factored further as products of elements of $\mathbb{R}[x]$).

Moreover, if $\mathbb{R}[x, y]$ is the set of all polynomials in x and y with real coefficients, a similar theorem from abstract algebra states that every polynomial, $p \in \mathbb{R}[x, y]$, can be written as,

$$p(x, y) = v_1(x, y)^{s_1} \cdots v_n(x, y)^{s_n},$$

where $s_i \in \mathbb{N}$ and $v_i(x, y), \cdots, v_n(x, y)$ are irreducible in the same sense as above.

We are interested in a different type of factorization of elements in $\mathbb{R}[x, y]$. If $p \in \mathbb{R}[x, y]$, we wish to write

$$p(x, y) = f(x)g(y),$$

where $f \in \mathbb{R}[x]$ and $g \in \mathbb{R}[y]$ but f and g are *not* necessarily irreducible in $\mathbb{R}[x]$ and $\mathbb{R}[y]$ respectively (we are not concerned about further factorization of f and g).

Although unconventional, this form of factorization has been seen before. The separable (but not necessarily polynomial) solutions to certain differential or partial differential equations such as the Laplace equation,

$$\Delta u = 0$$

or the wave equation,

$$u_{xx} - u_{yy} = 0,$$

are exactly of this form

$$u(x, y) = f(x)g(y).$$

Another factorization problem appears in Fourier analysis in the following form: Suppose that u a is bounded function on the unit circle with a complex Fourier series,

$$u \sim \sum_{n=-\infty}^{\infty} \widehat{u}(n)e^{in\theta},$$

where $\widehat{u}(n) = \frac{1}{2\pi}\int_0^{2\pi} u(e^{i\theta})e^{-in\theta}d\theta$ are the Fourier coefficients of u. When can we write

$$u = fg,$$

where f and g are bounded on the circle, and

$$f \sim \sum_{n=0}^{\infty} \widehat{f}(n)e^{in\theta}$$

and

$$g \sim \sum_{n=0}^{\infty} \widehat{g}(-n)e^{-in\theta}$$

[1]Gallian, J., *Contemporary Abstract Algebra*, 1998, Houghton Mifflin College Division, Boston.

are bounded functions? A theorem of Bourgain[2] states that this is possible if and only if

$$\int_0^{\pi} \log|f(e^{i\theta})| d\theta > -\infty.$$

The above polynomial factorization can be generalized to n dimensions. If $p \in \mathbb{R}[x_1, \cdots, x_n]$, we wish to write

$$p(x_1, \cdots, x_n) = f_1(x_1) \cdots f_n(x_n),$$

where $f_i \in \mathbb{R}[x_i]$ but the f_i's need *not* be necessarily irreducible in $\mathbb{R}[x_i]$. We formulate an algorithm to factor such polynomials and create a *Mathematica* program to carry out the computations.

We generalize the above factorization method by assuming $p \in \mathbb{R}_\infty[x, y]$, by which we mean the set of formal power series,

$$p(x, y) = \sum_{n,m=0}^{\infty} a_{n,m} x^n y^m.$$

After reviewing some basic notions about the domains of convergence of the formal power series, we will focus on the following problem. For a given $p \in \mathbb{R}_\infty[x, y]$, when can we write

$$p(x, y) = p_1(x) q_1(y),$$

where $p_1 \in \mathbb{R}_\infty[x]$ and $q_1 \in \mathbb{R}_\infty[y]$? We also generalize to factorizations of the form

$$p(x, y) = \sum_{j=1}^{r} p_j(x) q_j(y),$$

where $p_j \in \mathbb{R}_\infty[x]$ and $q_j \in \mathbb{R}_\infty[y]$. Such functions are called functions of *finite rank.* We then go on to determine the domain of convergence of such functions and explore the relationship between the rank and the domain of convergence.

[2]Bourgain, J, *A Problem of Douglas and Rudin*, Pacific Journal of Mathematics, (1986), 121, pp. 47-50.

2. Factorization of polynomials in two dimensions

We begin our examination of factorizing polynomials with some basic definitions.

Definition 2.1. The *degree* of a polynomial, $p(x,y) = a_{0,0} + ... + a_{m,n}x^m y^n$ is the highest power of x or y with non-zero coefficients.

Example 2.2.

(1) $\deg(2 + x + x^2) = 2$
(2) $\deg(2 + x^2 y + xy) = 2$
(3) $\deg(x^6 y^2 + y^4 x^5) = 6$

The notation in Definition 2.1 is not quite standard since others define the degree to be the highest number $m + n$. So for example, others define the degree of $p(x,y) = xy$ as 2, while we define it to be one. This non-standard notation will make things easier for us later on.

Definition 2.3. We say a polynomial $p(x,y)$ is of *rank one* if and only if $p(x,y) = p_1(x)p_2(y)$, where p_1, p_2 are polynomials.

Definition 2.4. For a polynomial

$$p(x,y) = \sum_{0 \le i,j \le n} a_{i,j} x^i y^j$$

of degree n we define the *coefficient matrix* $C(p)$ to be the $(n+1) \times (n+1)$ matrix whose i, j^{th} entry is $a_{i,j}$. Hence,

$$C(p) := \begin{pmatrix} a_{0,0} & \cdots & a_{0,n} \\ \vdots & \vdots & \vdots \\ a_{n,0} & \cdots & a_{n,n} \end{pmatrix}.$$

Recall from basic linear algebra, that the set of linear combinations of the column vectors of a $n \times m$ matrix A, is called the *column space* of A. Similarly, the set of all linear combinations of the row vectors is called the *row space* of A. The *rank* of a matrix, is the dimension of the column or row spaces, which, by a well known fact from linear algebra are the same[3].

Theorem 2.5. *A polynomial $p(x,y)$ is of rank one if and only if $C(p)$ is of rank one.*

Proof. Suppose, $p(x,y) = p_1(x)p_2(y)$. Then, since the function p is of degree n, the functions p_1 and p_2 must each be of at most degree n. Hence, $p_1 = \sum_{i=0}^{n} b_i x^i$ and $p_2 = \sum_{j=0}^{n} c_j y^j$ for some constants b_i and c_j where $i, j \in \{0, 1, \cdots, n\}$. Hence, if this factorization is possible, we write:

[3]Strong, G., *Introduction to Linear Algebra*, 1998, Wellesley Cambridge, Boston.

$$\begin{aligned} p(x,y) &= \sum_{0\le i,j\le n}' a_{i,j}x^iy^j \\ &\doteq p_1(x)p_2(y) \\ &= (\sum_{i=0}^{n} b_ix^i)(\sum_{j=0}^{n} c_jy^j) \\ &= \sum_{0\le i,j\le n} b_ic_jx^iy^j \end{aligned}$$

For $\doteq$ to be satisfied and hence for factorization to be possible, it is clear that we must set the coefficients of x and y in $p(x,y)$ equal to the coefficients of $p_1(x)p_2(y)$. Hence,

$$\begin{aligned} a_{0,0} &= b_0c_0 \\ a_{1,0} &= b_1c_0 \\ \dots & \quad \dots \\ a_{n,n} &= b_nc_n \end{aligned}$$

Or, simplifying this notation with matrices we can write:

$$\begin{pmatrix} a_{0,0} & \cdots & a_{0,n} \\ \vdots & \vdots & \vdots \\ a_{n,0} & \cdots & a_{n,n} \end{pmatrix} = \begin{pmatrix} b_0c_0 & \cdots & b_0c_n \\ \vdots & \vdots & \vdots \\ b_nc_0 & \cdots & b_nc_n \end{pmatrix}$$

Notice however, that each subsequent column in

$$\begin{pmatrix} b_0c_0 & \cdots & b_0c_n \\ \vdots & \vdots & \vdots \\ b_nc_0 & \cdots & b_nc_n \end{pmatrix}.$$

is a multiple of the first column; that is, the matrix is of rank one.
The above argument can be reversed and so we can conclude that p is of rank one exactly when the corresponding *coefficient matrix*, $C(p)$, is of rank one. □

Let us now try a simple example to demonstrate this theorem.

Example 2.6. Factor the polynomial,

$$p(x,y) = 8 + 12y + 16y^2 - 4x - 6xy - 8xy^2 + 6x^2 + 9x^2y + 12x^2y^2$$

as $p(x,y) = p_1(x)p_2(y)$, if possible.

First, let us write down the coefficient matrix of the above polynomial.

$$C = \begin{pmatrix} 8 & 12 & 16 \\ -4 & -6 & -8 \\ 6 & 9 & 12 \end{pmatrix}$$

(Notice that $deg(p) = 2$, and so $C(p)$ is a 3×3 matrix). Next, we column reduce the matrix C to,

$$C_{cr} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 \\ \frac{3}{4} & 0 & 0 \end{pmatrix}$$

But this means that C is a rank one matrix. So, according to the previous theorem we can factor the polynomial $p(x, y)$.

Notice, that since the matrix C is of rank one, each column in that matrix can be written as a product of a constant and the basis column vector, in the following fashion:

$$\begin{pmatrix} 8 & 12 & 16 \\ -4 & -6 & -8 \\ 6 & 9 & 12 \end{pmatrix} = \begin{pmatrix} 8(1) & 12(1) & 16(1) \\ 8(-\frac{1}{2}) & 12(-\frac{1}{2}) & 16(-\frac{1}{2}) \\ 8(\frac{3}{4}) & 12(\frac{3}{4}) & 16(\frac{3}{4}) \end{pmatrix}$$

But we have already seen this in the previous theorem:

$$\begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} \\ a_{1,0} & a_{1,1} & a_{1,2} \\ a_{2,0} & a_{2,1} & a_{2,2} \end{pmatrix} = \begin{pmatrix} b_0c_0 & b_0c_1 & b_0c_2 \\ b_1c_0 & b_1c_1 & b_1c_2 \\ b_2c_0 & b_2c_1 & b_2c_2 \end{pmatrix}$$

We can thus use this to solve for for the b_i's and c_i's This means that:

$$b_0 = 1 \quad c_0 = 8$$

$$b_1 = -\frac{1}{2} \quad c_1 = 12$$

$$b_2 = \frac{3}{4} \quad c_2 = 16$$

Finally using these values we can come up with the factorized expression,

$$\begin{aligned} p(x,y) &= 8 + 12y + 16y^2 - 4x - 6xy - 8y^2 \\ &\quad +6x^2 + 9x^2 + 12x^2y^2 \\ &= (b_0 + b_1x + b_2x^2)(c_0 + c_1y + c_2y^2) \\ &= (1 - \frac{1}{2}x + \frac{3}{4}x^2)(8 + 12y + 16y^2) \end{aligned}$$

The following is an alternate characterization of rank one polynomials involving partial differential equations.

Theorem 2.7. *Suppose $p(x, y)$ has no zeros. Then p is rank one if and only if*

$$\partial_y \left(\frac{\partial_x p}{p} \right) = 0$$

Proof. ($\Rightarrow$): Suppose,

$$p(x, y) = p_1(x)p_2(y).$$

Then,

$$\frac{\partial_x p}{p} = \frac{p_1'(x)p_2(y)}{p_1(x)p_2(y)} = \frac{p_1'(x)}{p_1(x)}.$$

Hence substituting $p_1'(x)/p_1(x)$ for $\partial_x p/p$,

$$\partial_y \left(\frac{\partial_x p}{p}\right) = \partial_y \left(\frac{p_1'(x)}{p_1(x)}\right) = 0$$

and so

$$\partial_y \left(\frac{\partial_x p}{p}\right) = 0.$$

($\Leftarrow$): If

$$\partial_y \left(\frac{\partial_x p}{p}\right) = 0$$

Integrating the above equation in terms of y, we obtain:

$$\frac{\partial_x p}{p} = q(x)$$

for some function q. Next, we integrate the above expression in terms of x.

$$\int \frac{\partial_x p}{p} dx = \int q(x) dx$$

And so,

$$\ln p(x,y) = G(x) + H(y)$$

Where, $G'(x) = q(x)$.

Exponentiating both sides of this expression we get,

$$p(x,y) = e^{G(x)+H(y)}$$

Hence

$$p(x,y) = e^{G(x)}e^{H(y)}$$

But this is only,

$$p(x,y) = p_1(x)p_2(y)$$

where $p_1(x) = e^{G(x)}$ and $p_2(y) = e^{H(y)}$. Note that since $p(x,y)$ is a polynomial, then p_1 and p_2 are also polynomials. But this means that $p(x,y)$ is of rank 1. □

Example 2.8. Consider the polynomial $p(x,y) = -1 - 3x - x^2 + y + 3xy + x^2y + 2y^2 + 6xy^2 + 2x^2y^2$. Note that,

$$\frac{\partial_x p}{p} = \frac{-3 - 2x + 3y + 2xy + 6y^2 + 4xy^2}{-1 - 3x - x^2 + y + 3xy + x^2y + 2y^2 + 6xy^2 + 2x^2y^2}$$

After some calculation,

$$\begin{aligned}\frac{\partial_x p}{p} &= \frac{-3-2x+3y+2xy+6y^2+4xy^2}{-1-3x-x^2+y+3xy+x^2y+2y^2+6xy^2+2x^2y^2} \\ &= \frac{(3+2x)(1+y)(-1+2y)}{(1+3x+x^2)(1+y)(-1+2y)} \\ &= \frac{(3+2x)}{(1+3x+x^2)}\end{aligned}$$

So now, we see that $\partial_x p/p$ is a function only in terms of x. Hence,

$$\partial_y \left(\frac{\partial_x p}{p}\right) = 0$$

According to our theorem, $p(x,y)$ is of rank one.

Is it possible to say something about polynomials of rank not equal one? Is it possible to factor such polynomials and if so, in what way? In order to answer these questions, we define the general rank of a polynomial.

Definition 2.9. We say that a polynomial p is of *rank* r if and only if there exits an integer $r > 0$ such that,

$$p(x,y) = \sum_{i=1}^{r} f_i(x)g_i(y),$$

where f_i,g_i are polynomials, and if r is the smallest integer to satisfy this condition.

To see why we require r to be the *smallest* integer to satisfy the above condition, consider the following example.

Example 2.10. Let $p(x,y) = xy - xy$. Notice that if we do not require r to be minimum, the rank of p is 2. In fact, if we try to loosen the definition, the rank of p can be any positive integer just by adding $(xy - xy) = 0$ to p. The correct way to calculate the rank of p, is to notice that $p(x,y) = 0$ and hence that the rank of p is zero.

Notice also, that this definition of rank, extends and complements the earlier definition of a rank one polynomial.

Theorem 2.11. *A polynomial $p(x,y)$ is of rank r if and only if the corresponding coefficient matrix, $C(p)$, is of rank r.*

Proof. If $r = 1$,

$$p(x,y) = p_1(x)q_1(y)$$

This however is simply the rank one case, which we have proved above. Therefore, we assume that $r \geq 2$.

($\Leftarrow$): We assume that C, the coefficient matrix of $p(x,y)$, is of rank r. By the definition of the rank of a matrix, C must have r basis vectors - that is, each column of the matrix C, taken as a vector, is a linear combination of given r column vectors. Let us assume, that these r column vectors are:

$$\{u_1, u_1, \cdots, u_r\}.$$

From the definition of the rank of a matrix, we know that C can be written as:

$$C = (\ c_1^1 u_1 + \cdots + c_1^r u_r \ \mid \ \cdots \ \mid \ c_n^1 u_1 + \cdots + c_n^r u_r\ )$$

where each entry represents column in C and n indicates the number of columns in the coefficient matrix and the c_i's are real numbers.

Notice now that C can be separated in the following way

$$\begin{aligned} C &= (\ c_1^1 u_1 + \cdots + c_1^r u_r \ \mid \ \cdots \ \mid \ c_n^1 u_1 + \cdots + c_n^r u_r\ ) \\ &= (\ c_1^1 u_1 \ \mid \ \cdots \ \mid \ c_n^1 u_1\ ) + \cdots + (\ c_1^r u_r \ \mid \ \cdots \ \mid \ c_n^r u_r\ ) \\ &= \sum_{i=1}^{r} (\ c_1^i u_i \ \mid \ \cdots \ \mid \ c_n^i u_i\ ) \end{aligned}$$

Each $(\ c_1^i u_i \ \mid \ \cdots \ \mid \ c_n^i u_i\ )$ is a rank one matrix, since each column is a linear combination of the first column. Hence, we use each rank one matrix, $(\ c_1^i u_i \ \mid \ \cdots \ \mid \ c_n^i u_i\ )$, and Theorem 2.5 to conclude that that matrix corresponds to a polynomial that can be factored into the product of one variable polynomials, $p_i(x)q_i(y)$. But since C is a sum of r such matrices, we conclude that,

$$p(x,y) = \sum_{i=1}^{r} p_i(x)q_i(y).$$

Since r is minimal, then p is of rank r.

($\Rightarrow$): The proof in this direction is almost identical to the one presented above. We begin by assuming,

$$p(x,y) = \sum_{i=1}^{r} p_i(x)q_i(y).$$

Since p is a sum of r rank 1 polynomials, each of those polynomials can be represented by a rank one matrix. This means that the coefficient matrix C, can be written as a sum of r rank one matrices. We write,

$$\begin{aligned} C &= \sum_{i=1}^{r} (\ c_1^i u_i \ \mid \ \cdots \ \mid \ c_n^i u_i\ ) \\ &= (\ c_1^1 u_1 \ \mid \ \cdots \ \mid \ c_n^1 u_1\ ) + \cdots + (\ c_1^r u_r \ \mid \ \cdots \ \mid \ c_n^r u_r\ ) \\ &= (\ c_1^1 u_1 + \cdots + c_1^r u_r \ \mid \ \cdots \ \mid \ c_n^1 u_1 + \cdots + c_n^r u_r\ ), \end{aligned}$$

where the u_i's are simply basis vectors and c_i's are constants.
Each column of C is hence a linear combination of the r basis vectors,

$$\{u_1, u_1, \cdots, u_r\}.$$

Hence C is a rank r matrix. □

Let us now consider an example of the above theorem.

Example 2.12. Let $p(x,y) = 1 + y^2 + xy + xy^2$. Let's use the above theorem to

(1) find its coefficient matrix C
(2) find the basis vectors of C
(3) find the rank of $p(x,y)$
(4) find $p(x,y)$'s factorization.

So,

(1)
$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

(2) Notice, that C can be written as,
$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$
It is now easy to see, that the (column) basis vector of the first matrix is $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and that of the second matrix is $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$. Hence, the basis for the column space of C is, $\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\}$

(3) Since the basis contains two vectors, we know that the rank of C (and hence of p) is 2.

(4) To solve this, either use Theorem 2.5, or - due to the easiness of the example - simply regroup $p(x,y)$. So,
$$p(x,y) = (1+y^2) + x(y+y^2)$$

We leave this section with an open question. Recall that a non-zero polynomial, $p(x,y)$, is of rank one if and only if
$$\partial_y \left(\frac{\partial_x p}{p}\right) = 0.$$
Is there a partial differential equation that if satisfied, is necessary and sufficient for a polynomial to be of rank r?

3. An Introduction to Real Analytic Function

We now wish to generalize the above results for polynomials, $p(x, y)$, to real analytic functions

$$f(x,y) = \sum_{m,n=0}^{\infty} a_{m,n} x^m y^n.$$

Before doing this however, we need to address the issue of convergence of these expressions. We begin with some standard results of single variable real analytic functions[4].

The formal expression,

$$\sum_{j=0}^{\infty} a_j (x-\alpha)^j$$

with either real or complex constant a_j's, is called a *power series* on the real line $\mathbb{R}$. We usually take the coefficients a_j to be real and there is no loss of generality in doing so. Before much can be done with this function, it is necessary to determine the nature of the set on which the power series converges.

Proposition 3.1. *Assume that the power series*

$$\sum_{j=0}^{\infty} a_j (x-\alpha)^j$$

converges at the value $x = c$. *Let* $R = |c - \alpha|$. *Then the series converges uniformly and absolutely on compact subsets of* $I = \{x : |x - \alpha| < R\}$.

Proof. We may take the compact subset of I to be $K = [\alpha - s, \alpha + s]$ for some number $0 < s < r$. For $x \in K$ it then holds that

$$\sum_{j=0}^{\infty} \left|a_j(x-\alpha)^j\right| = \sum_{j=0}^{\infty} \left|a_j(c-\alpha)^j\right| \cdot \left|\frac{x-\alpha}{c-\alpha}\right|^j.$$

In the sum on the right, the first expression in absolute values is bounded by some constant C (by the convergence hypothesis). The quotient in absolute values is not bigger than $L = \frac{s}{r} < 1$. The series on the right is thus dominated by

$$\sum_{j=0}^{\infty} C \cdot L^j.$$

Hence, this geometric series converges. So, by the Weierstrass M-test[5], the original series converges absolutely and uniformly on K. □

The above theorem says that the domain of convergence of a power series must be an interval. This interval may be bounded, as in the power series $\sum x^n$, or unbounded, as in the power series $\sum x^n/n!$.

Definition 3.2. The set on which

$$\sum_{j=0}^{\infty} a_j (x-\alpha)^j$$

[4]Krantz, S. G. and Parks, H. R., *A primer of real analytic functions.* Second edition. Birkhuser Advanced Texts: Basel Textbooks, Boston.

[5]Gaughan, E., *Introduction to Analysis,* 1998, Brooks-Cole, New York.

converges is an interval centered about α. This interval is termed the *interval of convergence.* The series will converge absolutely and uniformly on compact subsets of the interval of convergence. The *radius of convergence* of the interval is defined to be half its length.

We remind the reader of the following useful theorem of Hadamard[6].

Theorem 3.3. *If R is the radius of convergence of*

$$\sum_{n=0}^{\infty} a_n(x-\alpha)^n,$$

then,

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{\frac{1}{n}}}.$$

Whether convergence holds at the end points of the interval of convergence needs to be determined on case by case basis.

Definition 3.4. A function f, where $f : U \subseteq \mathbb{R} \to \mathbb{R}$, is said to be *real analytic at* α if the function f may be expressed as a power series on some interval of positive radius centered at α:

$$f(x) = \sum_{j=0}^{\infty} a_j(x-\alpha)^j.$$

We say the function is *real analytic on* $V \subseteq U$ if it is real analytic at each $\alpha \in V$.

Next, we present without proof the basic properties of real analytic functions.

Proposition 3.5. *Let,*

$$f(x) = \sum_{j=0}^{\infty} a_j(x-\alpha)^j$$

and

$$g(x) = \sum_{j=0}^{\infty} b_j(x-\alpha)^j$$

be two power series defining the functions $f(x)$ and $g(x)$ on the open intervals of convergence C_1 and C_2 respectively. Then, on their common domain, $C = C_1 \bigcap C_2$, it holds that,

(1) $f(x) \pm g(x) = \sum_{j=0}^{\infty}(a_j \pm b_j)(x-\alpha)^j$
(2) $f(x) \cdot g(x) = \sum_{m=0}^{\infty} \sum_{j+k=m}(a_j \cdot b_k)(x-\alpha)^m$
(3) *If $g \neq 0$ on C, $\exists$ an $h(x)$ on C such that, $h(x) = \frac{f(x)}{g(x)} = \sum_{j=0}^{\infty} d_j(x-\alpha)^j$, for some constants d_j.*

Proposition 3.6. *Let*

$$\sum_{j=0}^{\infty} a_j(x-\alpha)^j$$

be a power series with open interval of convergence C. Let $f(x)$ be the function defined by the series on the interval C. Then, the function f is continuous and has continuous, real analytic derivatives of all orders at α.

[6] Saff, E.B. and Snider, A.D., *Fundamentals of Complex Analysis with Applications to Engineering, Science, and Mathematics*, 1993, Prentice-Hall, New Jersey.

Using this proposition, it is easy to show that a real analytic function has a unique power series representation:

Corollary 3.7. *If the function f is written as a convergent power series on a given interval of positive radius centered at α,*

$$f(x) = \sum_{j=0}^{\infty} a_j(x-\alpha)^j,$$

then the coefficients of the power series can be obtained from the derivatives of the function by

$$a_n = \frac{f^n(\alpha)}{n!}.$$

Proof. To obtain this result, simply differentiate both sides of the above equation n-times and evaluate at α. Differentiation is possible by the previous proposition. □

We pause for a moment to point out that although real analytic functions are infinitely differentiable, the converse is not true. For example, the function

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0; \\ 0, & x = 0. \end{cases}$$

is infinitely differentiable. It is somewhat technical to show that $f^{(h)}(0)$ exists for all h, but nevertheless it can be done. Moreover,

$$f^{(h)}(0) = 0 \quad \forall h \geq 0$$

and so the power series of f about $x = 0$ is just zero. Thus, f does not equal its power series about $x = 0$ and so f is not real analytic - although it is infinitely differentiable. The real analytic functions are indeed a very special class of function.

The term *real analytic* comes from the fact that if the power series of

$$f(x) = \sum a_n(x-\alpha)^n$$

converges on $(\alpha - R, \alpha + R)$, then the series

$$f(z) = \sum a_n(z-\alpha)^n,$$

where $z = x + \imath y$ is a complex variable converges and is an analytic function on the ball $\{z \in \mathbb{C} : |z - \alpha| < R\}$. Conversely if f is any convergent function, with a power series which converges on $\{z \in \mathbb{C} : |z - \alpha| < R\}$ when $\alpha \in \mathbb{R}$, then $f(x)$ is a real analytic function with power series converging on $(\alpha - R, \alpha + R)$. We shall say more about this in later sections.

We now talk about real analytic functions of several variables. In order to generalize the power series to higher dimensions, we introduce the multi-index notation.

Definition 3.8. A *multi-index* μ, is an m-tuple $(\mu_1, \mu_2, \cdots, \mu_m)$ of non-negative integers. We write,

$$\Lambda(m) = \mathbb{N} \times \cdots \times \mathbb{N},$$

or alternatively,

$$\Lambda(m) = (\mathbb{N})^m.$$

Definition 3.9. For

$$\mu = (\mu_1, \mu_2, \cdots, \mu_m) \in \Lambda(m)$$

and

$$x = (x_1, x_2, \cdots, x_m) \in \mathbb{R}^m,$$

we define the following operations,

$$\begin{aligned}
\mu! &= \mu_1!\mu_2!\cdots\mu_m!, \\
|\mu| &= \mu_1 + \mu_2 + \cdots + \mu_m, \\
x^\mu &= x_1^{\mu_1} x_2^{\mu_2} \cdots x_m^{\mu_m}, \\
|x^\mu| &= |x_1|^{\mu_1} |x_2|^{\mu_2} \cdots |x_m|^{\mu_m}, \\
\frac{\partial^\mu}{\partial x^\mu} &= \frac{\partial^{\mu_1}}{\partial x_1^{\mu_1}} \frac{\partial^{\mu_2}}{\partial x_2^{\mu_2}} \cdots \frac{\partial^{\mu_m}}{\partial x_m^{\mu_m}}
\end{aligned}$$

And for,

$$\mu = (\mu_1, \mu_2, \ldots, \mu_m) \in \Lambda(m)$$

and

$$\nu = (\nu_1, \nu_2, \ldots, \nu_m) \in \Lambda(m),$$

we write,

$$\mu \leq \nu$$

if $\mu_j \leq \nu_j$ for $j = 1, 2, \ldots, m$.

Definition 3.10. The formal expression

$$\sum_{\mu \in \Lambda(m)} a_\mu (x - \alpha)^\mu,$$

with $\alpha \in \mathbb{R}^n$ and $a_\mu \in \mathbb{R}$ for each μ, is called a *power series in m variables.*

Definition 3.11. The power series

$$\sum_{\mu \in \Lambda(m)} a_\mu (x - \alpha)^\mu$$

is said to converge at x if there is a function $\phi : \mathbb{Z}^+ \to \Lambda(m)$ which is one-to-one and onto such that the series

$$\sum_{j=0}^{\infty} a_{\phi(j)} (x - \alpha)^{\phi(j)}$$

converges.

Definition 3.12. For a fixed power series $\sum_\mu a_\mu(x-\alpha)^\mu$, we set

$$C = \bigcup_{r>0}\{x \in \mathbb{R}^m : \sum_\mu |a_\mu(y-\alpha)^\mu| < \infty, \text{ all } |y-x| < r\}$$

This set is called the *domain of convergence.*

Definition 3.13. We say that a function $f : U \subset \mathbb{R}^m \to \mathbb{R}$ is called *real analytic* if for each $\alpha \in U$, the function f may be represented by a convergent power series in some neighborhood of α.

In a similar fashion to Proposition 3.5, it is relatively simple to prove the following:

Proposition 3.14. *Let $U, V \subset \mathbb{R}^m$ be open. If $f : U \to \mathbb{R}$ and $g : V \to \mathbb{R}$ are real analytic, then $f \pm g$, $f \cdot g$ are real analytic on $U \cap V$, and f/g is real analytic on $U \cap V \cap \{x : g(x) \neq 0\}$.*

Lemma 3.15. *Suppose that $g : (-a, a) \to \mathbb{R}$ is a real analytic function and $g(y) = 0$ on an open interval $I \subseteq (-a, a)$ then, $g \equiv 0$.*

Proof. If g is real analytic on $(-a, a)$, then we can write $g(x) = \sum_{n=0}^{\infty} a_n x^n$ on $(-a, a)$. Hence, we can say that $g(z) := \sum_{n=0}^{\infty} a_n z^n$ is analytic on $\{|z| < a\}$. The hypothesis states that the zeroes of g have an accumulation point on I and so, since the zeroes of an analytic function cannot have an accumulation point, g must be identically zero. □

Real analytic functions of one variable have domains of convergence equal to intervals. For several variable functions, the geometry of the domain of convergence is more complicated.

Definition 3.16. We say a set G in linear space, is called *convex* if for any two points $x, y \in G$, each point $z = \lambda x + (1-\lambda)y$, for $0 < \lambda < 1$, also belongs to G.

Example 3.17. A subset of $\mathbb{R}^2$ in the shape of pentagon is convex, whereas a subset of $\mathbb{R}^3$ in the shape of a donut is not.

Definition 3.18. For a set $G \subset \mathbb{R}^m$, we define $\log \parallel G \parallel$ as,

$$\log \parallel G \parallel = \{(\log|g_1|, \cdots, \log|g_m|) : g = (g_1, \cdots, g_m) \in SG\}.$$

The set G is said to be *logarithmically convex* if $\log \parallel G \parallel$ is a convex subset of of $\mathbb{R}^m$.

Before proving the next theorem, let us establish some facts.

Remark 3.19. For a fixed power series $\sum_\mu a_\mu(x-\alpha)^\mu$, we denote by B, the set of points $x \in R^m$ where $\sum_\mu |a_\mu||x-\alpha|^\mu$ is bounded. It is clear that if the power series converges at a point x, then $x \in B$. Furthermore, using a result called Abel's Lemma [7], it is possible to show that $Int(B) = C$, where C, is the domain of convergence of the power series. This information allows us to say something about the shape of the domain of convergence of a power series.

Theorem 3.20. *For a power series $\sum_\mu a_\mu x^\mu$, the domain of convergence C is logarithmically convex.*

[7]Krantz, S. G. and Parks, H. R., *A primer of real analytic functions.* Second edition. Birkhuser Advanced Texts: Basel Textbooks, Boston.

Proof. Fix two points $y = (y_1, \cdots, y_m) \in C$, $z = (z_1, \cdots, z_m) \in C$ and also let $0 \leq \lambda \leq 1$. Now, by the above remark, $y \in C$ means that $y \in Int(B)$. Hence, by the definition of an open set, for some $\epsilon > 0$, $(|y_1| + \epsilon, \cdots, |y_m| + \epsilon) \in B$. But, by the above remark, this means that there exists some constant L, such that,

$$|a_\mu||(|y_1| + \epsilon, \cdots, |y_m| + \epsilon)| \leq L.$$

Simplifying and rewriting, this becomes,

$$a_\mu \leq \frac{L}{\prod_{j=1}^m (|y_j| + \epsilon)^{\mu_j}}.$$

By the same process, we can replace ϵ by a smaller positive number and L by a larger number if necessary. Hence, without changing notation, we also have

$$a_\mu \leq \frac{L}{\prod_{j=1}^m (|z_j| + \epsilon)^{\mu_j}}.$$

Notice, that because we fixed y, z and λ, we can choose $\epsilon' > 0$, such that the following two expressions hold for $j = 1, \cdots, m$

$$(|y_j| + \epsilon)^\lambda \geq |y_j|^\lambda + \epsilon'$$

and

$$(|z_j| + \epsilon)^{1-\lambda} \geq |z_j|^{1-\lambda} + \epsilon'.$$

Then, we can choose $\sigma > 0$ so that for $j = 1, \cdots, m$,

$$(|y_j|^\lambda + \epsilon')(|z_j|^{1-\lambda} + \epsilon') \geq |y_j|^\lambda |x_j|^{1-\lambda} + \sigma,$$

holds. Putting these facts together, we conclude that,

$$|a_\mu| = |a_\mu|^\lambda |a_\mu|^{1-\lambda} \leq \frac{L}{\prod_{j=1}^m (|y_j|^\lambda |x_j|^{1-\lambda} + \sigma)^{\mu_j}}.$$

Thus, $(|y_1|^\lambda |z_1|^{1-\lambda}, \cdots, |y_m|^\lambda |z_m|^{1-\lambda}) \in Int(B) = C$, or equivalently

$$\lambda(\log|y_1|, \cdots, |y_m|) + (1 - \lambda)(\log(z_1), \cdots, \log|z_m|) \in \log ||C||.$$

That is, the domain of convergence, C, is logarithmically convex.

□

Example 3.21. Show that a square in $\mathbb{R}^2$, $S = \{(x, y) : a < x < b, c < y < d\}$ is logarithmically convex for some $a, b > 0$.

We want to show that

$$\log \| S \| = \{(u, v) = (\log|x|, \log|y|) : (x, y) \in S\},$$

is convex. Knowing that $a < x < b$ and $c < y < d$, we can write

$$\log|a| < \log|x| < \log|b|$$

and

$$\log|c| < \log|y| < \log|d|.$$

But this is just

$$\log|a| < u < \log|b|$$

and

$$\log|c| < v < \log|d|.$$

These restrictions define a square bounded by $\log|a|, \log|b|, \log|c|$ and $\log|d|$ in the u-v plane. But this means that $\log \parallel S \parallel$ is convex and hence S is logarithmically convex.

Example 3.22. Show that the domain of convergence of the power series $\sum_{n=0}^{\infty}(xy)^n = \frac{1}{1-xy}$, defined on $|xy| < 1$, is logarithmically convex.

Notice that any point within this domain will satisfy the inequality $|xy| < 1$. Hence the domain of convergence for the above power series is

$$S = \{(x, y) : |xy| < 1\}.$$

We want to show that

$$\log \parallel S \parallel = \{(u, v) = (\log|x|, \log|y|) : (x, y) \in S\},$$

is convex. Knowing that $|xy| < 1$, we write

$$\log|xy| < \log|1|,$$

which is simply,

$$\log|x| + \log|y| < 0.$$

Using the u-v notation we rewrite this as

$$v < -u.$$

This however, clearly implies that $\log \parallel S \parallel$ is convex (since the domain is simply the area under the $v = -u$ curve, in the u-v plane). Hence, S is logarithmically convex. As a side note, it is important to remember however, that even though the area is *logarithmically* convex, it is by no means convex.

With this new knowledge in tow, we now consider the factorization of multivariate real analytic functions.

4. Factorization of Multivariate Real Analytic Functions

In this section, we return to examining rank one functions, with the added constraint of real analyticity.

Consider first, a definition of rank one, real analytic functions.

Definition 4.1. A real analytic function with domain of convergence $C \subseteq \mathbb{R}^2$,

$$f(x_1, x_2) = \sum_{\mu} a_\mu x^\mu,$$

where $x = (x_1, x_2) \in \mathbb{R}^2$ and $\mu \in \Lambda(2)$, is said to be *rank one* if and only if

$$f(x_1, x_2) = g_1(x_1)g_2(x_2).$$

A clear question arises from this definition. Since f is real-analytic on its domain, is it necessarily true that g_1 and g_2 are real analytic on their respective domains as well? As it turns out, this conjecture is true.

Theorem 4.2. *Suppose that,*

$$f(x_1, x_2) = \sum_{\mu} a_\mu x^\mu$$

where $x = (x_1, x_2) \in \mathbb{R}^2$ *and* $\mu \in \Lambda(2)$ *has domain of convergence* C *and*

$$f(x_1, x_2) = g_1(x_1)g_2(x_2),$$

then, $C = (-a_1, a_1) \times (-a_2, a_2)$, *and* $g_1(x_1)$, $g_2(x_2)$ *are real analytic on* $(-a_1, a_1)$ *and* $(-a_2, a_2)$ *respectively.*

Proof. Expanding f, notice that

$$f(x_1, x_2) = \sum_{|\mu|=0}^{\infty} a_\mu x_1^{\mu_1} x_2^{\mu_2}$$

Now, choose a constant c, such that $\overrightarrow{P} = (x_1, c) \in C$, $c \neq 0$ and $g_i(c) \neq 0$, for $i = \{1, 2\}$. Notice that this last fact is guaranteed by Lemma 3.15.

Next, substitute $\overrightarrow{P}$ into f,

$$\begin{aligned} f(x_1, c) &= \sum_{|\mu|=0}^{\infty} a_\mu x_1^{\mu_1} c^{\mu_2} \\ &= g_i(x_1)g_2(c). \end{aligned}$$

Finally, remembering that $g_2(c) \neq 0$, divide both sides of

$$\sum_{|\mu|=0}^{\infty} a_\mu x_1^{\mu_1} c^{\mu_2} = g_i(x_1)g_2(c)$$

by $g_2(c)$, to obtain:

$$g_1(x_1) = \sum_{|\mu|=0}^{\infty} d_\mu x_1^{\mu_1},$$

where $d_\mu = a_\mu / g_2(c)$.

Proceed similarly for g_2. Hence, g_1 and g_2 are real analytic and can be represented by power series.

Since g_1 and g_2 can represented by an analytic power series on x_1 and x_2, the functions converge on the interval $(-a_1, a_1)$ and $(-a_2, a_2)$ respectively. Hence, $f = g_1(x_1)g_2(x_2)$ converges on $(-a_1, a_1) \times (-a_2, a_2)$. □

Having shown this theorem true for rank one, real analytic polynomials, it is interesting to see whether it will also hold for rank two, three or even r.

Before considering these cases let us first provide a detailed definition of rank r and let us derive two propositions that will be helpful in our proof.

Definition 4.3. A real analytic function, with domain of convergence $C \in \mathbb{R}^2$,

$$f(x_1, x_2) = \sum_{\mu} a_\mu x^\mu,$$

where $x = (x_1, x_2) \in \mathbb{R}^2$ and $\mu \in \Lambda(2)$, is said to be *rank r* if and only if

$$f(x_1, x_2) = \sum_{i=1}^{r} f_i(x_1) g_i(x_2).$$

Proposition 4.4. *Suppose that,*

$$f(x_1, x_2) = \sum_{\mu} a_\mu x^\mu$$

where $x = (x_1, x_2) \in \mathbb{R}^2$, has domain of convergence equal to C and is rank two, i.e.

$$f(x_1, x_2) = f_1(x_1)g_1(x_2) + f_2(x_1)g_2(x_2),$$

then, for some non-zero constants c, d, the following hold:

(1) $g_1(x_2) \neq c g_2(x_2) \qquad \forall x_2,$

(2) *the matrix,*

$$\begin{pmatrix} g_1(c) & g_2(c) \\ g_1(d) & g_2(d) \end{pmatrix}$$

is invertible.

Proof. (1) Proceed by contradiction. Assume that $g_1(x_2) = c g_2(x_2), \forall x_2$ and some constant c. Then,

$$\begin{aligned} f(x_1, x_2) &= f_1(x_1) c g_2(x_2) + f_2(x_1) g_2(x_2) \\ &= (g_2(x_2))(c f_1(x_1) + f_2(x_2)). \end{aligned}$$

Therefore $f(x_1, x_2)$ is rank one. But we already know that $f(x_1, x_2)$ is rank two. The contradiction is reached.
Therefore, g_1 is *not* a multiple of g_2.

(2) We will prove the second part using the above fact, that g_1 is *not* a multiple of g_2. Proceed by contradiction,
Suppose $g_1(c)g_2(d) - g_1(d)g_2(c) = 0$, for all c and d. Choose d such that $g_2(d) \neq 0$.[8]. Therefore, we rewrite the first statement as

$$g_1(c)g_2(d) = g_1(d)g_2(c),$$

Now, since $g_2(d) \neq 0$, divide both sides by $g_2(d)$ to obtain

$$g_1(c) = \frac{g_1(d)}{g_2(d)} g_2(c),$$

for all constants c. But notice that in the above statement, $\frac{g_1(d)}{g_2(d)}$ is just some constant, k. Therefore,

$$g_1(c) = kg_2(c),$$

for all c. This is a contradiction since g_1 is not a multiple of g_2. Therefore,

$$g_1(c)g_2(d) - g_1(d)g_2(c) \neq 0,$$

for *some* c and d. Also,

$$g_1(c)g_2(d) - g_1(d)g_2(c) = det \begin{vmatrix} g_1(c) & g_2(c) \\ g_1(d) & g_2(d) \end{vmatrix}.$$

Therefore,

$$det \begin{vmatrix} g_1(c) & g_2(c) \\ g_1(d) & g_2(d) \end{vmatrix} \neq 0$$

and hence,

$$\begin{pmatrix} g_1(c) & g_2(c) \\ g_1(d) & g_2(d) \end{pmatrix}$$

is invertible.

□

With this proposition in tow, we will now consider the rank two version of Theorem 4.2.

Theorem 4.5. *Suppose f is a real analytic, rank two function with domain of convergence $C \in \mathbb{R}^2$, i.e,*

$$f(x_1, x_2) = f_1(x_1)g_1(x_2) + f_2(x_1)g_2(x_2),$$

then, $C = L_1 \cap L_2$ where $L_i = (-a_i, a_i) \times (-b_i, b_i)$ and where $f_i(x)$ and $g_i(y)$ are defined on $(-a_i, a_i)$ and $(-b_i, b_i)$ respectively.

---

[8]To convince yourself of this fact consider the following argument. If $g_2(d) = 0$ for all d, then $f(x_1, x_2) = f_1(x)g_1(x_2)$ and f is rank one.. But we know f is rank two and so, $g_2(d) \neq 0$, for some d.

Proof. (Notice that we are effectively being asked to show that f_i and g_i are real analytic on their corresponding domains)

Expanding f, we obtain:

$$f(x_1, x_2) = \sum_{|\mu|=0}^{\infty} a_\mu x_1^{\mu_1} x_2^{\mu_2}$$

First, we show that f_1 and f_2 are real analytic. Then, by the same process it can be shown that g_1 and g_2 are real analytic.

Choose non-zero constants c and d, such that $\overrightarrow{P_1} = (x_1, c) \in C$, $\overrightarrow{P_2} = (x_1, d) \in C$ and such that the constants c and d satisfy Lemma 3.15.

Next, substitute $\overrightarrow{P_1}$ into f and call the resultant function $h_1(x_1)$,

$$\begin{aligned} f(x_1, c) &= \sum_{|\mu|=0}^{\infty} a_\mu x_1^{\mu_1} c^{\mu_2} \\ &= h_1(x_1) \end{aligned}$$

and $\overrightarrow{P_2}$ into f and call that function $h_2(x_1)$,

$$\begin{aligned} f(x_1, d) &= \sum_{|\mu|=0}^{\infty} a_\mu x_1^{\mu_1} d^{\mu_2} \\ &= h_2(x_1). \end{aligned}$$

Notice that $h_1(x_1)$ and $h_2(x_1)$ are real analytic. The following system of equations for $h_1(x_1)$ and $h_2(x_1)$ holds,

$$\begin{aligned} h_1(x_1) &= f(x_1, c) \\ &= f_1(x_1)g_1(c) + f_2(x_1)g_2(c) \end{aligned}$$

$$\begin{aligned} h_2(x_1) &= f(x_1, d) \\ &= f_1(x_1)g_1(d) + f_2(x_1)g_2(d). \end{aligned}$$

So, rewriting this in matrix notation, we obtain,

$$\begin{pmatrix} h_1(x_1) \\ h_2(x_1) \end{pmatrix} = \begin{pmatrix} g_1(c) & g_2(c) \\ g_1(d) & g_2(d) \end{pmatrix} \begin{pmatrix} f_1(x_1) \\ f_2(x_1) \end{pmatrix}$$

But, by second part of the previous proposition, we know that $\begin{pmatrix} g_1(c) & g_2(c) \\ g_1(d) & g_2(d) \end{pmatrix}$ is invertible.

So, letting

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} g_1(c) & g_2(c) \\ g_1(d) & g_2(d) \end{pmatrix}^{-1},$$

where A, B, C and D are constants, we see that,

$$\begin{aligned} \begin{pmatrix} f_1(x_1) \\ f_2(x_1) \end{pmatrix} &= \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} h_1(x_1) \\ h_2(x_1) \end{pmatrix} \\ &= \begin{pmatrix} Ah_1(x_1) + Bh_2(x_1) \\ Ch_1(x_1) + Dh_2(x_1) \end{pmatrix}. \end{aligned}$$

Hence, we can read off the values of f_1 and f_2. Therefore, $f_1(x_1) = Ah_1(x_1) + Bh_2(x_1)$ and $f_2(x_1) = Ch_1(x_1) + Dh_2(x_1)$. Since f_1 and f_2 are linear combinations of real analytic functions (recall that h_1 and h_2 are real analytic), we know that f_1 and f_2 must be real analytic themselves.

Similarly, g_1 and g_2 must be real analytic.

Since each f_i and g_i can represented by an analytic power series on the x_i's, the f_i's converge on $(-a_i, a_i)$ and the g_i's converge on $(-b_i, b_i)$. So, the product $f_i g_i$, converges on $L_i = (-a_i, a_i) \times (-b_i, b_i)$ and so $f_1(x_1)g_1(x_2) + f_2(x_1)g_2(x_2)$ converges on $L_1 \cap L_2$. Hence, $C = L_1 \cap L_2$. □

In fact, it is possible to prove the above theorem for real analytic polynomials of rank r in the same manner as above (except for tedious labelling).

Theorem 4.6. *Suppose f is a real analytic, rank r function with domain of convergence $C \in \mathbb{R}^2$,*

$$f(x_1, x_2) = f_1(x_1)g_1(x_2) + \ldots + f_r(x_1)g_r(x_2),$$

then, $C = L_1 \cap \ldots \cap L_r$, where $L_i = (-a_i, a_i) \times (-b_i, b_i)$ and $f_i(x)$ and $g_i(y)$ are defined on $(-a_i, a_i)$ and $(-b_i, b_i)$ respectively.

Remark 4.7. Notice that the above theorem tells us that if f is real analytic and of finite rank, then the domain of convergence of f, is a box.

Remark 4.8. Notice also, that the above theorem says that if a rank r function is real analytic, each component f_i and g_i, will also be real analytic.

5. Factorization of Complex Valued Real Analytic Functions

A very natural question to ask after seeing the theorems in the previous section, is whether their converses hold. In order to answer this question we must consider extending our factorization to complex valued functions that are analytic. We begin with some introductory definitions and theorems for complex functions of one variable.

Theorem 5.1. *If $f : G \subset \mathbb{C} \to \mathbb{C}$ is C^1, then f is analytic on G if*

$$\overline{\partial} f = 0,$$

where, $\overline{\partial} f = \frac{1}{2}(\partial_x + \imath\partial_y)f$.

Theorem 5.2. *The function f is analytic on $G \subset \mathbb{C}$, if and only if, at each point $a \in G$,*

$$f(z) = \sum_{n=0}^{\infty} b_n(z-a)^n,$$

for z in an open neighborhood of a for some b_n.

In other words, the above theorem says that the function f is analytic if and only if, f can be written locally as a convergent power series.

Theorem 5.3. *For an analytic function, the coefficients b_n can be computed as,*

$$a_n = \frac{f^n(a)}{n!}$$

Notice that an important relation does exist between analytic and real analytic functions. In fact, real analytic functions are nothing more than analytic functions restricted to the real number line.

Let us now consider analytic functions of higher dimensions.

Definition 5.4. We define $\overline{\partial}_{z_j}$, for $z_j = x_j + \imath y_j$, as

$$\overline{\partial}_{z_j} = \frac{1}{2}(\partial_{x_j} + \imath\partial_{y_j}),$$

for all $j \in \mathbb{N}^+$.

Theorem 5.5. *The C' function $f : G \subset \mathbb{C}^m \to \mathbb{C}$ is analytic on G if*

$$\overline{\partial}_{z_j} f = 0,$$

for all j.

Theorem 5.6. *We say that a complex function $f : G \subset \mathbb{C}^m \to \mathbb{C}$ is* analytic *if and only if, for all $a = (a_1, \cdots, a_m) \in G$, $z \in \mathbb{C}^m$ and $\mu \in \Lambda(m)$,*

$$f(z_1, \cdots, z_m) = \sum_{\mu} b_\mu (z-a)^\mu$$

Theorem 5.7. *A real analytic function $\hat{f} : C \subseteq \mathbb{R}^n \to \mathbb{R}$ is an analytic function $f : C' \subseteq \mathbb{C}^n \to \mathbb{C}$ with all the variables $(z_1, \ldots, z_n)$ restricted to the reals.*

It is exactly this fact that justifies introducing complex variables. We will now consider an example that will help us shed some light on why the converse of Theorem 4.2 (and by generalization Theorem 4.6) does not hold.

Example 5.8. Find the domain of convergence of $f(x) = e^{\frac{1}{1+x^2}}$, where $f(x) = \sum_{n=0}^{\infty} a_n z^n$.

Using regular methods of determining the radii of convergence, this problem would pose a challenge. However, if we simply notice that $f(x) = e^{\frac{1}{1+x^2}}$, is the restriction to the real numbers of $f(z) = e^{\frac{1}{1+z^2}}$, the problem becomes astonishingly easy. We recall that $f(z) = e^{\frac{1}{1+z^2}}$, is undefined at $z = \imath$ and $z = -\imath$. As such, the domain of convergence for $f(z)$ is $|z| < 1$. Remembering that $z = x + \imath y$ and letting $y = 0$, we automatically obtain the domain of convergence for $f(x)$, to be $|x| < 1$.

Having this result, we turn back to Theorem 4.2. We know from Theorem 4.2, that a real analytic finite rank function $f(x_1, x_2) = \sum_\mu a_\mu x^\mu$, converges on a box in $\mathbb{R}^2$. The question thus naturally arises, whether the opposite it true. That is, it remains to be shown that a real analytic function whose domain of convergence is a box, is of finite rank. Unfortunately as we shall show, this is *not* the case.

Example 5.9. Let $f(x,y) = \exp(\frac{1}{1+x^2}\frac{1}{1+y^2})$.
First, let us make sure that this is a real analytic function. To do this we need to show that the function $f(x,y) = \exp(\frac{1}{1+x^2}\frac{1}{1+y^2})$, is simply a restriction of the complex function $f(z,w) = \exp(\frac{1}{1+z^2}\frac{1}{1+w^2})$, where $z = x + \imath t$ and $w = y + \imath s$, to the real plane. But, by the same reasoning as in the above example, $f(z,w)$ is analytic on $D \times D = \{(z,w) \in \mathbb{C} \times \mathbb{C} : |z| < 1, |w| < 1\}$. Hence restricting $f(z,w)$ onto the real plane by setting $t = 0$ and $s = 0$, means that $f(x,y)$ is real analytic on $B = \{(x,y) \in \mathbb{R} \times \mathbb{R} : |x| < 1, |y| < 1\}$. Notice also that B, is simply a box in two dimensions.
Finally, we need to show that f is not of finite rank. This procedure is done by demonstration, rather than by strict mathematical proof - although a proof certainly seems to be achievable. We use *Mathematica* to calculate the power series of f to order 10, calling it $\hat{f}$ and noting that $\hat{f}$ has degree 10. Next, we place the coefficients of the resulting polynomial in a 11×11 coefficient matrix. We see that this matrix is made up of alternating zero and non-zero columns. In line with Theorem 2.11, we row reduce this matrix in order to find the rank. The results are encouraging. We see that the rank of this matrix is as high as it could be, judging from the number of non-zero columns. This matrix (and hence $\hat{f}$) is of rank 6. We repeat this operation with power series of 15, 20, 25 and 30. Each time, we see that the rank of the resultant coefficient matrix (and hence of $\hat{f}$ is

$$rank(\hat{f}) = (deg(\hat{f}))mod(2) + 1,$$

where $deg(\hat{f})$ is the degree of the given series expansion. It seems that this equation will be true for higher orders of $\hat{f}$. If the equation is indeed true, the following holds,

$$\lim_{deg(\hat{f}) \to \infty} rank(\hat{f}) = \infty.$$

However, note that as $deg(\hat{f}) \to \infty$, the approximation $\hat{f} \approx f$, becomes exact. That is, as $deg(\hat{f}) \to \infty$, $\hat{f} = f$, and so the rank of f becomes infinite.
Hence $f(x,y)$ is a function that converges on a box and is not of finite rank, disproving the converse of Theorem 4.2

6. Factorization of Polynomials in Three Dimensions

The following section examines the factorization of polynomials in three unknowns.

Before any new theorems can be proven, consider the following definitions.

Definition 6.1. We say a polynomial $p(x_1, x_2, x_3)$ is of *rank one* iff $p(x_1, x_2, x_3) = p_1(x_1)p_2(x_2)p_3(x_3)$, where p_1, p_2, p_3 are polynomials.

This is analogous to the two dimensional definition.

Definition 6.2. A $M_{n \times n \times n}$ *grid* is a $n \times n \times n$ three dimensional matrix, where an entry is written as $A_{i,j,k}$, for $1 \leq i, j, k \leq n$.

Definition 6.3. The i^{th} *face* or *slice* of M, the grid defined by $A_{i,j,k}$ for $1 \leq i, j, k \leq n$, is the matrix $F_{j,k}$, defined by $F_{j,k} := A_{i,j,k}$.

Definition 6.4. The j^{th} *layer* of M, the grid defined by $A_{i,j,k}$ for $1 \leq i, j, k \leq n$, is the matrix $L_{i,k}$, defined by $L_{i,k} := A_{i,j,k}$.

Let us attempt to visualize this concept in the following example.

Example 6.5. Consider a $2 \times 2 \times 2$ grid M, defined by the following:
$A_{1,1,1} = 1, A_{1,1,2} = 2, A_{1,2,1} = 2, A_{1,2,2} = 4, A_{2,1,1} = 3, A_{2,1,2} = 6, A_{2,2,1} = 6, A_{2,2,2} = 12$

The first face (or slice) of this grid is the matrix $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ and the second face is, $\begin{pmatrix} 3 & 6 \\ 6 & 12 \end{pmatrix}$. Furthermore, notice that the first layer of this grid is the matrix $\begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ and the second layer is, $\begin{pmatrix} 2 & 4 \\ 6 & 12 \end{pmatrix}$. Drawing this grid, can also be helpful.

Notice that it is possible to represent a three dimensional polynomial by a grid in a similar manner to the way the coefficients of the two-dimensional polynomial were written into a matrix.

Definition 6.6. The *degree* of a polynomial, $p(x_1, x_2, x_3) = a_{0,0} + ... + a_{m,n}x_1^k x_2^l x_3^m$ is the highest power of x_1, x_2 or x_3 with non-zero coefficients.

Definition 6.7. For a polynomial

$$p(x_1, x_2, x_3) = \sum_{0 \leq i,j,k \leq n} a_{i,j,k} x_1^i x_2^j x_3^k$$

of degree n, we define the *coefficient grid*, $C(p)$ to be the $(n+1) \times (n+1) \times (n+1)$ grid whose i, j, k^{th} entry is $a_{i,j,k}$. The i^{th} face of this grid is,

$$C_i(p) := \begin{pmatrix} a_{i,0,0} & \cdots & a_{i,0,n} \\ \vdots & \vdots & \vdots \\ a_{i,n,0} & \cdots & a_{i,n,n} \end{pmatrix}.$$

Definition 6.8. A $3 \times 3 \times 3$ coefficient grid, M, is said to be *rank one* iff the corresponding polynomial is rank one.

In two dimensions it was easy to show that a rank one coefficient matrix meant that the corresponding polynomial was rank one. In three dimensions, this is slightly more challenging. After all, we knew how to compute the rank of a matrix. How is it possible to find the rank of a grid? The following theorem answers these questions.

Theorem 6.9. *For a grid $G \in M_{3\times3\times3}$, the following are equivalent*

1. *G is rank one.*
2. *The polynomial $p(x_1, x_2, x_3)$ that corresponds to grid G is of rank one.*
3. *If F_1, F_2 and F_3 are the faces of the grid, then $Rank[(F_1 \mid F_2 \mid F_3)] = 1$ and $Rank[(F_1 \mid F_2 \mid F_3)^T] = 1$.*
4. *Every face and every layer of G is rank one.*

As a direct consequence of the definition, the first two statements are equivalent. It is also very easy to show that the last two statements are equivalent. Hence, all that remains to be proven, is the equivalency between the second and the third statements.

Proof. $(\Rightarrow)$: We assume that the polynomial $p(x_1, x_2, x_3)$ is rank one. Hence,

$$\begin{aligned} p(x_1,x_2,x_3) &= \sum_{0\le i,j,k\le n} d_{i,j,k} x_1^i x_2^j x_3^k \\ &= p_1(x_1)p_2(x_2)p_3(x_3) \\ &= (\sum_{i=0}^{n} a_i x_1^i)(\sum_{j=0}^{n} b_j x_2^j)(\sum_{k=0}^{n} c_j x_3^j) \\ &= \sum_{0\le i,j,k\le n} a_i b_j c_k x_1^i x_2^j x_3^k \end{aligned}$$

For notational clarity, we assume that $n = 2$. Now, setting coefficients equivalent and writing down the i^{th} face of the $3 \times 3 \times 3$ coefficient grid, we obtain

$$\begin{aligned} F_i &= \begin{pmatrix} d_{i,0,0} & d_{i,0,1} & d_{i,0,2} \\ d_{i,1,0} & d_{i,1,1} & d_{i,1,2} \\ d_{i,2,0} & d_{i,2,1} & d_{i,2,2} \end{pmatrix} \\ &= \begin{pmatrix} a_i b_0 c_0 & a_i b_0 c_1 & a_i b_0 c_2 \\ a_i b_1 c_0 & a_i b_1 c_1 & a_i b_1 c_2 \\ a_i b_2 c_0 & a_i b_2 c_1 & a_i b_2 c_2 \end{pmatrix} \end{aligned}$$

for $0 \le i \le 2$.
Clearly F_i is of rank one. Notice now, that

$$\begin{pmatrix} F_1 \\ - \\ F_2 \\ - \\ F_3 \end{pmatrix} = \begin{pmatrix} a_0b_0c_0 & a_0b_0c_1 & a_0b_0c_2 \\ a_0b_1c_0 & a_0b_1c_1 & a_0b_1c_2 \\ a_0b_2c_0 & a_0b_2c_1 & a_0b_2c_2 \\ a_1b_0c_0 & a_1b_0c_1 & a_1b_0c_2 \\ a_1b_1c_0 & a_1b_1c_1 & a_1b_1c_2 \\ a_1b_2c_0 & a_1b_2c_1 & a_1b_2c_2 \\ a_2b_0c_0 & a_2b_0c_1 & a_2b_0c_2 \\ a_2b_1c_0 & a_2b_1c_1 & a_2b_1c_2 \\ a_2b_2c_0 & a_2b_2c_1 & a_2b_2c_2 \end{pmatrix}$$

So, it is clear that $Rank[(\ F_1 \mid F_2 \mid F_3\ )^T] = 1$.
Also notice that,

$$(\ F_1 \mid F_2 \mid F_3\ ) =$$

$$\begin{pmatrix} a_0b_0c_0 & a_0b_0c_1 & a_0b_0c_2 & a_1b_0c_0 & a_1b_0c_1 & a_1b_0c_2 & a_2b_0c_0 & a_2b_0c_1 & a_2b_0c_2 \\ a_0b_1c_0 & a_0b_1c_1 & a_0b_1c_2 & a_1b_1c_0 & a_1b_1c_1 & a_1b_1c_2 & a_2b_1c_0 & a_2b_1c_1 & a_2b_1c_2 \\ a_0b_2c_0 & a_0b_2c_1 & a_0b_2c_2 & a_1b_2c_0 & a_1b_2c_1 & a_1b_2c_2 & a_2b_2c_0 & a_2b_2c_1 & a_2b_2c_2 \end{pmatrix}$$

It is clear that $Rank[(\ F_1 \mid F_2 \mid F_3\ )] = 1$.

$(\Rightarrow)$: Assume that $Rank[(\ F_1 \mid F_2 \mid F_3\ )] = 1$ and $Rank[(\ F_1 \mid F_2 \mid F_3\ )^T] = 1$.

Since $Rank[(\ F_1 \mid F_2 \mid F_3\ )^T] = 1$, there exist constants a_i, l, m for $0 \leq i \leq 8$ such that the following equality must hold,

$$\begin{pmatrix} F_1 \\ - \\ F_2 \\ - \\ F_3 \end{pmatrix} = \begin{pmatrix} a_0 & la_0 & ma_0 \\ a_1 & la_1 & ma_1 \\ a_2 & la_2 & ma_2 \\ -- & -- & -- \\ a_3 & la_3 & ma_3 \\ a_4 & la_4 & ma_4 \\ a_5 & la_5 & ma_5 \\ -- & -- & -- \\ a_6 & la_6 & ma_6 \\ a_7 & la_7 & ma_7 \\ a_8 & la_8 & ma_8 \end{pmatrix}.$$

But this means that,

$$(\ F_1 \mid F_2 \mid F_3\ ) = \begin{pmatrix} a_0 & la_0 & ma_0 & a_3 & la_3 & ma_3 & a_6 & la_6 & ma_6 \\ a_1 & la_1 & ma_1 & a_4 & la_4 & ma_4 & a_7 & la_7 & ma_7 \\ a_2 & la_2 & ma_2 & a_7 & la_7 & ma_7 & a_8 & la_8 & ma_8 \end{pmatrix}$$

But we also know, that $Rank[(\ F_1 \mid F_2 \mid F_3\ )] = 1$. Hence, it must be true that for some constants n and h, where $n \neq h \neq 0$, the following equalities hold

$$a_3 = na_0$$

$$a_4 = na_1$$

$$a_5 = na_2$$

and,

$$a_6 = ha_0$$

$$a_7 = ha_1$$

$$a_8 = ha_2$$

So, rewrite the faces F_1, F_2 and F_3 of the coefficient grid taking these equalities into account.

$$F_1 = \begin{pmatrix} a_0 & la_0 & ma_0 \\ a_1 & la_1 & ma_1 \\ a_2 & la_2 & ma_2 \end{pmatrix}$$

$$F_2 = \begin{pmatrix} na_0 & lna_0 & mna_0 \\ na_1 & lna_1 & mna_1 \\ na_2 & lna_2 & mna_2 \end{pmatrix}$$

$$F_2 = \begin{pmatrix} ha_0 & lha_0 & mha_0 \\ ha_1 & lha_1 & mha_1 \\ ha_2 & lha_2 & mha_2 \end{pmatrix}$$

Now, rewriting the grid into its polynomial representation, we get:

$p(x_1, x_2, x_3) =$
$a_0+na_0x_1+ha_0x_1^2+a_1x_2+na_1x_1x_2+ha_1x_1^2x_2+a_2x_2^2+na_2x_1x_2^2+ha_2x_1^2x_2^2+la_0x_3+$
$lna_0x_1x_3+hla_0x_1^2x_3+la_1x_2x_3+lna_1x_1x_2x_3+hla_1x_1^2x_2x_3+la_2x_2^2x_3+lna_2x_1x_2^2x_3+$
$hla_2x_1^2x_2^2x_3+ma_0x_3^2++mna_0x_1x_3^2+hma_0x_1^2x_3^2+ma_1x_2x_3^2+mna_1x_1x_2x_3^2+hma_1x_1^2x_2x_3^2+$
$ma_2x_2^2x_3^2 + mna_2x_1x_2^2x_3^2 + hmna_2x_1^2x_2^2x_3^2$
$= (1 + nx_1 + hx_1^2)(1 + \frac{a_1}{a_0}x_2 + \frac{a_2}{a_0}x_2^2)(a_0 + la_0x_3 + ma_0x_3^2)$

Therefore, the polynomial is of rank one.

Let us now consider polynomials of dimension three that are not of rank one. We define a polynomial of higher rank in three dimensions, in a similar manner to the two dimensional case.

Definition 6.10. We say that a polynomial p is of *rank* r if and only if there exits an r such that,

$$p(x_1, x_2, x_3) = \sum_{i=1}^{r} f_i(x_1)g_i(x_2)h_i(x_3)$$

and if r is the smallest integer to satisfy this condition.

Notice also, that this definition of rank, extends and complements the earlier definitions of rank one and rank r polynomials.

We had to leave the following as an open question. When $n = 2$, a polynomial is of rank r if and only if $C(p)$ is of rank r. How can this be generalized to grids?

Finally, it is quite possible and not at all conceptually difficult to extend the above theory for higher dimensions. Notice however, that although the above discussion forms a very neat and convenient theory of grids, it is not at all efficient from the computational point of view. Programs, that attempt to factorize polynomials in the above manner are large, unwieldily and slow. As such, we must develop a quicker and slicker method for factorizing multi-dimensional polynomials.

7. Algorithm and Examples in Two and N Dimensions

As we already mentioned, the theory of grids described in the previous section is difficult and unwieldily in its implementation as a computer algorithm. In this section we will describe an alternative algorithm. We shall first describe computational method for polynomials of two dimensions and later we will generalize this to n dimensions, by a surprisingly simple extension.

The two dimensional algorithm is based completely on Theorem 2.11 and Example 2.12. Let $p(x_1, x_2)$, be an 2 dimensional, rank one polynomial, of degree d, represented by the coefficient matric $C(p)$. Since, $p(x_1, x_2)$ is rank one, we know that $C(p)$ is rank one and so we can factor it as

$$p(x_1, x_2) = p_1(x_1)p_2(x_2).$$

We factorize by the following algorithm:

(1) Create coefficient matrix $C(p)$.
(2) Column reduce this matrix and save the leading (non-zero) column as the vector $\overrightarrow{\beta}$. (These will be the coefficients of $p_1(x_1)$).
(3) Save the top row of $C(p)$ in a vector, $\overrightarrow{\gamma}$. These will be the coefficients of the $p_2(x_2)$).

Example 2.12 follows this algorithm exactly.

In three and more dimensions, we simply use the above algorithm recursively, holding the other variables constant. Let us first consider an example before looking at the algorithm itself.

Example 7.1. Let p be the following rank polynomial

$$p(x_1, x_2, x_3) = 1 + 3\,x_1 + 2\,x_2 + 6\,x_1\,x_2 + 2\,x_3 + 6\,x_1\,x_3 + 4\,x_2\,x_3 + 12\,x_1\,x_2\,x_3.$$

We wish to factorize this polynomial as,

$$p(x_1, x_2, x_3) = p_1(x_1)p_2(x_2)p_3(x_3).$$

First, we want to express this polynomial in matrix form. However, we do *not* wish to use the complicated grid system. As such, let us simply assume for the moment that this polynomial is a function in two variables, x_1 and x_2 and that the third variable, x_3 is simply a constant. With this in mind, we write down the coefficient matrix to be:

$$\begin{pmatrix} 1 + 2\,x_3 & 2 + 4\,x_3 \\ 3 + 6\,x_3 & 6 + 12\,x_3 \end{pmatrix}.$$

So, the $(1, 1)$ entry represents the 'constant' term, $(1, 2)$ entry represents the x_2 term, the $(2, 1)$ entry represents the x_1 term and finally the $(2, 2)$ entry represents the x_1x_2 term. Let us now use the above algorithm. We column reduce the above matrix and obtain

$$\begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}.$$

Hence, our beta vector is,

$$\overrightarrow{\beta} = \{1, 3\}.$$

Next, we read off the top row of the initial coefficient matrix to obtain the gamma vector,

$$\overrightarrow{\gamma} = \{1 + 2\,x_3, 2 + 4\,x_3\}.$$

This information tells us that

$$p_1(x_1) = 1 + 3x_1$$

from the beta vector, and

$$q(x_2, x_3) = 1 + 2x_3 + (2 + 4x_3)x_2$$

from the gamma vector. Notice that we have managed to reduce the problem, and that we can now write $p(x_1, x_2, x_3)$ as,

$$\begin{aligned} p(x_1, x_2, x_3) &= p_1(x_1)q(x_2, x_3) \\ &= (1 + 3\,x_1)(1 + 2\,x_2 + 2\,x_3 + 4\,x_2\,x_3). \end{aligned}$$

Now, to find the full factorization of p, we only have to factorize q - a function of two variables. But we already know how to do this. We once again use the above algorithm to obtain,

$$q(x_1, x_2) = (1 + 2\,x_2)\ (1 + 2\,x_3)\,.$$

Putting these two facts together, gives us the final factorization for p,

$$p(x_1, x_2, x_3) = (1 + 3\,x_1)\,(1 + 2\,x_2)\,(1 + 2\,x_3)\,.$$

Hence, in general, the algorithm for $n \geq 3$ dimensional factorization of polynomial $p(x_1, \ldots x_n)$ is as follows,

(1) Create coefficient matrix $C(p)$ in terms of x_1 and x_2, holding the terms $x_3, \ldots, x_n$ constant.
(2) Column reduce this matrix and save the leading (non-zero) column as the vector $\overrightarrow{\beta}$. (These will be the coefficients of $p_1(x_1)$).
(3) Save the top row of $C(p)$ in a vector, $\overrightarrow{\gamma}$. This will be the coefficients of the new polynomial $q_1(x_2, \ldots, x_n)$).
(4) Repeat steps 1 through 3 using q_i as the new polynomial each time.

The above directions describe exactly the algorithm used in the construction of the *Mathematica* program *FactNDim*, that was created as part of this paper.

Recall that we mentioned creating an algorithm using the theory developed in the previous section. Although this is a far more complicated and involved algorithm we include it, for completeness sake.

8. Alternative Algorithm and Examples Three Dimensions

The following section examines the Mathematica algorithm that strictly follows from the grid theory proposed in one of the previous sections. This is in no way an efficient or optimal algorithm. It is presented to contrast with the previous algorithm and to demonstrate the way the theory could be used. For an efficient implementation refer to the previous section.

Let $p(x_1, x_2, x_3)$, be an 3 dimensional, rank one polynomial, of degree d, represented by the coefficient grid $G = \{A_1, \cdots, A_d\}$. Since, $p(x_1, x_2, x_3)$ is of rank one, we can can factor it as $p(x_1, x_2, x_3) = p_1(x_1)p_2(x_2)p_3(x_3)$. The following algorithm describes the method used to find the coefficients of each of the polynomials $p_i(x_i)$.

(1) Create an array of length two, call this the "answer matrix array" or AMA.
(2) In the first entry of AMA, create an array or length d, called AMA_1
(3) Enter the first non-zero element of each A_i into $AMA_{1,i}$. If, an entire face A_i is zero, enter zero into $AMA_{1,i}$.
(4) Find the index, j, of the first non-zero entry of the AMA_1 array.
(5) Enter the matrix A_j into the second entry of AMA called AMA_2.
(6) Create an array called *answer* of length three. The i^{th} entry of that array, will contain another array that will store the coefficients of x_i.
(7) Column reduce AMA_1 and enter the leading column into $answer_1$. These are the coefficients for x_1.
(8) Column reduce AMA_2 and enter the leading column into $answer_2$. These are the coefficients for x_2.
(9) Enter the first non-zero row of AMA_2 into $answer_3$. These are the coefficients for x_3.
(10) The *answer* array contains the desired coefficients.

We will see later that this algorithm generalizes to n dimensions. An AMA array will still be found. We will then column and row reduce each of the entries of AMA (except the last one - we will only column reduce *that* one), and enter the leading columns and rows of the reduced matrices into an *answer* array. There are however, some technical details that need to be explained about the creation of the AMA matrix in n-dimensions, so we will hold off with the explanation till later.

The simplest way to understand the workings of the mathematica methods and the above algorithm, is to examine some examples. As such, a simple example of the powerful algorithm that is used to solve the factorization problem is given. The example is then used as a tool for explaining individual methods.

Example 8.1. Notice that for polynomial, $p(x_1, x_2, x_3)$,

$$\begin{aligned} p(x_1,x_2,x_3) &= 1 + 3x_1 + 2x_2 + 6x_1x_2 + 2x_3 + 6x_1x_3 + 4x_2x_3 + 12x_1x_2x_3 \\ &= p_1(x_1)p_2(x_2)p_3(x_3) \\ &= (1+3x_1)(1+2x_2)(1+2x_3). \end{aligned}$$

First recognize that this polynomial can be represented by a 2×2 grid $G = \{A_1, Aa_2\}$, with faces $A_1 = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 3 & 6 \\ 6 & 12 \end{pmatrix}$.

Notice that this coefficient grid is of rank one[9] and hence the polynomial it represents can be factored.

Now, create a matrix (in this case an array) that contains the first non-zero entries of the two faces in the gird, A_1 and A_2. This array is simply $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$. Next, in accordance with the algorithm, find the *location* of first non-zero entry in this resultant array. In this case, the first non-zero entry is 1 and it is in the first slot of the $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ array. So, according to the algorithm, single out the *first* face of the grid, $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$. We have now found the two matrices that will determine the coefficients of the factored polynomials.

Create an "answer matrix array" (AMA), that will contain both these matrices. Hence,

$$AMA = \begin{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \end{pmatrix}$$

We will now column and row reduce each entry in the AMA matrix. Column reduce the first entry of AMA. Notice, that we get the same array, $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$. This is the coefficient matrix of $p_1(x_1)$. Hence, we can say that $p_1(x_1) = (1 + 3x_1)$. Now, row reduce the first entry of AMA. Notice, that we get the array, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. But this simply represents the polynomial $p = 1$. So we ignore its effect on the factorization.

Continue similarly with the second matrix.

Column reduce the second entry of AMA. Notice, that we get the matrix, $\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$. This is the coefficient matrix of $p_2(x_2)$. Hence, we can say that $p_2(x_2) = (1 + 2x_2)$.
Notice, that according to the algorithm, because we have reached the last entry of AMA, we do not row-reduce the matrix, but take the first row directly, as the coefficients of $p_3(x_3)$. As such, $p_3(x_3) = (1 + 2x_3)$.

To represent our answers neatly, we create an `answer` array. The i^{th} entry of that array, will contain another array that will store the coefficients of x_i. So,

$$answer = \begin{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \end{pmatrix}$$

We have successfully factored the polynomial. $p(x_1, x_2, x_3)$.

□

[9]Each slice and each layer of the grid is rank one.