Using Runge's Theorem to Determine Zariski Density of Integral Points in Two and Three Dimensions

Angelo Mekenkamp

Supervisor: Dr. Damaris Schindler

March 7, 2019

Abstract

The German mathematician Carl Runge (1856-1927) came up with a theorem that said that any Diophantine equation in two variables satisfying a certain set of conditions has only finitely many integral solutions. This thesis will provide a detailed proof of this theorem and some examples in which we can apply it. This proof makes use of two theorems from abstract algebra: The Symmetric Function Theorem and Newton-Puiseux's Theorem. The statement and proof of these theorems will also be given. This thesis will then introduce the Zariski Topology in all dimensions and show the strong connection between the notion of Zariski density in two dimensions and the property of having finitely or infinitely many integral solutions to a given Diophantine equation in two variables. The concept of Zariski density makes it possible to formulate generalizations to Runge's Theorem in higher variables. After this introduction there will be an attempt of the writer to generalize Runge's Theorem such that it can be applied to Diophantine equations of three variables.

Universiteit Utrecht

Acknowledgements

I would first like to thank my thesis supervisor Dr. Damaris Schindler of the Faculty of Science at Utrecht University. She came up with Runge's Theorem as subject for my thesis and guided me throughout this project with biweekly meetings that were both delightful and helpful.

Secondly I would like to thank Prof. dr. Gunther Cornelissen of the Faculty of Science at Utrecht University. He was the second reader of this thesis.

Finally, I must express gratitude towards Miriam Sterl. She read a huge part of my thesis and pointed out mistakes regarding my grammar and style.

Contents

1 Notation and Conventions

1.1 Notation

Notation	Meaning
$\mathbb{N}$	The set of natural numbers $\{1, 2, \ldots\}$
$\mathbb{N}_0$	The set of nonnegative integers $\{0, 1, \ldots\}$
$\mathbb{Z}$	The set of integers
$\mathbb{Q}$	The set of rational numbers
$\mathbb{R}$	The set of real numbers
$\mathbb{R}_{>0}$	The set of positive real numbers
$\mathbb{R}_{\geq 0}$	The set of nonnegative real numbers
$\mathbb{C}$	The set of complex numbers
$\overline{\mathbb{Q}}$	The set of algebraic numbers (as subset of $\mathbb{C}$)
$R[X_1, \ldots, X_n]$	The polynomial ring in n variables over the ring R
$R[X, X^{-1}]$	The ring of Laurent polynomials in X over the ring R
$R[[X]]$	See Definition 4.1
$R((X))$	See Definition 4.4
$k(X)$	See Definition 4.6
$R[[X_1, \ldots, X_n]]$	See Definition 4.22
$k((X^*))$	See Definition 4.16
$B = B_{\boldsymbol{t}}[[X_1, \ldots, X_n]]$	See Definition 4.27
$K_n = k[\{X_1, \ldots, X_n\}]$	See Definition 4.36
$k(\{X\})$	See Definition 4.45
$k(\{X^*\})$	See Definition 4.45
$k(((X^{-1})^*))$	See Definitions 4.16 and 4.52
$k(\{(X^{-1})^*\})$	See Definitions 4.45 and 4.54
$\deg_X F$	See Definition 3.7
$\mathrm{ht}(F)$	See Definition 3.8
$D(F)$	See Definitions 3.5,4.59 and 8.25
$\deg_\lambda F$	See Definition 3.10
$D_\lambda(F)$	See Definition 3.10
F_λ	See Definition 3.10
$\tilde{D}(F)$	See Definition 3.11
$\tilde{F}$	See Definition 3.11
$\deg_{\mu,\lambda}(F)$	See Definition 8.25
$D_{\mu,\lambda}(F)$	See Definition 8.25
$F_{\mu,\lambda}$	See Definition 8.25
$\|\|f\|\|_{\boldsymbol{t}}$	See Definition 4.25
$\mathrm{ord}_X f$	See Definition 4.57
$V(\mathcal{F})$	See Definition 7.1
$T(F)$	See Definition 7.10
$S(F)$	See Definition 7.10

1.2 Conventions

- When we talk about rings, we mean commutative rings that contain the multiplicative identity.
- When we talk about domains, we mean integral domains that contain the multiplicative identity.
- We will use multiple notations for elements in polynomial rings or any of its ring extensions. For example we can write an element in $R[X, Y]$ as f or $f(X, Y)$ or even as $f(Y)$. The choice of notation depends on how much emphasis we want to lay on these variables.

2 Introduction

Since the 16th century, many mathematicians have been studying the concepts of Diophantine equations. A *Diophantine equation* is a polynomial equation in two or more variables with integer coefficients. It is called *binary* if it has exactly two variables. One can write a binary Diophantine equation as $F(X, Y) = 0$, with $F \in \mathbb{Z}[X, Y]$. There are many interesting questions about this equation. This thesis will focus on the question: are there finitely or infinitely many possibilities for $x, y \in \mathbb{Z}$ such that $F(x, y) = 0$ holds? This of course depends on F itself, as the following examples show.

Example 2.1. *Let $F = aX + bY + c$, with $a, b, c \in \mathbb{Z}$. We split the problem into multiple cases:*

- *If the constants a, b and c are all zero, we find that $F(x, y) = 0$ holds for all $x, y \in \mathbb{Z}$.*
- *If a and b are zero and c is nonzero, we find that $F(x, y) = c \neq 0$ and therefore there are no $x, y \in \mathbb{Z}$ with $F(x, y) = 0$.*
- *If a and b are not both zero, we let g be the greatest common divisor of a and b. By Bézout's identity, there exist $x', y' \in \mathbb{Z}$, such that $ax' + by' = g$. If c is not a multiple of g, we see that for all $x, y \in \mathbb{Z}$ we have that*
$$F(x, y) \equiv c \not\equiv 0 \mod g,$$
which shows that $F(x, y) = 0$ has no integer solutions. If c is a multiple of g, there exists $d \in \mathbb{Z}$ such that $gd = -c$. We then find that $adx' + bdy' = gd = -c$, so $x = dx'$, $y = dy'$ is a solution for $F(x, y) = 0$. If we let $e \in \mathbb{Z}$ and take $x = dx' - be$ and $y = dy' + ae$, we again have
$$ax + by = adx' - abe + bdy' + abe = gd = -c.$$
As any $e \in \mathbb{N}$ gives a different solution for x and y, we have infinitely many integer solutions to $F(x, y) = 0$.

Example 2.2. *Let $F = X^2 - 3Y^2$. Suppose that x and $y \in \mathbb{Z}$ satisfy $F(x, y) = 0$. Then it follows that $x^2 = 3y^2$. So $3y^2$ must be a perfect square. Since 3 is no perfect square, we find that y must be zero and therefore x must be zero as well. We conclude that $(x, y) = (0, 0)$ is the only integer solution to $F(x, y) = 0$.*

Example 2.3. *Let $F = X^2 - 3Y^2 + 1$. Suppose there exists $x, y \in \mathbb{Z}$ with $F(x, y) = 0$. Then we must have*
$$0 \equiv x^2 - 3y^2 + 1 \equiv x^2 + 1 \mod 3,$$
which shows that $x^2 \equiv -1 \mod 3$ holds and that -1 is a square modulo 3. This clearly fails and we therefore see that there are no $x, y \in \mathbb{Z}$ with $F(x, y) = 0$.

Example 2.4 (Pell's equation)**.** *Let $F = X^2 - 3Y^2 - 1$. It can easily be seen that $(x, y) = (2, 1)$ satisfies $F(x, y) = 0$. By passing the equation $F(X, Y) = 0$ to the ring extension $\mathbb{Z}[\sqrt{3}][X, Y]$, we find*
$$(X + Y\sqrt{3})(X - Y\sqrt{3}) = 1. \tag{2.1}$$
If for any $n \in \mathbb{N}_0$ we take $x_n, y_n \in \mathbb{Z}$ such that $x_n + y_n\sqrt{3} = (2 + \sqrt{3})^n$, then we also have $x_n - y_n\sqrt{3} = (2 - \sqrt{3})^n$ and therefore we find that
$$(x_n + y_n\sqrt{3})(x_n - y_n\sqrt{3}) = (2 + \sqrt{3})^n(2 - \sqrt{3})^n = ((2 + \sqrt{3})(2 - \sqrt{3}))^n = 1^n = 1.$$
which shows that (x_n, y_n) also satisfies (2.1). Since $|2 + \sqrt{3}| \neq 1$, we see that these pairs are distinct for different n. So we find infinitely many different solutions to the equation $F(x, y) = 0$.

Example 2.5. *Let $F = X^2 - Y^3 + 1$. It can be seen easily that $(x, y) = (0, 1)$ satisfies $F(x, y) = 0$. It becomes a bit harder to prove that this is the only integral solution. A proof of this can be find in Problem 1.3. on page 8 in [13]. This proof uses basic algebraic number theory.*

These previous examples show that it is not always trivial or easy to prove whether some binary Diophantine equation has finitely or infinitely many integral solutions. There are however methods and theorems to help determine this for certain families of Diophantine equations. Section 3 will provide some tools that are used by some of these theorems. Carl Runge has proved in 1887 [12] that certain classes of Diophantine equations have only finitely many integral solutions. Runge's Theorem is the main theorem of this thesis. The statement of this theorem requires a number of definitions and results that are explained in detail further on in this thesis. The theorem is roughly as follows: Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be an irreducible polynomial. We can write $F(X, Y)$ as

$$F(X,Y) = \sum_{i=0}^{d_1} \sum_{j=0}^{d_2} a_{i,j} X^i Y^j$$

for smallest $d_1 \in \mathbb{N}_0$ and $d_2 \in \mathbb{N}_0$. Runge's Theorem requires d_1 and d_2 to be positive. We describe the sets

$$D(F) \coloneqq \{(i,j) \mid i \in \{0, \dots, d_1\}, j \in \{0, \dots, d_2\}, a_{i,j} \neq 0\}$$

and

$$D'(F) \coloneqq \{(i,j) \in D(F) \mid i/d_1 + j/d_2 = 1\}.$$

We then let

$$F'(X,Y) \coloneqq \sum_{(i,j) \in D'(F)} a_{i,j} X^i Y^j.$$

suppose that the equation

$$F(x, y) = 0$$

has infinitely many solutions with $x, y \in \mathbb{Z}$. Runge's Theorem then asserts that all $(i, j) \in D(F)$ satisfy $i/d_1 + j/d_2 \leq 1$ and that $F'(X, Y)$ is an integer times a power of an irreducible polynomial in $\mathbb{Z}[X, Y]$. We can, as a consequence of this theorem, prove that certain Diophantine equations only have finitely many integral solutions by simply proving that one of these two properties does not hold. We consider the following examples:

Example 2.6. *Consider the irreducible polynomial* $F(X, Y) = XY + 2X - Y - 7 \in \mathbb{Z}[X, Y]$*. We then have* $d_1 = d_2 = 1$*. We find* $D(F) = \{(1, 1), (1, 0), (0, 1), (0, 0)\}$*. For* $(i, j) = (1, 1)$ *we see that* $i/d_1 + j/d_2 = 2 > 1$*, so by Runge's Theorem we conclude that* $F(x, y) = 0$ *holds for only finitely many integers* $x, y \in \mathbb{Z}$*.*

Example 2.7. *Consider the irreducible polynomial* $F(X, Y) = Y^2 - X^2 - 2X - 5 \in \mathbb{Z}[X, Y]$*. We then have* $d_1 = d_2 = 2$*. We find* $D(F) = \{(0, 2), (2, 0), (1, 0), (0, 0)\}$ *and* $D'(F) = \{(0, 2), (2, 0)\}$*. We therefore have* $F'(X, Y) = Y^2 - X^2 = (Y + X)(Y - X)$*. We see that* F' *is not an integer times a power of an irreducible polynomial, so by Runge's Theorem we conclude that* $F(x, y) = 0$ *holds for only finitely many integers* $x, y \in \mathbb{Z}$*.*

The proof of Runge's Theorem uses Newton-Puiseux's Theorem and a corollary of the Symmetric Function Theorem. These are the main subject of Sections 4 and 5 respectively. The formal statement of Runge's Theorem with a proof can be found in Section 6. A rough sketch is as follows. We take an irreducible polynomial $F(X, Y) \in \mathbb{Z}[X, Y]$ and assume that there are infinitely many integral solutions to $F(x, y) = 0$. Newton-Puiseux's Theorem then shows the existence of a function $f(X)$, that is a generalization of a analytic function, such that $F(X, f(X)) = 0$ and such that there are infinitely many integer solutions to the equation $y = f(x)$. We then use the information from f, our corollary of the Symmetric Function Theorem, and some linear algebra to construct another polynomial $P(X, Y) \in \mathbb{Z}[X, Y]$ such that $P(X, f(X)) = 0$. In order to prove that $P(X, f(X)) = 0$, we first choose variables in such a way that $|P(x, f(x))| < 1/2$ holds for all $x \in \mathbb{C}$. For this we use the notion of convergence among others. As a consequence we then find that F is a factor of P in $\mathbb{Z}[X, Y]$ and this lets us find some properties of F regarding its terms. These properties of F are then the result of Runge's Theorem.
Runge's Theorem only says something about polynomials in two variables. In this thesis we investigate if we can generalize Runge's Theorem to polynomials in three variables. For this we first need to know what this generalization might say. Section 7 introduces the Zariski Topology and shows a strong connection between the notion of Zariski Density and the property of having finitely or infinitely many integer solutions to a given Diophantine equation in two variables. Because the notion of Zariski Density also exists for Diophantine Equations in three (or more) variables, we will try to generalize Runge's Theorem in such a way that we can relate this to Zariski Density. Two attempts at generalizing Runge's Theorem have been made by the author. These are Theorem 8.23 and Theorem 8.28. Theorem 8.28 can be used to show that the integer solutions to certain polynomials in three variables are not Zariski-dense within the zero locus of this corresponding polynomial. For example let $F(X, Y, Z) \coloneqq Z^2 - (X^2Y + 1)(X + Y)^2Y \in$

$\mathbb{Z}[X, Y, Z]$. We then can use Theorem 8.28 to see that the set of integer solutions to the Diophantine equation

$$F(x, y, z) = 0$$

is not Zariski-dense in the zero locus of F. The same can be said about $F'(X, Y, Z) = Z^3 - X^2Y(XY^5+2)$. These two examples are Example 8.30 and Example 8.31 respectively. Both Theorem 8.23 and Theorem 8.28 require a condition on the roots of the given polynomial that is about the notion of convergence. If conjecture 8.15 holds, it follows that these conditions automatically follow from the other conditions. So if conjecture 8.15 holds, we may know less about the given polynomial and still be able to apply these two theorems.

3 Basic Properties

3.1 Decomposition into irreducible factors

Any binary Diophantine equation can be rewritten as $F(X, Y) = 0$ for some $F(X, Y) \in \mathbb{Z}[X, Y]$. We want to know whether there are finitely or infinitely many integral solutions to such binary Diophantine equations. This subsection will show us that this problem can be reduced to the case where F is irreducible in $\mathbb{Z}[X, Y]$.
Suppose that $F \in \mathbb{Z}[X, Y]$ is the product of two polynomials $F_1, F_2 \in \mathbb{Z}[X, Y]$, so

$$F(X, Y) = F_1(X, Y)F_2(X, Y).$$

It can be seen that if $x, y \in \mathbb{C}$ satisfy either $F_1(x, y) = 0$ or $F_2(x, y) = 0$, it also satisfies $F(x, y) = 0$. Conversely, if $x, y \in \mathbb{C}$ satisfy $F(x, y) = 0$, either $F_1(x, y) = 0$ or $F_2(x, y) = 0$ must hold. In particular, we have that the integral solutions to $F(x, y) = 0$ are the integral solutions to $F_1(x, y) = 0$ combined with the integral solutions to $F_2(x, y) = 0$. We can also apply this reasoning when F is the product of multiple factors, in particular irreducible factors. This yields the following lemma:

Lemma 3.1. *Let $F \in \mathbb{Z}[X, Y]$ be a polynomial and let $F = \prod_{i=1}^{n} F_i$ be a factorization of F into $n \in \mathbb{N}$ polynomials in $\mathbb{Z}[X, Y]$. For every $x, y \in \mathbb{Z}$ we have $F(x, y) = 0$ if and only if there exists an $i \in \{1, \ldots, n\}$ such that $F_i(x, y) = 0$.*

Proof. Let $x, y \in \mathbb{Z}$. Suppose that $F(x, y) = 0$ holds. We then have that $\prod_{i=1}^{n} F_i(x, y) = 0$. Since $\mathbb{Z}$ is a domain, we must have a factor that is zero. So $F_i(x, y) = 0$ for some i. Conversely, if $F_i(x, y) = 0$, then $F(x, y) = \prod_{i=1}^{n} F_i(x, y) = 0$, as one of the factors of $F(x, y)$ is zero. □

Corollary 3.2. *Let $F \in \mathbb{Z}[X, Y]$ be a nonconstant polynomial. Then it has infinitely many integral solutions to $F(x, y) = 0$ if and only if there exists an irreducible polynomial $G \in \mathbb{Z}[X, Y]$ that is a factor of F in $\mathbb{Z}[X, Y]$ and has infinitely many solutions to $G(x, y) = 0$, where $x, y \in \mathbb{Z}$.*

Proof. Suppose that such G exists. As any solution to $G(x, y) = 0$ is also a solution to $F(x, y) = 0$, we get that F also has infinitely many integer solutions to $F(x, y) = 0$. Conversely, suppose that every irreducible factor G of F has only finitely many integer solutions to $G(x, y) = 0$. We let $F = \prod_{i=1}^{n} F_i$ be a factorization of irreducible polynomials in $\mathbb{Z}[X, Y]$. So every F_i has only finitely many integer solutions to $F_i(x, y) = 0$. According to the previous lemma, the (finite) union of these finite solutions, which also is of finite size, must consists of all the integer solutions to $F(x, y) = 0$. So F has only finitely many integer solutions to $F(x, y) = 0$. □

This corollary implies that in order to figure out if a polynomial $F \in \mathbb{Z}[X, Y]$ has infinitely or only finitely many integral solutions to $F(x, y) = 0$, we just need to check this for its irreducible factors. So we have reduced the problem to figuring out if irreducible polynomials have infinitely or only finitely many integral solutions. Later on we will use this fact and demand F to be irreducible.

3.2 Small solutions

In this subsection we will focus on 'small' solutions to binary Diophantine Equations. Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a polynomial in two variables. We can write F uniquely as

$$F(X,Y) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{i,j} X^i Y^j, \quad a_{i,j} \in \mathbb{Z}, \tag{3.1}$$

for the smallest possible $m, n \in \mathbb{N}_0$. For convenience in a later stage, we will also define $a_{i,j} = 0$ for all $i, j \in \mathbb{Q}$ that do not satisfy $i \in \{0, \ldots, m\}$ and $j \in \{0, \ldots, n\}$.

Lemma 3.3. *Let $F \in \mathbb{Z}[X,Y]$ be a polynomial and let $a \in \mathbb{Z}$ be an integer. There are infinitely many integers $b \in \mathbb{Z}$ such that $F(a,b) = 0$, if and only if $X - a$ is a factor of F in $\mathbb{Z}[X,Y]$.*

Proof. First suppose that $X-a$ is a factor of F. So there exists a polynomial $G(X,Y) \in \mathbb{Z}[X,Y]$ such that $F(X,Y) = (X-a) \cdot G(X,Y)$. We see that any $b \in \mathbb{Z}$ satisfies $F(a,b) = (a-a) \cdot G(a,b) = 0 \cdot G(a,b) = 0$. So indeed there are infinitely many integers $b \in \mathbb{Z}$ such that $F(a,b) = 0$. Conversely, suppose that there are infinitely many integers $b \in \mathbb{Z}$ such that $F(a,b) = 0$ holds. We have that $F(a,Y)$ lies in $\mathbb{Z}[Y]$. Since every nonzero polynomial in $\mathbb{Z}[Y]$ has only finitely many roots, we must have that $F(a,Y) = 0$ in $\mathbb{Z}[Y]$. We can write F as in (3.1) and by switching the summands, we get

$$F(X,Y) = \sum_{j=0}^{n} \left(\sum_{i=0}^{m} a_{i,j} X^i \right) Y^j.$$

From this we get that

$$0 = F(a,Y) = \sum_{j=0}^{n} \left(\sum_{i=0}^{m} a_{i,j} a^i \right) Y^j \in \mathbb{Z}[Y].$$

So the coefficient of Y^j in $F(a,Y)$ must be zero for every $j \in \{0, \ldots, n\}$. So we have for every such j that $\sum_{i=0}^{m} a_{i,j} a^i = 0$. So a is a root of the polynomial $\sum_{i=0}^{m} a_{i,j} X^i \in \mathbb{Z}[X]$. Because of this, we know that $X - a$ is a factor of $\sum_{i=0}^{m} a_{i,j} X^i$, hence also a factor of $\sum_{i=0}^{m} a_{i,j} X^i Y^j$. Since this holds for any $j \in \{0, \ldots, n\}$, we see that $X - a$ is also a factor of $F(X,Y)$. □

Corollary 3.4. *Let $F \in \mathbb{Z}[X,Y]$ and suppose that for every $a \in \mathbb{Z}$ we have that $F(X,Y)$ is not a multiple of the polynomial $X - a$. Let $R \in \mathbb{R}$ be any real number. Then there are only finitely many $x, y \in \mathbb{Z}$ such that $F(x,y) = 0$ and $|x| \leq R$ both hold.*

Proof. There are only finitely many $x \in \mathbb{Z}$ such that $|x| \leq R$ holds. We take such an element x. It has been given that $X - x$ is not a divisor of F, so by Lemma 3.3 we have only finitely many possibilities for $y \in \mathbb{Z}$ such that $F(x,y) = 0$ holds. From this we conclude that the equation $F(x,y) = 0$ has only finitely many integral solutions where $|x| \leq R$. □

From this we can see that any nonzero irreducible polynomial $F \in \mathbb{Z}[X,Y]$ that is not of the form $X - a$ for some $a \in \mathbb{Z}$ has only finitely many solutions $x, y \in \mathbb{Z}$ to $F(x,y) = 0$ such that x is 'small'. This corollary lets us choose what we consider small by changing R. By symmetric reasoning, if F is not of the form $Y - b$ for some $b \in \mathbb{Z}$, then it has only finitely many solutions $x, y \in \mathbb{Z}$ to $F(x,y) = 0$ such that y is small.

3.3 Newton dots, degrees and height

In this subsection we will introduce some concepts that we can apply to binary Diophantine equations. We will start with the Newton dots.

Definition 3.5. *Let $F \in \mathbb{Z}[X,Y]$ and write it as in* (3.1)*. We let*

$$D(F) \coloneqq \{(i,j) \in \{0, \ldots, m\} \times \{0, \ldots, n\} \mid a_{i,j} \neq 0\}$$

be the set of (two-dimensional) indices corresponding to nonzero coefficients of F. We will call these indices Newton dots of F and will sometimes view them as coordinates on the xy-plane.

For example, $D(X^3 + X^2Y^2 - 3XY^2 + 2Y + 2) = \{(3,0),(2,2),(1,2),(0,1),(0,0)\}$. These Newton dots have been plotted in Fig. 3.3. Note that $D(F) = \emptyset$ exactly when $F = 0$.

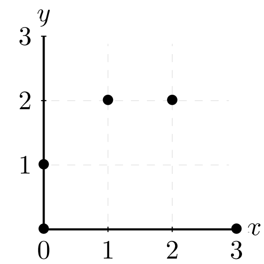


Figure 1: The Newton dots of the polynomial $X^3 + X^2Y^2 - 3XY^2 + 2Y + 2$.

Remark 3.6. *Any $F \in \mathbb{Z}[X,Y]$ can now uniquely be written as*

$$F = \sum_{(i,j)\in D(F)} a_{i,j} X^i Y^j, \tag{3.2}$$

where $a_{i,j} \in \mathbb{Z}$ is nonzero for each $(i,j) \in D(F)$. For convenience, we let $a_{i,j} = 0$ for every pair of rational numbers $(i,j) \in \mathbb{Q}^2$ with $(i,j) \notin D(F)$.

Next, we define the X-degree and the Y-degree of polynomials in $\mathbb{Z}[X,Y]$.

Definition 3.7. *Let $F \in \mathbb{Z}[X,Y]$ be a nonzero polynomial written as in* (3.1)*. We call $\deg_X F := m$ the degree of F in X and $\deg_Y F := n$ the degree of F in Y. We also call $\deg_X F$ the X-degree of F and $\deg_Y F$ the Y-degree of F. If F is the zero polynomial, we say that $\deg_X 0 = -\infty$ and that $\deg_Y 0 = -\infty$.*

If for example $F(X,Y) = 3X^2 - 2X^2Y^2 + XY^3 + Y - 1$, then $\deg_X F = m = 2$ and $\deg_Y F = n = 3$. The following definition is about the coefficients of polynomials.

Definition 3.8. *Let $F \in \mathbb{Z}[X,Y]$ and write it as in* (3.1)*. We call*

$$ht(F) := \max_{(i,j)\in\mathbb{Z}^2} |a_{i,j}|$$

the height of F, which is the maximum absolute value of its coefficients.

For example, if $F = 2X^2 - 5X + Y^2 - 4$, then $\mathrm{ht}(F) = \max\{|2|, |-5|, |1|, |-4|\} = 5$. Also, we have $\mathrm{ht}(0) = 0$.

We now have introduced enough concepts to prove that a certain family of polynomials has, as a Diophantine equation, only finitely many integer solutions.

Lemma 3.9. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a polynomial that is not the multiple of a polynomial of the form $X - a$ or $Y - b$ for any $a, b \in \mathbb{Z}$. Write F as in* (3.1)*. Suppose that $a_{m,n}$ is nonzero, then there are only finitely many integers $x, y \in \mathbb{Z}$ such that $F(x,y) = 0$ holds.*

Proof. Suppose that F is a constant in Z. We then have $F = a_{0,0} = a_{m,n} \neq 0$. Example 2.1 then directly shows us that there are no (hence only finitely many) integers $x, y \in \mathbb{Z}$ such that $F(x,y) = 0$ holds. Now suppose that F is not constant. Let $h := \mathrm{ht}(F)$ and $R := h \cdot ((m+1)(n+1) - 1)$. Note that $R \geq 1$ holds. Suppose that $x, y \in \mathbb{Z}$ satisfy $F(x,y) = 0$. Suppose additionally that $|x| > R$ and that $|y| > R$. We then have the following inequality for each $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$ with $(i,j) \neq (m,n)$:

$$|a_{i,j}x^iy^j| = |a_{i,j}||x|^i|y|^j \leq h|x|^i|y|^j \leq \begin{cases} h|x|^m|y|^{n-1} < hR^{-1}|x|^m|y|^n, & \text{if } 0 \leq j \leq n-1 \\ h|x|^{m-1}|y|^n < hR^{-1}|x|^m|y|^n, & \text{if } 0 \leq i \leq m-1 \end{cases}$$

By using the triangular inequality, we then find:

$$\begin{aligned}
|x|^m|y|^n &\leq |-a_{m,n}||x|^m|y|^n \\
&= |-a_{m,n}x^my^n| \\
&= |\sum_{i=0}^{m}\sum_{j=0}^{n-1} a_{i,j}x^iy^j + \sum_{i=0}^{m-1} a_{i,n}x^iy^n| \qquad \text{(This follows from } F(x,y)=0.\text{)} \\
&\leq \sum_{i=0}^{m}\sum_{j=0}^{n-1} |a_{i,j}x^iy^j| + \sum_{i=0}^{m-1} |a_{i,n}x^iy^n| \\
&< \sum_{i=0}^{m}\sum_{j=0}^{n-1} hR^{-1}|x|^m|y|^n + \sum_{i=0}^{m-1} hR^{-1}|x|^m|y|^n \\
&= (m+1)(n)(hR^{-1}|x|^m|y|^n) + m(hR^{-1}|x|^m|y|^n) \\
&= ((m+1)(n+1)-1)hR^{-1}|x|^m|y|^n \\
&= |x|^m|y|^n.
\end{aligned}$$

This leads to a contradiction, so there are no solutions $x, y \in \mathbb{Z}$ to $F(x,y) = 0$ with both $|x| > R$ and $|y| > R$. By Corollary 3.4, there are only finitely many integer solutions to $F(x,y) = 0$ where $|x| \leq R$. By symmetric reasoning, there are also only finitely many integer solutions to $F(x,y) = 0$ where $|y| \leq R$. We therefore conclude that there are indeed only finitely many possibilities for $x, y \in \mathbb{Z}$ such that $F(x,y) = 0$ holds. □

We used two important arguments in the previous lemma. First of all we have deduced that, for x and y big enough, we have a term that outweighs all other terms in size, from which we see that the terms can't add up to zero. Second of all we have used that there are only finitely many integer solutions to $F(x,y) = 0$ where either x and/or y is not big. Because of this, we want to find the terms of a polynomial that tend to outweigh other terms as x and y get big. This asks for the following notion of leading terms:

Definitions 3.10. *Let $F \in \mathbb{Z}[X,Y]$ be written as in (3.1). Let $\lambda \in \mathbb{R}_{>0}$. First, we call*

$$\deg_\lambda F := \max_{(i,j)\in D(F)} (i + \lambda j)$$

the λ-degree of F. Here we use the convention $\max\emptyset = -\infty$. In particular we call $\deg F := \deg_1 F$ the total degree of F. We then let

$$D_\lambda(F) := \{(i,j) \in D(F) \mid i + \lambda j = \deg_\lambda F\}$$

describe the indices of the (nonzero) terms of F with this maximal λ-degree. Finally, the λ-leading part of F, denoted by F_λ, is defined to be the sum over the terms of F with indices in $D_\lambda(F)$. So

$$F_\lambda := \sum_{(i,j)\in D_\lambda(F)} a_{i,j}X^iY^j.$$

Note that $D_\lambda(F) = \emptyset$ holds if and only if $F = 0$ and that we have $D(F_\lambda) = D_\lambda(F)$ for any $F \in \mathbb{Z}[X,Y]$ and any $\lambda \in \mathbb{R}_{>0}$.

Definitions 3.11. *Let $F \in \mathbb{Z}[X,Y]$. We define $\tilde{D}(F) := \bigcup_{\lambda\in\mathbb{R}_{>0}} D_\lambda(F)$. We then define the leading part of F, denoted by $\tilde{F}$, to be the sum over all the terms of F that are also terms in the λ-leading part of F for some $\lambda \in \mathbb{R}_{>0}$. So*

$$\tilde{F} := \sum_{(i,j)\in\tilde{D}(F)} a_{i,j}X^iY^j.$$

Example 3.12. *Let $F = X^2 + 2X - 3Y - 1$. We then find that $D(F) = \{(2,0),(1,0),(0,1),(0,0)\}$. For $\lambda = 2$ we find that $\deg_\lambda F = \max(2+2\cdot 0, 1+2\cdot 0, 0+2\cdot 1, 0+2\cdot 0) = 2$, that $D_\lambda(F) = \{(2,0),(0,1)\}$, and that $F_\lambda = X^2 - 3Y$. For $0 < \lambda < 2$ we find that $\deg_\lambda F = \max(2,1,\lambda,0) = 2$, that $D_\lambda(F) = \{(2,0)\}$, and that $F_\lambda = X^2$. For $2 < \lambda$ we find that $\deg_\lambda F = \max(2,1,\lambda,0) = \lambda$, that $D_\lambda(F) = \{(0,1)\}$, and that $F_\lambda = -3Y$. So we get $\tilde{D}(F) = \bigcup_{\lambda\in\mathbb{R}_{>0}} D_\lambda(F) = \{(2,0),(0,1)\}$. So we find $\tilde{F} = X^2 - 3Y$.*

Example 3.13. *Let* $F = Y^4 + XY^4 + X^2Y^3 + X^3Y^2 + XY + X^4$*. The Newton dots of* F *are given by* $D(F) = \{(0,4),(1,4),(2,3),(3,2),(1,1),(4,0)\}$*. We find that*

$$\deg_\lambda F = \max(4\lambda, 1+4\lambda, 2+3\lambda, 3+2\lambda, 1+\lambda, 4) = \begin{cases} 1+4\lambda & \text{if } 1 \le \lambda \\ 2+3\lambda & \text{if } \lambda = 1 \\ 3+2\lambda & \text{if } \frac{1}{2} \le \lambda \le 1 \\ 4 & \text{if } \lambda \le \frac{1}{2} \end{cases}$$

and therefore find the following:

	$D_\lambda(F)$	F_λ
$1 < \lambda$	$\{(1,4)\}$	XY^4
$\lambda = 1$	$\{(1,4),(2,3),(3,2)\}$	$XY^4 + X^2Y^3 + X^3Y^2$
$\frac{1}{2} < \lambda < 1$	$\{(3,2)\}$	X^3Y^2
$\lambda = \frac{1}{2}$	$\{(3,2),(4,0)\}$	$X^3Y^2 + X^4$
$\lambda < \frac{1}{2}$	$\{(4,0)\}$	X^4

So we have $\tilde{D}(F) = \bigcup_{\lambda \in \mathbb{R}_{>0}} D_\lambda(F) = \{(1,4),(2,3),(3,2),(4,0)\}$ *and* $\tilde{F} = XY^4 + X^2Y^3 + X^3Y^2 + X^4$*.*

Remark 3.14. *We can also visualize these examples on the xy-plane as can be seen in Fig. 3.14. We draw a point on coordinate* (i,j) *exactly when* $(i,j) \in D(F)$*. For any* $\lambda \in \mathbb{R}_{>0}$*, a point* $(i,j) \in D(F)$ *lies in* $D_\lambda(F)$*, exactly when no drawn points lie strictly above the line* $x + \lambda y = i + \lambda j$*. Furthermore, a point* $(i,j) \in D(F)$ *lies in* $\tilde{D}(F)$ *exactly when there is a line through* (i,j) *that is neither horizontal nor vertical, such that no drawn points lie strictly above this line. The slope of such a line corresponds to a* $\lambda \in \mathbb{R}_{>0}$ *such that* $(i,j) \in D_\lambda(F)$*.*

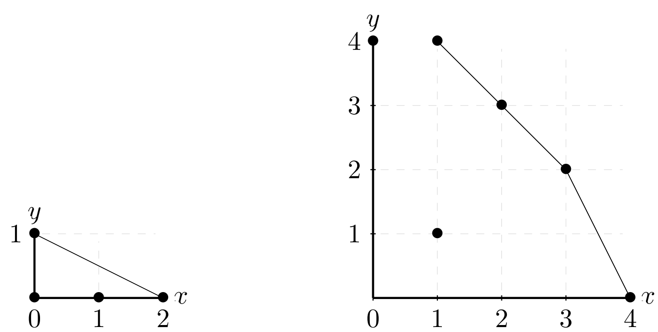


Figure 2: The Newton dots and slopes of the polynomials from Examples 3.12 and 3.13 respectively.

The λ-degree and the λ-leading part of binary polynomials satisfy a certain addition and product rule. These rules are given and proven in the following two lemmas.

Lemma 3.15. *Let* $F = G + H$*, with* $G, H \in \mathbb{Z}[X,Y]$*. Let* $\lambda \in \mathbb{R}_{>0}$*. Then* $\deg_\lambda F \le \max(\deg_\lambda G, \deg_\lambda H)$*. Additionally, if* $\deg_\lambda H < \deg_\lambda G$*, then* $F_\lambda = G_\lambda$*.*

Proof. We write $G = \sum_{(i,j)\in D(G)} b_{i,j} X^i Y^j$ and $H = \sum_{(i,j)\in D(H)} c_{i,j} X^i Y^j$ as in (3.2). We get

$$\begin{aligned} F &= G + H \\ &= \sum_{(i,j)\in D(G)} b_{i,j} X^i Y^j + \sum_{(i,j)\in D(H)} c_{i,j} X^i Y^j \\ &= \sum_{(i,j)\in D(G)\cup D(H)} (b_{i,j} + c_{i,j}) X^i Y^j. \end{aligned}$$

Suppose that there exists $(i,j) \in D(F)$. It then follows from the equation above that $(i,j) \in D(G)$ or $(i,j) \in D(H)$. Therefore $i + \lambda j \leq \deg_\lambda G$ or $i + \lambda j \leq \deg_\lambda H$. So $i + \lambda j \leq \max(\deg_\lambda G, \deg_\lambda H)$, hence

$$\deg_\lambda F = \max_{(i,j)\in D(F)} (i + \lambda j) \leq \max(\deg_\lambda G, \deg_\lambda H).$$

Now assume $\deg_\lambda H < \deg_\lambda G$. From this is follows that G is nonzero, so $D_\lambda(G)$ is nonempty. Let $(i,j) \in D_\lambda(G)$. Because $\deg_\lambda H < \deg_\lambda G = i + \lambda j$, we have $(i,j) \notin D(H)$. So the coefficient of X^iY^j in F must be $b_{i,j}$. Since this is nonzero we have $(i,j) \in D(F)$ and therefore get

$$\deg_\lambda G = i + \lambda j \leq \deg_\lambda F \leq \max(\deg_\lambda G, \deg_\lambda H) = \deg_\lambda G,$$

hence $\deg_\lambda F = \deg_\lambda G$. From this we also find $D_\lambda(F) = D_\lambda(G)$, so

$$F_\lambda = \sum_{(i,j)\in D_\lambda(F)} (b_{i,j} + c_{i,j})X^iY^j = \sum_{(i,j)\in D_\lambda(G)} (b_{i,j} + 0)X^iY^j = G_\lambda.$$

□

Lemma 3.16. *Let* $F = GH$, *with* $G, H \in \mathbb{Z}[X,Y]$. *Let* $\lambda \in \mathbb{R}_{>0}$. *Then* $\deg_\lambda F = \deg_\lambda G + \deg_\lambda H$. *Additionally, we have that* $F_\lambda = G_\lambda H_\lambda$.

Proof. Note that this is trivial if either $G = 0$ or $H = 0$ holds, so we will assume both to be nonzero. We write $G = \sum_{(i,j)\in D(G)} b_{i,j}X^iY^j$ and $H = \sum_{(i,j)\in D(H)} c_{i,j}X^iY^j$ as in (3.2). We then have

$$F = GH = \sum_{(i,j)\in D(G)} \sum_{(i',j')\in D(H)} (b_{i,j}c_{i',j'})X^{i+i'}Y^{j+j'}. \tag{3.3}$$

If $(i,j) \in D(G)$ and $(i',j') \in D(H)$, then $i + \lambda j \leq \deg_\lambda G$ and $i' + \lambda j' \leq \deg_\lambda H$. So $i + i' + \lambda(j + j') \leq \deg_\lambda G + \deg_\lambda H$. We see from (3.3) that any $(s,t) \in D(F)$ must satisfy $s = i + i'$ and $t = j + j'$ for at least one combination of pairs $(i,j) \in D(G)$, $(i',j') \in D(H)$. Such (s,t) then must satisfy $s + \lambda t \leq \deg_\lambda G + \deg_\lambda H$, hence

$$\deg_\lambda F = \max_{(s,t)\in D(F)} (s + \lambda t) \leq \deg_\lambda G + \deg_\lambda H. \tag{3.4}$$

We are now interested in the coefficients of exactly all terms of the form X^sY^t in F, where $s,t \in \mathbb{N}_0$ satisfy $s + \lambda t = \deg_\lambda G + \deg_\lambda H$. Suppose that $(i,j) \in D(G)$ and $(i',j') \in D(H)$. Also suppose that $(i,j) \notin D_\lambda(G)$, then

$$i + i' + \lambda(j + j') = i + \lambda j + i' + \lambda j' \leq i + \lambda j + \deg_\lambda H < \deg_\lambda G + \deg_\lambda H.$$

So in this case, the term $(b_{i,j}c_{i',j'})X^{i+i'}Y^{j+j'}$ does not contribute to any coefficient of our interest. The same can be said when $(i',j') \notin D_\lambda(H)$. So the coefficients in F that are of our interest are the same as the corresponding coefficients in the polynomial

$$\sum_{(i,j)\in D_\lambda(G)} \sum_{(i',j')\in D_\lambda(H)} (b_{i,j}c_{i',j'})X^{i+i'}Y^{j+j'} = G_\lambda H_\lambda.$$

For any $(p,q) \in D(G_\lambda H_\lambda)$ we have that the coefficient of X^pY^q in $G_\lambda H_\lambda$ is nonzero and therefore $(p,q) = (i + i', j + j')$ must hold for at least one combination of pairs $(i,j) \in D_\lambda(G), (i',j') \in D_\lambda(H)$. So $p + \lambda q = (i + i') + \lambda(j + j') = \deg_\lambda G + \deg_\lambda H$. So F has the same coefficient at X^pY^q as $G_\lambda H_\lambda$, which is nonzero, hence $\deg_\lambda G + \deg_\lambda H = p + \lambda q \leq \deg_\lambda F$. This combined with (3.4) gives our desired equality $\deg_\lambda F = \deg_\lambda G + \deg_\lambda H$
We have also found that the coefficient of X^pY^q in $G_\lambda H_\lambda$ is zero when $p + \lambda q \neq \deg_\lambda G + \deg_\lambda H$. As this also holds for F_λ, we see that all the coefficients in F_λ and $G_\lambda H_\lambda$ are pairwise the same. We therefore can conclude that $F_\lambda = G_\lambda H_\lambda$. □

3.4 Monomials

In this subsection we will focus on polynomials that consist of only one term.

Definition 3.17. *Let $n \in \mathbb{N}$ and let R be a ring. Let $F \in R[X_1, \ldots, X_n]$. We call F a monomial if it can be written as $F = aX_1^{p_1} \cdots X_n^{p_n}$ for some nonzero $a \in R$ and with $p_1, \ldots, p_n \in \mathbb{N}_0$.*

Notice that $F \in \mathbb{Z}[X, Y]$ is a monomial exactly when $D(F)$ contains precisely one element.

Lemma 3.18. *Let $n \in \mathbb{N}$ and let $F \in \mathbb{Z}[X_1, \ldots, X_n]$ be a monomial. Then any factor of F in $\mathbb{Z}[X_1, \ldots, X_n]$ is again a monomial.*

Proof. Let $G \in \mathbb{Z}[X_1, \ldots, X_n]$ be such a factor. Note that $\mathbb{Z}[X_1, \ldots, X_n]$ is a unique factorization domain. So G is a unit in $\mathbb{Z}[X_1, \ldots, X_n]$ (so a unit in $\mathbb{Z}$) times the product of zero or more irreducible factors of F. Since F is a monomial, we can write it as $X_1^{p_1} \cdots X_n^{p_n}$ for some nonzero $a \in \mathbb{Z}$ and with $p_1, \ldots, p_n \in \mathbb{N}_0$. We see that the irreducible factors of F are the prime divisors of a together with X_i (whenever $p_i > 0$) for any $i \in \{1, \ldots, n\}$. This shows us that $G = bX_1^{q_1} \cdots X_n^{q_n}$ for some nonzero $b \in \mathbb{Z}$ that divides a and for some $q_1, \ldots, q_n \in \mathbb{N}_0$ such that $q_i \leq p_i$ hold for all $i \in \{1, \ldots, n\}$. So G is indeed a monomial. ☐

Lemma 3.19. *Let $F \in \mathbb{Z}[X, Y]$. Suppose that F_λ is a monomial for all $\lambda \in \mathbb{R}_{>0}$. Then $\tilde{F}$ is a monomial as well.*

Proof. As there are only finitely many elements in $D(F)$, there are only finitely many elements in $\tilde{D}(F) \subset D(F)$. So there exist $\lambda_1 < \lambda_2 < \ldots < \lambda_r \in \mathbb{R}_{>0}$ for some smallest $r \in \mathbb{N}$, such that $\tilde{D}(F) = \bigcup_{i=1}^{r} D_{\lambda_i}(F)$. If $r = 1$, we immediately see that $\tilde{F} = F_{\lambda_1}$ is a monomial. We will prove by contradiction that the case $r \geq 2$ never happens. So suppose that $r \geq 2$. For each $t \in \{1, \ldots, r\}$ we have that $D_{\lambda_t}(F)$ consists of only one element (u_t, v_t) since F_{λ_t} is a monomial. Let $s, t \in \mathbb{N}$ satisfy $1 \leq s < t \leq r$. We have by minimality of r that $(u_s, v_s) \neq (u_t, v_t)$. So we have the inequalities $u_t + \lambda_s v_t < \deg_{\lambda_s}(F) = u_s + \lambda_s v_s$ and $u_s + \lambda_t v_s < \deg_{\lambda_t}(F) = u_t + \lambda_t v_t$. This yields

$$\lambda_s(v_t - v_s) < u_s - u_t < \lambda_t(v_t - v_s). \tag{3.5}$$

As $0 < \lambda_s < \lambda_t$, we have $v_t - v_s > 0$. We divide (3.5) by $v_t - v_s$ and find

$$\lambda_s < \frac{u_s - u_t}{v_t - v_s} < \lambda_t. \tag{3.6}$$

We also have found that $v_1 < v_2 < \ldots < v_r$. We define

$$\gamma \coloneqq \frac{u_1 - u_2}{v_2 - v_1}. \tag{3.7}$$

Then it follows from (3.6) that $0 < \lambda_1 < \gamma < \lambda_2$. Multiplying (3.7) by $v_2 - v_1$ and reordering terms gives us $u_1 + \gamma v_1 = u_2 + \gamma v_2$. This shows that if $D_\gamma(F)$ contains either (u_1, v_1) or (u_2, v_2), it must also contain the other, which is impossible since F_γ is a monomial. So $F_\gamma(F)$ can contain neither and therefore there must exist $(u, v) \in D_\gamma(F)$ with $u + \gamma v > u_2 + \gamma v_2$. Since $D_\gamma(F) \subset \tilde{D}(F)$, there is a $t \in \{1, \ldots, r\}$ with $(u, v) \in D_{\lambda_t}(F)$ and thus $(u, v) = (u_t, v_t)$. We already saw $t \neq 1, 2$, so $2 < t$. From (3.6) we derive

$$\gamma < \lambda_2 < \frac{u_2 - u_t}{v_t - v_2}.$$

We multiply by $v_t - v_2$ and reorder terms to get

$$u_t + \gamma v_t < u_2 + \gamma v_2,$$

which contradicts $(u_t, v_t) \in D_\gamma(F)$. ☐

Lemma 3.20. *Let $F \in \mathbb{Z}[X, Y]$ be a nonzero polynomial of X-degree $d_1 \in \mathbb{N}_0$ and Y-degree $d_2 \in \mathbb{N}_0$. Suppose that $\tilde{F}$ is a monomial. Then $(d_1, d_2) \in D(F)$ holds.*

Proof. Suppose $(d_1, d_2) \notin D(F)$. Since F is of X-degree d_1, then there must exist a highest $v \in \{1, \ldots, d_2 - 1\}$ such that $(d_1, v) \in D(F)$. Analogously there must exist a highest $u \in \{1, \ldots, d_1 - 1\}$ such that $(u, d_2) \in D(F)$. Let $\lambda > d_1$, then for any $(i, j) \in D(F)$ we have

$$i + \lambda j \leq \begin{cases} u + \lambda d_2 & \text{if } j = d_2. \\ d_1 + \lambda(d_2 - 1) < \lambda d_2 \leq u + \lambda d_2 & \text{if } j \leq d_2 - 1. \end{cases}$$

This shows that $(u, d_2) \in D_\lambda(F)$ when $\lambda > d_1$. In a similar way we can show that $(d_1, v) \in D_\lambda(F)$ when $\lambda < d_2^{-1}$. This shows that $(d_1, v), (u, d_2) \in \tilde{D}(F) = D(\tilde{F})$ and $\tilde{F}$ is therefore not a monomial, a contradiction. We conclude that $(d_1, d_2) \in D(F)$. □

4 Newton–Puiseux's Theorem

The proof of Runge's Theorem uses a theorem that says that a certain field extension of the polynomial field over an algebraically closed field is again algebraically closed. This theorem is called the Newton-Puiseux Theorem. In this section we will describe this field extension and prove this theorem.

4.1 Formal power series

There is a version of the proof that talks about formal power series and a version that talks about convergent power series. A proof of the formal case has been given by Bassel Mannaa and Thierry Coquand [10]. We will follow this proof throughout this subsection.

Note that each polynomial $F(X, Y) \in \mathbb{Z}[X, Y]$ can be viewed as a polynomial in Y with coefficients in $\mathbb{Z}[X]$. So we can view $\mathbb{Z}[X, Y]$ as $\mathbb{Z}[X][Y]$. We will often apply this fact. We begin by introducing some definitions and some easy lemmas about generalizations of polynomials.

A formal power series generalizes the concept of a polynomial in the sense that is does not require a highest exponent.

Definition 4.1. *Let R be a ring. A formal power series over R is of the form $\sum_{i=0}^{\infty} a_i X^i$, where the coefficients a_i lie in R. We call $R[[X]]$ the set of all these formal power series over R.*

Remark 4.2. *$R[[X]]$ is a ring, with canonical addition and multiplication. As every polynomial is a formal power series over R (where all but finitely many coefficients are zero), we may view $R[X]$ as a subset of $R[[X]]$.*

Lemma 4.3. *Let R be a ring. Let $f = \sum_{i=0}^{\infty} a_i X^i \in R[[X]]$ be a formal power series over R. Then f is a unit in $R[[X]]$ if and only if a_0 is a unit in R.*

Proof. First suppose that f is indeed a unit and that $g = \sum_{i=0}^{\infty} b_i X^i \in R[[X]]$ is its inverse. Then $a_0 b_0$, the constant term of fg, must be 1 and a_0 is thus indeed a unit in R. Conversely, suppose that a_0 is a unit in R. Let $b_0 = a_0^{-1} \in R$ and inductively define $b_t = -a_0^{-1} \sum_{i=1}^{t} a_i b_{t-i} \in R$ for $t \in \mathbb{N}$. The constant term of fg is $a_0 b_0 = 1$. For any $t \in \mathbb{N}$, the coefficient of X^t in fg is

$$\sum_{i=0}^{t} a_i b_{t-i} = a_0 b_t + \sum_{i=1}^{t} a_i b_{t-i} = a_0(-a_0^{-1} \sum_{i=1}^{t} a_i b_{t-i}) + \sum_{i=1}^{t} a_i b_{t-i} = 0,$$

which shows that $fg = 1$. Therefore, f is a unit in $R[[X]]$. □

A formal Laurent series generalizes the concept of a formal power series in the sense that some exponents may be negative.

Definition 4.4. *Let R be a ring. A formal Laurent series over R is of the form $\sum_{i=m}^{\infty} a_i X^i$, where the coefficients a_i lie in R and $m \in \mathbb{Z}$. We call $R((X))$ the set of all these formal Laurent series over R.*

Remark 4.5. *The set $R((X))$ is in fact a ring, with canonical addition and multiplication. If k is a field, then so is $k((X))$. In fact $k((X))$ is the quotient field of $k[[X]]$.*

Definition 4.6. *Let k be a field. We denote the subfield of $k((X))$ that is the fraction field of $k[X]$ by $k(X)$.*

Definition 4.7. *Let R be a ring. Let $f = \sum_{i=0}^{n} a_i X^i$ be a polynomial in $R[X]$ and let $f' = \sum_{i=1}^{n} i a_i X^{i-1}$ be its derivative. We call f separable (over R) if there exist $r, s \in R[X]$ such that*

$$rf + sf' = 1.$$

Lemma 4.8. *Let R be a ring and $f = gh$, with $g, h \in R[X]$. If f is separable over R, then so are g and h.*

Proof. If f is separable, we have $rf + sf' = 1$ for some $r, s \in R[X]$. By the product rule of derivatives, we find $f' = gh' + g'h$. This gives us

$$1 = rf + sf' = rgh + s(gh' + g'h) = (rg + sg')h + (sg)h'.$$

So h is separable. By symmetry of g and h, we have that g must be separable as well. □

Lemma 4.9. *Let R be a ring, $u \in R$ a unit, $a \in R$ any element of R and $f(X) \in R[X]$. If $f(X)$ is separable over R, then we have that $f(uY - a) \in R[Y]$ is separable over R as well.*

Proof. If f is separable, we have

$$r(X)f(X) + s(X)f'(X) = 1$$

for some $r(X), s(X) \in R[X]$. We take the derivative of $f(uY - a)$ with respect to Y and by the chain rule we find $f(uY - a)' = uf'(uY - a)$. We now have

$$r(uY - a)f(uY - a) + u^{-1}s(uY - a)f(uY - a)' = r(uY - a)f(uY - a) + s(uY - a)f'(uY - a) = 1.$$

Since $r(uY - a) \in R[Y]$ and $u^{-1}s(uY - a) \in R[Y]$, we conclude that $f(uY - a) \in R[Y]$ is separable over R. □

Lemma 4.10. *Let k be a field of characteristic zero. Let $f \in k[Y]$ be an irreducible polynomial. Then f is separable over k.*

Proof. Let $f' \in k[Y]$ be the derivative of f and let $h \in k[Y]$ be the greatest common divisor of f and f'. This means that there are $r, s \in k[Y]$ such that $rf + sf' = h$. If h is constant, we can divide by h (as h is never zero) and get $(h^{-1}r)f + (h^{-1}s)f' = 1$, which shows that f is separable. Now suppose that h is nonconstant. Since h divides f, there must exist $g \in k[Y]$ such that $f = gh$. Since f is irreducible, g must be a unit, hence a constant in k. This shows that the degree of h is the same as f and thus bigger than the degree of f'. This contradicts the fact that h divides g. □

Lemma 4.11. *Let R be a ring. We have that $R[[X]][Y] \subset R[Y][[X]]$ holds.*

Proof. If $F(X, Y) \in R[[X]][Y]$, we can write it as $F(X, Y) = \sum_{i=0}^{n} a_i(X)Y^i$, for some $n \in \mathbb{N}_0$ and $a_i(X) \in R[[X]]$. Any $a_i(X)$ can then be written as $\sum_{j=0}^{\infty} a_{i,j} X^j$, with $a_{i,j} \in R$, so we find

$$F(X, Y) = \sum_{i=0}^{n} \sum_{j=0}^{\infty} a_{i,j} X^j Y^i = \sum_{j=0}^{\infty} \sum_{i=0}^{n} a_{i,j} Y^i X^j = \sum_{j=0}^{\infty} \left(\sum_{i=0}^{n} a_{i,j} Y^i \right) X^j,$$

which indeed shows that $F(X, Y) \in R[Y][[X]]$. □

Lemma 4.12. *Let R be a ring. Let $F(X, Y) \in R[Y][[X]]$ be nonzero. We can write $F = \sum_{i=0}^{\infty} a_i(Y)X^i$, where $a_i(Y) \in R[Y]$. Suppose there exists a lowest $n \in \mathbb{N}$ such that all polynomials $a_i(Y)$ have degree at most n, then F lies in $R[[X]][Y]$ and has degree n as a polynomial in Y.*

Proof. Suppose that this lowest n exists. Then we can write $a_i(Y) = \sum_{j=0}^{n} a_{i,j} Y^j$ for each i. Since n is the lowest, there must exist $s \in \mathbb{N}_0$ such that $a_s(Y)$ has degree n, so $a_{s,n}$ is nonzero. We get

$$F(X, Y) = \sum_{i=0}^{\infty} \sum_{j=0}^{n} a_{i,j} Y^j X^i = \sum_{j=0}^{n} \left(\sum_{i=0}^{\infty} a_{i,j} X^i \right) Y^j,$$

which shows that $F(X, Y) \in R[[X]][Y]$. The degree of F in Y is n, as $\sum_{i=0}^{\infty} a_{i,n} X^i$ is nonzero since $a_{s,n}$ is nonzero.

□

An important part of the proof of Newton-Puiseux's Theorem is Hensel's Lemma. There exist many different versions and formulations of Hensel's Lemma. We will use two versions. The first one can be applied to formal power series and is the following lemma. The second version of Hensel's Lemma 4.44 can be applied only to convergent power series. Recall that a polynomial in one variable over a ring is called *monic* if its leading coefficient is one.

Lemma 4.13 (Hensel's Lemma)**.** *Let R be a ring and $F(X,Y) \in R[[X]][Y]$ a monic polynomial in Y of degree $n \geq 2$, with coefficients in $R[[X]]$. Suppose that there exist $G_0, H_0 \in R[Y]$ of nonzero degrees r, s respectively, such that $F(0,Y) = G_0H_0$ and such that $r + s = n$. Suppose that there also exist $G^*, H^* \in R[Y]$ such that $G_0H^* + H_0G^* = 1$ holds. Then there exist $G, H \in R[[X]][Y]$, whose degree as polynomials in Y are r, s respectively, such that $F = GH$.*

Proof. By Lemma 4.11, F also lies in $R[Y][[X]]$, so we can write it as $F(X,Y) = \sum_{i=0}^{\infty} F_i(Y)X^i$, where $F_i \in R[Y]$. Since F is monic of degree n in Y, we can see that $\deg(F_i) < n$ for all $i > 0$ and $\deg(F_0) = n$. We define $G \coloneqq \sum_{i=0}^{\infty} G_i(Y)X^i$ and $H \coloneqq \sum_{i=0}^{\infty} H_i(Y)X^i$, where we still need to define $G_i(Y), H_i(Y) \in R[Y]$ for all $i > 0$. We want to have $F = GH$, so for each $q \in \mathbb{N}_0$, we want that both F and GH have the same coefficient at X^q. This gives the equation $F_q = \sum_{i=0}^{q} G_iH_{q-i}$ for each such q. This already holds for $q = 0$, as $F_0 = F(0,Y) = G_0H_0$. To make sure that this also holds for higher q, we will define G_q and H_q by induction on q. In this process we will also see that the degrees of G_q and H_q will be strictly smaller than r and s, respectively. Let $q \in \mathbb{N}$. Suppose that we have defined G_i and H_i for all $1 \leq i < q$, such that $\deg(G_i) < r$ and $\deg(H_i) < s$. Let

$$U_q \coloneqq F_q - \sum_{i=1}^{q-1} G_iH_{q-i}.$$

Since $\deg(F_q) < n$ holds and

$$\deg(G_iH_{q-i}) \leq \deg(G_i) + \deg(H_{q-i}) < r + s = n$$

holds for every $1 \leq i < q$, we find that $\deg(U_q) < n$. By using Euclidean division on polynomials, we find that there exist polynomials E_q and H_q in $R[Y]$, with $\deg(H_q) < \deg(H_0) = s$, such that $U_qH^* = E_qH_0 + H_q$. We then have

$$\begin{aligned}
U_q &= U_q \cdot 1 \\
&= U_q(G_0H^* + H_0G^*) \\
&= G_0U_qH^* + U_qH_0G^* \\
&= G_0(E_qH_0 + H_q) + U_qH_0G^* \\
&= H_0(G_0E_q + U_qG^*) + G_0H_q.
\end{aligned}$$

This gives

$$\deg(H_0(G_0E_q + U_qG^*)) = \deg(U_q - G_0H_q) < \max(\deg(U_q), \deg(G_0H_q)) < n.$$

Because F is monic in Y, so is $F_0 = F(0,Y) = G_0H_0$. Since $\deg(G_0) + \deg(H_0) = \deg(F_0)$, we see that the leading coefficient of H_0 is a unit in R and thus not a zero divisor in R. Therefore we have

$$\deg(H_0(G_0E_q + U_qG^*)) = \deg(H_0) + \deg(G_0E_q + U_qG^*).$$

We let $G_q \coloneqq E_qG_0 + G^*U_q$ and then find

$$r + \deg(G_q) = \deg(H_0) + \deg(G_0E_q + U_qG^*) = \deg(H_0(G_0E_q + U_qG^*)) < n,$$

hence $\deg(G_q) < n - r = s$. So we have defined H_q, G_q such that $\deg(G_q) < r$ and $\deg(H_q) < s$.

Moreover, we have

$$\begin{aligned}F_q &= U_q + \sum_{i=1}^{q-1} G_i H_{q-i} \\ &= H_0(G_0 E_q + U_q G^*) + G_0 H_q + \sum_{i=1}^{q-1} G_i H_{q-i} \\ &= H_0 G_q + G_0 H_q + \sum_{i=1}^{q-1} G_i H_{q-i} \\ &= \sum_{i=0}^{q} G_i H_{q-i}.\end{aligned}$$

So in our process of defining G and H, we saw that F and GH have the same coefficient at X^q for every q. Therefore $F = GH$. Also, by Lemma 4.12, G and H lie in $R[[X]][Y]$ and have degree r, s as a polynomial in Y. □

The proof of Newton-Puiseux's Theorem has been divided in three parts. The first part applies Hensel's Lemma and is as follows.

Lemma 4.14. *Let k be an algebraically closed field of characteristic zero. Let $F(X,Y) \in k[[X]][Y]$ be a polynomial in Y over $k[[X]]$, which is monic of degree $n \geq 2$ and separable over $k((X))$. Then there exist $m \in \mathbb{N}$ and $G, H \in k[[T]][Y]$ both nonconstant as polynomial over Y, such that*

$$F(T^m, Y) = G(T,Y)H(T,Y).$$

Proof. We can write $F(X,Y) = \sum_{i=0}^{n-1} a_i(X)Y^i + Y^n$, where $a_i(X) \in k[[X]]$.
We start with the special case where $a_{n-1}(X)$ is zero and where $a_t(0)$ is nonzero for some $0 \leq t < n-1$. This gives us $F(0,Y) = \sum_{i=0}^{n-2} d_i Y^i + Y^n$, where $d_i = a_i(0)$. In particular we see that d_t is nonzero. Let $a \in k$ be a root of $F(0,Y) \in k[Y]$. Suppose that a is the only root, then we must have $F(0,Y) = (Y-a)^n$. As d_t is nonzero, a can not be zero, but if a is nonzero, we see that the coefficient of Y^{n-1} in $(Y-a)^n$ is $\pm n a^{n-1}$, which is nonzero and contradicts the fact that the coefficient of Y^{n-1} in $F(0,Y)$ is zero. So there must be other roots apart from a. Suppose that a is a root of multiplicity p. So $(Y-a)^p$ divides $F(0,Y)$, hence their exists nonconstant $H \in k[Y]$ such that $F(0,Y) = (Y-a)^p H(Y)$ where $Y-a$ is not a factor of H. Therefore $(Y-a)^p$ and $H(Y)$ share no common factors, hence their greatest common divisor is 1. This assures the existence of $H^*, G^* \in k[Y]$ such that $(Y-a)^p H^* + HG^* = 1$. We can apply Hensel's Lemma 4.13 and indeed find $G, H \in k[[X]][Y]$ both nonconstant as polynomial over Y, such that $F(X,Y) = G(X,Y)H(X,Y)$. So in this case the Lemma holds (we have $m = 1$). Note that in this case we have not used the fact that F is separable.
We now drop the assumption that $a_t(0)$ is nonzero for some $0 \leq t < n-1$. We can write $a_i(X) = \sum_{j=0}^{\infty} a_{i,j} X^j$, with $a_{i,j} \in k$. Consider the set $\{(n-i)/j \mid 0 \leq i < n-1, j \in \mathbb{N}, a_{i,j} \neq 0\}$. The set cannot be empty as this would imply $F(X,Y) = Y^n$, which contradicts the fact that F is separable. The set has a largest element: there are only finitely many possibilities for $i < n-1$ and for any such i with $a_i(X)$ nonzero, we see that $(n-i)/j' > (n-i)/j$ holds exactly when $j' < j$. So we take the lowest $j' \in \mathbb{N}$ with $a_{i',j'} \neq 0$ corresponding to an i' with $0 \leq i' < n-1$ and $a_{i'}(X)$ nonzero, such that $(n-i')/j'$ is the largest. Now let $m = n - i'$ and $p = j'$. Since

$$F(X,Y) = \sum_{i=0}^{n-2} a_i(X)Y^i + Y^n = \sum_{i=0}^{n-2}\sum_{j=0}^{\infty} a_{i,j}X^j Y^i + Y^n,$$

we get

$$F(T^m, T^p Z) = \sum_{i=0}^{n-2}\sum_{j=0}^{\infty} a_{i,j} T^{jm+ip} Z^i + T^{pn} Z^n.$$

For any $a_{i,j}$ nonzero, we have $(n-i)/j \leq m/p$, so $pn - pi \leq mj$, hence $pn \leq jm + ip$. This means that T^{np} divides $a_{i,j}T^{jm+ip}Z^i$ and therefore also $F(T^m, T^p Z)$. This shows that there exists a factorization

$$F(T^m, T^p Z) = T^{np}\left(\sum_{i=0}^{n-2} c_i(T) Z^i + Z^n\right)$$

for some $c_i \in k[T]$. Also, we have $c_{i'}(0) = a_{i',j'} \neq 0$. So we can apply the special case to the polynomial $\sum_{i=0}^{n-2} c_i(T)Z^i + Z^n$ to know that there exist $G_1(T,Z), H_1(T,Z) \in k[[T]][Z]$ both nonconstant as polynomials over Z, such that

$$\sum_{i=0}^{n-2} c_i(T)Z^i + Z^n = G_1(T,Z)H_1(T,Z). \tag{4.1}$$

Let $l, q \in \mathbb{N}_0$ be the Z-degree of $G_1(T,Z)$ and $H_1(T,Z)$ respectively. Then we see from (4.1) that $l + q = n$. Also we find that if $G(T,Y) \coloneqq T^{lp}G_1(T,Y/T^p)$ and $H(T,Y) \coloneqq T^{qp}H_1(T,Y/T^p)$, then G and H both lie in $k[[T]][Y]$ and are both nonconstant as polynomials over Y. This gives us the desired equality

$$G(T,Y)H(T,Y) = T^{lp+qp}G_1(T,Y/T^p)H_1(T,Y/T^p) = T^{np}\left(\sum_{i=0}^{n-2} c_i(T)(Y/T^p)^i + (Y/T^p)^n\right) = F(T^m,Y).$$

At last we will now look at the general case where $a_{n-1}(X)$ may be nonzero. We consider $F(X, Z - a_{n-1}/n) \in k[[X]][Z]$. According to Lemma 4.9, this is also separable over $k((X))$. A quick inspection shows that this polynomial is also monic of degree n when viewed as a polynomial over Z, and that its coefficient at Z^{n-1} is zero. So we can apply the less general case and find $F(T^m, Z - a_{n-1}/n) = G_1(T,Z)H_(T,Z)$ for some $G_1, H_1 \in k[[T]][Z]$ nonconstant as polynomials over Z. We can take $G(T,Y) = G_1(T, Y + a_{n-1}/n)$ and $H(T,Y) = H_1(T, Y + a_{n-1}/n)$ and find

$$G(T,Y)H(T,Y) = G_1(T,Y+a_{n-1}/n)H_1(T,Y+a_{n-1}/n) = F(T^m,(Y+a_{n-1}/n)-a_{n-1}/n) = F(T^m,Y),$$

which finishes our proof. □

The previous lemma showed that $F(T^m, Y)$ is the factorization of two nonconstant polynomials over $k[[T]]$ for a suitable $m \in \mathbb{N}$. The following lemma uses this to show that $F(T^m, Y)$ can actually be fully factorized into linear polynomials over $k[[T]]$ for a (perhaps different) suitable $m \in \mathbb{N}$.

Lemma 4.15. *Let k be an algebraically closed field of characteristic zero. Let $F(X,Y) \in k[[X]][Y]$ be a polynomial in Y over $k[[X]]$, which is monic of degree $n \geq 1$ and separable over $k((X))$. Then there exist $m \in \mathbb{N}$ and $f_i(T) \in k[[T]]$ such that*

$$F(T^m, Y) = \prod_{i=1}^{n} (Y - f_i(T)).$$

Proof. Note that this is trivial for $n = 1$. We will prove this by induction on n. Suppose that the lemma holds for all $n_0 < n$. By Lemma 4.14 we see that there exist $m_0 \in \mathbb{N}$ and $G, H \in k[[T_0]][Y]$ both nonconstant as polynomials over Y, such that

$$F(T_0^{m_0}, Y) = G(T_0,Y)H(T_0,Y). \tag{4.2}$$

Let r, s be the degrees in Y of G and H, respectively, let a be the coefficient of Y^r in G and b the coefficient of Y^s in H. Since F is monic, we get from (4.2) that $r + s = n$ and that $1 = ab$. We will replace G by bG and H by aH and see that (4.2) still holds and that additionally G and H are monic. Since $F(X,Y)$ is separable over $k((X))$, $F(T_0^{m_0}, Y)$ must be separable over $k((T_0^{m_0}))$ and therefore also over the bigger field $k((T_0))$. By Lemma 4.8, we find that G and H are also separable over $k((T_0))$. Since $r, s < n$, we can use the induction hypothesis to find $m_1, m_2 \in \mathbb{N}$ and $g_i(T_1) \in k[[T_1]]$ and $h_i(T_2) \in k[[T_2]]$ such that $G(T_1^{m_1}, Y) = \prod_{i=1}^{r} (Y - g_i(T_1))$ and $H(T_2^{m_2}, Y) = \prod_{i=1}^{s} (Y - h_i(T_2))$ hold. We now take $m = m_0m_1m_2$ and we take $f_i(T) = g_i(T^{m_2})$ for $i \leq r$ and $f_i(T) = h_{i-r}(T^{m_1})$ for $i > r$. We then indeed find

$$\begin{aligned}
F(T^m, Y) &= F((T^{m_1 m_2})^{m_0}, Y) \\
&= G(T^{m_1 m_2}, Y)H(T^{m_1 m_2}, Y) \\
&= G((T^{m_1})^{m_2}, Y)H((T^{m_2})^{m_1}, Y) \\
&= \prod_{i=1}^{r} (Y - g_i(T^{m_2})) \prod_{i=1}^{s} (Y - h_i(T^{m_1})) \\
&= \prod_{i=1}^{r} (Y - f_i(T)) \prod_{i=1}^{s} (Y - f_{i+r}(T)) \\
&= \prod_{i=1}^{n} (Y - f_i(T)).
\end{aligned}$$

□

Definition 4.16. *Let k be a field. We call $k((X^*)) \coloneqq \bigcup_{m \in \mathbb{N}} k((X^{1/m}))$ the field of formal Puiseux series (in X) over k.*

We are now ready to prove Newton-Puiseux's Theorem.

Theorem 4.17 (Newton-Puiseux's theorem)**.** *Let k be an algebraically closed field of characteristic zero. Then $k((X^*))$ is also algebraically closed.*

Proof. First, let $F \in k((X^*))[Y]$ be a nonconstant polynomial in Y over $k((X^*))$. We need to show that this polynomial contains a root in $k((X^*))$. Since it contains a root exactly when one of its irreducible components does, we may assume F to be irreducible. We can write $F(X, Y) = \sum_{i=0}^{n} a_i(X)Y^i$, with $a_i(X) \in k((X^*))$. We may replace F by $(a_n(X))^{-1}F$, as multiplying by a unit preserves roots and irreducibility. We will therefore assume that F is monic, so $a_n(X) = 1$. Since

$$a_i(X) \in k((X^*)) = \bigcup_{m \in \mathbb{N}} k((X^{1/m})),$$

there are m_i such that $a_i(X) \in k((X^{1/m_i}))$. If we now take $m = \prod_{i=0}^{n} m_i$, we get that $a_i(X) \in k((X^{1/m_i})) \subset k((X^{1/m}))$. Let

$$G(T, Y) = F(T^m, Y). \tag{4.3}$$

Then we see that G lies in $k((T))[Y]$. Note that $G(T, Y)$ is monic. It is also irreducible in $k((T))[Y]$, since F is irreducible in $k((X^*))[Y]$. We can write $G(T, Y) = \sum_{i=1}^{n} g_i(T)Y^i$. We let $b(T) \in k[[T]]$ be the product of the denominator of each $g_i(T)$. We now find that

$$G(T, (b(T))^{-1}Z) = (b(T))^{-n}H(T, Z), \tag{4.4}$$

with $H(T, Z) \coloneqq \sum_{i=1}^{n} g_i(T)b(T)^{n-i}Z^i$. Since the denominator of $g_i(T)$ divides $b(T)$, we get $g_i(T)b(T)^{n-i} \in k[[T]]$ for each i, so $H(T, Z) \in k[[T]][Z]$. Lemma 4.10 tells us that G is separable over $k((T))$. Lemma 4.9 then tells us that $G(T, (b(T))^{-1}Z) \in k((T))[Z]$ is separable over $k((T))$, hence by Lemma 4.8, so is $H(T, Z)$. As $H(T, Z)$ is also monic and nonconstant, we can use Lemma 4.15 to see that there exist $l \in \mathbb{N}$ and $c(S) \in k[[S]]$ such that $H(S^l, c(S)) = 0$. This combined with (4.4) gives $G(S^l, b(S^l)^{-1}c(S)) = 0$. This combined with (4.3) gives $F(S^{lm}, b(S^l)^{-1}c(S)) = 0$. We replace S by $X^{1/lm}$ and get

$$F(X, b(X^{1/m})^{-1}c(X^{1/lm})) = 0.$$

Since $b(X^{1/m})^{-1}c(X^{1/lm})$ lies in $k((X^{1/lm}))$, it also lies in $k((X^*))$, so F does have a root that lies in $k((X^*))$. We therefore see that $k((X^*))$ is indeed algebraically closed.

□

4.2 Convergent power series

The subfield consisting of convergent Puiseux series over the complex numbers is also algebraically closed. That is the second version of the Newton/Puiseux Theorem that we will prove. In order to do this, we need to have a different version of Hensel's Lemma. A proof of this has been given (in German) by H. Grauert and R. Remmert[5]. In this subsection we will show this proof, after which we will prove that the named subfield is indeed algebraically closed.

We start by introducing some definitions and easy lemmas that will be needed for this version of Hensel's Lemma.

Definition 4.18. *Let k be a field. We will call a function $|.| : k \to \mathbb{R}$ a norm when it satisfies the following properties for all $a, b \in k$:*

- $|a| \geq 0$*, and* $|a| = 0$ *exactly when* $a = 0$.
- $|ab| = |a| \cdot |b|$.
- $|a + b| \leq |a| + |b|$.

If k has a norm, then k is a metric space where the metric is given by $d(a, b) \coloneqq |a - b|$.

Example 4.19. *Any subfield of the complex numbers $\mathbb{C}$ has the canonical norm given by $|a + bi| = \sqrt{a^2 + b^2}$.*

Definition 4.20. *Let k be a field with a norm. We will call k a complete field if k is complete as a metric space , i.e. when every Cauchy sequence with elements in k has a limit in k.*

Example 4.21. *The fields $\mathbb{R}$ and $\mathbb{C}$ are complete fields with their canonical norms.*

Recall that for any ring R, the set of formal power series over R, called $R[[X]]$, is again a ring. So we can also look at $R[[X_1]][[X_2]]$, the ring of formal power series over $R[[X_1]]$. We can repeat this to get the ring $R[[X_1]][[X_2]] \cdots [[X_n]]$ for any $n \in \mathbb{N}$. Its elements can be written in the form

$$\sum_{i_1=0}^{\infty} \sum_{i_2=0}^{\infty} \cdots \sum_{i_n=0}^{\infty} a_{i_1,i_2,\ldots,i_n} X_1^{i_1} X_2^{i_2} \cdots X_n^{i_n},$$

with $a_{i_1,i_2,\ldots,i_n} \in R$. We will write this shorthanded as

$$\sum_{\boldsymbol{i} \in \mathbb{N}_0^n} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}.$$

Note that for any $f(X_1, \ldots, X_n) \in R[[X_1]][[X_2]] \cdots [[X_n]]$ and any permutation $X_{\sigma_1}, \ldots, X_{\sigma_n}$ of the variables $X_1, \ldots, X_n$, we still have $f(X_{\sigma_1}, \ldots, X_{\sigma_n}) \in R[[X_1]][[X_2]] \cdots [[X_n]]$.

Definition 4.22. *Let R be a ring. We write $R[[X_1, \ldots, X_n]]$ for $R[[X_1]][[X_2]] \cdots [[X_n]]$ and call it the ring of formal power series over R, in n variables.*

Lemma 4.3 can be generalized in higher dimensions, as suggested by the following lemma.

Lemma 4.23. *Let R be a ring and $n \in \mathbb{N}$. Let $f = \sum_{\boldsymbol{i} \in \mathbb{N}_0^n} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in R[[X_1, \ldots, X_n]]$. We have that f is a unit in $R[[X_1, \ldots, X_n]]$ if and only if $a_{\boldsymbol{0}} = a_{(0,\ldots,0)}$ is a unit in R.*

Proof. We prove this by induction on n. We have already proved the case where $n = 1$ in Lemma 4.3. Now suppose the statement holds for $n - 1$. Note that as $R[[X_1, \ldots, X_n]] = R[[X_1, \ldots, X_{n-1}]][[X_n]]$, we can view f as a power series in X_n over the ring $R[[X_1, \ldots, X_{n-1}]]$. We can thus write $f = \sum_{i=0}^{\infty} b_i X_n^i$, with $b_i \in R[[X_1, \ldots, X_{n-1}]]$. By Lemma 4.3, f is a unit in $R[[X_1, \ldots, X_{n-1}]][[X_n]]$ if and only if b_0 is a unit in $R[[X_1, \ldots, X_{n-1}]]$. Note that $b_0 = \sum_{\boldsymbol{j} \in \mathbb{N}_0^{n-1}} d_{\boldsymbol{j}} \boldsymbol{X'}^{\boldsymbol{j}}$ with $d_{\boldsymbol{j}} = d_{(j_1,\ldots,j_{n-1})} = a_{(j_1,\ldots,j_{n-1},0)}$ and $\boldsymbol{X'}^{\boldsymbol{j}} = X_1^{j_1} X_2^{j_2} \cdots X_{n-1}^{j_{n-1}}$. By the induction hypothesis, b_0 is a unit in $R[[X_1, \ldots, X_{n-1}]]$ if and only if $d_{(0,\ldots,0)}$ is a unit in R. Since $d_{(0,\ldots,0)} = a_{(0,\ldots,0)}$, we see that indeed f is a unit in $R[[X_1, \ldots, X_n]]$ if and only if $a_{(0,\ldots,0)}$ is a unit in R. □

Definition 4.24. *Let R be a ring and $n \in \mathbb{N}$. Let $f = \sum_{\boldsymbol{i} \in \mathbb{N}_0^n} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in R[[X_1, \ldots, X_n]]$. For any $j \in \mathbb{N}_0$, let*

$$f_j \coloneqq \sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^n \\ i_1 + \ldots + i_n = j}} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}.$$

We then call f_j the j-th homogeneous component of f. We can write any such f as $\sum_{j=0}^{\infty} f_j$. This series is called the sum of its homogeneous components.

Definition 4.25. *Let k be a field with a norm $|\cdot|$, and let $n \in \mathbb{N}$. Let $f = \sum_{\boldsymbol{i} \in \mathbb{N}_0^n} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in k[[X_1, \dots, X_n]]$ and let $\boldsymbol{t} = (t_1, \dots, t_n) \in \mathbb{R}^n_{\geq 0}$ be an n-tuple of nonnegative real numbers. We define the series*

$$||f||_{\boldsymbol{t}} \coloneqq \sum_{\boldsymbol{i} \in \mathbb{N}_0^n} |a_{\boldsymbol{i}}| \boldsymbol{t}^{\boldsymbol{i}} \in \mathbb{R} \cup \{\infty\}, \quad \textit{with } \boldsymbol{t}^{\boldsymbol{i}} = t_1^{i_1}, \dots, t_n^{i_n}.$$

Lemma 4.26. *Let k be a field with a norm, let $n \in \mathbb{N}$ and let $\boldsymbol{t} = (t_1, \dots, t_n) \in \mathbb{R}^n_{>0}$ be an n-tuple of positive real numbers. The function $||.||_{\boldsymbol{t}} : k[[X_1, \dots, X_n]] \to \mathbb{R} \cup \{\infty\}$ described above satisfies the following properties for all $f, g \in k[[X_1, \dots, X_n]]$ and $a \in k$:*

- *$||f||_{\boldsymbol{t}} \geq 0$, and $||f||_{\boldsymbol{t}} = 0$ exactly when $f = 0$.*
- *$||af||_{\boldsymbol{t}} = |a| \cdot ||f||_{\boldsymbol{t}}$.*
- *$||f + g||_{\boldsymbol{t}} \leq ||f||_{\boldsymbol{t}} + ||g||_{\boldsymbol{t}}$.*
- *$||fg||_{\boldsymbol{t}} \leq ||f||_{\boldsymbol{t}} \cdot ||g||_{\boldsymbol{t}}$.*

Proof. The first three properties follow directly from applying the properties of the norm of k to the definition of $||.||_{\boldsymbol{t}}$, so only the last property remains to be proved. We take $h \coloneqq fg$ and write $f = \sum_{p=0}^{\infty} f_p$, and $g = \sum_{q=0}^{\infty} f_q$ and $h = \sum_{s=0}^{\infty} h_s$, as the sum of their homogeneous components. For each $s \in \mathbb{N}_0$, we then have

$$h_s = \sum_{\substack{p,q \in \mathbb{N}_0 \\ p+q=s}} f_p g_q.$$

We write $f = \sum_{\boldsymbol{i} \in \mathbb{N}_0^n} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}$ and $g = \sum_{\boldsymbol{j} \in \mathbb{N}_0^n} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{j}}$. For each $p, q \in \mathbb{N}_0$, we find the inequality

$$\begin{aligned}
||f_p g_q||_{\boldsymbol{t}} &= ||\Bigg(\sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^n \\ i_1 + \ldots + i_n = p}} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \Bigg) \Bigg(\sum_{\substack{\boldsymbol{j} \in \mathbb{N}_0^n \\ j_1 + \ldots + j_n = q}} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{j}} \Bigg) ||_{\boldsymbol{t}} \\
&= || \sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^n \\ i_1 + \ldots + i_n = p}} \sum_{\substack{\boldsymbol{j} \in \mathbb{N}_0^n \\ j_1 + \ldots + j_n = q}} a_{\boldsymbol{i}} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{i}+\boldsymbol{j}} ||_{\boldsymbol{t}} \\
&\leq \sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^n \\ i_1 + \ldots + i_n = p}} \sum_{\substack{\boldsymbol{j} \in \mathbb{N}_0^n \\ j_1 + \ldots + j_n = q}} ||a_{\boldsymbol{i}} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{i}+\boldsymbol{j}}||_{\boldsymbol{t}} \\
&= \sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^n \\ i_1 + \ldots + i_n = p}} \sum_{\substack{\boldsymbol{j} \in \mathbb{N}_0^n \\ j_1 + \ldots + j_n = q}} |a_{\boldsymbol{i}} b_{\boldsymbol{j}}| \boldsymbol{t}^{\boldsymbol{i}+\boldsymbol{j}} \\
&= \sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^n \\ i_1 + \ldots + i_n = p}} \sum_{\substack{\boldsymbol{j} \in \mathbb{N}_0^n \\ j_1 + \ldots + j_n = q}} |a_{\boldsymbol{i}}| \cdot |b_{\boldsymbol{j}}| \boldsymbol{t}^{\boldsymbol{i}+\boldsymbol{j}} \\
&= \Bigg(\sum_{\substack{\boldsymbol{i} \in \mathbb{N}_0^n \\ i_1 + \ldots + i_n = p}} |a_{\boldsymbol{i}}| \boldsymbol{t}^{\boldsymbol{i}} \Bigg) \Bigg(\sum_{\substack{\boldsymbol{j} \in \mathbb{N}_0^n \\ j_1 + \ldots + j_n = q}} |b_{\boldsymbol{j}}| \boldsymbol{t}^{\boldsymbol{j}} \Bigg) \\
&= ||f_p||_{\boldsymbol{t}} \cdot ||g_q||_{\boldsymbol{t}}.
\end{aligned}$$

This gives us

$$\begin{aligned}
||fg||_{\boldsymbol{t}} &= \sum_{s=0}^{\infty} ||h_s||_{\boldsymbol{t}} \\
&= \sum_{s=0}^{\infty} || \sum_{\substack{p,q\in\mathbb{N}_0 \\ p+q=s}} f_p g_q ||_{\boldsymbol{t}} \\
&\leq \sum_{s=0}^{\infty} \sum_{\substack{p,q\in\mathbb{N}_0 \\ p+q=s}} ||f_p g_q||_{\boldsymbol{t}} \\
&\leq \sum_{s=0}^{\infty} \sum_{\substack{p,q\in\mathbb{N}_0 \\ p+q=s}} ||f_p||_{\boldsymbol{t}} \cdot ||g_q||_{\boldsymbol{t}} \\
&= \sum_{p=0}^{\infty} ||f_p||_{\boldsymbol{t}} \cdot \sum_{q=0}^{\infty} ||g_q||_{\boldsymbol{t}} \\
&= ||f||_{\boldsymbol{t}} \cdot ||g||_{\boldsymbol{t}},
\end{aligned}$$

which indeed shows us that $||fg||_{\boldsymbol{t}} \leq ||f||_{\boldsymbol{t}} \cdot ||g||_{\boldsymbol{t}}$. □

Definition 4.27. *Let k be a field with a norm, let $n \in \mathbb{N}$ and let $\boldsymbol{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n_{>0}$. We then let*

$$B_{\boldsymbol{t}}[[X_1, \ldots, X_n]] \coloneqq \{f \in k[[X_1, \ldots, X_n]] \mid ||f||_{\boldsymbol{t}} < \infty\}.$$

Remark 4.28. *Throughout this subsection we will let k be a complete field, $n \in \mathbb{N}$ a natural number and $\boldsymbol{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n_{>0}$. We also will use the abbreviations B for $B_{\boldsymbol{t}}[[X_1, \ldots, X_n]]$ and B' for $B_{\boldsymbol{t}'}[[X_1, \ldots, X_{n-1}]]$, where $\boldsymbol{t}' = (t_1, \ldots, t_{n-1})$.*

If $f \in B$, we can write it as $f = \sum_{i=0}^{\infty} f_i X_n^i$, with $f_i \in k[[X_1, \ldots, X_{n-1}]]$. These f_i satisfy $||f_i||_{t'} < \infty$, as the following lemma shows.

Lemma 4.29. *Let $f = \sum_{i=0}^{\infty} f_i X_n^i \in B$. Then $f_i \in B'$ for all $i \in \mathbb{N}_0$.*

Proof. We have that $||f||_{\boldsymbol{t}} = \sum_{i=0}^{\infty} ||f_i||_{\boldsymbol{t}'} t_n^i < \infty$. So for each $i \in \mathbb{N}_0$, we find $||f_i||_{\boldsymbol{t}'} t_n^i < \infty$, hence $||f_i||_{\boldsymbol{t}'} < \infty$. So $f_i \in B'$ for all $i \in \mathbb{N}_0$. □

Lemma 4.30. *We have the inclusion $B'[X_n] \subset B$.*

Proof. Let $f = \sum_{i=0}^{m} f_i X_n^i \in B'[X_n]$. As $B' \subset k[[X_1, \ldots, X_{n-1}]]$, it must hold that $f \in k[[X_1, \ldots, X_n]]$. Furthermore $f_i \in B'$ holds, hence $||f_i||_{\boldsymbol{t}'} < \infty$ holds for all f_i. So $||f_i||_{\boldsymbol{t}'} t_n^i < \infty$, hence

$$||f||_{\boldsymbol{t}} = \sum_{i=0}^{m} ||f_i||_{\boldsymbol{t}'} t_n^i < \infty,$$

which shows that $f \in B$ holds and that $B'[X_n]$ indeed lies within B. □

The properties from Lemma 4.26 show us that B is closed under multiplication and addition. We also note that $k \subset B$. In fact, B is a metric space where the metric is given by $d(f, g) = ||f - g||_{\boldsymbol{t}}$. The following lemma shows us that B is also complete as a metric space.

Lemma 4.31. *The metric space B is complete.*

Proof. Let $(f_j)_{j\in\mathbb{N}}$, with $f_j \in B$ be a Cauchy sequence. So for any real $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $||f_j - f_p||_{\boldsymbol{t}} < \varepsilon$ holds for all $j, p > N$. We write $f_j = \sum_{\boldsymbol{i}\in\mathbb{N}_0^n} a_{j,\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}$, with $a_{j,\boldsymbol{i}} \in k$. For any $j, p \in \mathbb{N}$ we have $f_j - f_p = \sum_{\boldsymbol{i}\in\mathbb{N}_0^n} (a_{j,\boldsymbol{i}} - a_{p,\boldsymbol{i}}) \boldsymbol{X}^{\boldsymbol{i}}$, hence $||f_j - f_p||_{\boldsymbol{t}} = \sum_{\boldsymbol{i}\in\mathbb{N}_0^n} |a_{j,\boldsymbol{i}} - a_{p,\boldsymbol{i}}| \boldsymbol{t}^{\boldsymbol{i}}$. Let $\boldsymbol{i} \in \mathbb{N}_0^n$. It then follows that $|a_{j,\boldsymbol{i}} - a_{p,\boldsymbol{i}}| \boldsymbol{t}^{\boldsymbol{i}} \leq ||f_j - f_p||_{\boldsymbol{t}}$. In order to show that $(a_{j,\boldsymbol{i}})_{j\in\mathbb{N}}$ is a Cauchy sequence in k, let $\varepsilon' > 0$. Let $\varepsilon \coloneqq \varepsilon' \boldsymbol{t}^{\boldsymbol{i}}$, then there exists $N \in \mathbb{N}$ such that $||f_j - f_p||_{\boldsymbol{t}} < \varepsilon$ hold for all $j, p > N$. For any such j and p, we now have $|a_{j,\boldsymbol{i}} - a_{p,\boldsymbol{i}}| \leq ||f_j - f_p||_{\boldsymbol{t}} / (\boldsymbol{t}^{\boldsymbol{i}}) < \varepsilon / (\boldsymbol{t}^{\boldsymbol{i}}) = \varepsilon'$. So $(a_{j,\boldsymbol{i}})_{j\in\mathbb{N}}$ is indeed a Cauchy sequence in k. Since k is a complete field, the limit $a_{\boldsymbol{i}} \coloneqq \lim_{j\to\infty} a_{j,\boldsymbol{i}}$ exists and lies in k. Now let $f = \sum_{\boldsymbol{i}\in\mathbb{N}_0^n} a_{\boldsymbol{i}}\boldsymbol{X}^{\boldsymbol{i}}$. We want to show that $f = \lim_{j\to\infty} f_j$ and that $f \in B$. Let $\varepsilon > 0$ be given. Then there exists $N \in \mathbb{N}$ such that

$$\sum_{\boldsymbol{i}\in\mathbb{N}_0^n} |a_{p,\boldsymbol{i}} - a_{j,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} = ||f_p - f_j||_{\boldsymbol{t}} < \varepsilon/2, \quad \text{for all } j,p > N.$$

For any $\boldsymbol{i} = (i_1,\ldots,i_n) \in \mathbb{N}_0^n$, we define $|\boldsymbol{i}| := i_1 + \ldots + i_n$. Then for all $s \in \mathbb{N}$ and $j,p > N$,

$$\begin{aligned}
\sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{\boldsymbol{i}} - a_{j,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} &= \sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{\boldsymbol{i}} - a_{p,\boldsymbol{i}} + a_{p,\boldsymbol{i}} - a_{j,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} \\
&\le \sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{\boldsymbol{i}} - a_{p,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} + \sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{p,\boldsymbol{i}} - a_{j,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} \\
&\le \sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{\boldsymbol{i}} - a_{p,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} + \sum_{\boldsymbol{i}\in\mathbb{N}_0^n} |a_{p,\boldsymbol{i}} - a_{j,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} \\
&< \sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{\boldsymbol{i}} - a_{p,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} + \varepsilon/2
\end{aligned}$$

If we choose p to be high enough, we get

$$\sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{\boldsymbol{i}} - a_{p,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} \le \varepsilon/2,$$

hence

$$\sum_{\substack{\boldsymbol{i}\in\mathbb{N}_0^n\\|\boldsymbol{i}|\le s}} |a_{\boldsymbol{i}} - a_{j,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} < \varepsilon$$

for all $s \in \mathbb{N}$ and $j > N$. We can thus take the limit for $s \to \infty$ and find

$$||f - f_j||_{\boldsymbol{t}} = \sum_{\boldsymbol{i}\in\mathbb{N}_0^n} |a_{\boldsymbol{i}} - a_{j,\boldsymbol{i}}|\boldsymbol{t}^{\boldsymbol{i}} \le \varepsilon$$

for all $j > N$. This shows that $f = \lim_{j\to\infty} f_j$. Furthermore, we have

$$||f||_{\boldsymbol{t}} = ||f - f_N + f_N||_{\boldsymbol{t}} \le ||f - f_N||_{\boldsymbol{t}} + ||f_N||_{\boldsymbol{t}} \le \varepsilon + ||f_N||_{\boldsymbol{t}} < \infty.$$

We therefore conclude that $f \in B$. So B is indeed complete as a metric space. □

Lemma 4.32. *Let $a \in \mathbb{R}_{>0}$ and let $\varepsilon \in \mathbb{R}$ with $0 \le \varepsilon < 1$. For each $j \in \mathbb{N}_0$, let $f_j \in B$ such that $||f_j||_{\boldsymbol{t}} \le a\varepsilon^j$. Then the limit $\sum_{j=0}^{\infty} f_j$ exists and lies in B.*

Proof. Note that $\sum_{j=0}^{\infty} \varepsilon^j = 1/(1-\varepsilon)$ since $|\varepsilon| < 1$. For any $l \in \mathbb{N}_0$, we define $h_l := \sum_{j=0}^{l} f_j$. Let $\varepsilon' > 0$

and take $N = \lceil \log_\varepsilon(\varepsilon'(1-\varepsilon)a^{-1}) \rceil$, where $\lceil . \rceil$ is the ceiling function. For any $l'' \geq l' > N$ we find

$$\begin{aligned} ||h_{l''} - h_{l'}||_{\mathfrak{t}} &= || \sum_{j=l'+1}^{l''} f_j ||_{\mathfrak{t}} \\ &\leq \sum_{j=l'+1}^{l''} ||f_j||_{\mathfrak{t}} \\ &\leq \sum_{j=l'+1}^{l''} a\varepsilon^j \\ &= \varepsilon^{l'+1} a \sum_{j=0}^{l''-l'-1} \varepsilon^j \\ &\leq \varepsilon^N a \sum_{j=0}^{\infty} \varepsilon^j \\ &= \varepsilon^N a \frac{1}{1-\varepsilon} \\ &\leq \varepsilon'. \end{aligned}$$

This shows that $(h_j)_{j\in\mathbb{N}_0}$ is a Cauchy sequence, and since B is complete by Lemma 4.31, we find that its limit $\sum_{j=0}^{\infty} f_j$ exists and lies in B. □

Lemma 4.33. *Let $f \in B$ such that $||1-f||_{\mathfrak{t}} < 1$. Then f is a unit in B.*

Proof. With $a = 1$ and $\varepsilon = ||1-f||_{\mathfrak{t}}$, notice that $||(1-f)^j||_{\mathfrak{t}} \leq (||1-f||_{\mathfrak{t}})^j = a\varepsilon^j$. We can apply Lemma 4.32 to find that $\sum_{j=0}^{\infty}(1-f)^j$ exists and lies in B. We find that

$$f\sum_{j=0}^{\infty}(1-f)^j = (1-(1-f))\sum_{j=0}^{\infty}(1-f)^j = \sum_{j=0}^{\infty}(1-f)^j - \sum_{j=0}^{\infty}(1-f)^{j+1} = (1-f)^0 = 1.$$

So f is indeed a unit in B. □

A big ingredient for the proof of Hensel's Lemma for convergent power series is the Weierstrass preparation theorem. This theorem follows indirectly from the Weierstrass Division Theorem, which is as follows:

Lemma 4.34 (Weierstrass Division Theorem)**.** *Let $g = \sum_{i=0}^{\infty} g_i X_n^i \in B$. Let $b \in \mathbb{N}$ and suppose that g_b is a unit in B' such that*

$$||X_n^b - g g_b^{-1}||_{\mathfrak{t}} \leq \varepsilon t_n^b,$$

for some ε that lies in the open interval $(0,1)$. Let $f = \sum_{i=0}^{\infty} f_i X_n^i \in B$. Then there exist unique $q \in B$ and $r \in B'[X_n]$ with $\deg_{X_n} r < b$, such that

$$f = qg + r.$$

We also have the inequality

$$||X_n^b g_b q - \sum_{i=b}^{\infty} f_i X_n^i||_{\mathfrak{t}} \leq \frac{\varepsilon}{1-\varepsilon}||f||_{\mathfrak{t}}. \tag{4.5}$$

If additionally f, g lie in $B'[X_n]$, with $\deg_{X_n} g = b$, then this q also lies in $B'[X_n]$ and $\deg_{X_n} q = \deg_{X_n} f - b$ (except when $\deg_{X_n} f - b < 0$, in which case we trivially have $q = 0$).

Proof. First we need some notation. For any $h = \sum_{i=0}^{\infty} h_i X_n^i \in B$, we define $\hat{h} \in B'[X_n]$ and $\tilde{h} \in B$ as follows:

$$\hat{h} := \sum_{i=0}^{b-1} h_i X_n^i, \quad \tilde{h} := \sum_{i=b}^{\infty} h_i X_n^{i-b}.$$

So we have $h = \hat{h} + X_n^b \tilde{h}$ and $\deg_{X_n} \hat{h} < b$.
Because

$$\begin{aligned} ||h||_{\boldsymbol{t}} &= \sum_{i=0}^{\infty} ||h_i||_{\boldsymbol{t}'} t_n^i \\ &= \sum_{i=0}^{b-1} ||h_i||_{\boldsymbol{t}'} t_n^i + \sum_{i=b}^{\infty} ||h_i||_{\boldsymbol{t}'} t_n^i \\ &= ||\hat{h}||_{\boldsymbol{t}} + t_n^b \sum_{i=b}^{\infty} ||h_i||_{\boldsymbol{t}'} t_n^{i-b} \\ &= ||\hat{h}||_{\boldsymbol{t}} + t_n^b ||\tilde{h}||_{\boldsymbol{t}}, \end{aligned}$$

we find that $||\hat{h}||_{\boldsymbol{t}} \leq ||h||_{\boldsymbol{t}}$ and that $||\tilde{h}||_{\boldsymbol{t}} \leq t_n^{-b} ||h||_{\boldsymbol{t}}$ must hold.

We now let $v_0 \coloneqq f$ and inductively define $v_{j+1} \coloneqq (X_n^b - g g_b^{-1}) \tilde{v}_j$ for every $j \in \mathbb{N}_0$. For any $j \in \mathbb{N}_0$, we find the inequality

$$\begin{aligned} ||v_{j+1}||_{\boldsymbol{t}} &\leq ||X_n^b - g g_b^{-1}||_{\boldsymbol{t}} \cdot ||\tilde{v}_j||_{\boldsymbol{t}} \\ &\leq \varepsilon t_n^b \cdot t_n^{-b} ||v_j||_{\boldsymbol{t}} \\ &= \varepsilon ||v_j||_{\boldsymbol{t}}. \end{aligned}$$

So by repeating this inequality, we find $||v_j||_{\boldsymbol{t}} \leq \varepsilon^j ||f||_{\boldsymbol{t}}$ for all $j \in \mathbb{N}_0$. By using Lemma 4.32 with $a = ||f||_{\boldsymbol{t}}$, we find that $\sum_{j=0}^{\infty} v_j \in B$. By applying this lemma in a similar way, we also find that

$$q \coloneqq g_b^{-1} \sum_{j=0}^{\infty} \tilde{v}_j \in B, \quad \text{and that } r \coloneqq \sum_{j=0}^{\infty} \hat{v}_j \in B'[X_n].$$

Because $\deg_{X_n} \hat{v}_j < b$ for all $j \in \mathbb{N}_0$, it follows that $\deg_{X_n} r < b$. We now find

$$f = v_0 = \sum_{j=0}^{\infty} (v_j - v_{j+1}) = \sum_{j=0}^{\infty} (\hat{v}_j + X_n^b \tilde{v}_j - (X_n^b - g g_b^{-1}) \tilde{v}_j) = \sum_{j=0}^{\infty} (g g_b^{-1} \tilde{v}_j + \hat{v}_j) = qg + r.$$

We also find the inequality

$$\begin{aligned} ||X_n^b g_b q - \sum_{i=b}^{\infty} f_i X_n^i||_{\boldsymbol{t}} &= ||X_n^b \left(\sum_{j=0}^{\infty} \tilde{v}_j - \tilde{f} \right)||_{\boldsymbol{t}} \\ &\leq ||X_n^b||_{\boldsymbol{t}} \cdot || \sum_{j=1}^{\infty} \tilde{v}_j ||_{\boldsymbol{t}} \\ &\leq t_n^b \sum_{j=1}^{\infty} ||\tilde{v}_j||_{\boldsymbol{t}} \\ &\leq t_n^b \sum_{j=1}^{\infty} t_n^{-b} ||v_j||_{\boldsymbol{t}} \\ &\leq \sum_{j=1}^{\infty} \varepsilon^j ||f||_{\boldsymbol{t}} \\ &= ||f||_{\boldsymbol{t}} \cdot \varepsilon \sum_{j=0}^{\infty} \varepsilon^j \\ &= \frac{\varepsilon}{1-\varepsilon} ||f||_{\boldsymbol{t}}. \end{aligned}$$

To prove uniqueness of q and r, we first look at the special case where $f = 0$, so $qg + r = 0$. We define $h \coloneqq g \cdot g_b^{-1} - X_n^b$. By assumption we have $||h||_{\boldsymbol{t}} \leq \varepsilon t_n^b$. We also have $g = g_b(X_n^b + h)$, so

$$q g_b X_n^b + q g_b h + r = 0,$$

hence $qg_bX_n^b + r = -qg_bh$. Note that $\deg_{X_n} r < b$, so $||qg_bX_n^b + r||_{\boldsymbol{t}} = ||qg_bX_n^b||_{\boldsymbol{t}} + ||r||_{\boldsymbol{t}}$. We therefore find

$$M := ||qg_b||_{\boldsymbol{t}} \cdot t_n^b = ||qg_bX_n^b||_{\boldsymbol{t}} \leq ||qg_bX_n^b + r||_{\boldsymbol{t}} = ||qg_bh||_{\boldsymbol{t}} \leq ||qg_b||_{\boldsymbol{t}} \cdot \varepsilon t_n^b = \varepsilon M.$$

Since $\varepsilon \in (0,1)$, we must have that $M = 0$, hence $||qg_b||_{\boldsymbol{t}} = 0$, and therefore $qg_b = 0$, hence $q = 0$ and $r = -qg = 0$. So in the special case we see that q and r are uniquely determined. For general $f \in B$ we consider $q' \in B$ and $r' \in B'[X_n]$ that satisfy both $\deg_{X_n} r' < b$ and $f = q'g + r'$. This gives $0 = f - f = (q - q')g + (r - r')$. Because we have uniqueness in the special case, we must have $q - q' = 0$ and $r - r' = 0$. So $q = q'$ and $r = r'$ and we therefore also have uniqueness in the general case.

Suppose additionally that f, g lie in $B'[X_n]$, with $\deg_{X_n} g = b$. Let $m := \deg_{X_n} f$ and suppose $m \geq b$. We now have that gg_b^{-1} is monic as a polynomial in X_n. Therefore $\deg_{X_n}(X_n^b - gg_b^{-1}) < b$. Any $h \in B'[X_n]$ satisfies $\deg_{X_n} \tilde{h} \leq \deg_{X_n} h - b$, as can be seen by the definition of $\tilde{h}$. We therefore have $\deg_{X_n} \tilde{h}(X_n^b - gg_b^{-1}) < \deg_{X_n} h$. From this we can conclude by induction that $\deg_{X_n} v_j \leq m$ and $\deg_{X_n} \tilde{v}_j \leq m - b$ for all $j \in \mathbb{N}_0$. And thus $\deg_{X_n} q \leq m - b$ as we can see by definition of q. If $\deg_{X_n} q < m - b$, then

$$m = \deg_{X_n} f = \deg_{X_n}(qg + r) \leq \min(\deg_{X_n}(qg), \deg_{X_n} r) < m,$$

which leads to a contradiction. Hence $\deg_{X_n} q = m - b$, which finishes our proof. □

The Weierstrass Division Theorem will be used twice in the following lemma.

Lemma 4.35. *Let $g = \sum_{i=0}^{\infty} g_i X_n^i \in B$. Let $b \in \mathbb{N}$ and suppose that g_b is a unit in B' such that*

$$||X_n^b - gg_b^{-1}||_{\boldsymbol{t}} \leq \varepsilon t_n^b,$$

for some ε that lies in the open interval $(0, \frac{1}{2})$. Then there exist a unit $e \in B$ and a monic polynomial $\omega \in B'[X_n]$ of degree b, such that

$$g = e \cdot \omega.$$

If additionally g lies in $B'[X_n]$, then so does e and in that case we have $\deg_{X_n} e = \deg_{X_n} g - b$.

Proof. We apply Lemma 4.34 with $g = g$ and $f = X_n^b$. We therefore know that there exist $q \in B$ and $r \in B'[X_n]$ with $\deg_{X_n} r < b$ such that

$$X_n^b = q \cdot g + r. \tag{4.6}$$

Equation (4.5) then translates translates to

$$||X_n^b g_b q - X_n^b||_{\boldsymbol{t}} \leq \frac{\varepsilon}{1-\varepsilon} t_n^b.$$

Hence

$$||g_b q - 1||_{\boldsymbol{t}} \leq \frac{\varepsilon}{1-\varepsilon} < \frac{\frac{1}{2}}{1-\frac{1}{2}} = 1.$$

We can now apply Lemma 4.33 and discover that $g_b \cdot q$ is a unit in B. Since g_b is a unit in B, so is q. We can now take this unit $e := q^{-1} \in B$ and also define $\omega := X_n^b - r$. This combined with (4.6) gives us the desired equation $g = e \cdot \omega$. Suppose additionally that g lies in $B'[X_n]$, with $\deg_{X_n} g = m$. Notice that ω is indeed a monic polynomial of degree b. We can apply Lemma 4.34 with $f = g, g = \omega$. We already know $g = e \cdot \omega + 0$, and since this pair $(e, 0)$ is unique by this lemma and since g, ω lie in $B'[X_n]$ with $\deg_{X_n} \omega = b$, we get the additional fact that $e \in B'[X_n]$ with $\deg_{X_n} e = \deg_{X_n} g - b$, which ends our proof. □

For the Weierstrass Preparation Theorem and Hensel's Lemma, we need the definition of convergent Power series.

Definition 4.36. *Let $n \in \mathbb{N}$ and let $f \in k[[X_1, \ldots, X_n]]$. We say that f is a convergent power series (in n variables) if there exists $\boldsymbol{t} \in \mathbb{R}_{>0}^n$ such that $f \in B_{\boldsymbol{t}}$. We name*

$$k[\{X_1, \ldots, X_n\}] := \bigcup_{\boldsymbol{t} \in \mathbb{R}_{>0}^n} B_{\boldsymbol{t}}$$

the set of convergent power series. We also will use the shorthand notation $K = K_n := k[\{X_1, \ldots, X_n\}]$ and $K' = K_{n-1}$.

Remark 4.37. *It can be quickly seen that the above set K_n is in fact a subring of $k[[X_1, \dots, X_n]]$. For any $f \in K_n$, we can choose to write $f = \sum_{\boldsymbol{i} \in \mathbb{N}_0^n} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}$, with $a_{\boldsymbol{i}} \in k$. We can also choose to write $f = \sum_{i=0}^{\infty} a_i X_n^i$, with $a_i \in k[[X_1, \dots, X_{n-1}]]$. Note that in this last form, since $f \in B_{\boldsymbol{t}}$ for some $\boldsymbol{t} \in \mathbb{R}_{>0}^n$, we have $a_i \in B'_{\boldsymbol{t}}$, hence $a_i \in K_{n-1}$.*

Lemma 4.38. *Let $f \in k[[X_1, \dots, X_n]]$ be a convergent power series. Then there exists a positive real number $R \in \mathbb{R}_{>0}$ such that the series $f(x_1, \dots, x_n)$ converges in k for all $x_1, \dots, x_n \in k$ with $|x_j| < R$ for all $j \in \{1, \dots, n\}$.*

Proof. Since $f = \sum_{(i_1,\dots,i_n) \in \mathbb{N}_0^n} a_{i_1,\dots,i_n} X_1^{i_1} \cdots X_n^{i_n}$ is convergent, there exists $\boldsymbol{t} = (t_1, \dots, t_n) \in \mathbb{R}_{>0}^n$ such that $||f||_{\boldsymbol{t}} < \infty$. Now take $R \coloneqq \min(t_1, \dots, t_n)$ and let $x_1, \dots, x_n \in k$ with $|x_j| < R$ for all $j \in \{1, \dots, n\}$. We have for any $(i_1, \dots, i_n) \in \mathbb{N}_0^n$ that

$$|a_{i_1,\dots,i_n} x_1^{i_1} \cdots x_n^{i_n}| = |a_{i_1,\dots,i_n}| \cdot |x_1|^{i_1} \cdots |x_n|^{i_n} < |a_{i_1,\dots,i_n}| t_1^{i_1} \cdots t_n^{i_n}.$$

This shows us that

$$\sum_{(i_1,\dots,i_n) \in \mathbb{N}_0^n} |a_{i_1,\dots,i_n} x_1^{i_1} \cdots x_n^{i_n}| < \sum_{(i_1,\dots,i_n) \in \mathbb{N}_0^n} |a_{i_1,\dots,i_n}| t_1^{i_1} \cdots t_n^{i_n} = ||f||_{\boldsymbol{t}} < \infty.$$

Since k is complete, we conclude that $f(x_1, \dots, x_n) = \sum_{(i_1,\dots,i_n) \in \mathbb{N}_0^n} a_{i_1,\dots,i_n} x_1^{i_1} \cdots x_n^{i_n}$ converges in k. ☐

Definition 4.39. *Let $g = \sum_{i=0}^{\infty} g_i X_n^i \in K$. Let $b \in \mathbb{N}$. We say that g is X_n-generic of order b if $g_i(\mathbf{0}) = 0$ for all $i < b$ and $g_b(\mathbf{0}) \neq 0$.*

Note that $g_i(\mathbf{0}) = g_i(0, \dots, 0)$ is the constant coefficient in $g_i(X_1, \dots, X_{n-1}) \in k[[X_1, \dots, X_{n-1}]]$, when we view g_i as a formal power series in $n-1$ variables.

Lemma 4.40. *Let $g = \sum_{i=0}^{\infty} g_i X_n^i \in K$. Let $b \in \mathbb{N}$. Suppose that g is X_n-generic of order b. Let $\varepsilon > 0$. Then there exist $\delta_n > 0$ and a function $\delta : (0, \delta_n) \to (0, \delta_n)$ such that the following holds for any $t = (t_1, \dots, t_n) \in \mathbb{R}_{>0}^n$ with $t_n < \delta_n$ and $t_1, \dots, t_{n-1} < \delta(t_n)$:*

- *g lies in $B_{\boldsymbol{t}}$.*
- *g_b is a unit in $B'_{\boldsymbol{t}}$.*
- *$||X_n^b - g \cdot g_b^{-1}||_{\boldsymbol{t}} < \varepsilon t_n^b$.*

Proof. Since g is convergent, it must lie in some B_s for some $s \in \mathbb{R}_{>0}^n$. Note that any $u \in \mathbb{R}_{>0}^n$ satisfying $u_i < s_i$ for all i also satisfies $||g||_u < ||g||_s$ and thus $g \in B_u$. Also note that for any $h(X_1, \dots, X_n) \in K$ we have $\lim_{\boldsymbol{t} \to \mathbf{0}} ||h||_{\boldsymbol{t}} = h(\mathbf{0})$. As the constant term of $g_b(\mathbf{0})^{-1} g_b - 1$ is zero, we find $\lim_{\boldsymbol{t} \to \mathbf{0}} ||g_b(\mathbf{0})^{-1} g_b - 1||_{\boldsymbol{t}} = 0$. So we can take a positive $\eta < \min_{i \in \{1,\dots,n\}} s_i$ that is small enough such that

$$||g_b(\mathbf{0})^{-1} g_b - 1||_{\boldsymbol{t}} < 1$$

for all $\boldsymbol{t} \in \mathbb{R}_{>0}^n$ with $t_1, \dots, t_n < \eta$. For any such $\boldsymbol{t}$ we have $g_b(\mathbf{0})^{-1} g_b \in B_{\boldsymbol{t}}$ and we then can apply Lemma 4.33 to conclude that $g_b(\mathbf{0})^{-1} g_b$ is a unit in $B_{\boldsymbol{t}}$. So g_b is a unit in $B_{\boldsymbol{t}}$. We can now consider $g g_b^{-1} = \sum_{i=0}^{\infty} g_b^{-1} g_i X_n^i$. We let $c_i = g_i g_b^{-1}$ for all $i \in \mathbb{N}_0$. In particular we have $c_b = 1$. We can choose δ_n with $0 < \delta_n < \eta$ small enough such that

$$||\sum_{i=b+1}^{\infty} c_i X_n^i||_{\boldsymbol{t}} = \sum_{i=b+1}^{\infty} ||c_i||_{\boldsymbol{t}'} \cdot t_n^i = t_n^b \sum_{i=b+1}^{\infty} ||c_i||_{\boldsymbol{t}'} \cdot t_n^{i-b} \leq \frac{\varepsilon}{2} t_n^b$$

holds for all $\boldsymbol{t} \in \mathbb{R}_{>0}^n$ with $t_1, \dots, t_n < \delta_n$. Now note that $c_i(\mathbf{0}) = g_i(\mathbf{0}) g_b^{-1}(\mathbf{0}) = 0 g_b^{-1}(\mathbf{0}) = 0$ for all $i < b$. So $\lim_{\boldsymbol{t}' \to \mathbf{0}} ||c_i||_{\boldsymbol{t}'} = 0$ for such i. This means that we can find $\delta(t_n) \in \mathbb{R}$ with $0 < \delta(t_n) < \delta_n$, such that

$$||c_i||_{\boldsymbol{t}'} < \frac{1}{b} \frac{\varepsilon}{2} t_n^{b-i}$$

for all $\boldsymbol{t} \in \mathbb{R}_{>0}^n$ with $t_n < \delta_n$ and $t_1, \dots, t_{n-1} < \delta(t_n)$. For the same $\boldsymbol{t}$, this leads to

$$||\sum_{i=0}^{b-1} c_i X_n^i||_{\boldsymbol{t}} = \sum_{i=0}^{b-1} ||c_i||_{\boldsymbol{t}'} \cdot t_n^i \leq \sum_{i=0}^{b-1} \frac{1}{b} \frac{\varepsilon}{2} t_n^{b-i} t_n^i = \frac{\varepsilon}{2} t_n^b.$$

We now have for all $\boldsymbol{t} \in \mathbb{R}^n_{>0}$ with $t_n < \delta_n$ and $t_1, \ldots, t_{n-1} < \delta(t_n)$ that g lies in $B_{\boldsymbol{t}}$, that g_b is a unit in $B_{\boldsymbol{t}'}$, and that

$$\begin{aligned}
||X_n^b - g \cdot g_b^{-1}||_{\boldsymbol{t}} &= ||X_n - \sum_{i=0}^{\infty} c_i X_n^i||_{\boldsymbol{t}} \\
&= ||X_n^b - \sum_{i=0}^{b-1} c_i X_n^i - c_b X_n^b - \sum_{i=b+1}^{\infty} c_i X_n^i||_{\boldsymbol{t}} \\
&= ||\sum_{i=0}^{b-1} c_i X_n^i||_{\boldsymbol{t}} + ||\sum_{i=b+1}^{\infty} c_i X_n^i||_{\boldsymbol{t}} \\
&< \varepsilon t_n^b.
\end{aligned}$$

□

Definition 4.41. *Let $\omega = X_n^b + a_1 X_n^{b-1} + \ldots + a_b \in K'[X_n]$ be a monic polynomial over K' in X_n of degree $b \in \mathbb{N}_0$. We call ω a Weierstrass polynomial of degree b if $a_i(\mathbf{0}) = 0$ for all $i \in \{1, \ldots, b\}$.*

We are now ready to prove the Weierstrass Preparation Theorem.

Lemma 4.42 (Weierstrass Preparation Theorem)**.** *Let $g = \sum_{i=0}^{\infty} g_i X_n^i \in K$. Let $b \in \mathbb{N}$. Suppose that g is X_n-generic of order b. Then there exist a Weierstrass polynomial $\omega \in K'[X_n]$ of degree b and a unit $e \in K$ such that $g = e \cdot \omega$. Additionally, if $g \in K'[X_n]$, then $e \in K'[X_n]$ with $\deg_{X_n} e = \deg_{X_n} g - b$.*

Proof. We choose $\varepsilon = 1/4$ and apply Lemma 4.40. We find δ_n and $\delta : (0, \delta_n) \to (0, \delta_n)$ such that for any $\boldsymbol{t} \in \mathbb{R}^n_{>0}$ with $t_n < \delta_n$ and $t_1, \ldots, t_{n-1} < \delta(t_n)$ it follows that g lies in $B_{\boldsymbol{t}}$, that g_b is a unit in $B_{\boldsymbol{t}'}$ and that $||X_n^b - g \cdot g_b^{-1}||_{\boldsymbol{t}} < 1/4 t_n^b$. We pick such a $\boldsymbol{t}$ and apply Lemma 4.35. This gives us the monic polynomial $\omega \in B'_{\boldsymbol{t}}[X_n]$ and a unit $e \in B_{\boldsymbol{t}}$ such that $g = e \cdot \omega$. Since $B_{\boldsymbol{t}}$ is a subring of K, e is also a unit in K. As ω is a monic polynomial of degree b over $B'_{\boldsymbol{t}}$, it is also a monic polynomial of degree b over K'. Notice that in the case where $g \in K'[X_n]$, it also follows from Lemma 4.35 that $e \in B_{\boldsymbol{t}'}[X_n] \subset K'[X_n]$ with $\deg_{X_n} e = \deg_{X_n} g - b$. It remains for us to prove that ω is a Weierstrass polynomial. We write $\omega = X_n^b + a_1 X_n^{b-1} + \ldots + a_b$. As $e \in K \subset k[[X_1, \ldots, X_n]]$ is a unit, we can apply Lemma 4.23 to see that $e(\mathbf{0}) \neq 0$. We have

$$g(\mathbf{0}, X_n) = \sum_{i=0}^{\infty} g_i(\mathbf{0}) X_n^i = \sum_{i=b}^{\infty} g_i(\mathbf{0}) X_n^i = X_n^b \left(\sum_{i=b}^{\infty} g_i(\mathbf{0}) X_n^{i-b} \right).$$

So the coefficients of X_n^i with $i \in \{0, \ldots, b-1\}$ in the formal power series

$$g(\mathbf{0}, X_n) = e(\mathbf{0}, X_n)(X_n^b + a_1(\mathbf{0}) X_n^{b-1} + \ldots + a_b(\mathbf{0}))$$

are zero. We can now inductively compare these coefficients to conclude that

$$a_1(\mathbf{0}) = a_2(\mathbf{0}) = \ldots = a_b(\mathbf{0}) = 0.$$

This shows that ω is indeed a Weierstrass polynomial. □

The following lemma is in fact a version of Hensel's Lemma, but not exactly the version we want. We do however require this lemma for our desired version.

Lemma 4.43. *Suppose that k is a complete algebraically closed field. Let $\omega(\boldsymbol{X}, Y) = \omega(X_1, \ldots, X_n, Y) \in K_n[Y]$ be any monic polynomial of degree $b > 0$. Let the factorization of $\omega(\mathbf{0}, Y) \in k[Y]$ be given by*

$$\omega(\mathbf{0}, Y) = (Y - c_1)^{b_1} \cdots (Y - c_t)^{b_t}$$

such that the constants $c_1, \ldots, c_t \in k$ are all distinct. Then there exist $\omega_1(\boldsymbol{X}, Y), \ldots, \omega_t(\boldsymbol{X}, Y) \in K_n[Y]$ that are monic as polynomials in Y such that $\omega(\boldsymbol{X}, Y) = \omega_1(\boldsymbol{X}, Y) \cdots \omega_t(\boldsymbol{X}, Y)$ and such that $\omega_i(\mathbf{0}, Y) = (Y - c_i)^{b_i}$ and $\deg_Y \omega_i(\boldsymbol{X}, Y) = b_i$ hold for any $i \in \{1, \ldots, t\}$.

Proof. We prove this by induction on t. If $t = 1$, we can trivially take $\omega_1 = \omega$. So now suppose that $t > 1$ and that the statement is true for $t-1$. We let $q(\boldsymbol{X}, Y) \coloneqq \omega(\boldsymbol{X}, Y + c_t)$. Note that q is also monic in Y. We find

$$q(\mathbf{0}, Y) = \omega(\mathbf{0}, Y + c_t) = (Y - (c_1 - c_t))^{b_1} \cdots (Y - (c_{t-1} - c_t)^{b_{t-1}} Y^{b_t}.$$

Note that we can view $K_n[Y]$ as a subring of K_{n+1}. We see that $q(\boldsymbol{X}, Y)$ is Y-generic of order b_t. We apply the Weierstrass Preparation Theorem 4.42 on $q(\boldsymbol{X}, Y)$ to find a Weierstrass polynomial $q_t \in K_n[Y]$ of degree b_t and a polynomial $e' \in K_n[Y]$ that is a unit in K_{n+1} and of degree $b - b_t$ in Y, such that $q(\boldsymbol{X}, Y) = e'(\boldsymbol{X}, Y) \cdot q_t(\boldsymbol{X}, Y)$. Because q_t and q are monic polynomials in Y, so is e'. We have $q_t(\mathbf{0}, Y) = Y^{b_t}$ as q_t is a Weierstrass polynomial of degree b_t. By Lemma 4.23 we have that $e'(\mathbf{0}, 0)$ is a unit in k, hence $e'(\mathbf{0}, 0) \neq 0$. We now define $\omega_t(\boldsymbol{X}, Y) \coloneqq q_t(\boldsymbol{X}, Y - c_t)$ and $\omega'(\boldsymbol{X}, Y) \coloneqq e'(\boldsymbol{X}, Y - c_t)$. It then follows that

$$\omega(\boldsymbol{X}, Y) = \omega(\boldsymbol{X}, Y - c_t + c_t) = q(\boldsymbol{X}, Y - c_t) = e'(\boldsymbol{X}, Y - c_t) \cdot q_t(\boldsymbol{X}, Y - c_t) = \omega'(\boldsymbol{X}, Y)\omega_t(\boldsymbol{X}, Y),$$

and that

$$\omega_t(\mathbf{0}, Y) = q_t(\mathbf{0}, Y - c_t) = (Y - c_t)^{b_t}.$$

Since $e'(\boldsymbol{X}, Y)$ is monic in Y, so is $\omega'(\boldsymbol{X}, Y)$. We also have

$$(Y - c_1)^{b_1} \cdots (Y - c_t)^{b_t} = \omega(\mathbf{0}, Y) = \omega'(\mathbf{0}, Y) \cdot \omega_t(\mathbf{0}, Y) = \omega'(\mathbf{0}, Y) \cdot (Y - c_t)^{b_t}.$$

So $\omega'(\mathbf{0}, Y) = (Y - c_1)^{b_1} \cdots (Y - c_{t-1})^{b_{t-1}}$. By our induction hypothesis, we find monic polynomials $\omega_1, \ldots, \omega_{t-1} \in K_n[Y]$ such that $\omega'(\boldsymbol{X}, Y) = \omega_1(\boldsymbol{X}, Y) \cdots \omega_{t-1}(\boldsymbol{X}, Y)$ and such that $\omega_i(\mathbf{0}, Y) = (Y - c_i)^{b_i}$ for all $i \in \{1, \ldots, t-1\}$. This gives the desired result. □

We will now prove the version of Hensel's Lemma that is about convergent power series.

Lemma 4.44 (Hensel's Lemma)**.** *Suppose that k is a complete algebraically closed field. Let $F(\boldsymbol{X}, Y) = F(X_1, \ldots, X_n, Y) \in K_n[Y]$ be a monic polynomial in Y of degree $n \geq 2$, with coefficients in K_n. Suppose that there exist $G_0(Y), H_0(Y) \in k[Y]$ of nonzero degree r, s respectively, such that $F(\mathbf{0}, Y) = G_0(Y)H_0(Y)$ and such that $r + s = n$. Suppose that there also exist $G^*, H^* \in k[Y]$ such that $G_0H^* + H_0G^* = 1$ holds. Then there exist $G, H \in K_n[Y]$, whose degree as polynomials in Y are r, s respectively, such that $F = GH$.*

Proof. First note that the polynomials $G(\mathbf{0}, Y)$ and $H(\mathbf{0}, Y)$ share no common root, for if $a \in k$ is such a common root, we would have

$$1 = G_0(a)H^*(a) + H_0(a)G^*(a) = 0 \cdot H^*(a) + 0 \cdot G^*(a) = 0,$$

a contradiction. We first assume G_0 and H_0 to be monic. Now let $\prod\limits_{i=1}^{t} (Y - c_i)^{b_i}$ and $\prod\limits_{i=t+1}^{p} (Y - c_i)^{b_i}$, with $c_i \in k$ and $b_i \in \mathbb{N}$ be the factorization of G_0 and H_0 respectively. Note that $r = \deg_Y G = \sum_{i=1}^{t} b_i$ and $s = \deg_Y H = \sum_{i=t+1}^{p} b_i$. Note that we may assume that $c_i \neq c_j$ holds for all $i \neq j$. So $F(0, Y) = \prod\limits_{i=1}^{p} (Y - c_i)^{b_i}$ and we apply Lemma 4.43 to find the existence of $F_i(\boldsymbol{X}, Y) \in K_n[Y]$ for each $i \in \{1, \ldots, p\}$, such that $F_i(0, Y) = (Y - c_i)^{b_i}$ and $\deg_Y F_i(\boldsymbol{X}, Y) = b_i$ hold for all such i, and such that $F(\boldsymbol{X}, Y) = F_1(\boldsymbol{X}, Y) \cdots F_p(\boldsymbol{X}, Y)$. We now take $G(\boldsymbol{X}, Y) \coloneqq F_1(\boldsymbol{X}, Y) \cdots F_t(\boldsymbol{X}, Y)$ and $H(\boldsymbol{X}, Y) \coloneqq F_{t+1}(\boldsymbol{X}, Y) \cdots F_p(\boldsymbol{X}, Y)$. We find $\deg_Y G(\boldsymbol{X}, Y) = \sum_{i=1}^{t} b_i = r$ and $\deg_Y H(\boldsymbol{X}, Y) = \sum_{i=t+1}^{p} b_i = s$. This gives the desired result. □

With the convergent version of Hensel's Lemma proven, we can now look at the convergent version of Newton-Puiseux's theorem. We will now restrict ourselves to the case where $n = 1$. So $K = K_n = k[\{X\}]$. We will also define the field of convergent Puiseux series.

Definition 4.45. *Let k be a complete field. We denote the quotient field of $k[\{X\}]$ by $k(\{X\})$. We call $k(\{X^*\}) \coloneqq \bigcup\limits_{m=1}^{\infty} k(\{X^{1/m}\})$ the field of convergent Puiseux series (in X) over k. Since $k[\{X\}] \subset k[[X]]$, we may and will view $k(\{X\})$ as subset of $k((X))$, the field of formal Puiseux series.*

Theorem 4.46 (Newton-Puiseux)**.** *Suppose that k is a complete algebraically closed field of characteristic zero. Then the field of convergent Puiseux series $k(\{X^*\})$ is also algebraically closed.*

Proof. Let $F \in k(\{X^*\})[Y]$ be a nonconstant polynomial in Y over $k(\{X^*\})$. We need to show that this polynomial contains a root in $k(\{X^*\})$. Note that $k(\{X^*\})[Y] \subset k((X^*))[Y]$ and that $k((X^*))$ is algebraically closed by Theorem 4.17. So there exists $f(X) \in k((X^*))$ such that $F(X, f(X)) = 0$. To see that this $f(X)$ actually also lies in $k(\{X^*\})$, we can look at Subsection 4.1 and replace $k[[X]]$ by $K[\{X\}]$, $k((X))$ by $k(\{X\})$ and $k((X^*))$ by $k(\{X^*\})$. We get that the statements are still true except for Hensel's Lemma 4.13. We replace that one by the convergent version of Hensel's Lemma 4.44 and indeed find that $f(X) \in k(\{X^*\})$ holds. $\square$

As a direct consequence, we now see that the field $\mathbb{C}(\{X^*\})$ is algebraically closed. Because $\mathbb{C}(\{X^*\})$ and $\overline{\mathbb{Q}}((X^*))$ are both subsets of $\mathbb{C}((X^*))$, we can say something useful about their intersection.

Definition 4.47. *We call the set*

$$\overline{\mathbb{Q}}(\{X^*\}) \coloneqq \overline{\mathbb{Q}}((X^*)) \cap \mathbb{C}(\{X^*\})$$

the field of convergent Puiseux series over $\overline{\mathbb{Q}}$.

This set is actually also an algebraically closed field. This follows from the following lemma:

Lemma 4.48. *Let H be a field and let F and K be two algebraically closed subfields of H. Let $M = F \cap K$. Then M is also an algebraically closed subfield of H.*

Proof. It can be easily seen that M is a subfield of H as it contains 1 and is closed under addition, subtraction, multiplication and division. To show that it is also algebraically closed, let $f \in M[X]$ be a nonconstant monic polynomial of degree $n \in \mathbb{N}$. Because $f \in F[X]$ holds and because F is algebraically closed, we can factorize f as $f = (X - a_1) \cdots (X - a_n)$ where $a_1, \ldots, a_n \in F$ are uniquely determined (up to order). This factorization is also a factorization in H as $F \subset H$. We can, in a similar way, factorize f as $f = (X - b_1) \cdots (X - b_n)$ where $b_1, \ldots, b_n \in K$ are uniquely determined (up to order of the factors). This second factorization is also a factorization in H. Since H is a field, it is a unique factorization domain, so these two factorizations must be equal to one another (up to units in H and up to the order of the factors). So for any $i \in \{1, \ldots, n\}$, there exist a nonzero $c_i \in H$ and $j \in \{1, \ldots, n\}$ such that

$$(X - a_i) = c_i(X - b_j) = c_i X - c_i b_j.$$

This gives us $c_i = 1$ and $b_j = a_i$. So $a_i \in K$, hence $a_i \in F \cap K = M$. This shows us that $f = (X - a_1) \cdots (X - a_n)$ is also a factorization in M. So M is indeed an algebraically closed field as any nonconstant monic polynomial over M factorizes into linear monic polynomials over M. $\square$

Corollary 4.49. *The set $\overline{\mathbb{Q}}(\{X^*\})$ is an algebraically closed subfield of $\mathbb{C}((X^*))$.*

Proof. Because $\overline{\mathbb{Q}}$ is an algebraically closed field of characteristic zero, we can apply Theorem 4.17 to see that $\overline{\mathbb{Q}}((X^*))$ is algebraically closed. Since $\mathbb{C}$ is a complete algebraically closed field of characteristic zero (with the canonical norm), we can apply Theorem 4.46 to see that $\mathbb{C}(\{X^*\})$ is algebraically closed. These two fields are subfields of $\mathbb{C}((X^*))$, so by the previous lemma we find that its intersection $\overline{\mathbb{Q}}(\{X^*\})$ must be an algebraically closed subfield of $\mathbb{C}((X^*))$ as well. $\square$

4.3 Puiseux expansions at infinity

Let $f(X) \in \mathbb{C}[\{X\}]$ be a convergent power series in X over $\mathbb{C}$. We saw in Lemma 4.38 that $f(x)$ converges for $x \in \mathbb{C}$ when its norm $|x|$ is small enough. For Runge's Theorem we are actually interested in values $x \in \mathbb{C}$ whose norm are big enough. In this subsection we solve this problem by making use of the inverted variable X^{-1} and look at Puiseux series in X^{-1}, or as we will (re)name them, Puiseux series at infinity (in X).

Lemma 4.50. *Let k be an algebraically closed field of characteristic zero. Let $F(X, Y) \in k[X, X^{-1}][Y]$ be an n-th degree polynomial in Y over $k[X, X^{-1}]$ with $n \in \mathbb{N}$. We can then factorize $F(X, Y)$ as*

$$F(X, Y) = g(X) \prod_{i=1}^{n} (Y - f_i(X)), \tag{4.7}$$

for some $f_1(X), \ldots, f_n(X) \in k(((X^{-1})^*))$ and $g(X) \in k[X, X^{-1}]$. If k is a complete field, then we also have $f_1(X), \ldots, f_n(X) \in k(\{(X^{-1})^*\})$.

Proof. We can write $F(X,Y) = \sum_{i=0}^{n} g_i(X)Y^i$, with $g_i(X) \in k[X, X^{-1}]$. We let $T \coloneqq X^{-1}$ be the inverse of the formal variable X. We then see that $g_i(T^{-1}) \in k[T^{-1}, T]$ holds for all $i \in \{1, \dots, n\}$. We then define $F'(T,Y) \coloneqq F(T^{-1}, Y) = \sum_{i=0}^{n} g_i(T^{-1})Y^i$. We see that $F'(T,Y) \in [T, T^{-1}][Y] \subset k((T^*))[Y]$. By Newton-Puiseux's Theorem 4.17 we have that $k((T^*))$ is algebraically closed. So we can write

$$F'(T,Y) = g'(T) \prod_{i=1}^{n} (Y - f_i'(T))$$

for some $g'(T), f_1'(T), \dots, f_n'(T) \in k((T^*))$. We then have

$$F(X,Y) = F'(X^{-1}, Y) = g'(X^{-1}) \prod_{i=1}^{n} (Y - f_i'(X^{-1})).$$

We take $f_i(X) \coloneqq f_i'(X^{-1}) \in k(((X^{-1})^*))$ and $g(X) \coloneqq g' \in k(((X^{-1})^*))$ and find (4.7). We only need to check that $g(X)$ lies in $k[X, X^{-1}]$. This immediately follows from the fact that $g(X)$ is the leading coefficient of $F(X,Y)$, when we view $F(X,Y)$ as a polynomial in Y over $k[X, X^{-1}]$. Now suppose that k is also a complete field. We then use the convergent version of Newton-Puiseux's Theorem 4.46 to see that the elements $f_1'(T), \dots, f_n'(T)$ also lie in $k(\{T^*\})$. From the same reasoning it then follows that $f_1(X), \dots, f_n(X) \in k(((X^{-1})^*))$. □

Definition 4.51. *We call the roots $f_1(X), \dots, f_n(X)$ from (4.7) the Puiseux expansions at infinity of F.*

Definition 4.52. *Let k be an algebraically closed field of characteristic zero. We then call $k(((X^{-1})^*))$ the field of formal Puiseux series at infinity (in X) over k.*

Remark 4.53. *Note that $k(((X^{-1})^*)) = \bigcup_{e=1}^{\infty} k(((X^{-1})^{1/e})) = \bigcup_{e=1}^{\infty} k((X^{-1/e}))$. Any $f(X) \in k(((X^{-1})^*))$ can thus be written as*

$$f(X) = \sum_{n=-m}^{\infty} a_n X^{-n/e}, \tag{4.8}$$

with $a_n \in k$, $m \in \mathbb{Z}$ and $e \in \mathbb{N}$. If f is nonzero, we let e and m be the smallest possible values such that we still can write $f(X)$ in the form of (4.8). In that case we have $a_{-m} \neq 0$.

Definition 4.54. *Let k be an algebraically closed complete field of characteristic zero. We then call $k(\{(X^{-1})^*\})$ the field of convergent Puiseux series at infinity (in X) over k.*

Remark 4.55. *We have that $k((T^*))$ and $k(((X^{-1})^*))$ are field isomorphic, with the canonical isomorphism given by $f(T) \mapsto f(X^{-1})$. Because of this we immediately see that $k(((X^{-1})^*))$ is also an algebraically closed field. The same also holds for $k(\{(X^{-1})^*\})$ by similar reasoning.*

The following lemma shows that convergent Puiseux series at infinity actually converge for values whose norms are big enough.

Lemma 4.56. *Let k be a complete field and let $f(X) \in k(\{(X^{-1})^*\})$. Then there exists an $R \in \mathbb{R}_{\geq 0}$ such that the series $f(x)$ converges in k for all $x \in k$ satisfying $|x| > R$.*

Proof. Any $g(T) \in k(\{T\})$ can be written as $T^{-m} \cdot h(T)$, where $m \in \mathbb{N}$ and $h(T) \in k[\{T\}]$. By the one dimensional case of Lemma 4.38, there exists $R' \in \mathbb{R}_{>0}$ such that the series $h(t)$ converges in k for all $t \in k$ satisfying $|t| < R'$. Since t^{-m} converges for all nonzero $t \in k$, we see that the product $t^{-m} \cdot h(t)$ converges for all nonzero $t \in k$ satisfying $|t| < R'$. Now let $f(X) \in k(\{(X^{-1})^*\})$. Then there exists $e \in \mathbb{N}$ with $f(X) \in k(\{X^{-1/e}\})$. This shows us that there exists $g(T) \in k(\{T\})$ with $f(X) = g(X^{-1/e})$. Let $R' \in \mathbb{R}_{>0}$ be such that the series $g(t)$ converges for all nonzero $t \in k$ satisfying $|t| < R'$. Take $R \coloneqq R'^{-e}$. For any $x \in k$ with $|x| > R$ we then have

$$0 < |x^{-1/e}| = \left(\frac{1}{|x|}\right)^{1/e} < R^{-1/e} = R'.$$

So $f(x) = g(x^{-1/e})$ converges for all $x \in k$ satisfying $|x| > R$. □

We will now define the order of Puiseux series at infinity.

Definition 4.57. *Let k be a field. Let $f(X) \in k(((X^{-1})^*))$ be a nonzero formal Puiseux series at infinity over k. We can then write $f(X)$ as in (4.8) for some smallest $m \in \mathbb{Z}$, $e \in \mathbb{N}$ and with $a_n \in k$ for all $n \in \{-m, -m+1, \ldots, \infty\}$. We then define the order of $f(X)$ in X to be $ord_X f := m/e$. We also define $ord_X 0 = -\infty$.*

Note that any polynomial in X over $\mathbb{C}$ is also a Puiseux series at infinity in X over $\mathbb{C}$ and that its X-degree and order are always the same. The order of Puiseux series at infinity also satisfies some nice rules, as the following lemma shows.

Lemma 4.58. *The following rules hold for any $f(X), g(X) \in k(((X^{-1})^*))$:*

- $ord_X(f+g) \leq \max(ord_X f, ord_X g)$.
- $ord_X(f \cdot g) = ord_X f + ord_X g$.
- *If f is nonzero, then* $ord_X(1/f) = -ord_X f$.

Proof. The proof of the first two rules are similar to the proof of Lemmas 3.15 and 3.16 and can be proven in a very straightforward way. The third rule follows from $0 = \text{ord}_X(1) = \text{ord}_X(f \cdot (1/f)) = \text{ord}_X f + \text{ord}_X(1/f)$. Note that if $f(X)$ is also a nonzero polynomial in X over k, then we have $\text{ord}_X f = \deg_X f$. □

4.4 Coefficients of Puiseux expansions

Let k be a field of characteristic zero and let $\overline{k}$ be its algebraic closure. Throughout this subsection we let $F(X,Y) \in k[X,Y]$ be an irreducible polynomial of Y-degree $d_2 \geq 1$, and we let $f(X) = \sum_{n=-m}^{\infty} a_n X^{-n/e} \in \overline{k}((X^{-1/e}))$ be any of its Puiseux expansions at infinity. Our aim in this subsection is to show that there exists a subfield $l \subset \overline{k}$ that is a finite field extension of k, such that all coefficients a_n lie in l. The proof of this is inspired by a different proof to Newton-Puiseux's Theorem 4.17 with the restriction that the groundfield is $\mathbb{C}$ [1]. Since we already have proved the more general Newton-Puiseux's theorem, we have adapted this proof such that it contains only what we still want to prove and such that it also works in cases where $k \neq \mathbb{C}$.

First we will generalize the definition of a Newton dot such that it also works with fractional exponents of X:

Definition 4.59. *Let K be a field and let $G(X,Y) \in K((X^{-1/e}))[Y]$. We can write G as*

$$G(X,Y) = \sum_{i=-t}^{\infty} \sum_{j=0}^{d_2} c_{i/e,j} X^{i/e} Y^j,$$

for some $t \in \mathbb{Z}$ and $d_2 \in \mathbb{N}_0$ and with $c_{i/e,j} \in K$. We define $c_{a,b} = 0$ for all $(a,b) \in \mathbb{Q} \times \mathbb{Z}$ where $c_{a,b}$ is still undefined. We then define

$$D(G) := \{(i/e, j) \in \mathbb{Q} \times \mathbb{Z} \mid c_{i/e,j} \neq 0\}.$$

Remark 4.60. *Using a similar proof of Lemmas 3.15 and 3.16, we can see that these lemmas not only hold for polynomials in $\mathbb{Z}[X,Y]$, but also for polynomials in the ring $K((X^{-1/e}))[Y]$. So we have in for all $G, H \in K((X^{-1/e}))[Y]$ and $\lambda \in \mathbb{R}_{>0}$:*

- $\deg_\lambda(G+H) \leq \max(\deg_\lambda G, \deg_\lambda H)$.
- *If* $\deg_\lambda G > \deg_\lambda H$, *then* $(G+H)_\lambda = G_\lambda$.
- $\deg_\lambda(GH) \leq \deg_\lambda G + \deg_\lambda H$.
- $(GH)_\lambda = G_\lambda H_\lambda$.

Now let $h(X) = \sum_{n=-m}^{N} a_n X^{-n/e} \in \overline{k}((X^{-1/e}))$, for some $N \in \mathbb{N}$. So $h(X)$ consists of the highest order terms of $f(X)$, where $f(X)$ has been defined at the start of this subsection. Furthermore, let $G(X,Y) := F(X, Y + h(X)) \in \overline{k}((X^{-1/e}))[Y]$. We are interested in its Newton dots. First of all we let b/e be the largest order among the coefficients of G, where G is viewed as a polynomial in Y over $\overline{k}((X^{-1/e}))$. Secondly, we observe that G is of the same Y-degree as F. Now let $(i/e, j)$ be a Newton dot of G. So $cX^{i/e}Y^j$ is a term of G for some nonzero constant $c \in \overline{k}$. We therefore find that $i/e \leq b/e$ and $0 \leq j \leq d_2$ must hold. We want to show that the value b/e is independent of N. To do this, we first need to look at the following lemma.

Lemma 4.61. *Let $G(X,Y)$ be defined as above and let $L \leq 0$ be an integer. Let $g(X) \in \overline{k}((X^{-1/e}))$ be a formal Puiseux series at infinity whose order is less than L/e. Let $i \in \mathbb{Z}$ and $j \in \mathbb{N}_0$ such that $i/e \geq b/e + L/e$. Then the coefficient of $X^{i/e}Y^j$ in $G(X, Y + g(X))$ is the same as the coefficient of $X^{i/e}Y^j$ in $G(X,Y)$.*

Note that this last statement implies that $(i/e, j)$ is a Newton dot of $G(X, Y + g(X))$ if and only if it is a Newton dot of $G(X,Y)$.

Proof. We can write $G(X,Y) = \sum_{t=0}^{d_2} q_t(X)Y^t$ where all $q_t(X) \in \overline{k}((X^{-1/e}))$ are of order at most b/e. This gives us

$$G(X, Y + g(X)) = \sum_{t=0}^{d_2} q_t(X)(Y + g(X))^t = \sum_{t=0}^{d_2} q_t(X)Y^t + \sum_{t=0}^{d_2} q_t(X) \sum_{s=1}^{t} \binom{t}{s} g(X)^s Y^{t-s}. \tag{4.9}$$

The second part of the right-hand side consists of terms of the form $q_t(X)\binom{t}{s}g(X)^sY^{t-s}$. The order of the coefficient of any such term satisfies:

$$\begin{aligned}
\mathrm{ord}_X(q_t(X)\binom{t}{s}g(X)^s) &= \mathrm{ord}_X(q_t(X)) + \mathrm{ord}_X(\binom{t}{s}) + s \cdot \mathrm{ord}_X(g(X)) \\
&< b/e + 0 + s \cdot L/e \\
&\leq b/e + L/e \\
&\leq i/e.
\end{aligned}$$

This shows us that the second part of the right-hand side of (4.9) does not contribute to the coefficient of $X^{i/e}Y^j$. So only the first part may contribute. Since this first part is exactly $G(X,Y)$, we see that $G(X,Y)$ and $G(X, Y + g(X))$ indeed have the same coefficient at $X^{i/e}Y^j$. □

We now are able to prove that b/e is independent of N. Suppose we had taken a larger integer $N' > N$. We then have corresponding $h'(X) := \sum_{n=-m}^{N'} a_n X^{-n/e}$ and $G'(X,Y) := F(X, Y + h'(X))$. Let $g(X) := h'(X) - h(X) = \sum_{n=N+1}^{N'} a_n X^{-n/e}$. Then we see that the order of $g(X)$ is smaller than $-N/e$, hence smaller than zero. So according to the previous lemma, $G(X,Y)$ and $G(X, Y + g(X)) = F(X, Y + h(X) + g(X)) = G'(X,Y)$ have the same Newton dots of the form $(i/e, j)$ where $i/e \geq b/e$. This shows that $G'(X,Y)$ also has b/e as boundary. So b/e is indeed independent of N.

As $F(X,Y)$ is irreducible, $f(X)$ must be a root of single multiplicity. Therefore, $F(X, Y + f(X)) \in \overline{k}((X^{-1/e}))[Y]$ must have $Y = 0$ as a root of single multiplicity. From this we can see that $F(X, Y + f(X))$ has, as a polynomial in Y, a zero constant term and a nonzero linear term (the coefficient of Y). So $F(X, Y + f(X))$ has no Newton dots on the x-axis, but it does have at least one Newton dot of the form $(p/e, 1)$ with $p \in \mathbb{Z}$. We let p be the largest number with this property, so p/e is the order of the coefficient of Y in $F(X, Y + f(X))$. Note that p/e is also independent of N.

We now demand $N \in \mathbb{N}$ to be large enough, such that $N/e > b/e - p/e$ holds. Let $q(X) := \sum_{n=N+1}^{\infty} a_n X^{-n/e} \in \overline{k}((X^{-1/e}))$. We then see that $G(X, Y + q(X)) = F(X, Y + h(X) + q(X)) = F(X, Y + f(X))$. We have that $q(X)$ is of order less than $-N/e$. So by Lemma 4.61, any $(i/e, j) \in \mathbb{Q} \times \mathbb{N}_0$ with $i/e \geq b/e - N/e$ is a Newton dot of $G(X,Y)$ exactly when it is a Newton dot of $G(X, Y + q(X)) = F(X, Y + f(X))$. From $N/e > b/e - p/e$ we get $p/e > b/e - N/e$. So $(p/e, 1)$ is also a Newton dot of $G(X,Y)$ and every $(a, 1) \in \mathbb{Q} \times \mathbb{N}_0$ with $a > p/e$ is no Newton dot of $G(X,Y)$. Also, the coefficient of $X^{p/e}Y^1$ in $G(X,Y)$, which we will call d, is equal to the coefficient of $X^{p/e}Y^1$ in $F(X, Y + f(X))$. In particular we see that this coefficient is independent on N and nonzero.

We can now prove a lemma that is similar to the previous one.

Lemma 4.62. *Let $G(X,Y)$ be as above (with N large enough, such that $N/e > b/e - p/e$ holds) and let $L \in \mathbb{Z}$ be a nonpositive integer such that $L/e \leq p/e - b/e$ holds. Let $g(X) \in \overline{k}((X^{-1/e}))$ be a formal Puiseux series at infinity whose order is less than L/e. Let $i \in \mathbb{Z}$ satisfy $i/e \geq p/e + L/e$. Then the coefficient of $X^{i/e}Y^0$ in $G(X, Y + g(X))$ is the same as the coefficient of $X^{i/e}Y^0$ in $G(X,Y)$.*

Proof. We can write $G(X,Y) = \sum_{t=0}^{d_2} q_t(X)Y^t$ where all $q_t(X)$ are of order at most b/e and where $q_1(X)$ is of order p/e. Note that we are only interested in the coefficients whose corresponding Newton dots lie on the x-axis. We therefore might as well look at $G(X,0)$ and $G(X,g(X))$ as the variable Y contributes only to terms with Newton dots that lie above the x-axis. We find

$$G(X,g(X)) = \sum_{t=0}^{d_2} q_t(X)(g(X))^t = q_0(X) + q_1(X)g(X) + \sum_{t=2}^{d_2} q_t(X)(g(X))^t. \tag{4.10}$$

The third part of the right-hand side of (4.10) consists of terms of the form $q_t(X)(g(X))^t$ with $t \geq 2$. The order of any such term satisfies:

$$\begin{aligned}\operatorname{ord}_X(q_t(X)(g(X))^t) &= \operatorname{ord}_X(q_t(X)) + t \cdot \operatorname{ord}_X(g(X)) < b/e + t \cdot L/e \\ &\leq b/e + 2L/e \leq b/e + L/e + p/e - b/e = p/e + L/e.\end{aligned}$$

The order of the second part of the right-hand side of (4.10) satisfies

$$\operatorname{ord}_X(q_1(X)g(X)) = \operatorname{ord}_X(q_1(X)) + \operatorname{ord}_X(g(X)) < p/e + L/e.$$

This shows us that the second and third part of the right-hand side of (4.9) do not contribute to the coefficient of $X^{i/e}$. So only the first part may contribute. Since this first part is exactly $G(X,0)$, we are done. □

We now consider $q(X) = \sum_{n=N+1}^{\infty} a_n X^{-n/e}$ again. We have that $q(X)$ is of order less than $-N/e < p/e - b/e$. So by Lemma 4.62, any $(i/e, 0) \in \mathbb{Q} \times \mathbb{N}_0$ with $i/e \geq p/e - N/e$ is a Newton dot of $G(X,Y)$ exactly when it is a Newton dot of $G(X, Y + q(X)) = F(X, Y + f(X))$. We saw that $F(X, Y + f(X))$ has no Newton dots on the x-axis, so we find that $G(X,Y) = F(X, Y + \sum_{n=-m}^{N} a_n X^{-n/e})$ has no Newton dots of the form $(i/e, 0)$ with $i/e \geq p/e - N/e$. If we now substitute N by $N+1$, we find that $G'(X,Y) := F(X, Y + \sum_{n=-m}^{N+1} a_n X^{-n/e})$ has no Newton dot of the form $(i/e, 0)$ with $i/e \geq p/e - (N+1)/e$. So in particular we have that the coefficient of $X^{p/e-(N+1)/e}$ in $G'(X,Y)$ is zero. Note that $G'(X,Y) = G(X, Y + a_{N+1}X^{-(N+1)/e})$.

Consider the elements $a_{-m}, a_{-m+1}, \ldots, a_N$. As they all lie in the algebraic closure of k and there are only finitely many of them, there exists a subfield l of $\overline{k}$ that is a finite field extension of k and that contains all these elements (for example the field $l = k(a_{-m}, a_{-m+1}, \ldots, a_N)$). If we can prove that a_{N+1} also lies in l, then by induction we have that all coefficients a_n of $f(X)$ lie in l.

Because we have $F(X,Y) \in k[X,Y] \subset l[X,Y]$ and $h(X) = \sum_{n=-m}^{N} a_n X^{-n/e} \in l((X^{-1/e}))$, we must have $G(X,Y) = F(X, Y + h(X)) \in l((X^{-1/e}))[Y]$. Let us write $G(X,Y) = \sum_{t=0}^{d_2} q_t(X) Y^t$ with $q_t(X) \in l((X^{-1/e}))$. The coefficient of $X^{p/e-(N+1)/e}$ in $G'(X,Y)$, which is zero, is the same as the coefficient of $X^{p/e-(N+1)/e}$ in $G'(X,0)$, because every term with Y does not contribute to the coefficient of $X^{p/e-(N+1)/e}$. We have

$$\begin{aligned}G'(X,0) = G(X, a_{N+1}X^{-(N+1)/e}) &= \sum_{t=0}^{d_2} q_t(X)(a_{N+1}X^{-(N+1)/e})^t \\ &= q_0(X) + a_{N+1}q_1(X)X^{-(N+1)/e} + \sum_{t=2}^{d_2} q_t(X)(a_{N+1}X^{-(N+1)/e})^t.\end{aligned} \tag{4.11}$$

We are interested in the coefficient of $X^{p/e-(N+1)/e}$ in each term of the right-hand side of (4.11). For each $t \in \{2, \ldots, d_2\}$ we have

$$\begin{aligned}&\operatorname{ord}_X(q_t(X)(a_{N+1}X^{-(N+1)/e})^t) = \operatorname{ord}_X(q_t(X)) - t(N+1)/e \\ &\leq b/e - 2(N+1)/e < b/e - N/e - 2/e - b/e + p/e < p/e - (N+1)/e,\end{aligned}$$

so the coefficient of $X^{p/e-(N+1)/e}$ in the last term on the right-hand side of (4.11) is zero. Let $c_N \in l$ be the coefficient of $X^{p/e-(N+1)/e}$ in $q_0(X)$. We also defined $d \in l$ to be the coefficient of $X^{p/e}$ in $q_1(X)$, which is nonzero and also the coefficient of $X^{p/e}Y^1$ in $F(X, Y + f(X))$. We then see from (4.11) that the coefficient of $X^{p/e-(N+1)/e}$ in $G'(X,0)$ is $c_N + a_{N+1} \cdot d$. Since this coefficient was zero, this gives us the equation $c_N + a_{N+1} \cdot d = 0$, hence $a_{N+1} = -c_N/d$. We therefore have found that $a_{N+1} \in l$. By induction we find that all coefficients a_n lie in l. We have thus proven the following two lemmas (for the first one we have $N' = b/e - p/e$):

Lemma 4.63. *Let k be a field of characteristic zero and $\overline{k}$ an algebraic closure of k. Let $F(X,Y) \in k[X,Y]$ be an irreducible polynomial with positive Y-degree. Let $f(X) = \sum_{n=-m}^{\infty} a_n X^{-n/e} \in \overline{k}((X^{-1/e}))$ be any Puiseux expansion at infinity of $F(X,Y)$. Then there exist $N' \in \mathbb{N}$ and $p \in \mathbb{Z}$ such that for all $N > N'$ we have that $a_{N+1} = -c_N/d$, where $c_N \in \overline{k}$ is the coefficient of $X^{p/e-(N+1)/e}$ in the formal Puiseux series at infinity $F(X, \sum_{n=-m}^{N} a_n X^{-n/e}) \in \overline{k}((X^{-1/e}))$ and where $d \in \overline{k}$ is some coefficient of $F(X, Y + \sum_{n=-m}^{N} a_n X^{-n/e}) \in \overline{k}((X^{-1/e}))[Y]$. We have that d is independent of N.*

Lemma 4.64. *Let k be a field of characteristic zero and $\overline{k}$ an algebraic closure of k. Let $F(X,Y) \in k[X,Y]$ be an irreducible polynomial with positive Y-degree. Let $f(X) = \sum_{n=-m}^{\infty} a_n X^{-n/e} \in \overline{k}((X^{-1/e}))$ be any Puiseux expansion at infinity of $F(X,Y)$. Then there exists a finite field extension of k that is included in $\overline{k}$ and contains all the coefficients a_n of $f(X)$.*

In the previous lemma we demand F to be irreducible. We may remove this restriction, as the following lemma suggests:

Lemma 4.65. *Let k be a field of characteristic zero and $\overline{k}$ an algebraic closure of k. Let $F(X,Y) \in k[X,Y]$ be a polynomial with positive Y-degree. Let $f(X) = \sum_{n=-m}^{\infty} a_n X^{-n/e} \in \overline{k}((X^{-1/e}))$ be any Puiseux expansion at infinity of $F(X,Y)$. Then there exists a finite field extension of k that is included in $\overline{k}$ and contains all the coefficients a_n of $f(X)$.*

Proof. We decompose F into irreducible factors $F_1(X,Y), \ldots, F_t(X,Y) \in k[X,Y]$. So $F(X,Y) = F_1(X,Y) \cdots F_t(X,Y)$. Since $F(X, f(X)) = 0$, we have that $F_i(X, f(X)) = 0$ must hold for some $i \in \{1, \ldots, t\}$. So $f(X)$ is also a Puiseux expansion at infinity of $F_i(X,Y)$. We apply the previous lemma with this F_i, which yields the desired result. □

Corollary 4.66. *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a polynomial with positive Y-degree. Let*

$$f(X) = \sum_{n=-m}^{\infty} a_n X^{-n/e} \in \overline{Q}((X^{-1/e}))$$

be any Puiseux expansion at infinity of $F(X,Y)$. Then there exists an algebraic number field $l \subset \overline{\mathbb{Q}}$ that contains all the coefficients a_n of $f(X)$.

5 The Symmetric Function Theorem

The proof of Runge's Theorem uses a corollary from the Symmetric Function Theorem. In this section we will prove this theorem, where we follow the proof given by K. Conrad [3]. First we must give some definitions and small lemmas again.

Definition 5.1. *Let R be a ring. A polynomial $f(X_1, \ldots, X_n) \in R[X_1, \ldots, X_n]$ is called symmetric if the equation*

$$f(X_{\sigma_1}, \ldots, X_{\sigma_n}) = f(X_1, \ldots, X_n)$$

holds for any permutation σ of $\{1, \ldots, n\}$.

Example 5.2. *The polynomials $S_1 := X_1 + \ldots + X_n$ and $S_n := X_1 X_2 \cdots X_n$ are symmetric polynomials; after any permutation of the variables one can rearrange them back again as the variables know a commutative addition and multiplication. In fact, for any $k \in \{1, \ldots, n\}$ it also holds that*

$$S_k := \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} X_{i_1} \cdots X_{i_k} \tag{5.1}$$

is a symmetric polynomial.

Remark 5.3. *Note that $T^n - S_1 T^{n-1} + S_2 T^{n-2} - \ldots + (-1)^n S_n$ is the expansion of the factorization $(T - X_1)(T - X_2) \cdots (T - X_n)$.*

We will define the lexicographic order on n-tuples of non-negative integers.

Definition 5.4. *Let $\boldsymbol{a} = (a_1, , \ldots, a_n)$ and $\boldsymbol{b} = (b_1, \ldots, b_n)$ be two n-tuples in $\mathbb{N}_0^n$ such that $\boldsymbol{a} \neq \boldsymbol{b}$. We say that $\boldsymbol{a} < \boldsymbol{b}$ holds if $a_i < b_i$ holds for the lowest $i \in \{1, \ldots, n\}$ that satisfies $a_i \neq b_i$.*

It is called the lexicographic ordering as a dictionary uses the same ordering on every two words by comparing the first letter that is different. It can easily be seen that this is a well-ordering and therefore we can apply complete induction to it. Also, it can be seen that we have the implication

$$\boldsymbol{i} < \boldsymbol{j} \implies \boldsymbol{i} + \boldsymbol{k} < \boldsymbol{j} + \boldsymbol{k} \tag{5.2}$$

for any $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k} \in \mathbb{N}_0^n$, where the n-tuples have componentwise addition.

Remark 5.5. *Let R be a ring and $f \in R[X_1, \ldots, X_n]$. We can write it as*

$$f(X_1, \ldots, X_n) = \sum_{i_1, \ldots, i_n \in \mathbb{N}_0} a_{i_1, \ldots, i_n} X_1^{i_1} \cdots X_n^{i_n},$$

where the sum has only finitely many nonzero terms. We will write it shorthanded as $\sum_{\boldsymbol{i}} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}$, where $\boldsymbol{i} = (i_1, \ldots, i_n)$ iterates over $\mathbb{N}_0^n$ and where $\boldsymbol{X}^{\boldsymbol{i}} = X^{i_1} \cdots X^{i_n}$. Note that if $\boldsymbol{i}, \boldsymbol{j}$ are two multi-indices, we have $\boldsymbol{X}^{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{j}} = \boldsymbol{X}^{\boldsymbol{i}+\boldsymbol{j}}$.

Definition 5.6. *Let R be a ring and $f = \sum_{\boldsymbol{i}} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}} \in R[X_1, \ldots, X_n]$ a nonzero polynomial. The multidegree of f is $mdeg(f) := \max\{\boldsymbol{i} \mid a_{\boldsymbol{i}} \neq 0\} \in N_0^n$, where we use the lexicographic ordering of N_0^n. If $\boldsymbol{p} = mdeg(f)$, then we call $a_{\boldsymbol{p}} T^{\boldsymbol{p}}$ the leading term of f, and $lead(f) := a_{\boldsymbol{p}}$ the leading coefficient of f.*

Example 5.7. *Let $k \in \{1, \ldots, n\}$ and consider S_k from (5.1). Since S_k is the sum of all distinct products of k different variables, each multi-index corresponding to a nonzero coefficient of S_k is an n-tuple that is 1 at k entries and is 0 at the other $n - k$ entries. By comparing all those in the lexicographic order, we find that the multi-index, where the first k entries are 1 and the others 0, is the highest of them all. Therefore, $mdeg(S_k) = (1, \ldots, 1, 0, \ldots, 0)$, where the first k entries are 1 and where the last $n - k$ entries are 0.*

Lemma 5.8. *Let R be a domain and $f, g \in R[X_1, \ldots, X_n]$ both nonzero. Then $mdeg(fg) = mdeg(f) + mdeg(g)$ and $lead(fg) = lead(f)lead(g)$.*

Proof. If $\boldsymbol{p}, \boldsymbol{q}$ are the multidegree of f and g respectively, then we can write $f = a_{\boldsymbol{p}} \boldsymbol{X}^{\boldsymbol{p}} + \sum_{\boldsymbol{i}<\boldsymbol{p}} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}$ and $g = b_{\boldsymbol{q}} \boldsymbol{X}^{\boldsymbol{q}} + \sum_{\boldsymbol{j}<\boldsymbol{q}} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{j}}$. This gives us

$$\begin{aligned} fg &= \left(a_{\boldsymbol{p}} \boldsymbol{X}^{\boldsymbol{p}} + \sum_{\boldsymbol{i}<\boldsymbol{p}} a_{\boldsymbol{i}} \boldsymbol{X}^{\boldsymbol{i}}\right)\left(b_{\boldsymbol{q}} \boldsymbol{X}^{\boldsymbol{q}} + \sum_{\boldsymbol{j}<\boldsymbol{q}} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{j}}\right) \\ &= a_{\boldsymbol{p}} b_{\boldsymbol{q}} \boldsymbol{X}^{\boldsymbol{p}+\boldsymbol{q}} + \sum_{\boldsymbol{j}<\boldsymbol{q}} a_{\boldsymbol{p}} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{p}+\boldsymbol{j}} + \sum_{\boldsymbol{i}<\boldsymbol{p}} a_{\boldsymbol{i}} b_{\boldsymbol{q}} \boldsymbol{X}^{\boldsymbol{i}+\boldsymbol{q}} + \sum_{\boldsymbol{i}<\boldsymbol{p}} \sum_{\boldsymbol{j}<\boldsymbol{q}} a_{\boldsymbol{i}} b_{\boldsymbol{j}} \boldsymbol{X}^{\boldsymbol{i}+\boldsymbol{j}}. \end{aligned}$$

Since R is a domain, $a_{\boldsymbol{p}} b_{\boldsymbol{q}} \neq 0$, hence fg has a nonzero term of multidegree $\boldsymbol{p} + \boldsymbol{q}$. Now every other term has multidegree $\boldsymbol{p} + \boldsymbol{j}$, $\boldsymbol{i} + \boldsymbol{q}$ or $\boldsymbol{i} + \boldsymbol{j}$, where $\boldsymbol{i} < \boldsymbol{p}$ and $\boldsymbol{j} < \boldsymbol{q}$. By (5.2), we see that all these multidegrees are smaller than the multidegree $\boldsymbol{p} + \boldsymbol{q}$. This shows us that $\text{mdeg}(fg) = \boldsymbol{p} + \boldsymbol{q} = \text{mdeg}(f) + \text{mdeg}(g)$ and $\text{lead}(fg) = a_{\boldsymbol{p}} b_{\boldsymbol{q}} = \text{lead}(f)\text{lead}(g)$. ☐

Lemma 5.9. *Let R be a ring and $f \in R[X_1, \ldots, X_n]$ nonzero. Let $\boldsymbol{p} = (p_1, \ldots, p_n)$ be the multidegree of f. If f is symmetric, then $p_i \geq p_j$ for all $i, j \in \mathbb{N}$ satisfying $1 \leq i < j \leq n$.*

Proof. This is trivial for $n = 1$, so we will assume $n > 1$. By contradiction, suppose that $p_i < p_j$ for some $1 \leq i < j \leq n$. Let g be $f(X_1, \ldots, X_n)$, but with the variables X_i, X_j swapped. Since $a_{\boldsymbol{p}}$ is nonzero by definition of $\boldsymbol{p}$, so must the coefficient in g corresponding to

$$\boldsymbol{q} := (p_1, \ldots, p_{i-1}, p_j, p_{i+1}, \ldots, p_{j-1}, p_i, p_{j+1}, \ldots, p_n),$$

where $\boldsymbol{q}$ is the multi-index that is $\boldsymbol{p}$, but then with the entries in the i-th and j-th spot swapped. We see that $\boldsymbol{p} < \boldsymbol{q}$ holds. Since f is symmetric, a swap of variables makes no difference, hence $f = g$. This tells us that the coefficient $a_{\boldsymbol{q}}$ in f is also nonzero, but that contradicts $\text{mdeg}(f) = \boldsymbol{p} < \boldsymbol{q}$ by definition of $\text{mdeg}(f)$, which ends our proof. ☐

We are now able to prove the Symmetric Function Theorem.

Theorem 5.10 (Symmetric Function Theorem)**.** *Let R be a domain and $f \in R[X_1, \ldots, X_n]$. If f is symmetric, then it also lies in $R[S_1, \ldots, S_n]$.*

Proof. This is trivial for $f = 0$, so we will assume f to be nonzero. We will prove the statement by induction on the multidegree $\boldsymbol{p} = (p_1, \ldots, p_n)$ of f. First we notice that any f having the lowest multidegree $\mathrm{mdeg}(f) = \mathbf{0} = (0, \ldots, 0)$ must be a constant in R, which directly yields $f \in R[S_1, \ldots, S_n]$. Now let $\boldsymbol{p} \neq \mathbf{0}$ and suppose that every nonzero symmetric polynomial $g \in R[X_1, \ldots, X_n]$ with $\mathrm{mdeg}(g) < \boldsymbol{p}$ lies in $R[S_1, \ldots, S_n]$. We want to prove that every symmetrical polynomial $f \in R[X_1, \ldots, X_n]$ with $\mathrm{mdeg}(f) = \boldsymbol{d}$ lies in $R[S_1, \ldots, S_n]$. If no such f exists, we are immediately done. Otherwise, we will take such f and we can then write $f = a_{\boldsymbol{p}} T^{\boldsymbol{p}} + \sum_{\boldsymbol{i} < \boldsymbol{p}} a_{\boldsymbol{i}} T^{\boldsymbol{i}}$, with $a_{\boldsymbol{p}} \in R$ nonzero. By Lemma 5.9, we find that $p_i > p_{i+1}$ holds for all $i \in \{1, \ldots, n-1\}$. So with $d_i := p_{i+1} - p_i$ for such i and $d_n := p_n$, we find that $d_i \geq 0$ for all $i \in \{1, \ldots, n\}$. Additionally, we see that $\sum_{i=j}^{n} d_i = p_j$ for all $j \in \{1, \ldots, j\}$. Now consider the polynomial $S_1^{d_1} S_2^{d_2} \cdots S_n^{d_n}$. We find by Lemma 5.8 in combination with Example 5.7 that

$$\begin{aligned}
\mathrm{mdeg}(S_1^{d_1} S_2^{d_2} \cdots S_n^{d_n}) &= d_1 \mathrm{mdeg}(S_1) + d_2 \mathrm{mdeg}(S_2) + \ldots + d_n \mathrm{mdeg}(S_n) \\
&= (d_1 + \ldots + d_n, d_2 + \ldots + d_n, \ldots, d_n) \\
&= (p_1, \ldots, p_n) \\
&= \boldsymbol{p}.
\end{aligned}$$

So if we take $g := a_{\boldsymbol{p}} S_1^{d_1} S_2^{d_2} \cdots S_n^{d_n}$, then

$$\mathrm{mdeg}(g) = \mathrm{mdeg}(a_{\boldsymbol{p}}) + \mathrm{mdeg}(S_1^{d_1} S_2^{d_2} \cdots S_n^{d_n}) = \mathbf{0} + \boldsymbol{p} = \boldsymbol{p},$$

and by Lemma 5.8, we have

$$\mathrm{lead}(g) = \mathrm{lead}(a_{\boldsymbol{p}}) \prod_{i=1}^{n} \mathrm{lead}(S_i)^{d_i} = a_{\boldsymbol{p}}.$$

So f and g have the same multidegree and the same leading coefficient, so their leading terms kill eachother in $f - g$. If $f - g = 0$, then certainly $f = g \in R[S_1, \ldots, S_n]$. Otherwise, we have that $f - g$ consists of terms with multidegree smaller than $\boldsymbol{p}$, so $\mathrm{mdeg}(f - g) < \boldsymbol{p}$. By Example 5.2, we see that g is a symmetric function as it is the product of symmetrical functions S_k. Hence $f - g$ is also a symmetric function. By our induction hypothesis, $f - g \in R[S_1, \ldots, S_n]$ holds, as $f - g$ is symmetric of multidegree strict less than $\boldsymbol{p}$. Since g also lies in $R[S_1, \ldots, S_n]$, so does $f = (f - g) + g$. This finishes our proof. $\square$

The following two corollaries from the Symmetric Function Theorem will be used in the proof of Runge's Theorem.

Corollary 5.11. *Let R' be a domain and let K be the field of fractions of R'. Let*

$$f = Y^n + b_1 Y^{n-1} + \ldots + b_n \in R'[Y] \subset K[Y]$$

be a monic polynomial of degree $n \geq 1$. Let L be a field extension over K, in which f completely factors into linear polynomials. So $f(Y) = (Y - a_1)(Y - a_2) \cdots (Y - a_n)$, where the roots $a_1, \ldots, a_n$ of $f(Y)$ all lie in L. Let R be a domain that includes R'. Let $g(Z) \in R[Z]$ be another polynomial, then the product $\prod_{i=1}^{n} g(a_i)$ lies in R.

Proof. Note that $G(X_1, \ldots, X_n) := \prod_{i=1}^{n} g(X_i)$ is a symmetric polynomial in $R[X_1, \ldots, X_n]$. By the Symmetric Function Theorem 5.10, it follows that $G(X_1, \ldots, X_n) \in R[S_1, \ldots, S_n]$. If we expand the polynomial $(Y - X_1)(Y - X_2) \cdots (Y - X_n)$, we get $Y^n - S_1 Y^{n-1} + S_n Y^{n-2} - \ldots + (-1)^n S_n$. So S_k is $(-1)^k$ times the coefficient of Y^{n-k} in the polynomial $(Y - X_1)(Y - X_2) \cdots (Y - X_n)$. If we now substitute X_i for a_i, we see that S_k, which depends on $X_1, \ldots, X_n$, becomes $(-1)^k$ times the coefficient of Y^{n-k} in the polynomial

$$(Y - a_1)(Y - a_2) \cdots (Y - a_n) = f(Y) = Y^n + b_1 Y^{n-1} + \ldots + b_n.$$

So the S_k become $(-1)^k b_k$. Since $b_k \in R' \subset R$, we get $\prod_{i=1}^{n} g(a_i) \in R[-b_1, b_2, -b_3, \ldots, (-1)^n b_n] \subset R$. $\square$

Corollary 5.12. *Let R' be a domain and K its fraction field, and let*

$$f = Y^n + b_1 Y^{n-1} + \ldots + b_n \in R'[Y] \subset K[Y]$$

be a monic polynomial of degree $n \geq 1$ such that b_n is a unit in R'. Let L be a field extension over K in which f has the factorization $f(Y) = (Y - a_1)(Y - a_2) \cdots (Y - a_n)$, where the roots $a_1, \ldots, a_n$ of $f(Y)$ all lie in L. Let R be a domain that includes R'. Let $g(Z) \in R[Z, Z^{-1}]$, then the product $\prod_{i=1}^{n} g(a_i)$ lies in R.

Proof. We can write $g(Z) = Z^{-k}h(Z)$, for some $k \in \mathbb{N}$ and $h(Z) \in R[Z]$. Because b_n is a unit in R', we have $(\prod_{i=1}^{n} a_i)^{-k} = (\pm b_n)^{-k} \in R' \subset R$. By the previous corollary we have that $\prod_{i=1}^{n} h(a_i)$ lies in R, hence

$$\prod_{i=1}^{n} g(a_i) = \left(\prod_{i=1}^{n} a_i\right)^{-k} \prod_{i=1}^{n} h(a_i) \in R.$$

□

6 Runge's Theorem

This section will focus on Runge's Theorem which is the main theorem of this thesis. First we give the statement after which we will look at examples that demonstrates the power of this theorem. After that we will give a proof of this theorem.

6.1 Runge's Theorem

Carl Runge has proved in 1887 [12] that certain classes of Diophantine equations have only finitely many integral solutions. Runge's Theorem (whose formulations differ in the literature) is as follows:

Theorem 6.1 (Runge's Theorem [12]). *Let $F(X,Y) = \sum_{i=0}^{d_1} \sum_{j=0}^{d_2} a_{i,j} X^i Y^j \in \mathbb{Z}[X,Y]$ be a polynomial of positive X-degree d_1 and positive Y-degree d_2. Suppose that $F(X,Y)$ is irreducible in $\mathbb{Q}[X,Y]$, and suppose that the equation*

$$F(x,y) = 0$$

has infinitely many solutions with $x, y \in \mathbb{Z}$. Then the following properties hold:

1. *In the xy-plane, no point of $D(F)$ lies above the line connecting $(d_1, 0)$ and $(0, d_2)$;*
2. *For $\lambda = d_1/d_2$, the λ-leading part of F satisfies*

$$\sum_{(i,j)\in D_\lambda(F)} a_{i,j} X^i Y^j = ap^k,$$

where $a \in \mathbb{Z}, k \in \mathbb{N}$ and $p = p(X,Y) \in \mathbb{Z}[X,Y]$ is an irreducible polynomial

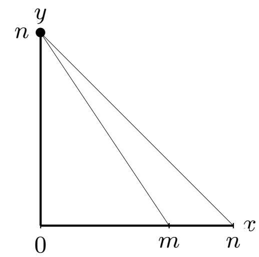
One can see from this theorem that every irreducible F that does not satisfy both properties has only finitely many integral solutions for the equation $F(x,y) = 0$. Before we prove this theorem, we will investigate in what ways we can use this theorem. The proof of a particular case of Runge's Theorem 6.1 has been given by Martin Klazar [8]. It can also be proven as a corollary of Runge's Theorem. The particular case is as follows:

Corollary 6.2 (Runge, a particular case). *Let $F \in \mathbb{Z}[X,Y]$ be a nonzero and irreducible polynomial in $\mathbb{Q}[X,Y]$. Suppose that $\deg F = n$ and that $\deg_Y F = n$ for some integer $n \geq 2$. Let $f(X,Y) \coloneqq F_1(X,Y)$ be the leading degree n form of $F(X,Y)$. Suppose that $f(1,Y)$ is reducible in $\mathbb{Q}[Y]$ and has only simple roots. Then the equation*

$$F(x,y) = 0$$

has only finitely many solutions $x, y \in \mathbb{Z}$.

Proof. If we want to use Theorem 6.1, we must prove that $m \coloneqq \deg_X F \geq 1$. Suppose that $m = 0$, then $f(1,Y) = f(X,Y) = cY^n$ for some nonzero constant c, which has $Y = 0$ as root of higher multiplicity, which leads to a contradiction and shows that indeed $m \geq 1$.
Suppose that there are infinitely many integer solutions to $F(x,y) = 0$. We then have that F must satisfy the two properties of Runge's Theorem 6.1. We have $m \leq \deg F = n$. We will consider two cases. First assume that $m < n$, then the only lattice point (i,j), with $i,j \geq 0$ and $i+j = n$, that does not lie above the line connecting $(m,0)$ and $(0,n)$ is the point $(0,n)$, as suggested by the picture below.



By property 1 of Runge's Theorem 6.1, we therefore have that $f(X,Y) = cY^n$, which again leads to a contradiction. Now assume that $m = n$. We can then use property 2 of Runge's Theorem 6.1, and write

$$f(X,Y) = \sum_{(i,j)\in D_1(F)} a_{i,j} X^i Y^j = ap^k$$

for some $0 \neq a \in \mathbb{Z}$, $k \in \mathbb{N}$ and an irreducible polynomial $p = p(X,Y) \in \mathbb{Z}[X,Y]$. In particular, we have

$$f(1,Y) = ap(1,Y)^k.$$

We can't have $k > 1$, as $f(1,Y)$ only has simple roots and therefore can't have a non-constant polynomial as a factor of higher multiplicity. So we must have $k = 1$. Because the n-form $f(X,Y)$ is the homogenization of $f(1,Y)$, we find that $f(X,Y)$ is reducible in $\mathbb{Q}[X,Y]$. This however contradicts with the fact that $f(X,Y)$ is an integer times p, where p is irreducible in $\mathbb{Z}[X,Y]$. We conclude that $F(x,y) = 0$ has only finitely many integer solutions. □

One might wonder if such polynomials as described above actually exist. For this the example found in [8] will be considered.

Example 6.3. *Consider the equation of the form*

$$Y^n = (aX)^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0,$$

where $n \geq 2$ and $a, a_i \in \mathbb{Z}$, with $a \neq 0$. If the polynomial on the right side is not a d-th power in $\mathbb{Z}[X]$ for any divisor $d \geq 2$ of n, then the equation holds for only finitely many integers $x, y \in \mathbb{Z}$.

Proof. Let $F = Y^n - ((aX)^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0) \in \mathbb{Z}[X,Y]$. We see that $n = \deg F = \deg_Y F \geq 2$ and that $f(X,Y) = Y^n - (aX)^n$. So

$$f(1,Y) = Y^n - a^n = (Y-a)(Y^{n-1} + aY^{n-2} + \ldots + a^{n-2}Y + a^{n-1})$$

is reducible, whose solutions are of the form $a \cdot \zeta$, where ζ is any n-th root of unity. So we have n different solutions since a is nonzero, so $f(1,Y)$ has only simple roots. Obviously F is nonzero, so it only remains for us to prove that F is irreducible in $\mathbb{Q}[X,Y]$. This follows from the fact that the right-hand side is not a d-th power in $\mathbb{Z}[X]$ for any divisor $d \geq 2$ of n. We can therefore apply the particular case of Runge 6.2 and see that $F(x,y) = 0$ indeed has finitely many integral solutions. □

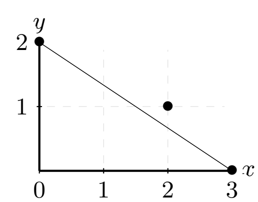
Let us look at some more examples where we can apply Runge's Theorem.

Example 6.4. *The equation*

$$F(x,y) := x^3 + x^2y + y^2 = 0$$

has only finitely many integral solutions.

Proof. We will use Runge's Theorem 6.1. $F(X,Y)$ is irreducible in $\mathbb{Z}[X,Y]$, so by Gauss's lemma, it is also irreducible in $\mathbb{Q}[X,Y]$. We look at the xy-plane, where we draw the Newton dots of $F(X,Y)$. We also draw a line between points $(0, \deg_Y F) = (0,2)$ and $(\deg_X F, 0) = (3,0)$. We get the following:



It can be seen that a Newton dot does in fact lie above the line. This violates property 1 of Runge's Theorem 6.1. We can conclude that $F(x, y) = 0$ holds for only finitely many integers x and y.
An alternative proof that does not use Runge's Theorem 6.1 is as follows. Let $x, y \in \mathbb{Z}$ and suppose that $x^3 + x^2y + y^2 = 0$ holds. By the quadratic formula we get

$$y = \frac{-x^2 \pm \sqrt{x^4 - 4x^3}}{2}.$$

Since y is an integer, $x^4 - 4x^3 = x^2(x^2 - 4x)$ must be a perfect square. Therefore $x^2 - 4x$ must be a perfect square as well, so there exists $a \in \mathbb{Z}$ such that $x^2 - 4x = a^2$. By using the quadratic formula again we get

$$x = \frac{4 \pm \sqrt{16 + 4a^2}}{2} = 2 \pm \sqrt{4 + a^2}.$$

Since x is an integer, $4 + a^2$ must also be a perfect square. So a^2 must differ 4 with another perfect square. Because large perfect squares do not lie close to other perfect squares, a must be small. A quick inspection shows that $a = 0$ is the only possibility. This gives the two possibilities $x = 0$ and $x = 4$. If $x = 0$, we have $0 = F(0, y) = y^2$, hence $y = 0$. If $x = 4$, we have

$$y = \frac{-x^2 \pm \sqrt{x^4 - 4x^3}}{2} = \frac{-16 \pm \sqrt{256 - 256}}{2} = -8.$$

So $(x, y) = (0, 0)$ and $(x, y) = (4, -8)$ are the only two integer solutions to $F(x, y) = 0$.

□

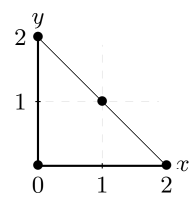
The previous example shows us that applying Runge's Theorem is an easier way to prove that the polynomial has only finitely many intergral solutions. Another example is the following:

Example 6.5. *The equation*

$$F(x, y) := x^2 - xy - 2y^2 + 2 = 0$$

has only finitely many integral solutions.

Proof. Again, $F(X, Y)$ is irreducible in $\mathbb{Q}[X, Y]$. Just like the previous example, we draw points on all coordinates (i, j), where $F(X, Y)$ contains a term of the form $a_{i,j}X^iY^j$, where $a_{i,j} \neq 0$. We also draw a line between points $(0, \deg_Y F) = (0, 2)$ and $(\deg_X F, 0) = (2, 0)$. We get the following:



It can be seen that no point lies above the line. So property 1 of Runge's Theorem 6.1 is not violated. We notice that the λ-leading part of F, with $\lambda = \deg_X F / \deg_Y F = 2/2$ is

$$F_\lambda(X, Y) = X^2 - XY - 2Y^2 = (X + Y)(X - 2Y).$$

This is a product of two irreducible factors that are not equal to eachother up to a constant factor. This violates property 2 of Runge's Theorem 6.1. We can conclude that $F(x, y) = 0$ holds for only finitely many integers x and y.
One could also have seen this by noticing that the equation $F(x, y) = 0$ can be rewritten as $(x + y)(x - 2y) = -2$. If x and y are integral solutions to this equation, both factors on the left hand side must be integers as well. Because the product of these two factors is -2, there are the following 4 possibilities.

$(x + y)$	$(x - 2y)$	x	y
-2	1	-1	-1
-1	2	0	-1
1	-2	0	1
2	-1	1	1

So there are exactly four, hence only finitely many, integral solutions to $F(x, y) = 0$. □

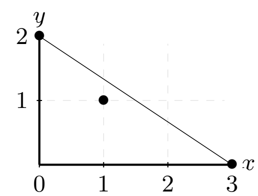
Let us now look at an example which does have infinitely many integral solutions.

Example 6.6. *The equation*

$$F(x, y) := x^3 + xy + y^2 = 0$$

has infinitely many integral solutions.

Proof. Similar with the previous example. We look at the xy-plane, where we draw the Newton dots of F. We also draw a line between points $(0, \deg_Y F) = (0, 2)$ and $(\deg_X F, 0) = (3, 0)$. We get the following:



We notice that no points lie above line ℓ, so property 1 of Runge's Theorem 6.1 has been satisfied by F. Also, with $\lambda = \deg_X F / \deg_Y F = 3/2$, we see that the λ-leading part of F, which is $F_\lambda = X^3 + Y^2$, is irreducible. So also property 2 of Runge's Theorem 6.1 has been satisfied by F, where $a = k = 1$ and $p = X^3 + Y^2$. So we can't use Runge's Theorem to conclude that $F(x, y) = 0$ has only finitely many integral solutions. As it happens this polynomial does have infinitely many integral solutions. We see this because $F(-k(k+1), k(k+1)^2) = 0$ and $F(-k(k+1), -k^2(k+1)) = 0$ hold for any integer k, as can be checked by a simple evaluation. □

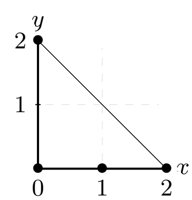
It is not the case that if an irreducible polynomial $F \in \mathbb{Z}[X, Y]$ does satisfy both properties of Runge's Theorem 6.1, that it must have infinitely many integral solutions to $F(x, y) = 0$. The following example shows this.

Example 6.7. *The equation*

$$F(x, y) := x^2 - 2x + y^2 + 1 = 0$$

has only finitely many integral solutions.

Proof. Again we translate the terms of F to points on the xy-plane and draw the line connecting $(\deg_X F, 0)$ and $(0, \deg_Y F)$. We get the following:



We notice that no points lie above this line and for $\lambda = \deg_X F / \deg_Y F = 1$, we see that the λ-leading part of F, which is $F_\lambda = X^2 + Y^2$, is irreducible over $\mathbb{Z}[X]$. So both properties of Theorem 6.1 have been satisfied by F. But the equation $F(X, Y) = 0$ can be rewritten as $(x-1)^2 + y^2 = 0$. Suppose that $x \in \mathbb{Z}$ and $y \in \mathbb{Z}$ satisfy this equation. Since a sum of perfect squares can only be zero if all perfect squares are zero, we get that $x - 1 = 0$ and $y = 0$. This shows that we have only one integral solution, namely $(x, y) = (1, 0)$. □

6.2 Proof of Runge's Theorem

Lemma 6.8. *Let R be a domain. Let $g(Y) = \sum_{i=0}^{s} a_i Y^i \in R[Y]$ be an irreducible polynomial of degree $s \in \mathbb{N}$ that is no multiple of Y. Let $m, e \in \mathbb{N}$ with $\gcd(m, e) = 1$. Let $G(X, Y) = X^{sm} g(Y^e X^{-m}) \in R[X, Y]$. Then G is irreducible in $R[X, Y]$.*

Proof. Since this statement is trivial when g is constant, we assume that g is nonconstant. We have

$$G(X,Y) = X^{sm}\sum_{i=0}^{s} a_i X^{-im}Y^{ie} = \sum_{i=0}^{s} a_i X^{m(s-i)}Y^{ie}. \tag{6.1}$$

Note that for $\lambda \coloneqq m/e$ we have $m(s-i)+\lambda ie = ms$ for every $i \in \{0,\dots,s\}$. This shows us that $G_\lambda = G$. We also see by the definition of G that it is not divisible by a nonunit in R, because neither is g. Suppose that G is not irreducible in $R[X,Y]$, then we can write $G(X,Y) = F(X,Y)H(X,Y)$ for two nonconstant polynomials $F, H \in R[X,Y]$. We have $G = G_\lambda = F_\lambda H_\lambda$ and $\deg_\lambda G = \deg_\lambda F + \deg_\lambda H$ by Remark 4.60. Note that we can use this remark as $R[X,Y]$ is the subset of $K((X^{-1}))[Y]$, where K is the field of fractions of R. We can write $F_\lambda = \sum_{(i,j)\in D_\lambda(F)} b_{i,j}X^iY^j$ and $H_\lambda = \sum_{(k,l)\in D_\lambda(H)} c_{k,l}X^kY^l$ with $b_{i,j}, c_{k,l} \in R$. For each $(i,j) \in D_\lambda(F)$ and $(k,l) \in D_\lambda(H)$ we have $i + \lambda j = \deg_\lambda F$ and $k + \lambda l = \deg_\lambda H$, hence $(i+k)+\lambda(j+l) = \deg_\lambda G = ms$. We multiply this last equation by e and get $(i+k)e+m(j+l) = mes$. Since $\gcd(m,e) = 1$, this shows us that $i + k \equiv 0 \mod m$ and that $j + l \equiv 0 \mod e$. For any other $(i',j') \in D_\lambda(F)$, we also have $i' + k \equiv 0 \mod m$ and $j' + l \equiv 0 \mod e$. We therefore have $i = i' \equiv 0 \mod m$ and $j = j' \equiv 0 \mod e$ for any $(i,j), (i',j') \in D_\lambda(F)$. Since $g(Y)$ is no multiple of Y, we in particular have $a_0 \neq 0$. We also have $a_s \neq 0$ since g is of degree s. We use this to see from (6.1) that X and Y do not divide G and therefore also do not divide F_λ. So in particular we must have $(\deg_F, 0) \in D_\lambda(F)$ and $(0, \lambda^{-1}\deg_F) \in D_\lambda(F)$. This shows us that for all $(i,j) \in D_\lambda(F)$ we have $i \equiv 0 \mod m$ and $j \equiv 0 \mod e$. So any nonzero term $b_{i,j}X^iY^j$ of F_λ can be written as $b_{i,j}(X^m)^a(Y^e)^b$ for some $a, b \in \mathbb{N}$. We thus have $F_\lambda(X,Y) = F'(X^m,Y^e)$ for some $F'(X,Y) \in R[X,Y]$. In the same way we have $H_\lambda(X,Y) = H'(X^m,Y^e)$ for some $H'(X,Y) \in R[X,Y]$. We now have the following:

$$F'(1,Y^e)H'(1,Y^e) = F_\lambda(1,Y)H_\lambda(1,Y) = G(1,Y) = g(Y^e).$$

We substitute Y^e by Y and get:

$$g(Y) = F'(1,Y)H'(1,Y).$$

We have $F'(1,Y) = F_\lambda(1,Y^{1/e}) = \sum_{(i,j)\in D_\lambda(F)} b_{i,j}Y^{j/e}$, which is nonconstant since F_λ is nonconstant. In the same way we have that $H'(1,Y)$ is nonconstant, which shows that $g(Y)$ is reducible. This is a contradiction from which we may conclude that G is indeed irreducible in $R[X,Y]$. ☐

The next theorem has been described by Hilliker and Straus [7]. We will follow their proof here.

Theorem 6.9. *Let $f(X) \in K(\{(X^{-1})^*\})$ be a nonzero convergent Puiseux series at infinity over an algebraic number field $K \subset \overline{\mathbb{Q}}$ of degree $s \coloneqq [K:\mathbb{Q}]$. Write f as in (4.8). So*

$$f(X) = \sum_{n=-m}^{\infty} a_n X^{-n/e},$$

with $m \in \mathbb{Z}$, $e \in \mathbb{N}$, and where all coefficients a_n lie in K such that $a_{-m} \neq 0$. Then their exists nonzero $P(X,Y) \in \mathbb{Z}[X,Y]$ such that the following properties hold:

- $\deg_Y P \leq se$.
- *If $x \in \mathbb{Z}$ and $f(x) \in \mathbb{Z}$, then $P(x, f(x)) = 0$.*
- *P_λ is a monomial for every $\lambda \in \mathbb{R}_{>0}$ with $\lambda \neq m/e$.*
- *If $\lambda = m/e > 0$, then P_λ is a constant multiple of a power of X times a power of an irreducible polynomial in $\mathbb{Z}[X,Y]$.*
- *If the set $\{x \in \mathbb{Z} \mid f(x) \in \mathbb{Z}\}$ is of infinite cardinality, then $P(X, f(X)) = 0$ in $K(\{X^{-1/e}\})$.*

Proof. Let $M \in \mathbb{N}_0$ be the cardinality of the finite set

$$B \coloneqq \{(i,j) \in \mathbb{Z}^2 \mid 1 \leq j < se, -jm/e \leq i < 0\}. \tag{6.2}$$

Let $N \in \mathbb{N}$ satisfy

$$N \geq Me + (se-1)m.$$

Since K is an algebraic number field, there exists a monic irreducible polynomial $G(X) \in \mathbb{Q}[X]$ of degree s which has a root $\theta_1 \in K$ that generates K. So $K = \mathbb{Q}(\theta_1)$. Let L be the splitting field of $G(X)$ over K. So there exist $\theta_2, \dots, \theta_s \in L$ such that $G(X) = (X-\theta_1)\cdots(X-\theta_s)$ in $L[X]$. Any $c \in K$ can be uniquely written as $c_0 + c_1\theta_1 + \ldots + c_{s-1}\theta_1^{s-1}$ for some $c_0, \ldots, c_{s-1} \in \mathbb{Q}$. We denote the polynomial $c_0 + c_1 W + \ldots + c_{s-1}W^{s-1} \in \mathbb{Q}[W]$ by $g_c(W)$. For $\sigma \in \{1, \ldots, s\}$, we call

$$c^{(\sigma)} \coloneqq g_c(\theta_\sigma) = c_0 + c_1\theta_\sigma + \ldots + c_{s-1}\theta_\sigma^{s-1} \in L$$

the σ-th conjugate of c in K. In particular we have $c^{(1)} = c$.

Let $\zeta \in \mathbb{C}$ be a primitive e-th root of unity. This means that $\zeta^e = 1$ and that the elements $\zeta^0, \ldots, \zeta^{e-1}$ are all different solutions to $Y^e = 1$. It follows that $Y^e - X^{-1}$, a polynomial in Y over $\mathbb{Q}[X, X^{-1}]$, factorizes over the field extension $\mathbb{Q}(\zeta)(X^{-1/e})$ as

$$Y^e - X^{-1} = (Y - \zeta^0 X^{-1/e}) \cdots (Y - \zeta^{e-1}X^{-1/e}).$$

Notice that X^{-1} is a unit in $\mathbb{Q}[X, X^{-1}]$. We now consider the Laurent polynomial:

$$Y - \sum_{n=-m}^{N} g_{a_n}(W)V^n \in \mathbb{Q}[X, X^{-1}][Y][W][V, V^{-1}].$$

By Corollary 5.12 we find that

$$\prod_{\mathscr{E}=0}^{e-1} \left(Y - \sum_{n=-m}^{N} g_{a_n}(W)(\zeta^{\mathscr{E}} X^{-1/e})^n \right) \in \mathbb{Q}[X, X^{-1}][Y][W].$$

By Corollary 5.11 we find that

$$\prod_{\sigma=1}^{s} \prod_{\mathscr{E}=0}^{e-1} \left(Y - \sum_{n=-m}^{N} g_{a_n}(\theta_\sigma)(\zeta^{\mathscr{E}} X^{-1/e})^n \right) \in \mathbb{Q}[X, X^{-1}][Y].$$

We thus find for each $\beta = 0, \ldots, M$ that

$$F(X, Y; \beta) \coloneqq X^\beta \prod_{\sigma=1}^{s} \prod_{\mathscr{E}=0}^{e-1} \left(Y - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^n \right) \in \mathbb{Q}[X, X^{-1}][Y]. \tag{6.3}$$

If we write $F(X, Y; \beta) = \sum_{(i,j)\in D(F(X,Y;\beta))} b_{\beta,i,j}X^iY^j$,we see from (6.3) that if $(i, j) \in D(F(X, Y; \beta))$, then $0 \le j \le se$ and $i \in \mathbb{Z}$ must hold. The terms $b_{\beta,i,j}X^iY^j$ of $F(X, Y; \beta)$ can thus be divided into the following three categories:

- Terms $b_{\beta,i,j}X^iY^j$ with $i \ge 0$.
- Terms $b_{\beta,i,j}X^iY^j$ with $i < 0$ and $i + jm/e \le -1/e$.
- Terms $b_{\beta,i,j}X^iY^j$ with $i < 0$ and $i + jm/e \ge 0$.

Note that the terms in the first category lie in $\mathbb{Q}[X, Y]$ and that the terms in the second category satisfy $\mathrm{ord}_X(b_{\beta,i,j}X^i(f(X))^j) \le -1/e$. The terms in the third category satify $j \ne 0$, as otherwise $i + jm/e = i \ge 0$ would contradict $i < 0$. They also satify $j \ne se$, as it would otherwise follow from (6.3) that $i = \beta \ge 0$, which again would contradict $i < 0$. We can conclude that the terms in the third category satisfy $(i, j) \in B$ as defined in (6.2). It can be seen that the terms in the first and second category do not satisfy $(i, j) \in B$. We can add terms of the same category together and get

$$F(X, Y; \beta) = P(X, Y; \beta) + S(X, Y; \beta) + \sum_{(i,j)\in B} b_{\beta,i,j}X^iY^j,$$

with $P(X, Y; \beta) \in \mathbb{Q}[X, Y]$ and $S(X, Y; \beta) \in \mathbb{Q}[X, X^{-1}][Y]$ with $\mathrm{ord}_X(S(X, f(X); \beta)) \le -1/e$. We want to find $c_0, \ldots, c_M \in \mathbb{Z}$, not all zero, such that $\sum_{\beta=0}^{M} c_\beta b_{\beta,i,j} = 0$ for all $(i, j) \in B$. This is a homogeneous system of $|B| = M$ linear equations in $M + 1$ variables. It therefore indeed has a solution. We may scale these $c_0, \ldots, c_M$ by multiplying with any nonzero integer and still find a suitable solution. We scale by a nonzero integer such that $c_\beta \cdot P(X, Y; \beta) \in \mathbb{Z}[X][Y]$ for all $\beta \in \{0, \ldots, M\}$. This gives us

$$\sum_{\beta=0}^{M} c_\beta F(X, Y; \beta) = \sum_{\beta=0}^{M} c_\beta P(X, Y; \beta) + \sum_{\beta=0}^{M} c_\beta S(X, Y; \beta),$$

hence

$$Q(X,Y) \coloneqq \sum_{\beta=0}^{M} c_\beta P(X,Y;\beta) = \sum_{\beta=0}^{M} c_\beta F(X,Y;\beta) - \sum_{\beta=0}^{M} c_\beta S(X,Y;\beta) \in \mathbb{Z}[X,Y].$$

We are interested in the order of $Q(X,f(X)) \in K(\{(X^{-1})^*\})$. First note that

$$f(X) - \sum_{n=-m}^{N} a_n^{(\sigma)} (\zeta^{\mathscr{E}} X^{-1/e})^n \in \mathbb{C}((X^{-1/e}))$$

holds for any $\mathscr{E} \in \{0,\dots,e-1\}$ and $\sigma \in \{1,\dots,s\}$. We see that

$$\begin{aligned}\operatorname{ord}_X\left(f(X) - \sum_{n=-m}^{N} a_n^{(\sigma)} (\zeta^{\mathscr{E}} X^{-1/e})^n\right) &\le \max\left(\operatorname{ord}_X f(X), \operatorname{ord}_X\left(\sum_{n=-m}^{N} a_n^{(\sigma)} (\zeta^{\mathscr{E}} X^{-1/e})^n\right)\right) \\ &= \max(m/e, m/e) \\ &= m/e.\end{aligned}$$

In the case where $\mathscr{E} = 0$ and $\sigma = 1$ we find

$$\operatorname{ord}_X\left(f(X) - \sum_{n=-m}^{N} a_n^{(1)} (\zeta^0 X^{-1/e})^n\right) = \operatorname{ord}_X\left(\sum_{n=N+1}^{\infty} a_n X^{-n/e}\right) \le -(N+1)/e.$$

Together this gives us

$$\operatorname{ord}_X(F(X,f(X);\beta)) \le \beta + (se-1)m/e - (N+1)/e \le M + (se-1)m/e - N/e - 1/e \le -1/e.$$

We therefore have

$$\begin{aligned}\operatorname{ord}_X(Q(X,f(X))) &= \operatorname{ord}_X\left(\sum_{\beta=0}^{M} c_\beta F(X,f(X);\beta) - \sum_{\beta=0}^{M} c_\beta S(X,f(X);\beta)\right) \\ &\le \max_\beta(\operatorname{ord}_X(F(X,f(X);\beta)), \operatorname{ord}_X(S(X,f(X);\beta))) \\ &\le \max_\beta(-1/e, -1/e) \\ &< 0.\end{aligned}$$

Since $f(X)$ is a convergent Puiseux series at infinity, so is $Q(X,f(X))$. So there exists $R \in \mathbb{R}$ such that $Q(x,f(x))$ converges in $\mathbb{C}$ for all $x \in \mathbb{Z}$ with $|x| > R$. Because $\operatorname{ord}_X(Q(X,f(X))) < 0$, we see that $\lim_{x\to\pm\infty} Q(x,f(x)) = 0$. So there exists $R' > R$ such that for all $x \in \mathbb{Z}$ with $|x| > R'$, we have $|Q(x,f(x))| < \frac{1}{2}$. There are only finitely many $x \in \mathbb{Z}$ with $|x| \le R'$, so there exists a nonzero polynomial $Q'(X) \in \mathbb{Z}[X]$ with $Q'(X) = 0$ for all such x. We now take

$$P = P(X,Y) \coloneqq Q'(X)Q(X,Y)$$

and prove the properties.

First of all, we see that $F(X,Y;\beta)$ is of Y-degree se for all $\beta \in \{0,\dots,M\}$. Because $P(X,Y;\beta)$ consists of terms from $F(X,Y;\beta)$, it has a Y-degree that is at most se. Therefore $Q(X,Y)$ also has Y-degree at most se. The same can then be said about $P(X,Y)$ as $Q'(X)$ is of zero Y-degree.

Now let $x \in \mathbb{Z}$ such that $f(x)$ converges and lies in $\mathbb{Z}$. If $x \le |R'|$, then

$$P(x,f(x)) = Q'(x)Q(x,f(x)) = 0 \cdot Q(x,f(x)) = 0.$$

And if $|x| > R'$, then $|Q(x,f(x))| < \frac{1}{2}$. Since $Q \in \mathbb{Z}[X,Y]$, we also have $Q(x,f(x)) \in \mathbb{Z}$. This combined gives us $Q(x,f(x)) = 0$, hence $P(x,f(x)) = 0$.

Now let $\lambda \in \mathbb{R}_{>0}$ with $\lambda \ne m/e$. Note that since $Q'(X) \in \mathbb{Z}[X,Y]$ is a nonzero polynomial without terms with Y as a factor, we must have $Q'_\lambda(X) = cX^t$ for some nonzero $c \in \mathbb{Z}$ and $t \in \mathbb{N}_0$. We want to find P_λ using Remark 4.60. Let γ be the largest element in $\{0,\dots,M\}$ such that $c_\gamma \neq 0$. First suppose that $\lambda > m/e$. This gives

$$\left(Y - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^n\right)_\lambda = Y,$$

for any $\sigma \in \{1,\dots,s\}$ and $\mathscr{E} \in \{0,\dots,e-1\}$. Therefore we have

$$F(X,Y;\beta)_\lambda = (X^\beta)_\lambda \prod_{\sigma=1}^{s}\prod_{\mathscr{E}=0}^{e-1}\left(Y - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^n\right)_\lambda = X^\beta Y^{se}.$$

So $X^\beta Y^{se} \in \mathbb{Q}[X,Y]$ is the term of F_β that is of the highest λ-degree. This gives us

$$P(X,Y;\beta)_\lambda = X^\beta Y^{se},$$

and therefore we see that

$$P_\lambda = cX^t\left(\sum_{\beta=0}^{M} c_\beta P(X,Y;\beta)\right)_\lambda = cX^t c_\gamma P(X,Y,Z;\gamma)_\lambda = cc_\gamma X^{\gamma+t}Y^{se},$$

which indeed is a monomial. Note that this additionally shows that P is nonzero since its λ-leading part is nonzero.

Now suppose that $\lambda < m/e$. From this, we can deduce that $m > 0$. This gives

$$\left(Y - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^n)_\lambda = -a_{-m}^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^{-m}\right)$$

for any $\sigma \in \{1,\dots,s\}$ and $\mathscr{E} \in \{0,\dots,e-1\}$. Therefore we have

$$\begin{aligned}
F(X,Y;\beta)_\lambda &= (X^\beta)_\lambda \prod_{\sigma=1}^{s}\prod_{\mathscr{E}=0}^{e-1}\left(Y - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^n\right)_\lambda \\
&= X^\beta \prod_{\sigma=1}^{s}\prod_{\mathscr{E}=0}^{e-1} -a_{-m}^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^{-m} \\
&= X^\beta \prod_{\sigma=1}^{s} \pm (a_{-m}^{(\sigma)})^e X^m \\
&= d_\beta X^{\beta+ms},
\end{aligned}$$

where $d_\beta \in \mathbb{Q}$ is nonzero, since a_{-m} is nonzero. We find in a similar way as in the previous case

$$P(X,Y;\beta)_\lambda = d_\beta X^{\beta+ms},$$

and therefore we see that

$$P_\lambda = cX^t c_\gamma P(X,Y,Z;\gamma)_\lambda = cc_\gamma d_\gamma X^{\gamma+ms+t},$$

which also is a monomial.

Now suppose that $\lambda = m/e$. From this, we can deduce that $m > 0$. This gives

$$\left(Y - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^n\right)_\lambda = Y - a_{-m}^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^{-m}.$$

for any $\sigma \in \{1,\dots,s\}$ and $\mathscr{E} \in \{0,\dots,e-1\}$. So

$$P_\lambda = cX^t c_\gamma X^\gamma \prod_{\sigma=1}^{s}\prod_{\mathscr{E}=0}^{e-1}(Y - a_{-m}^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^{-m}) \in \mathbb{Q}[X,X^{-1}][Y] \subset \mathbb{Q}(X)[Y].$$

Let $m_1, e_1 \in \mathbb{N}$ be such that $\gcd(m_1, e_1) = 1$ and $m/e = m_1/e_1$. Let $h(Y) \in \mathbb{Z}[Y]$ be the minimum polynomial of $a_{-m}^{e_1}$. So $h(Y)$ is irreducible in $Z[Y]$ and contains the nonzero element ${a_{-m}}^{e_1}$ as a root. Therefore we see in particular that $h(Y)$ is no multiple of Y. Let $b \in \mathbb{Z}$ be the leading coefficient of h and s_1 the degree of h. Since h is irreducible, we have by Lemma 6.8 that

$$G(X,Y) \coloneqq X^{m_1 s_1} h(Y^{e_1} X^{-m_1}) \in \mathbb{Z}[X,Y]$$

is irreducible in $\mathbb{Z}[X,Y]$, and therefore also irreducible in $\mathbb{Q}(X)[Y]$. For any $\sigma \in \{1,\ldots,s\}$ and $\mathscr{E} \in \{0,\ldots,e-1\}$ we have

$$\begin{aligned}
G(X, {a_{-m}}^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^{-m}) &= X^{m_1 s_1} h(({a_{-m}}^{(\sigma)}(\zeta^{\mathscr{E}} X^{-1/e})^{-m})^{e_1} X^{-m_1}) \\
&= X^{m_1 s_1} h(({a_{-m}}^{(\sigma)})^{e_1} \zeta^{-m e_1 \mathscr{E}} X^{m e_1/e - m_1}) \\
&= X^{m_1 s_1} h(({a_{-m}}^{e_1})^{(\sigma)} \zeta^{-m_1 e \mathscr{E}}) \\
&= X^{m_1 s_1} h(({a_{-m}}^{e_1})^{(\sigma)} (\zeta^{e})^{-m_1 \mathscr{E}}) \\
&= X^{m_1 s_1} h(({a_{-m}}^{e_1})^{(\sigma)}) \\
&= X^{m_1 s_1} h({a_{-m}}^{e_1})^{(\sigma)} \\
&= X^{m_1 s_1} 0^{(\sigma)} \\
&= 0.
\end{aligned}$$

Since all roots of P_λ, when viewed as a polynomial in Y, are roots of the irreducible polynomial G, we conclude that P_λ has no irreducible factors in $\mathbb{Q}(X)[Y]$ besides G. Since G is of Y-degree $s_1 e_1$ and P_λ of Y-degree se, it must follow that $s_1 e_1$ divides se and that $P_\lambda = dG^{se/(s_1 e_1)}$ for some $d \in \mathbb{Q}(X)$. By looking at the leading coefficient of P_λ and $G^{se/(s_1 e_1)}$ when viewed as polynomials in Y, we see that $d = cc_\gamma X^{\gamma+t} b^{-se/(s_1 e_1)}$. So indeed P_λ is a constant ($cc_\gamma b^{-se/(s_1 e_1)}$) times a power of X times a power of an irreducible polynomial in $\mathbb{Z}[X,Y]$.

Now suppose that the set $\{x \in \mathbb{Z} | f(x) \in \mathbb{Z}\}$ is of infinite cardinality. Because $f(X) \in K(\{X^{-1/e}\})$ is an algebraic element over the subfield $\mathbb{Q}(X^{-1}) \subset K(\{X^{-1/e}\})$, so is $P(X, f(X))$. So there exists a nonconstant polynomial $S''(X,Y) \in \mathbb{Q}(X^{-1})[Y]$ with $S''(X, P(X, f(X))) = 0 \in K(\{X^{-1/e}\})$. We may assume S'' to lie in $\mathbb{Z}[X^{-1}, Y]$ by multiplying it with the common denominators of the coefficients of S'', when we view S'' as a polynomial in Y. We then let $S'(X,Y)$ be $X^q S''(X,Y)$, where $q \in \mathbb{N}_0$ is large enough such that $S'(X,Y) \in \mathbb{Z}[X,Y]$. So we also have $S'(X, P(X, f(X))) = 0$. We then let $S \in \mathbb{Z}[X,Y]$ be an irreducible factor of S' that also has $P(X, f(X))$ as root. Suppose that $x \in \mathbb{Z}$ satisfies $f(x) \in \mathbb{Z}$. Then we have $P(x, f(x)) = 0$ and thus $0 = S(x, P(x, f(x))) = S(x, 0)$. We see that x is a root of the polynomial $S(X, 0) \in \mathbb{Z}[X]$. Because the set $\{x \in \mathbb{Z} | f(x) \in \mathbb{Z}\}$ is of infinite cardinality, we see that the polynomial $S(X,0)$ has infinitely many roots. This implies that $S(X,0) = 0$, so Y is a divisor of $S(X,Y)$. Since $S(X,Y)$ was irreducible in $\mathbb{Z}[X,Y]$, we have $S = \pm Y$ and therefore $0 = S(X, P(X, f(X))) = \pm P(X, f(X))$. So we may conclude that indeed $P(X, f(X)) = 0 \in K(\{X^{-1/e}\})$. □

We are now ready to give the proof of Runge's Theorem. We will follow the proof described by Hilliker and Straus [7] here.

Proof Runge's Theorem. If $\deg F = 1$, we have $d_1 = d_2 = 1$ and we can write $F(X,Y) = aX + bY + c$ for some $a, b, c \in \mathbb{Z}$ such that $a \neq 0$ and $b \neq 0$. Property 1 of Runge's Theorem then immediately follows. We have $\lambda = d_1/d_2 = 1$, and the λ-leading part of F is $aX + bY$. Since this is irreducible in $\mathbb{Q}[X,Y]$, we have that the λ-leading part of F is indeed an integer times (a power of) an irreducible polynomial in $\mathbb{Z}[X,Y]$, which shows that Property 2 of Runge's Theorem also holds. Now suppose that $\deg F > 1$. We in particular have $F \in \overline{Q}[X, X^{-1}][Y]$. Let $f_1(X), \ldots, f_{d_2}(X) \in \overline{Q}(\{(X^{-1})^*\})$ be the Puiseux expansions at infinity of F. These expansions are all nonzero as otherwise Y would divide F. By Lemma 4.50, we can factorize F as

$$F(X,Y) = g(X) \prod_{i=1}^{d_2} (Y - f_i(X)),$$

where $g(X) \in \mathbb{Z}[X]$ is the leading coefficient of F, when F is viewed as a polynomial in Y. By Corollary 4.66, we also have that all f_i lie in $K(\{(X^{-1})^*\})$ for some algebraic number field K. Since all Puiseux expansions at infinity of F are convergent and there are only finitely many of them, there exists $R \in \mathbb{R}$ such that $f_i(x)$ converges for all $x \in \mathbb{C}$ with $|x| > R$. Let $a \in \mathbb{Z}$. Since F is irreducible in $\mathbb{Z}[X, Y]$, we have that F is a multiple of $X - a$ only when $F = \pm(X - a)$. Since this would contradict $\deg F > 1$, we see that F is not a multiple of $X - a$ for any $a \in \mathbb{Z}$. By the same reasoning it follows that F is not a multiple of $Y - b$ for any $b \in \mathbb{Z}$. By Corollary 3.4 we see that there are only finitely many integral solutions to $F(x, y) = 0$ with $|x| \leq R$. So there must be infinitely many integral solutions to $F(x, y) = 0$ with $|x| > R$. For each such x, y we have

$$0 = F(x, y) = g(x) \prod_{i=1}^{d_2} (y - f_i(x)).$$

So either $g(x) = 0$ holds, or there must exist $i \in \{1, \ldots, d_2\}$ such that $y = f_i(x)$ holds for these x and y. Since g is a nonzero polynomial, it has only finitely many roots. So there must be an $i \in \{1, \ldots, d_2\}$ such that $y = f_i(x)$ holds for infinitely many x, y satisfying $F(x, y) = 0$ and $|x| > R$. This f_i lies in $K(\{X^{-1/e}\})$ for some $e \in \mathbb{N}$. We can apply Theorem 6.9 to find nonzero $P(X, Y) \in \mathbb{Z}[X, Y]$ such that $P(X, f_i(X)) = 0$ in $K(\{X^{-1/e}\})$. When we view P and F as polynomials in Y over $\mathbb{Z}[X]$, we see that F is irreducible and that one of its roots is also a root of P. Because of this we have that F is a factor of P in $\mathbb{Z}[X, Y]$. So $P = FH$ for some $H \in \mathbb{Z}[X, Y]$. By Lemma 3.16 we have $P_\lambda = F_\lambda H_\lambda$ for all $\lambda \in \mathbb{R}_{>0}$. Suppose that P_λ is a monomial for all $\lambda \in \mathbb{R}_{>0}$. By Lemma 3.18 we then have that F_λ is a monomial as well for all $\lambda \in \mathbb{R}_{>0}$. We then have that $\tilde{F}$ is a monomial by Lemma 3.19. We then apply Lemma 3.20 to see that $(d_1, d_2) \in D(F)$. We may in that case apply Lemma 3.9 and conclude that there are only finitely many integral solutions to $F(X, Y) = 0$, a contradiction. Therefore, there exists $\lambda_0 \in \mathbb{R}_{>0}$ such that P_{λ_0} is not a monomial. By Theorem 6.9 we see that λ_0 is uniquely determined and that P_{λ_0} is a constant multiple of a power of X times a power of an irreducible polynomial $G(X, Y) \in \mathbb{Z}[X, Y]$. Since F_{λ_0} is a factor of P_{λ_0}, we see that F_{λ_0} is also a constant multiple of a power of X times a power of $G(X, Y)$. Since F_{λ_0} is not a monomial, we have that $G(X, Y)$ is not a monomial. Because F_λ is no monomial for only one value $\lambda \in \mathbb{R}_{>0}$, we must have $\lambda = d_1/d_2$. We now fix this λ to be d_1/d_2. We now reverse the roles of X and Y: let $F'(X, Y) \coloneqq F(Y, X)$. We see that $F'(X, Y)$ is of positive X-degree d_2 and of positive Y-degree d_1. We also have $\deg F' = \deg F$ and there are infinitely many integral solutions $x, y \in \mathbb{Z}$ to $F'(x, y) = 0$. By applying the same reasoning as with F, we see that with $\lambda' = \deg_X F' / \deg_Y F' = d_2/d_1 = \lambda^{-1}$ we have that $F'_{\lambda'} = F_\lambda$ is a constant multiple of a power of Y times a power of an irreducible polynomial $G'(X, Y) \in \mathbb{Z}[X, Y]$ that is not a monomial. Since $\mathbb{Z}[X, Y]$ is a unique factorization domain we conclude that $G(X, Y) = G'(X, Y)$ and that F_λ is a constant times a power of this irreducible $G(X, Y)$. This is the second property of Runge's theorom. For the first property we notice that since $G(X, Y)$ is irreducible and not a monomial, it is not divisible by X and therefore there exists $a \in \mathbb{N}$ with $(a, 0) \in D_\lambda(G)$. Furthermore, because F_λ is a constant times a power of G, we have that $(ka, 0) \in D_\lambda(F)$ for some $k \in \mathbb{N}$. We have the restriction $ka \leq d_1$ since F is of X-degree d_1 and $D_\lambda(F) \subset D(F)$. For any $(i, j) \in D(F)$ we now have $i + \lambda j \leq \deg_\lambda(F) = ka + 0\lambda \leq d_1$. This can be rewritten as $d_2 i + d_1 j \leq d_1 d_2$, which shows that no point of $D(F)$ lies above the line connecting $(d_1, 0)$ and $(0, d_2)$. This is the first property of Runge's Theorem. □

7 Zariski Density

In order to generalize Runge's Theorem such that it works with polynomials in three variables, we need to know what this generalized statement could look like. In the case with two variables we talked about having infinitely or only finitely many integral solutions. With one extra variable one could expect a lot more solutions. Having infinitely many integral solutions does not directly mean having 'enough' integral solutions anymore. To make sense of this we need to look at the Zariski topology and about the notion of Zariski-density. The first subsection will introduce the Zariski topology in any finite dimension. The second subsection will show the connection between the notion of Zariski-density and the property of having infinitely many integer solutions to a binary Diophantine equation. As a consequence we can view the property of being Zariski-dense as criterion for having enough integral solutions to such Diophantine equation. Because the Zariski topology is also defined in three variables, we can apply the same criterion. The third and fourth subsection will give some tools for determining Zariski-density that works for any number of variables. The tools will come in handy in our attempt to generalize Runge's Theorem.

7.1 Zariski topology

We will here define the Zariski topology on algebraic varieties in the classical way. Those who are interested in the Zariski topology defined on the spectrum of a ring may want to look at the book on Commutative algebra by David Eisenbud [4].

Definition 7.1. *Let k be an algebraically closed field and $n \in \mathbb{N}$. Let $\mathcal{F} \subset k[X_1, \ldots, X_n]$ be a set of polynomials in n variables over k. We call*

$$V(\mathcal{F}) := \{(x_1, \ldots, x_n) \in k^n \mid f(x_1, \ldots, x_n) = 0 \quad \forall f \in \mathcal{F}\}$$

the variety of $\mathcal{F}$.

Definition 7.2. *Let k be an algebraically closed field and $n \in \mathbb{N}$. We call a subset $S \subset k^n$ Zariski-closed if there exists $\mathcal{F} \subset k[X_1, \ldots, X_n]$ such that $S = V(\mathcal{F})$.*

Lemma 7.3 (Zariski Topology)**.** *Let k be an algebraically closed field and $n \in \mathbb{N}$. Then the Zariski-closed subsets of k^n define a topology on k^n.*

Proof. Note that $V(\emptyset) = k^n$ and that $V(\{1\}) = \emptyset$. So k^n and $\emptyset$ are Zariski-closed.

Now let $S_1, S_2 \subset k^n$ be Zariski-closed subsets. So there exists $\mathcal{F}_1, \mathcal{F}_2 \subset k[X_1, \ldots, X_n]$ such that $V(\mathcal{F}_1) = S_1$ and $V(\mathcal{F}_2) = S_2$. Take $\mathcal{F} = \{f_1 \cdot f_2 \mid f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2\}$. We want to prove that $S_1 \cup S_2 = V(\mathcal{F})$, which shows that the union of any two Zariski-closed subsets is also Zariski-closed. Let $(x_1, \ldots, x_n) \in S_1 \cup S_2$ and $f \in \mathcal{F}$. We can write f as $f_1 f_2$ with $f_1 \in \mathcal{F}_1, f_2 \in \mathcal{F}_2$. We have $f_1(x_1, \ldots, x_n) = 0$ or $f_2(x_1, \ldots, x_n) = 0$ since $(x_1, \ldots, x_n) \in S_1 \cup S_2 = V(\mathcal{F}_1) \cup V(\mathcal{F}_2)$. So $f(x_1, \ldots, x_n) = 0$ holds and therefore we have that $(x_1, \ldots, x_n) \in V(\mathcal{F})$.
Conversely let $(x_1, \ldots, x_n) \in V(\mathcal{F})$. If $(x_1, \ldots, x_n)$ lies in S_1, then certainly $(x_1, \ldots, x_n) \in S_1 \cup S_2$. If it does not lie in S_1 there must exist $f_1 \in \mathcal{F}_1$ with $f_1(x_1, \ldots, x_n) \neq 0$. Now let $f_2 \in \mathcal{F}_2$. Then $f_1 f_2 \in \mathcal{F}$, hence $f_1(x_1, \ldots, x_n) f_2(x_1, \ldots, x_n) = 0$. Since k is a field and therefore a domain, it must follow that $f_2(x_1, \ldots, x_n) = 0$ for all $f_2 \in \mathcal{F}_2$, hence $(x_1, \ldots, x_n) \in V(\mathcal{F}_2) = S_2 \subset S_1 \cup S_2$. So we have that the union of any two Zariski-closed subsets is indeed also Zariski-closed.

Now let I be a (possibly infinite) set of indices and let $S_i \subset k^n$ be a Zariski-closed subset for each $i \in I$. So there exists $\mathcal{F}_i \subset k[X_1, \ldots, X_n]$ such that $S_i = V(\mathcal{F}_i)$. Let

$$\mathcal{F} := \bigcup_{i \in I} \mathcal{F}_i.$$

We want to prove that

$$S := \bigcap_{i \in I} S_i = V(\mathcal{F}),$$

which shows that the intersection of arbitrary many Zariski-closed subsets is again Zariski-closed. Let $(x_1, \ldots, x_n) \in S$ and let $f \in \mathcal{F}$. Then there exists $i \in I$ with $f \in \mathcal{F}_i$. Since $(x_1, \ldots, x_n) \in S_i = V(\mathcal{F}_i)$, we have $f(x_1, \ldots, x_n) = 0$. So indeed $(x_1, \ldots, x_n) \in V(\mathcal{F})$.
Conversely let $(x_1, \ldots, x_n) \in V(\mathcal{F})$ and let $i \in I$. Any $f \in \mathcal{F}_i$ also lies in $\mathcal{F}$, hence $f(x_1, \ldots, x_n) = 0$ hold for all such f, which shows that $(x_1, \ldots, x_n) \in V(\mathcal{F}_i) = S_i$. This holds for any $i \in I$, hence $(x_1, \ldots, x_n) \in S$. So the intersection of arbitrary many Zariski-closed subsets is indeed again Zariski-closed.

So the property of being Zariski-closed satisfies all axioms of being a topology. So we indeed have found a topology on k^n given by the Zariski-closed subsets of k^n. □

Remark 7.4. *Note that any subset $T \subset k^n$ is again a topology with the induced subspace topology. If T is Zariski-closed, then its Zariski-closed subsets are exactly all Zariski-closed subsets of k^n that are included in T.*

The following lemma shows that finite sets are always Zariski-closed.

Lemma 7.5. *Let k be an algebraically closed field and $n \in \mathbb{N}$. Any finite subset of k^n is Zariski-closed.*

Proof. We already saw that the empty set is Zariski-closed. Because the finite union of Zariski-closed subsets of k^n is again Zariski-closed, it suffices to prove that every subset consisting of exactly one element is Zariski-closed. So suppose that $S = \{(x_1, \ldots, x_n)\}$ for some $(x_1, \ldots, x_n) \in k^n$. Let

$$\mathcal{F} = \{X_1 - x_1, \ldots, X_n - x_n\} \subset k[X_1, \ldots, X_n].$$

It can easily be seen that $S = V(\mathcal{F})$. So S is indeed Zariski-closed, which finishes our proof. □

7.2 Connection with earlier problem

In order to show the connection between Zariski-density and having infinitely many integral solutions to binary Diophantine equations, we first need to prove some lemmas and define the notion of Zariski-density.

Lemma 7.6. *Let k be an algebraically closed field and $n \in \mathbb{N}$. Let $F, G \in k[X_1, \ldots, X_n]$ both nonzero, such that F and G share no nonconstant factors in $k[X_1, \ldots, X_n]$. Then there exists nonzero $H \in k[X_1, \ldots, X_{n-1}]$ such that the points $(x_1, \ldots, x_n) \in k^n$ satisfying $F(x_1, \ldots, x_n) = G(x_1, \ldots, x_n) = 0$ also satisfy $H(x_1, \ldots, x_{n-1}) = 0$.*

Proof. Let $r, s \in \mathbb{N}_0$ be the X_n-degree of F and G respectively. We will prove this lemma by complete induction on (r, s), using the lexicographic ordering (See Definition 5.4). First suppose that $r = 0$. We then have $F \in k[X_1, \ldots, X_{n-1}]$, so we simply take $H = F$ and are immediately done. Now suppose that $(r, s) \geq (1, 0)$. If $s < r$, we have $(s, r) < (r, s)$ in the lexicographic ordering. Since the problem is symmetric in G and F, we can swap these and find H by the induction hypothesis. So we may now assume that $r \leq s$. When viewing F and G as polynomials over X_n, let $A \in k[X_1, \ldots, X_{n-1}]$ and $B \in k[X_1, \ldots, X_{n-1}]$ be their respective leading coefficients. We take

$$G' := AG - BY^{s-r}F.$$

Note that $s' := \deg_{X_n} G' < s$. Let $D \in k[X_1, \ldots, X_n]$ be the greatest common divisor of F and G'. Then D must also divide AG. Since F and G share no common nonconstant factors, D must divide A. So we have in particular that $D \in k[X_1, \ldots, X_{n-1}]$. This also shows us that G' is nonzero, because otherwise we would have $D = \gcd(F, G') = \gcd(F, 0) = F \notin k[X_1, \ldots, X_{n-1}]$. We have that G'/D and F/D are nonzero polynomials in $k[X_1, \ldots, X_n]$ that share no nonconstant factors in $k[X_1, \ldots, X_n]$. Since $(\deg_{X_n}(F/D), \deg_{X_n}(G'/D)) = (\deg_{X_n} F, \deg_{X_n} G') = (r, s') < (r, s)$, we can use the induction hypothesis and find nonzero $H' \in k[X_1, \ldots, X_{n-1}]$ that satifies $H'(x_1, \ldots, x_{n-1}) = 0$ for all $(x_1, \ldots, x_n) \in k^n$ such that $(F/D)(x_1, \ldots, x_n) = (G'/D)(x_1, \ldots, x_n) = 0$. Now take $H := DH' \in k[X_1, \ldots, X_{n-1}]$, which is nonzero as D and H' are nonzero as well. Let $(x_1, \ldots, x_n) \in k^n$ such that $F(x_1, \ldots, x_n) = G(x_1, \ldots, x_n) = 0$. Then it follows from the definition of G' that $G'(x_1, \ldots, x_n) = 0$. If $D(x_1, \ldots, x_{n-1}) = 0$ then we directly have $H(x_1, \ldots, x_{n-1}) = 0$. If $D(x_1, \ldots, x_{n-1}) \neq 0$, we have $(F/D)(x_1, \ldots, x_n) = (G'/D)(x_1, \ldots, x_n) = 0$ and therefore $H'(x_1, \ldots, x_{n-1}) = 0$, hence $H(x_1, \ldots, x_{n-1}) = 0$. □

Lemma 7.7. *Let k be an algebraically closed field and let $F(X, Y) \in k[X, Y]$ be an irreducible polynomial. Let $S = V(\{F\}) \subset k^2$ be the set of ordered pairs $x, y \in k$ such that $F(x, y) = 0$. Then the Zariski-closed subsets of S are exactly the set S combined with all finite subsets of S.*

Proof. It is trivial that S is a Zariski-closed subset of S. By Lemma 7.5 and Remark 7.4 we see that any finite subset of S is Zariski-closed in S. We want to show that there are no other Zariski-closed subsets of S. Let $S_1 \subsetneq S$ be a proper Zariski-closed subset of S. We want to prove that S_1 contains only finitely many elements. Since S_1 is Zariski-closed, there exists $\mathcal{F}_1 \subset k[X, Y]$ such that $S_1 = V(\mathcal{F}_1)$. Let $(x_0, y_0) \in S$ be an element that does not lie in $S_1 = V(\mathcal{F}_1)$. Then there must exist $G \in \mathcal{F}_1 \subset k[X, Y]$ such that $G(x_0, y_0) \neq 0$. Since $F(x_0, y_0) = 0$, we have that G is not a multiple of the irreducible polynomial F. So G and F share no nonconstant factors in $k[X, Y]$. By Lemma 7.6, there exists nonzero $H(X) \in k[X]$ such that if $x, y \in k$ satisfy $F(x, y) = G(x, y) = 0$ that $H(x) = 0$ must follow. Since H is nonzero, only finitely many $x \in k$ satisfy $H(x) = 0$. Now suppose that $(x, y) \in S_1 = V(\mathcal{F}_1)$. It then must follow that $G(x, y) = 0$. We also have $F(x, y) = 0$ because $(x, y) \in S$. Therefore we must have $H(x) = 0$. Now suppose there exists infitely many pairs $(x, y) \in k^2$ with $(x, y) \in S_1$. There are finitely many possibilities for x, so there must exists $a \in k$ such that $(a, y) \in S_1$ holds for infinitely many possibilities for $y \in k$. So $G(a, y) = 0$ and $F(a, y) = 0$ hold for infinitely many possibilities $y \in k$. Since Lemma 3.3 also holds, with an analogue proof, when $\mathbb{Z}$ gets replaced by any field, we conclude that $(X - a)$ is a factor of both G and F, which is a contradiction. So S_1 must consist of only finitely many elements. □

We now define the concept of being Zariski-dense.

Definition 7.8. *Let k be an algebraically closed field and $n \in \mathbb{N}$. Let $T \subset k^n$ be subspace of k^n with the induced Zariski topology. Let S be any subset of T. We call S Zariski-dense in T if T is the only Zariski-closed subset of T that includes S.*

Lemma 7.9. *Let k be an algebraically closed field and $n \in \mathbb{N}$. Let $T \subset k^n$ be subspace of k^n with the induced Zariski topology. Let S be any subset of T. Let S' be any subset of S. If S' is Zariski-dense in T, then so is S.*

Proof. Suppose that S is not Zariski-dense in T. Then there must exist a proper subset $R \subsetneq T$ that is Zariski-closed in T and that includes S. Therefore, we have that R includes S' as well. As a consequence we directly see that S' is not Zariski-dense in T. □

Definitions 7.10. *let $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in \mathbb{C}[X_1, \ldots, X_n]$ be a polynomial. We define the sets*

$$T(F) := \{(x_1, \ldots, x_n) \in \mathbb{C}^n \mid F(x_1, \ldots, x_n) = 0\}$$

and

$$S(F) := \{(x_1, \ldots, x_n) \in \mathbb{Z}^n \mid F(x_1, \ldots, x_n) = 0\}.$$

We call $T(F)$ the (total) zero locus of F and $S(F)$ the integral zero locus of F.

Note that the zero locus of any polynomial is exactly the variety of the set that contains only this polynomial. So we have $T(F) = V(\{F\})$ for all $F \in \mathbb{C}[X_1, \ldots, X_n]$. We are often interested whether $S(F)$ is Zariski-dense in $T(F)$ for some $F(X_1, \ldots, X_n) \in \mathbb{C}[X_1, \ldots, X_n]$.

Lemma 7.11. *Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be a nonconstant polynomial, then we have the proper inclusion $S(F) \subsetneq T(F)$.*

Proof. We immediately see from the definition of $S(F)$ and $T(F)$ that $S(F) \subset T(F)$, so it rest us to prove the this inclusion is proper. Because F is nonconstant we may without loss of generality assume that F is of positive Y-degree. Write F as a polynomial in Y, so $F(X, Y) = \sum_{i=0}^{n} a_i(X)Y^i$, where $n \in \mathbb{N}$ is the Y-degree of F and where $a_i(X) \in \mathbb{Z}[X]$ for all $i \in \{0, \ldots, n\}$. In particular we have $a_n(X) \neq 0 \in \mathbb{Z}[X]$, so $a_n(X)$ has only finitely many complex roots. Let $x \in \mathbb{C}$ be a complex number that is not an integer and also not a root of $a_n(X)$. This shows that $F(x, Y) = \sum_{i=0}^{n} a_i(x)Y^i$ is a polynomial in $\mathbb{C}[Y]$ of degree n and therefore has n complex roots counting multiplicity. Let $y \in \mathbb{C}$ be such a root. We then have $F(x, y) = 0$ with $(x, y) \in \mathbb{C}^2$ and $(x, y) \notin \mathbb{Z}^2$. This shows that $(x, y) \in T(F)$ and $(x, y) \notin S(F)$, hence $S(F) \subsetneq T(F)$. □

The following lemma gives the connection between Zariski-density and having infinitely many integral solutions to a binary Diophantine equation.

Lemma 7.12. *Let $F(X, Y) \in \mathbb{Z}[X, Y]$ be an irreducible polynomial. Then $S := S(F)$ is Zariski-dense in $T := T(F)$ if and only if there are infinitely many integral solutions to $F(x, y) = 0$.*

Proof. Suppose that there are only finitely many integral solutions to $F(x, y) = 0$. Then S is a finite subset of T and by Lemma 7.7 also a Zariski-closed subset of T. We can see from Lemma 7.11 that $S \neq T$. This shows that T is not the only Zariski-closed subset of T that includes S. So S is not Zariski-dense in T.
Now suppose that there are infinitely many integral solutions to $F(x, y) = 0$. Then S is an infinite subset of T. Suppose that S is not Zariski-dense in T. Then there must exist a proper Zariski-closed R of T that includes S. By Lemma 7.7 we see that R must be a finite subset, which contradicts the fact that R includes the infinite set S. So S is indeed Zariski-dense in T. □

7.3 Vanishing integral zero locus on a coprime polynomial

In this subsection we are interested in the connection between Zariski-density of a certain set and the existence of a second polynomial that vanishes on this set. The following lemma directly tells something about this.

Lemma 7.13. *Let k be an algebraically closed field and $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in k[X_1, \ldots, X_n]$ and let*

$$T \coloneqq \{(x_1, \ldots, x_n) \in k^n \mid F(x_1, \ldots, x_n) = 0\} = V(\{F\}) \subset k^n$$

be the variety of $\{F\}$. Let $S \subsetneq T$ be a proper subset of T. Then S is Zariski-dense in T if and only if there does not exist $G(X_1, \ldots, X_n) \in k[X_1, \ldots, X_n]$ such that $G(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in S$ and such that there is some $(x_1, \ldots, x_n) \in T$ with $G(x_1, \ldots, x_n) \neq 0$.

Proof. Suppose that such G does exist. Let $R = T \cap V(\{G\}) \subset k^n$. Then R is Zariski-closed in T as it is the intersection of a Zariski-closed subset with T. We have $S \subset V(\{G\})$ and thus $S \subset R$. Because there is some $(x_1, \ldots, x_n) \in T$ satisfying $G(x_1, \ldots, x_n) \neq 0$, we have $R \subsetneq T$. So there exists a Zariski-closed proper subset of T that includes S. This shows that S is not Zariski-dense in T.
Conversely, suppose that S is not Zariski-dense in T, then we have $S \subset R \subsetneq T$ for some proper Zariski-closed subset R of T. Since R is Zariski-closed, there exists $\mathcal{F} \subset k[X_1, \ldots, X_n]$ such that $R = V(\mathcal{F})$. Because R is a proper subset of T, there must exist $(x_1, \ldots, x_n) \in T$ such that $G(x_1, \ldots, x_n) \neq 0$ for some $G \in \mathcal{F}$. From $S \subset R = V(\mathcal{F})$, we conclude that $G(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in S$. □

Suppose that a polynomial $F \in \mathbb{C}[X_1, \ldots, X_n]$ is of zero X_n-degree. If we want to know whether the integral zero locus of F is Zariski-dense in the total zero locus, we might as well view F as a polynomial in $\mathbb{C}[X_1, \ldots, X_{n-1}]$. The following lemma shows this.

Lemma 7.14. *Let $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in \mathbb{C}[X_1, \ldots, X_n]$. Let $T \coloneqq T(F)$ and $S \coloneqq S(F)$. Suppose that $\deg_{X_n} F = 0$. Let $F'(X_1, \ldots, X_{n-1}) = F(X_1, \ldots, X_{n-1}, 0) \in \mathbb{C}[X_1, \ldots, X_{n-1}]$. Let*

$$T' \coloneqq T(F') = \{(x_1, \ldots, x_{n-1}) \in \mathbb{C}^{n-1} \mid F'(x_1, \ldots, x_{n-1}) = 0\}$$

and

$$S' = S(F') = \{(x_1, \ldots, x_{n-1}) \in \mathbb{Z}^{n-1} \mid F'(x_1, \ldots, x_{n-1}) = 0\}.$$

Then S is Zariski-dense in T if and only if S' is Zariski-dense in T'.

Proof. Suppose that S is not Zariski-dense in T. By Lemma 7.13 there then must exist $G(X_1, \ldots, X_n) \in \mathbb{C}[X_1, \ldots, X_n]$ such that $G(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in S$ and such that there exists $(a_1, \ldots, a_n) \in T$ with $G(a_1, \ldots, a_n) \neq 0$. Note that $\deg_{X_n} F = 0$ implies that $F(x_1, \ldots, x_{n-1}, 0) = F(x_1, \ldots, x_{n-1}, x_n)$ hold for all $x_1, \ldots, x_n \in \mathbb{C}$. Now let

$$G'(X_1, \ldots, X_{n-1}) \coloneqq G(X_1, \ldots, X_{n-1}, a_n) \in \mathbb{C}[X_1, \ldots, X_{n-1}].$$

Furthermore, let $(x_1, \ldots, x_{n-1}) \in S'$. We then have

$$F(x_1, \ldots, x_{n-1}, a_n) = F(x_1, \ldots, x_{n-1}, 0) = F'(x_1, \ldots, x_{n-1}) = 0,$$

hence $(x_1, \ldots, x_{n-1}, a_n) \in S$. Therefore we have

$$G'(x_1, \ldots, x_{n-1}) = G(x_1, \ldots, x_{n-1}, a_n) = 0,$$

so G' vanishes at all elements in S'. We also have $G'(a_1, \ldots, a_{n-1}) = G(a_1, \ldots, a_n) \neq 0$, while

$$F'(a_1, \ldots, a_{n-1}) = F(a_1, \ldots, a_{n-1}, 0) = F(a_1, \ldots, a_{n-1}, a_n) = 0.$$

So G' does not vanish at the element $(a_1, \ldots, a_{n-1}) \in T'$. By Lemma 7.13 we see that S' is not Zariski-dense in T'.
Conversely suppose that S' is not Zariski-dense in T'. We use Lemma 7.13 and find $G'(X_1, \ldots, X_{n-1}) \in \mathbb{C}[X_1, \ldots, X_{n-1}]$ such that G' vanishes on S' but not on T', hence there exists $(a_1, \ldots, a_{n-1}) \in T'$ with $G'(a_1, \ldots, a_{n-1}) \neq 0$. Let $G(X_1, \ldots, X_n) = G'(X_1, \ldots, X_{n-1}) \in \mathbb{C}[X_1, \ldots, X_n]$. For any $(x_1, \ldots, x_n) \in S$ we have

$$F'(x_1, \ldots, x_{n-1}) = F(x_1, \ldots, x_{n-1}, 0) = F(x_1, \ldots, x_{n-1}, x_n) = 0,$$

so $(x_1, \ldots, x_{n-1}) \in S'$. This gives us $G(x_1, \ldots, x_n) = G'(x_1, \ldots, x_{n-1}) = 0$, which shows that G vanishes on S. We have $F(a_1, \ldots, a_{n-1}, 0) = F'(a_1, \ldots, a_{n-1}) = 0$ and $G(a_1, \ldots, a_{n-1}, 0) = G'(a_1, \ldots, a_{n-1}) \neq 0$. So G does not vanish at the element $(a_1, \ldots, a_{n-1}, 0) \in T$. We once again use Lemma 7.13 and see that S is not Zariski-dense in T. □

The following theorem was discovered by David Hilbert and is called Hilbert's Nullstellensatz. We will use this theorem, but will not give its proof. A proof of this has been given by David Eisenbud [4].

Theorem 7.15 (Hilbert's Nullstellensatz)**.** *Let k be an algebraically closed field and $n \in \mathbb{N}$. Let $\mathcal{F} \subset k[X_1, \ldots, X_n]$ and $G(X_1, \ldots, X_n) \in k[X_1, \ldots, X_n]$ be such that $G(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in V(\mathcal{F})$. Then the ideal in $k[X_1, \ldots, X_n]$ that is generated by elements in $\mathcal{F}$ contains some power of G.*

Lemma 7.16. *Let $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ be a nonconstant polynomial and $G(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ a nonzero polynomial such that they do not share a common nonunit as factor in $\mathbb{Z}[X_1, \ldots, X_n]$. Then there exists $(x_1, \ldots, x_n) \in \mathbb{C}^n$ such that $F(x_1, \ldots, x_n) = 0$ and $G(x_1, \ldots, x_n) \neq 0$.*

Proof. Suppose that such $(x_1, \ldots, x_n)$ does not exist. So $G(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in V(\{F\}) \subset \mathbb{C}^n$. From Hilbert's Nullstellensatz 7.15 it then follows that there exists $m \in \mathbb{N}$ such that G^m lies in the ideal in $\mathbb{C}[X_1, \ldots, X_n]$ that is generated by F. So F divides G^m in $\mathbb{C}[X_1, \ldots, X_n]$. Let $f \in \mathbb{C}[X_1, \ldots, X_n]$ be an irreducible divisor of F. Then f is also a divisor of G^m. Since f is nonconstant, there exists $i \in \{1, \ldots, n\}$ with $d_i := \deg_{X_i} f > 0$. We may assume without loss of generality that $d_n > 0$. So when we view f as a polynomial in X_n, it is nonconstant. Therefore there exists a field extension k over the fractionfield of $\mathbb{C}[X_1, \ldots, X_{n-1}]$ that contains a root of f. Since f divides F and G in $\mathbb{C}[X_1, \ldots, X_n]$, we see that this root is also a root of F and G. Therefore F and G can't be coprime in $\mathbb{Z}[X_1, \ldots, X_n]$. This contradicts the fact that F and G share no nonunits as factors in $\mathbb{Z}[X_1, \ldots, X_n]$, so there indeed must exist $(x_1, \ldots, x_n) \in \mathbb{C}^n$ such that $F(x_1, \ldots, x_n) = 0$ and $G(x_1, \ldots, x_n) \neq 0$. □

As a corollary, we have found another connection between Zariski-density of a certain set and the existence of a second polynomial that vanishes on this set.

Corollary 7.17. *Let $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ be a nonconstant irreducible polynomial. Let $T = T(F)$. Let S be a subset of T. Suppose that we find nonzero $G(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ such that G is no multiple of F in $\mathbb{Z}[X_1, \ldots, X_n]$ and such that $G(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in S$. Then S is not Zariski-dense in T.*

Proof. Since F is irreducible and not a divisor of G in $\mathbb{Z}[X_1, \ldots, X_n]$, they do not share a common nonconstant factor in $\mathbb{Z}[X_1, \ldots, X_n]$. By Lemma 7.16 we find $(x_1, \ldots, x_n) \in \mathbb{C}^n$ such that $F(x_1, \ldots, x_n) = 0$ and $G(x_1, \ldots, x_n) \neq 0$. We now apply Lemma 7.13 and conclude that S is Zariski-dense in T. □

The following lemma shows that the integral zero locus of a nonzero polynomial in $\mathbb{Z}[X_1, \ldots, X_n]$ is Zariski-dense in the total zero locus of this polynomial if and only if it holds for all nonconstant irreducible factors of F. This reduces the question of Zariski-density to irreducible polynomials.

Lemma 7.18. *Let $n \in \mathbb{N}_0$. Let $F(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ be nonzero. Let*

$$F = c \cdot F_1^{a_1} \cdots F_m^{a_m}$$

be a factorization of F in $\mathbb{Z}[X_1, \ldots, X_n]$, where $c \in \mathbb{Z}$ is a nonconstant integer, and where $F_1, \ldots, F_m$ are pairwise coprime nonconstant irreducible factors in $\mathbb{Z}[X_1, \ldots, X_n]$. So we have $a_i \in \mathbb{N}$ for all $i \in \{1, \ldots, m\}$. It then holds that $S(F)$ is Zariski-dense in $T(F)$ if and only if $S(F_i)$ is Zariski-dense in $T(F_i)$ for all $i \in \{1, \ldots, m\}$.

Proof. We derive the following from the definitions of the zero locus and the integral zero locus:

$$S(F) = \bigcup_{i=1}^{m} S(F_i)$$

and

$$T(F) = \bigcup_{i=1}^{m} T(F_i).$$

Now first suppose that $S(F)$ is not Zariski-dense in $T(F)$. By Lemma 7.13 there then must exists $G(X_1, \ldots, X_n) \in \mathbb{C}[X_1, \ldots, X_n]$ such that G vanishes on $S(F)$, but not on $T(F)$. So there exists $i \in \{1, \ldots, m\}$ such that G does not vanish on $T(F_i)$. Since G does vanish on $S(F_i) \subset S(F)$, we can apply Lemma 7.13 again to find that $S(F_i)$ is not Zariski-dense in $T(F_i)$. Conversely, suppose that there exists $i \in \{1, \ldots, m\}$ such that $S(F_i)$ is not Zariski-dense in $T(F_i)$. We apply Lemma 7.13 to find the existence of $G(X_1,\dots,X_n)\in\mathbb{C}[X_1,\dots,X_n]$ such that G vanishes on $S(F_i)$, but not on $T(F_i)$. In particular this means that G is no multiple of the irreducible factor F_i and G and F_i are therefore coprime. Now define

$$G' \coloneqq G\cdot\prod_{\substack{j=1\\j\neq i}}^{m}F_j.$$

We see that G' and F_i are coprime in $\mathbb{Z}[X_1,\dots,X_n]$. From Lemma 7.16 it then follows that there exists $(x_1,\dots,x_n)\in\mathbb{C}^n$ such that $F_i(x_1,\dots,x_n)=0$ and $G'(x_1,\dots,x_n)\neq 0$. This shows that G' does not vanish on $T(F_i)$ and therefore especially does not vanish on $T(F)$. Because F_j vanishes on $S(F_j)$ for each $j\in\{1,\dots,m\}$ and since G vanishes on $S(F_i)$, we find that G' vanishes on their union $\bigcup_{i=1}^{m}S(F_i)=S(F)$. We can apply Lemma 7.13 again and conclude that $S(F)$ is not Zariski-dense in $T(F)$. □

7.4 Partitioning integral zero locus

In order to determine whether the integral zero locus of an irreducible polynomial in $\mathbb{Z}[X_1,\dots,X_n]$ is Zariski-dense in its total zero locus, we may divide these integer solutions over finitely many subsets and determine whether these subsets are Zariski-dense in the total zero locus. We will show this in this subsection. For this we first need to prove some lemmas.

Lemma 7.19. *Let $n\in\mathbb{N}_0$. Let $F\in\mathbb{Z}[X_1,\dots,X_n]$ be nonzero. Let $G,H\in\mathbb{C}[X_1,\dots,X_n]$ such that $F=GH$. Then there exists nonzero $\gamma\in\mathbb{C}$ and an algebraic number field $k\subset\overline{\mathbb{Q}}$ such that $\gamma G\in k[X_1,\dots,X_n]$ and such that $\gamma^{-1}H\in k[X_1,\dots,X_n]$.*

Proof. We prove this by induction on n. If $n=0$, we have $F\in\mathbb{Z}[X_1,\dots,X_0]=\mathbb{Z}$ and $G,H\in\mathbb{C}[X_1,\dots,X_0]=\mathbb{C}$ with $F=GH$. Note that G and H also must be nonzero. We take $\gamma=H$ and $k=\mathbb{Q}$ and the results will then follow directly. We now assume $n>0$ and that this lemma holds for $n-1$. Since $\mathbb{C}$ is an algebraically closed field of characteristic zero, we have by Newton-Puiseux's Theorem 4.17 that $\mathbb{C}((X_1^*))$ is also algebraically closed. Since this is also of characteristic zero, we can use this theorem multiple times to find that $\mathbb{C}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))$ is algebraically closed. The same holds if we replace $\mathbb{C}$ with the algebraically closed subfield $\overline{Q}$. Let $d_n\in\mathbb{N}_0$ be the X_n-degree of F. We can then write

$$F=f\prod_{i=1}^{d_n}(X_n-f_i),$$

where $f\in\mathbb{Z}[X_1,\dots,X_{n-1}]$ is the leading coefficient of F, when viewed as a polynomial in X_n, and where $f_1,\dots,f_{d_n}\in\overline{Q}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))$. For every $i\in\{1,\dots,d_n\}$ we have that (X_n-f_i), being a linear polynomial, is irreducible in $\mathbb{C}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n]$. We also have that f is a unit in $\mathbb{C}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n]$. Since $\mathbb{C}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n]$ is a unique factorization domain that includes $\mathbb{C}[X_1,\dots,X_n]$, we may assume that

$$G=g\prod_{i=1}^{t}(X_n-f_i),\quad\text{and that}\quad H=h\prod_{i=t+1}^{d_n}(X_n-f_i),$$

for some $t\in\{0,\dots,d_n\}$ and $g,h\in\mathbb{C}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))$ such that $f=gh$. Because g and h are the leading coefficients of G and H respectively, when viewed as polynomials in X_n, we must have $g,h\in\mathbb{C}[X_1,\dots,X_{n-1}]$. We apply the induction hypothesis on $f=gh$ and find $\gamma\in\mathbb{C}$ and an algebraic number field $k'\subset\overline{\mathbb{Q}}$ such that $\gamma g\in k'[X_1,\dots,X_{n-1}]$ and such that $\gamma^{-1}h\in k'[X_1,\dots,X_{n-1}]$. In particular we have $\gamma g\in\overline{\mathbb{Q}}[X_1,\dots,X_{n-1}]$. We also have $\prod_{i=1}^{t}(X_n-f_i)\in\overline{\mathbb{Q}}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n]$. So from this we find

$$\gamma G\in\overline{\mathbb{Q}}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n]\subset\mathbb{C}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n].$$

We also have

$$\gamma G\in\mathbb{C}[X_1,\dots,X_n]\subset\mathbb{C}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n].$$

This combined shows us that

$$\gamma G\in\overline{\mathbb{Q}}((X_1^*))((X_2^*))\cdots((X_{n-1}^*))[X_n]\cap\mathbb{C}[X_1,\dots,X_n]=\overline{\mathbb{Q}}[X_1,\dots,X_n].$$

So γG is a polynomial in n variables with coefficients in $\overline{\mathbb{Q}}$. The same follows for $\gamma^{-1}H$. Since G and H both have only finitely many of these coefficients in $\overline{\mathbb{Q}}$, there exists an algebraic number field $k \subset \overline{\mathbb{Q}}$ that contains all these coefficients. So $\gamma G \in k[X_1, \ldots, X_n]$ and $\gamma^{-1}H \in k[X_1, \ldots, X_n]$.

□

Corollary 7.20. *Let $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ be a nonconstant polynomial. If F is reducible in $\mathbb{C}[X_1, \ldots, X_n]$, it is also reducible in $k[X_1, \ldots, X_n]$ for some algebraic number field $k \subset \overline{\mathbb{Q}}$.*

Proof. Suppose that F is reducible in $\mathbb{C}[X_1, \ldots, X_n]$. Then we have some factorization $F = GH$, where $G, H \in \mathbb{C}[X_1, \ldots, X_n]$ are both nonunits, hence nonconstant. We apply Lemma 7.19 and find nonzero $\gamma \in \mathbb{C}$ and an algebraic number field $k \subset \overline{\mathbb{Q}}$ such that $G' := \gamma G \in k[X_1, \ldots, X_n]$ and such that $H' := \gamma^{-1}H \in k[X_1, \ldots, X_n]$. We have $F = G'H'$ and both G', H' are nonconstant, and thus nonunits, in $k[X_1, \ldots, X_n]$. This shows that F is also reducible in $k[X_1, \ldots, X_n]$. □

Lemma 7.21. *Let $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ be an irreducible polynomial. Suppose that F is reducible in $\mathbb{C}[X_1, \ldots, X_n]$. Then $S(F)$ is not Zariski-dense in $T(F)$.*

Proof. By Corollary 7.20 we see that F is reducible in $k[X_1, \ldots, X_n]$ for some algebraic number field k. Let $F(X_1, \ldots, X_n) = G(X_1, \ldots, X_n)H(X_1, \ldots, X_n)$ be a proper factorization in $k[X_1, \ldots, X_n]$. Then we have for all $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ that $F(x_1, \ldots, x_n) = 0$ if and only if $G(x_1, \ldots, x_n) = 0$ or $H(x_1, \ldots, x_n) = 0$. So $S(F) = S(G) \cup S(H)$. Let α be a primitive element of k. So $k = \mathbb{Q}(\alpha)$. Let $s = [k : \mathbb{Q}]$ be the degree of k over $\mathbb{Q}$. We then can write G as

$$G(X_1, \ldots, X_n) = \sum_{i_1=0}^{d_1} \cdots \sum_{i_n=0}^{d_n} \sum_{j=0}^{s-1} a_{i_1,\ldots,i_n,j} \alpha^j X_1^{i_1} \cdots X_n^{i_n}$$

with $a_{i_1,\ldots,i_n,j} \in \mathbb{Q}$ and $d_1, \ldots, d_n \in \mathbb{N}_0$. We can then write

$$G(X_1, \ldots, X_n) = G_0(X_1, \ldots, X_n) + G_1(X_1, \ldots, X_n)\alpha + \ldots + G_{s-1}(X_1, \ldots, X_n)\alpha^{s-1},$$

with

$$G_j(x_1, \ldots, x_n) = \sum_{i_1=0}^{d_1} \cdots \sum_{i_n=0}^{d_n} a_{i_1,\ldots,i_n,j} X_1^{i_1} \cdots X_n^{i_n} \in \mathbb{Q}[X_1, \ldots, X_n]$$

for each $j \in \{0, \ldots, s-1\}$. So the equation $G(x_1, \ldots, x_n) = 0$ with $x_1, \ldots, x_n \in \mathbb{Z}$ is equivalent to the combined equations $G_0(x_1, \ldots, x_n) = 0, \ldots, G_{s-1}(x_1, \ldots, x_n) = 0$. If F divides all G_j in $k[X_1, \ldots, X_n]$, then F divides G, which contradicts the fact that $F = GH$ was a proper factorization. So F does not divide G_j in $k[X_1, \ldots, X_n]$ for some $j \in \{0, \ldots, s-1\}$. There exists nonzero $a \in \mathbb{Z}$ such that $G'(X_1, \ldots, X_n) := a \cdot G_j \in \mathbb{Z}[X_1, \ldots, X_n]$. Since a is a unit in k, the nonconstant F does not divide G' in $k[X_1, \ldots, X_n]$ and therefore also not in $\mathbb{Z}[X_1, \ldots, X_n]$. Note that $G(x_1, \ldots, x_n) = 0$ implies $G'(x_1, \ldots, x_n) = 0$ for any $x_1, \ldots, x_n \in \mathbb{Z}$. In a similar way we find $H'(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$, that is no multiple of F in $\mathbb{Z}[X_1, \ldots, X_n]$ and satisfies $H'(x_1, \ldots, x_n) = 0$ for all $x_1, \ldots, x_n \in \mathbb{Z}$ that satisfies $H(x_1, \ldots, x_n) = 0$. Since F is irreducible in $\mathbb{Z}[X_1, \ldots, X_n]$, we have that $G'H'$ is no multiple of F and vanishes on $S(F) = S(G) \cup S(H)$. We can thus apply Corollary 7.17 to conclude that $S(F)$ is not Zariski-dense in $T(F)$. □

We can now show that we may split the problem of determining Zariski-density into finitely many subsets. The exact statement is as follows:

Lemma 7.22. *Let $n \in \mathbb{N}$. Let $F(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$ be an irreducible polynomial. Let $m \in \mathbb{N}$. Let $S_1, \ldots, S_m$ be subsets of $S(F)$. Then $S' := \bigcup_{i=1}^{m} S_i$ is Zariski-dense in $T(F)$ if and only if S_i is Zariski-dense in $T(F)$ for some $i \in \{1, \ldots, m\}$.*

Proof. First suppose that S' is not Zariski-dense in $T(F)$. Then we have by Lemma 7.9 that all of its subsets $S_1, \ldots, S_m$ are also not Zariski-dense in $T(F)$.
Conversely, suppose that $S_1, \ldots, S_m$ are all not Zariski-dense in $T(F)$. We assume that S' is Zariski-dense in $T(F)$ and will reach a contradiction. Since $S' \subset S(F)$ is Zariski-dense in $T(F)$, we have by Lemma 7.9 that $S(F)$ is Zariski-dense in $T(F)$ as well. Lemma 7.21 then tells us that F must be irreducible in $\mathbb{C}[X_1, \ldots, X_n]$. Let $i \in \{1, \ldots, m\}$. Since S_i is not Zariski-dense in $T(F)$, we can apply Lemma 7.13 to find $G_i(X_1, \ldots, X_n) \in \mathbb{C}[X_1, \ldots, X_n]$ such that $G_i(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in S_i$ but not for

all $(x_1, \ldots, x_n) \in T(F)$. In particular this shows that G_i is not a multiple of F in $\mathbb{C}[X_1, \ldots, X_n]$. Since F is irreducible in $\mathbb{C}[X_1, \ldots, X_n]$, F and G_i share no common factor in $\mathbb{C}[X_1, \ldots, X_n]$. This shows that the polynomial $G \coloneqq \prod_{i=1}^{s} G_i \in \mathbb{C}[X_1, \ldots, X_n]$ satisfies $G(x_1, \ldots, x_n) = 0$ for all $(x_1, \ldots, x_n) \in \bigcup_{i=1}^{m} S_i = S'$ and that G and F do not share a common factor in $\mathbb{C}[X_1, \ldots, X_n]$. Since F is irreducible in $\mathbb{C}[X_1, \ldots, X_n]$, this means that no power of G lies in the ideal generated by F so by Hilbert's Nullstellensatz 7.15 we see that there must exist $(x_1, \ldots, x_n) \in V(\{F\}) = T(F)$ with $G(x_1, \ldots, x_n) \neq 0$. We can now apply Lemma 7.13 to conclude that S' is not Zariski-dense in $T(F)$; a contradiction. □

8 Generalizing Runge's Theorem to Three Variables

In this section we are interested in whether the integral zero locus of a polynomial in $\mathbb{Z}[X, Y, Z]$ is Zariski-dense in the total zero locus. In the first subsection we will provide some examples of polynomials which we already can solve quite easily. In the second and third subsection we will study the roots of these polynomials. This will help us in our attempt in the fourth subsection to generalize Runge's Theorem. In this attempt, we see that we do need to be familiar with the roots of this polynomial, when viewed as a polynomial in Z over $\mathbb{Z}[X, Y]$ and that these roots are of a certain form.

8.1 First look

We start by showing a family of polynomials whose integral zero locus is Zariski-dense in its total zero locus.

Example 8.1. *Let $H(X, Y) \in \mathbb{Z}[X, Y]$. Let $F(X, Y, Z) \coloneqq Z - H(X, Y)$. Then $S(F)$ is Zariski-dense in $T(F)$.*

Proof. Suppose that this is not the case. By Lemma (7.13) we then find $G(X, Y, Z) \in \mathbb{C}[X, Y, Z]$ such that G vanishes on $S(F)$, but not on $T(F)$. So there exists $x_0, y_0, z_0 \in \mathbb{Z}$ with $F(x_0, y_0, z_0) = 0$ and $G(x_0, y_0, z_0) \neq 0$. Now consider the polynomial

$$K(X, Y) \coloneqq G(X, Y, H(X, Y)) \in \mathbb{C}[X, Y].$$

Let $a \in \mathbb{Z}$ and $b \in \mathbb{Z}$ be any integers. We have $F(a, b, H(a, b)) = 0$ and $H(a, b) \in \mathbb{Z}$, which shows us that $(a, b, H(a, b)) \in S(F)$. We therefore have $K(a, b) = G(a, b, H(a, b)) = 0$. Since Lemma 3.3 also holds, with an analogue proof, when $\mathbb{Z}$ gets replaced by any field, we see that $X - a$ is a factor of $K(X, Y)$ in $\mathbb{C}[X, Y]$. Since this holds for all $a \in \mathbb{Z}$ and since all these factors are unique up to multiplication with units in $\mathbb{C}[X, Y]$, we can use the fact that $\mathbb{C}[X, Y]$ is a unique factorization domain to conclude that $K(X, Y)$ must be the zero polynomial. We have $0 = F(x_0, y_0, z_0) = z_0 - H(x_0, y_0)$ and therefore

$$0 = K(x_0, y_0) = G(x_0, y_0, H(x_0, y_0)) = G(x_0, y_0, z_0),$$

a contradiction. So $S(F)$ must indeed be Zariski-dense in $T(F)$. □

In the previous section we saw some lemmas that are useful to determine whether $S(F)$ is Zariski-dense in $T(F)$ for some $F \in \mathbb{Z}[X, Y, Z]$. We see for example by Lemma 7.18 that we may assume F to be irreducible. If F is of zero Z-degree, we see by Lemma 7.14 that we have reduced the problem to the problem of Zariski-density in two dimensions. Because of this we may from now on assume that F is of positive Z-degree. In a similar way we may assume that F is also of positive X-degree and of positive Y-degree. There is even a stronger way to reduce the problem to the two-dimensional case. The next lemma shows this.

Lemma 8.2. *Let $F(X, Y, Z) \in \mathbb{Z}[X, Y, Z]$ be a nonconstant irreducible polynomial. Suppose that we can write*

$$F(X, Y, Z) = H(F_1(X, Y, Z), F_2(X, Y, Z)),$$

where $H(U, V) \in \mathbb{Z}[U, V]$ is a polynomial in two variables and where $F_1, F_2 \in \mathbb{Z}[X, Y, Z]$ are polynomials such that either F_1 or F_2 is not a constant in $\mathbb{Z}$ plus a multiple of F in $\mathbb{Z}[X, Y, Z]$. Suppose that the equation $H(u, v) = 0$ holds for only finitely many $u, v \in \mathbb{Z}$. Then $S(F)$ is not Zariski-dense in $T(F)$.

Proof. Without loss of generality we may assume that F_1 is not a constant in $\mathbb{Z}$ plus a multiple of F in $\mathbb{Z}[X,Y,Z]$. We will look at the set

$$M = \{u \in \mathbb{Z} \mid \exists v \in \mathbb{Z}, H(u,v) = 0\}.$$

Note that M is a finite set because the equation $H(u,v) = 0$ has only finitely many integral solutions. So we can define the polynomial

$$G(X,Y,Z) = \prod_{u \in M} (F_1(X,Y,Z) - u) \in \mathbb{Z}[X,Y,Z].$$

Let $(x,y,z) \in S(F)$. Then we find

$$0 = F(x,y,z) = H(F_1(x,y,z), F_2(x,y,z)).$$

So we have $F_1(x,y,z) = u$ for some $u \in M$. This shows that $G(x,y,z) = 0$, so G vanishes on $S(F)$. Suppose that G is a multiple of F in $\mathbb{Z}[X,Y,Z]$. Because F is irreducible, we see that $F_1 - u$ must then be a multiple of F for some $u \in M$. This shows that F_1 is the constant u plus a multiple of F, a contradiction. As a consequence, we see that G is not a multiple of F. We then may conclude by Corollary 7.17 that $S(F)$ is not Zariski-dense in $T(F)$. □

We can use this lemma for the following two examples

Example 8.3. *Consider the polynomial of the form*

$$H(X,Y) \coloneqq Y^n - ((aX)^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0),$$

where $n \geq 2$ and $a, a_i \in \mathbb{Z}$, with $a \neq 0$. Suppose that the polynomial $(aX)^n + a_{n-1}X^{n-1} + \ldots + a_1X + a_0$ is not a d-th power in $\mathbb{Z}[X]$ for any divisor $d \geq 2$ of n. Let

$$F(X,Y,Z) \coloneqq H(X+Z, Y-Z).$$

Then $S(F)$ is not Zariski-dense in $T(F)$.

Proof. Example 6.3 tells us that the equation $H(x,y) = 0$ has only finitely many integral solutions. The proof of this example also tells us that $H(X,Y)$ is irreducible in $\mathbb{Z}[X,Y]$. As a consequence, we have that $F(X,Y) = H(X+Z, Y-Z)$ is irreducible in $\mathbb{Z}[X,Y,Z]$. Because F is of X-degree two and $X+Z$ is of X-degree one, we see that $X+Z$ certainly is not a constant plus a multiple of F. The result then automatically follows from Lemma 8.2. □

Example 8.4. *Let $F \in \mathbb{Z}[X,Y,Z]$ be the polynomial given by*

$$F(X,Y,Z) \coloneqq 3XZ - 6YZ + X^2 - 2Y^2 - XY + 2.$$

Then $S(F)$ is not Zariski-dense in $T(F)$.

Proof. When we view F as a polynomial in Z over $\mathbb{Z}[X,Y]$, we see that F is a linear polynomial. Its two coefficients $3(X-2Y)$ and $X^2 - 2Y^2 - XY + 2$ are coprime in $\mathbb{Z}[X,Y]$ and we therefore see that F is irreducible in $\mathbb{Z}[X,Y,Z]$. We define

$$H(X,Y) \coloneqq X^2 - 2Y^2 - XY + 2.$$

The equation $H(x,y) = 0$ has only finitely many integral solutions as has been showed in Example 6.5. We find

$$\begin{aligned} H(X+2Z, Y+Z) &= (X+2Z)^2 - 2(Y+Z)^2 - (X+2Z)(Y+Z) + 2 \\ &= X^2 + 4XZ + 4Z^2 - 2Y^2 - 4YZ - 2Z^2 - XY - XZ - 2YZ - 2Z^2 + 2 \\ &= X^2 + 3XZ - 2Y^2 - 6YZ - XY + 2 \\ &= F(X,Y,Z). \end{aligned}$$

Because $X + 2Z$ is not a constant plus a multiple of F, we can use Lemma 8.2 to conclude that $S(F)$ is not Zariski-dense in $T(F)$. □

The following example is about Pythagorean triples. These are integers $a, b, c \in \mathbb{N}$ such that $a^2+b^2=c^2$.

Example 8.5. *Let $F(X,Y,Z) = Z^2 - X^2 - Y^2$. Then $S(F)$ is Zariski-dense in $T(F)$.*

Proof. Suppose that $S(F)$ is not Zariski-dense in $T(F)$. By Lemma 7.13 there then exists $G(X,Y,Z) \in \mathbb{C}[X,Y,Z]$ such that G vanishes on $S(F)$, but not on $T(F)$. We therefore see that G is no multiple of F in $\mathbb{C}[X,Y,Z]$. Because X^2+Y^2 is not a square in $\mathbb{C}[X,Y]$, we see that $F(X,Y,Z)$ is irreducible in $\mathbb{C}[X,Y,Z]$. This shows that G and F share no nonconstant factors in $\mathbb{C}[X,Y,Z]$. We then can apply Lemma 7.6 to find the existence of nonzero $H(X,Y) \in \mathbb{C}[X,Y]$ such that $H(x,y) = 0$ for all $x,y,z \in S(F)$. Because $(3,4,5) \in S(F)$, we see that $H(3,4) = 0$. So H cannot be a constant polynomial. We may without loss of generality assume that $d_1 \coloneqq \deg_X H$ is positive. Let $p \in \mathbb{N}$ and let $t \in \{0,\ldots,d_1\}$. Let $m = 2^{p+2d_1+1-t}$ and $n = 2^t$. So m and n are integers. We see that

$$F(m^2-n^2, 2mn, m^2+n^2) = (m^2+n^2)^2 - (m^2-n^2)^2 - (2mn)^2 = 0.$$

So $(m^2-n^2, 2mn, m^2+n^2) \in S(F)$ and therefore we have $H(m^2-n^2, 2mn) = 0$. We find

$$\begin{aligned} H(m^2-n^2, 2mn) &= H(2^{2(p+2d_1+1-t)} - 2^{2t}, 2^{p+2d_1+1-t+t+1}) \\ &= H(2^{2t}(2^{2p+4d_1+2-4t}-1), 2^{p+2d_1+2}). \end{aligned}$$

Let $b = 2^{p+2d_1+2}$. We then see for any $t \in \{0,\ldots,d_1\}$ that $2^{2t}(2^{2p+4d_1+2-4t}-1)$ is a root of the polynomial $H(X,b) \in \mathbb{C}[X]$. Since $H(X,Y)$ is of X-degree d_1, we have that $H(X,b)$ is of X-degree at most d_1. We see that $2^{2t}(2^{2p+4d_1+2-4t}-1)$ is different for each $t \in \{0,\ldots,d_1\}$ because they all contain different powers of 2 in their prime decomposition. So $H(X,b)$ contains at least d_1+1 different roots. From this we conclude that $H(X,b) = 0$. Since Lemma 3.3 also holds, with an analogue proof, when $\mathbb{Z}$ gets replaced by any field, we see that $Y - b$ is a factor of H. Because $b = 2^{p+2d_1+2}$ is different for each $p \in \mathbb{N}$, we see that H has infinitely many different factors in $\mathbb{C}[X,Y]$. From this we conclude that $H(X,Y) = 0$, which leads to a contradiction. So $S(F)$ must indeed be Zariski-dense in $T(F)$. □

We end this subsection by an example that makes uses of Pell's equation 2.4:

Example 8.6. *Let $F(X,Y,Z) = Z^2 - (X^2-1)(Y^2-1)$. Then $S(F)$ is Zariski-dense in $T(F)$.*

Proof. The start of the proof is similar as in the previous example. Suppose that $S(F)$ is not Zariski-dense in $T(F)$. By Lemma 7.13 there then exists $G(X,Y,Z) \in \mathbb{C}[X,Y,Z]$ such that G vanishes on $S(F)$, but not on $T(F)$. We therefore see that G is no multiple of F in $\mathbb{C}[X,Y,Z]$. Because $(X^2-1)(Y^2-1)$ is not a square in $\mathbb{C}[X,Y]$, we see that $F(X,Y,Z)$ is irreducible in $\mathbb{C}[X,Y,Z]$. This shows that G and F share no nonconstant factors in $\mathbb{C}[X,Y,Z]$. We then can apply Lemma 7.6 to find the existence of nonzero $H(X,Y) \in \mathbb{C}[X,Y]$ such that $H(x,y) = 0$ for all $x,y,z \in S(F)$. From Example 2.4 we see that there are infinitely many different integers $x_n \in \mathbb{Z}$ such that there exists $y_n \in \mathbb{Z}$ such that $x_n^2 - 1 = 3y_n^2$. For each $n \in \mathbb{N}_0$ we pick such a x_n and a corresponding y_n. Now let $i \in \mathbb{N}_0$ and let $j \in \mathbb{N}_0$. We then have

$$F(x_i, x_j, 3y_iy_j) = (3y_iy_j)^2 - (x_i^2-1)(x_j^2-1) = 9y_i^2y_j^2 - (3y_i)(3y_j) = 0.$$

So $(x_i, x_j, 3y_iy_j) \in S(F)$ and therefore we have $H(x_i, x_j) = 0$. We then can then use Lemma 3.3 to conclude for each $i \in \mathbb{N}_0$ that $X - x_i$ is a factor of H. Because x_i is different for each $i \in \mathbb{N}_0$, we see that H has infinitely many different factors in $\mathbb{Z}[X,Y]$. From this we conclude that $H(X,Y) = 0$, which leads to a contradiction. So $S(F)$ must indeed be Zariski-dense in $T(F)$. □

8.2 The form of its roots

We will inspire ourselves by Runge's Theorem and attempt to generalize that proof to the case of three variables. For this we will need to look at the roots of polynomials in Z over $\mathbb{Z}[X,Y]$. Here we are interested in the form of such roots. We will use Newton-Puiseux's Theorem 4.17 twice. First we note that since $\overline{\mathbb{Q}}$ is an algebraically closed field of characteristic zero, the formal Puiseux field $\overline{\mathbb{Q}}((X^*))$ must be algebraically closed as well. Since it is a field extension over $\overline{\mathbb{Q}}$, it is also of characteristic zero. So by using Newton-Puiseux's Theorem again, we see that the field $\overline{\mathbb{Q}}((X^*))((Y^*))$ is also algebraically closed. Note that the fields $\overline{\mathbb{Q}}(((X^{-1})^*))(Y^*))$, $\overline{\mathbb{Q}}((X^*))(((Y^{-1})^*))$ and $\overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ are also all algebraically closed by similar reasons. Note that $\mathbb{Z}[X,Y]$ can be embedded in all these four fields.

So any polynomial in $\mathbb{Z}[X,Y,Z]$ of positive Z-degree factorizes completely into linear factors over for example the field $\overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$. Let $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$. We can then write f as

$$f(X,Y) = \sum_{n=-m}^{\infty} a_n(X)Y^{-n/e}, \tag{8.1}$$

for some $m \in \mathbb{Z}$, $e \in \mathbb{N}$ and with $a_n(X) \in \overline{\mathbb{Q}}(((X^{-1})^*))$. Each of these $a_n(X)$ can then be written as

$$a_n(X) = \sum_{l=-p_n}^{\infty} a_{n,l}X^{-l/q_n},$$

for some $p_n \in \mathbb{Z}$, $q_n \in \mathbb{N}$ and with $a_{n,l} \in \overline{\mathbb{Q}}$. Notice that the denominator of the exponents of X may differ for each $a_n(X)$. We also see that the order of each $a_n(X)$ may differ. And since there are infinitely many coefficients $a_n(X)$, there may not be a shared maximum for the order of each $a_n(X)$. We therefore see that the elements in $\overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ can be quite complex. We can ask ourselves if the roots in $\overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ of polynomials in $\mathbb{Z}[X,Y][Z]$ can be written in a more easy way. This question will be dealt with in this subsection.

First of all we show that these roots can not necessarily be written in the nice way mentioned in the following Remark, where we provide a counterexample.

Remark 8.7. *Let $F \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ be a root of F. We can not generally write f as*

$$f(X,Y) = \sum_{n=-m}^{\infty}\sum_{l=-p}^{\infty} a_{n,l}X^{-l/q}Y^{-n/e}, \tag{8.2}$$

for some $m,p \in \mathbb{Z}$, $e,q \in \mathbb{N}$ and with $a_{n,l} \in \overline{\mathbb{Q}}$.

Example 8.8. *We prove the above statement by looking at the example $F(X,Y,Z) := (Y-X)Z - Y$. Since the Z-degree of F is 1, there is only one root $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$. We see that this root must be*

$$f(X,Y) = \frac{Y}{Y-X} = \frac{1}{1-\frac{X}{Y}}.$$

We substitute $W = X/Y$ in the identity $(1-W)^{-1} = 1 + W + W^2 + \dots$ and find

$$f(X,Y) = \frac{1}{1-\frac{X}{Y}} = 1 + XY^{-1} + X^2Y^{-2} + \dots.$$

We see that f can't be written as in (8.2)*, because there exists no upper boundary p/q on the exponent of X for the terms of $f(X,Y)$.*

Sometimes however, we can write the roots as in (8.2), as can be seen by the following example

Example 8.9. *Let $F(X,Y,Z) := (XY-1)Z - XY$. Since the Z-degree of F is 1, there is only one root $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$. We see that this root must be*

$$f(X,Y) = \frac{XY}{XY-1} = \frac{1}{1-\frac{1}{XY}}.$$

We substitute $W = 1/(XY)$ in the identity $(1-W)^{-1} = 1 + W + W^2 + \dots$ and find

$$f(X,Y) = \frac{1}{1-\frac{1}{XY}} = 1 + X^{-1}Y^{-1} + X^{-2}Y^{-2} + \dots.$$

So we can write f as in (8.2)*, with $e = q = 1$, $m = p = 0$ and where $a_{n,l} = 1$ exactly when $n = l$ and $a_{n,l} = 0$ otherwise.*

We do know however that all terms of any root in $\overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ of a polynomial in $\mathbb{Z}[X,Y,Z]$ have a common denominator in the exponent of X. The proof is given below.

Lemma 8.10. *Let $F \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ be a root of F. Then we also have*

$$f(X,Y) \in \overline{\mathbb{Q}}((X^{-1/q}))((Y^{-1/e})),$$

for some $e, q \in \mathbb{N}$.

Proof. By the definition of formal Puiseux series we immediately have $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))((Y^{-1/e}))$ for some $e \in \mathbb{N}$. We can write

$$f(X,Y) = \sum_{n=-m}^{\infty} a_n(X) Y^{-n/e}$$

for some $m \in \mathbb{Z}$ and with $a_n(X) \in \overline{\mathbb{Q}}(((X^{-1})^*))$. Take the field $k = \mathbb{Q}(X^{-1})$. Then we have $F \in k[Y,Z]$, where F is of positive Z-degree. Since k is a subfield of the algebraically closed field $\overline{\mathbb{Q}}(((X^{-1})^*))$, there must exist $\overline{k} \subset \overline{\mathbb{Q}}(((X^{-1})^*))$ such that $\overline{k}$ is an algebraic closure of k. By Lemma 4.65, there then exists a subfield $l \subset \overline{k} \subset \overline{\mathbb{Q}}(((X^{-1})^*))$ that is a finite field extension of k such that l contains all coefficients $a_n(X)$. Since any finite field extension of characteristic zero is also a simple field, there exists $\alpha \in l$ such that $l = k(\alpha)$. Since $\alpha \in \overline{\mathbb{Q}}(((X^{-1})^*))$, we also have $\alpha \in \overline{\mathbb{Q}}((X^{-1/q}))$ for some $q \in \mathbb{N}$. This combined with $k \subset \overline{\mathbb{Q}}((X^{-1/q}))$ gives $l = k(\alpha) \subset \overline{\mathbb{Q}}((X^{-1/q}))$. In particular all $a_n(X)$ lie in $\overline{\mathbb{Q}}((X^{-1/q}))$, which shows that indeed $f(X,Y) \in \overline{\mathbb{Q}}((X^{-1/q}))((Y^{-1/e}))$ holds. □

We can find an even smaller field that contains $f(X,Y)$; We may replace $\overline{\mathbb{Q}}$ by an algebraic number field.

Lemma 8.11. *Let $F \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ be a root of F. Then we also have*

$$f(X,Y) \in h((X^{-1/q}))((Y^{-1/e})),$$

for some $e, q \in \mathbb{N}$ and some algebraic number field $h \subset \overline{\mathbb{Q}}$.

Proof. We adopt all notation in the proof of the previous lemma. Since $k(\alpha)$ is a finite field extension over $k = \mathbb{Q}(X^{-1})$, α must be algebraic over k. Let $s = [k(\alpha) : k]$ be the degree of this field extension. Then there exists a nonzero polynomial

$$g(X,Y) = g_0(X) + g_1(X)Y + \ldots + g_s(X)Y^s \in \mathbb{Q}(X^{-1})[Y]$$

such that $g(X, \alpha(X)) = 0$ in $\overline{\mathbb{Q}}((X^{-1/q}))$. Since $\mathbb{Q}(X^{-1})$ is a fraction field of $\mathbb{Z}[X]$, we can multiply g with the common denominator $d(X)$ of $g_0(X), \ldots, g_s(X)$ to get $g'(X,Y) := d(X)g(X,Y) \in \mathbb{Z}[X,Y]$. This gives

$$g'(X, \alpha(X)) = d(X)g(X, \alpha(X)) = d(X) \cdot 0 = 0.$$

So $\alpha(X)$ is a Puiseux expansion at infinity of g'. By Corollary 4.66 we see that $\alpha(X) \in h((X^{-1/q}))$ for some algebraic number field h. Since $k \subset h((X^{-1/q}))$ holds, we also have $l = k(\alpha) \subset h((X^{-1/q}))$. In particular all $a_n(X)$ lie in $h((X^{-1/q}))$, which shows that $f(X,Y) \in h((X^{-1/q}))((Y^{-1/e}))$ indeed holds. □

Let $F(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree and let

$$f(X,Y) = \sum_{n=-m}^{\infty} a_n(X) Y^{-n/e} \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$$

be any of its roots. If F is as in Example 8.8 we can take $e = 1$, $m = 0$ and $a_n(X) = X^n$ for all $n \in \mathbb{N}_0$. This example shows us that there does not necessarily exist a maximum on the order of the elements $a_{-m}(X), a_{-m+1}(X), \ldots \in \overline{\mathbb{Q}}(((X^{-1})^*))$. We do however notice in this example that $\operatorname{ord}_X a_n(X) = \operatorname{ord}_X X^n = n$ for all $n \in \mathbb{N}_0$, so the order of $a_n(X)$ grows linear with respect to n. We will show for the general case that the order of $a_n(X)$ does not grow faster with respect to n than in a linear way. For this we first need the following lemma.

Lemma 8.12. *Let $F \in \mathbb{Z}[X,Y,Z]$. Write F as*

$$F(X,Y,Z) = \sum_{i=0}^{d_1}\sum_{j=0}^{d_2}\sum_{k=0}^{d_3} c_{i,j,k}X^iY^jZ^k,$$

for some $d_1, d_2, d_3 \in \mathbb{N}_0$ and with $c_{i,j,k} \in \mathbb{Z}$. Let $p \in \mathbb{Z}$, $e, N \in \mathbb{N}$ and $a_{-m}, a_{-m+1}, \ldots, a_N \in \overline{\mathbb{Q}}(((X^{-1})^*))$. Let c_N be the coefficient of $Y^{p/e-(N+1)/e}$ in $F(X, Y, \sum_{n=-m}^{N} a_n(X)Y^{-n/e})$, when it is viewed as a formal Puiseux series at infinity in Y over the field $\overline{\mathbb{Q}}(((X^{-1})^*))$. Let $b \in \mathbb{R}$. Suppose that

$$i + \sum_{t=1}^{k} ord_X a_{n_t} \leq b$$

holds for any $i, j, k \in \mathbb{N}_0$ such that $i \leq d_1, j \leq d_2, k \leq d_3$ and for any $n_1, \ldots, n_k \in \{-m, -m+1, \ldots, N\}$ such that $j + \sum_{t=1}^{k} -n_t/e = p/e - (N+1)/e$. Then we have that $ord_X c_N \leq b$.

Proof. For this we will use Lemma 4.58 repeatedly. Consider

$$F(X,Y,\sum_{n=-m}^{N} a_nY^{-n/e}) = \sum_{i=0}^{d_1}\sum_{j=0}^{d_2}\sum_{k=0}^{d_3} c_{i,j,k}X^iY^j(\sum_{n=-m}^{N} a_nY^{-n/e})^k.$$

We can also write this as finite sum of terms that are of the form

$$c_{i,j,k}X^iY^j\prod_{t=1}^{k} a_{n_t}Y^{-n_t/e}$$

with $i, j, k \in \mathbb{N}_0$ such that $i \leq d_1, j \leq d_2, k \leq d_3$ and such that $c_{i,j,k} \neq 0$ and with $n_1, \ldots, n_k \in \{-m, -m+1, \ldots, N\}$. Note that these terms are monomials when viewed as formal Puiseux series at infinity in Y over the field $\overline{\mathbb{Q}}(((X^{-1})^*))$. We see that the exponent of Y in such terms are of the form $j\sum_{t=1}^{k} -n_t/e$. So c_N is the sum of terms that are of the form

$$c_{i,j,k}X^i\prod_{t=1}^{k} a_{n_t}$$

with $i, j, k \in \mathbb{N}_0$ such that $i \leq d_1, j \leq d_2, k \leq d_3$ and such that $c_{i,j,k} \neq 0$ and with $n_1, \ldots, n_k \in \{-m, -m+1, \ldots, N\}$ such that $j + \sum_{t=1}^{k} -n_t/e = p/e - (N+1)/e$. Any such term satisfies

$$\mathrm{ord}_X\left(c_{i,j,k}X^i\prod_{t=1}^{k} a_{n_t}\right) = \mathrm{ord}_X(c_{i,j,k}) + i + \sum_{t=1}^{k}\mathrm{ord}_X a_{n_t} = i + \sum_{t=1}^{k}\mathrm{ord}_X a_{n_t} \leq b.$$

So therefore we also have $\mathrm{ord}_X c_N \leq b$. □

Lemma 8.13. *Let $F \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$ be a root of F. Write f as*

$$f(X,Y) = \sum_{n=-m}^{\infty} a_n(X)Y^{-n/e}$$

for some $m \in \mathbb{Z}$, $e \in \mathbb{N}$ and with $a_n = a_n(X) \in \overline{\mathbb{Q}}(((X^{-1})^*))$. Then there exists $\gamma \in \mathbb{R}$ and $\delta \in \mathbb{R}_{\geq 0}$ such that $ord_X a_n \leq \gamma + n\delta$ for all $n \in \{-m, -m+1, \ldots\}$.

Proof. We can write

$$F(X,Y,Z) = \sum_{i=0}^{d_1}\sum_{j=0}^{d_2}\sum_{k=0}^{d_3} c_{i,j,k}X^iY^jZ^k,$$

for some $d_1, d_2, d_3 \in \mathbb{N}_0$ and with $c_{i,j,k} \in \mathbb{Z}$. We can use Lemma 4.63. This shows the existence of $N' \in \mathbb{N}$ and $p \in \mathbb{Z}$ such that for all $N \in \mathbb{N}$ with $N > N'$, we have that $a_{N+1} = -c_N/d$, where $d \in \overline{\mathbb{Q}}(((X^{-1})^*))$ is nonzero and independent on N, and where $c_N \in \overline{\mathbb{Q}}(((X^{-1})^*))$ is the coefficient of $Y^{p/e-(N+1)/e}$ in

$F(X,Y,\sum_{n=-m}^{N} a_n(X)Y^{-n/e})$, when it is viewed as a formal Puiseux series at infinity in Y over the field $\overline{\mathbb{Q}}(((X^{-1})^*))$. Let

$$A \coloneqq \max(N', d_2e - p + 1 + (d_3 - 1)|m|) \in \mathbb{N}.$$

Since $a_{-m}, a_{-m+1}, \ldots, a_A \in \overline{\mathbb{Q}}(((X^{-1})^*))$ are only finitely many elements, there exists $M \in \mathbb{N}_0$ such that $\mathrm{ord}_X a_n \leq M$ for all $n \in \{-m, -m+1, \ldots, A\}$. Now let

$$\delta \coloneqq \max(1, d_1 + d_3M - \mathrm{ord}_X d) \in \mathbb{N}_0.$$

This gives us the inequality $d_1 + d_3M \leq \delta + \mathrm{ord}_X d$. Let $N \in \mathbb{N}$ be such that $N \geq A$. We let $S(N)$ be the following statement:
$S(N)$: Let $i, j, k \in \mathbb{N}_0$ such that the inequalities $i \leq d_1, j \leq d_2, k \leq d_3$ all hold. Let $n_1, \ldots, n_k \in \{-m, -m+1, \ldots, N\}$ such that

$$j + \sum_{t=1}^{k} -n_t/e = p/e - (N+1)/e. \tag{8.3}$$

Then the following inequality holds:

$$i + \sum_{t=1}^{k} \mathrm{ord}_X a_{n_t} \leq (N + 1 - A)\delta + \mathrm{ord}_X d. \tag{8.4}$$

Note that for any $N \geq A$ such that $S(N)$ holds, we can apply Lemma 8.12 to see that $\mathrm{ord}_X c_N \leq (N+1-A)\delta + \mathrm{ord}_X d$ holds, and therefore $\mathrm{ord}_X a_{N+1} = \mathrm{ord}_X(c_N/d) = \mathrm{ord}_X c_N - \mathrm{ord}_X d \leq (N+1-A)\delta$ as well. We will prove by induction that $S(N)$ holds for all $N \geq A$. Suppose first that $N = A$. Then let $i, j, k \in \mathbb{N}_0$ such that $i \leq d_1$, $j \leq d_2$ and such that $k \leq d_3$ and let $n_1, \ldots, n_k \in \{-m, -m+1, \ldots, N\}$. We then have

$$i + \sum_{t=1}^{k} \mathrm{ord}_X a_{n_t} \leq d_1 + \sum_{t=1}^{k} \mathrm{ord}_X M = d_1 + kM \leq d_1 + d_3M \leq \delta + \mathrm{ord}_X d = (N+1-A)\delta + \mathrm{ord}_X d.$$

So (8.4) holds in particular and the statement $S(A)$ is therefore true. Now suppose that $N \geq A + 1$ and as our induction hypothesis assume that the statement $S(\overline{N})$ holds for all $\overline{N} \in \mathbb{N}$ with $A \leq \overline{N} < N$. This then implies that $\mathrm{ord}_X a_n \leq (n - A)\delta$ for all $n \in \{A+1, \ldots, N\}$. We again let $i, j, k \in \mathbb{N}_0$ be such that $i \leq d_1$, $j \leq d_2$ and such that $k \leq d_3$ hold and again let $n_1, \ldots, n_k \in \{-m, -m+1, \ldots, N\}$ be such that (8.3) holds. Note that because of symmetry we may assume that $n_1 \leq \ldots \leq n_k$. Define $n_0 = -\infty$ and $n_{k+1} = \infty$. Let $t' \in \{0, \ldots, k\}$ be such that we have

$$n_0 \leq \ldots \leq n_{t'} \leq A < A + 1 \leq n_{t'+1} \leq \ldots \leq n_{k+1}.$$

We then find

$$\begin{aligned}
i + \sum_{t=1}^{k} \mathrm{ord}_X a_{n_t} &\leq d_1 + \sum_{t=1}^{t'} \mathrm{ord}_X a_{n_t} + \sum_{t=t'+1}^{k} \mathrm{ord}_X a_{n_t} \\
&\leq d_1 + \sum_{t=1}^{t'} M + \sum_{t=t'+1}^{k} (n_t - A)\delta \\
&= d_1 + t'M + \left(\sum_{t=t'+1}^{k} n_t - (k - t')A \right)\delta \\
&\leq d_1 + d_3M + \left(\sum_{t=t'+1}^{k} n_t - (k - t')A \right)\delta \\
&\leq \left(\sum_{t=t'+1}^{k} n_t - (k - t')A + 1 \right)\delta + \mathrm{ord}_X d.
\end{aligned}$$

If we now can prove

$$\sum_{t=t'+1}^{k} n_t - (k-t')A + 1 \leq N + 1 - A,$$

then we have indeed found that (8.4) holds. To prove this, we consider three cases. First suppose that $t' = k$. This gives

$$\sum_{t=t'+1}^{k} n_t - (k-t')A + 1 = 1 \leq N + 1 - A.$$

Now suppose that $t' = k - 1$. This gives

$$\sum_{t=t'+1}^{k} n_t - (k-t')A + 1 = n_k - A + 1 \leq N + 1 - A.$$

Now suppose that $t' \leq k - 2$. We rewrite (8.3) to

$$\sum_{t=1}^{k} n_t = je - p + N + 1$$

We then have

$$\begin{aligned}
\sum_{t=t'+1}^{k} n_t - (k-t')A + 1 &= je - p + N + 1 - \sum_{t=1}^{t'} n_t - \delta(k-t')A + 1 \\
&\leq d_2 e - p + N + 1 - \sum_{t=1}^{t'} -m - 2A + 1 \\
&\leq d_2 e - p + N + 1 + t'm - 2A + 1 \\
&\leq d_2 e - p + N + 1 + (k-1)|m| - 2A + 1 \\
&\leq d_2 e - p + N + 1 + (d_3 - 1)|m| - 2A + 1 \\
&\leq N - A + 1 \\
&= N + 1 - A.
\end{aligned}$$

So all three case result in (8.4). This shows that the statement $S(N)$ holds. By induction we have now found that $\mathrm{ord}_X a_n \leq (n-A)\delta$ for all $n \in \mathbb{N}$ with $n \geq A + 1$. Now let

$$\gamma := \max(0, \mathrm{ord}_X a_{-m} - (-m)\delta, \mathrm{ord}_X a_{-m+1} - (-m+1)\delta, \ldots, \mathrm{ord}_X a_A - (A)\delta).$$

This then gives

$$\mathrm{ord}_X a_n \leq \gamma + n\delta$$

for all $n \in \{-m, -m+1, \ldots\}$. □

8.3 Notion of convergence

Now that we are familiar with the form of any root $f(X,Y)$ of a nonconstant polynomial in Z over $\mathbb{Z}[X,Y]$, we will focus on the notion of convergence of such a root $f(X,Y)$. We want to know if there exists $R \in \mathbb{R}$ such that $f(x,y)$ converges in $\mathbb{C}$ for all $x, y \in \mathbb{C}$ with $|x|, |y| > R$. This is not always the case. Let us again look at our example

$$f(X,Y) = \frac{1}{1 - \frac{X}{Y}} = 1 + XY^{-1} + X^2Y^{-2} + \ldots,$$

which is the root of the polynomial $F(X,Y,Z) := (Y-X)Z - Y$. Since the series $1 + w + w^2 + \ldots$ converges only for all $w \in \mathbb{C}$ with $|w| < 1$, we see that $f(x,y)$ converges only for all $x, y \in \mathbb{C}$ with $|x/y| < 1$, or in other words, when $|x| < |y|$. So $f(x,y)$ does not converge when $|x|$ is bigger than $|y|$. We can however say something useful about convergence. First of all we have by Corollary 4.49 that $\overline{\mathbb{Q}}(\{X^*\})$ is an algebraically closed subfield of $\overline{\mathbb{Q}}((X^*))$. In the same way we have that

$\overline{\mathbb{Q}}(\{(X^{-1})^*\})$ is an algebraically closed subfield of $\overline{\mathbb{Q}}(((X^{-1})^*))$. So by Newton-Puiseux's Theorem 4.17 we have that $\overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ is an algebraically closed subfield of $\overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$. Since $\overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ includes $\mathbb{Z}[X,Y,Z]$, we see that for any nonzero $F(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ and any of its roots $f(X,Y) \in \overline{\mathbb{Q}}(((X^{-1})^*))(((Y^{-1})^*))$, that $f(X,Y)$ also lies in $\overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$. So if we write

$$f(X,Y) = \sum_{n=-m}^{\infty} a_n(X) Y^{-n/e}$$

for some $m \in \mathbb{Z}$ and with $a_n(X) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})$, we see that $a_n(X)$ is a convergent Puiseux series at infinity for all $n \in \{-m, -m+1, \ldots\}$. So for each such n, there exists $R_n \in \mathbb{R}$ such that $a_n(x)$ converges in $\mathbb{C}$ for any $x \in \mathbb{C}$ with $|x| > R_n$. We want to know whether we also may assume that R_n is independent on n. This is not trivial if $R = \max_{n \in \{-m,-m+1,\ldots\}}(R_n)$ does not exist as real number. Fortunately, this is indeed the case, as the following lemma shows.

Lemma 8.14. *Let $F \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ be a root of F. Write f as*

$$f(X,Y) = \sum_{n=-m}^{\infty} a_n(X) Y^{-n/e}$$

for some $m \in \mathbb{Z}$, $e \in \mathbb{N}$ and with $a_n(X) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})$. Then there exists $R \in \mathbb{R}$ such that $a_n(x)$ converges in $\mathbb{C}$ for all $n \in \{-m, -m+1, \ldots\}$ and for all $x \in \mathbb{C}$ with $|x| > R$.

Proof. We may assume F to be irreducible in $\mathbb{Z}[X,Y,Z]$, since its root $f(X,Y)$ must also be the root of one of the irreducible factors of F. Since F is irreducible in $\mathbb{Z}[X,Y,Z]$, it is also irreducible in $\mathbb{Q}(X)[Y,Z]$ by Gauss's Lemma. let $k = \mathbb{Q}(X)$ and take the algebraic closure $\overline{k}$ of k that is included in the algebraically closed field $\overline{\mathbb{Q}}(\{(X^{-1})^*\})$. We make use of Lemma 4.63. This shows the existence of $N' \in \mathbb{N}$ such that for all $N \in \mathbb{N}$ with $N > N'$, we have that $a_{N+1} = -c_N/d$, where $d \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})$ is independent on N and where $c_N \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})$ is one of the coefficients in $F(X,Y,\sum_{n=-m}^{N} a_n(X)Y^{-n/e})$, when it is viewed as formal Puiseux series at infinity in Y over the field $\overline{\mathbb{Q}}(\{(X^{-1})^*\})$. Since $a_{-m}(X), a_{m-1}(X), \ldots, a_{N'}(X)$ and $d^{-1}(X) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})$ are finitely many convergent Puiseux series at infinity, there exists $R \in \mathbb{R}$ such that $a_{-m}(x), a_{m-1}(x), \ldots, a_{N'}(x)$ and $d^{-1}(x)$ converge (in $\mathbb{C}$) for all $x \in \mathbb{C}$ with $|x| > R$. We now want to prove that for all $N \in \mathbb{N}$ with $N > N'$ it also holds that $a_N(x)$ converges for all $x \in \mathbb{C}$ with $|x| > R$. We do this by induction on N. Suppose that $a_{-m}(x), a_{m-1}(x), \ldots, a_N(x)$ converge for all $x \in \mathbb{C}$ with $|x| > R$. We have $a_{N+1}(X) = c_N(X) \cdot d^{-1}(X)$, where $c_N(X)$ is one of the coefficients in $F(X,Y,\sum_{n=-m}^{N} a_n(X)Y^{-n/e})$, when it is viewed as formal Puiseux series at infinity in Y over the field $\overline{\mathbb{Q}}(\{(X^{-1})^*\})$. This shows that $c_N(X)$ lies in the ring $\mathbb{Z}[X, a_{-m}(X), a_{m-1}(X), \ldots, a_N(X)]$. So $c_N(X)$ is the finite sum of products whose factors lie in $\mathbb{Z}[X] \cup \{a_{-m}(X), a_{m-1}(X), \ldots, a_N(X)\}$. This shows that $c_N(x)$ converges for all $x \in \mathbb{C}$ with $|x| > R$. Since the same hold for $d^{-1}(x)$, we see that $a_{N+1}(x)$ also converges for all $x \in \mathbb{C}$ with $|x| > R$. So by induction we see that $a_n(x)$ indeed converges in $\mathbb{C}$ for all $n \in \{-m, -m+1, \ldots\}$ and for all $x \in \mathbb{C}$ with $|x| > R$. ☐

If a root $f(X,Y)$ can be written as (8.2), we suspect that there does exist $R \in \mathbb{R}$ such that $f(x,y)$ converges. This is the following conjecture.

Conjecture 8.15. *Let $F \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ be a root of F. Suppose that we can write f as*

$$f(X,Y) = \sum_{n=-m}^{\infty} \sum_{l=-p}^{\infty} a_{n,l} X^{-l/q} Y^{-n/e}, \tag{8.5}$$

for some $m,p \in \mathbb{Z}$, $e,q \in \mathbb{N}$ and with $a_{n,l} \in \overline{\mathbb{Q}}$. Then we suspect that there exists $R \in \mathbb{R}$ such that $f(x,y)$ converges in $\mathbb{C}$ for all $x,y \in \mathbb{C}$ with $|x| > R$ and $|y| > R$.

The reason behind this conjecture is as follows: By Lemma 8.14 we see that there exists $R_1 \in \mathbb{R}$ such that $\sum_{l=-p}^{\infty} a_{n,l} x^{-l/q}$ converges in $\mathbb{C}$ for all $n \in \{-m, -m+1, \ldots\}$ and for all $x \in \mathbb{C}$ with $|x| > R_1$. For such x and n we then have

$$\sum_{l=-p}^{\infty} |a_{n,l} x^{-l/q}| < \infty.$$

By reversing the summands of (8.5), we get

$$f(X,Y) = \sum_{l=-p}^{\infty} \sum_{n=-m}^{\infty} a_{n,l} Y^{-n/e} X^{-l/q},$$

which shows that $f(X,Y) \in \overline{\mathbb{Q}}(((Y^{-1})^*))(((X^{-1})^*))$. Since $\overline{\mathbb{Q}}(\{(Y^{-1})^*\})(((X^{-1})^*))$ is an algebraically closed subfield of $\overline{\mathbb{Q}}(((Y^{-1})^*))(((X^{-1})^*))$ and since $F \in \overline{\mathbb{Q}}(\{(Y^{-1})^*\})(((X^{-1})^*))[Z]$, we find that $f(X,Y)$ also lies in $\overline{\mathbb{Q}}(\{(Y^{-1})^*\})(((X^{-1})^*))$. We use Lemma 8.14 again and find $R_2 \in \mathbb{R}$ such that $\sum_{n=-m}^{\infty} a_{n,l} y^{-n/e}$ converges in $\mathbb{C}$ for all $l \in \{-p, -p+1, \ldots\}$ and $y \in \mathbb{C}$ with $|y| > R_2$. For such y and l we then have

$$\sum_{n=-m}^{\infty} |a_{n,l} y^{-n/e}| < \infty. \tag{8.6}$$

Let $x_0 \in \mathbb{C}$ satisfy $|x_0| > R_1$. Then we have

$$f(x_0, Y) = \sum_{n=-m}^{\infty} \left(\sum_{l=-p}^{\infty} a_{n,l} x_0^{-l/q} \right) Y^{-n/e} \in \mathbb{C}((Y^{-1/e}))$$

and $F(x_0, Y, f(x_0, Y)) = 0$. So $f(x_0, Y)$ is a root of the polynomial $F(x_0, Y, Z)$. Because $\mathbb{C}$ is complete, we find by Lemma 4.50 that $f(x_0, Y) \in \mathbb{C}(\{Y^{-n/e}\})$. By Lemma 4.56 there then exists $R_3 \in \mathbb{R}$ such that

$$f(x_0, y) = \sum_{n=-m}^{\infty} \left(\sum_{l=-p}^{\infty} a_{n,l} x_0^{-l/q} \right) y^{-n/e}$$

converges in $\mathbb{C}$ for all $y \in \mathbb{C}$ with $|y| > R_3$. For such y we thus have

$$\sum_{n=-m}^{\infty} |\left(\sum_{l=-p}^{\infty} a_{n,l} x_0^{-l/q} \right) y^{-n/e}| < \infty. \tag{8.7}$$

This combined with (8.6) and (8.7) makes us suspect that there exists $R > \max(R_1, R_2, R_3)$ such that

$$\sum_{n=-m}^{\infty} \sum_{l=-p}^{\infty} |a_{n,l} x^{-l/q} y^{-n/e}| < \infty$$

for all $x \in \mathbb{C}$ and $y \in \mathbb{C}$ with $|x| > R$ and $|y| > R$. If we assume that Conjecture 8.15 holds, we get the following consequence:

Lemma 8.16. *Let $F(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ be a root of F. Write f as*

$$f(X,Y) = \sum_{n=-m}^{\infty} a_n(X) Y^{-n/e}$$

for some $m \in \mathbb{Z}$, $e \in \mathbb{N}$ and with $a_n(X) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})$. Assume that Conjecture 8.15 holds. Then there exists $\lambda, R \in \mathbb{R}$ such that $f(x,y)$ converges in $\mathbb{C}$ for all $x, y \in \mathbb{C}$ with $|x| > R$ and $|y| \geq |x|^\lambda$.

Proof. Lemma 8.10 tells us that there exists $q \in \mathbb{N}$ such that $a_n(X) \in \overline{\mathbb{Q}}((X^{-1/q}))$ for all $n \in \{-m, -m+1, \ldots\}$. Lemma 8.13 tells us that there exists $\gamma \in \mathbb{R}$ and $\delta \in \mathbb{R}_{\geq 0}$ such that $\mathrm{ord}_X a_n \leq \gamma + n\delta$ for all $n \in \{-m, -m+1, \ldots\}$. Combined, this shows for each n that we can write $a_n(X)$ as

$$a_n(X) = \sum_{l=-(\gamma+n\delta)q}^{\infty} a_{n,l} X^{-l/q},$$

with $a_{n,l} \in \overline{\mathbb{Q}}$. So we can write f as

$$f(X,Y) = \sum_{n=-m}^{\infty} \sum_{l=-(\gamma+n\delta)q}^{\infty} a_{n,l} X^{-l/q} Y^{-n/e}.$$

If $\delta = 0$, we have written f as in (8.5) with $p = \gamma q$ and we can then use Conjecture 8.15 to see that there exists $R \in \mathbb{R}$ such that $f(x, y)$ converges in $\mathbb{C}$ for all $x, y \in \mathbb{C}$ with $|x| > R$ and $|y| > R$. We can take $\lambda = 1$ and get the desired result. Now suppose that $\delta > 0$. We apply the substitutions $l' = l + n\delta q$, $a'_{n,l'} = a_{n,l}$ and $Y' = YX^{-e\delta}$. This gives us

$$\begin{aligned} f'(X, Y') &:= f(X, Y'X^{e\delta}) \\ &= \sum_{n=-m}^{\infty} \sum_{l'=-\gamma q}^{\infty} a'_{n,l'} X^{-(l'-n\delta q)/q} (Y'X^{e\delta})^{-n/e} \\ &= \sum_{n=-m}^{\infty} \sum_{l'=-\gamma q}^{\infty} a'_{n,l'} X^{-l'/q} Y'^{-n/e}. \end{aligned}$$

From $F(X, Y, f(X, Y)) = 0$ in $\overline{Q}((X^{-1/q}))((Y^{-1/e}))$, it follows that $F(X, Y'X^{e\delta}, f(X, Y'X^{e\delta})) = 0$ in $\overline{Q}((X^{-1/q}))((Y'^{-1/e}))$. So $f'(X, Y')$ is a root of the polynomial $F'(X, Y', Z) := F(X, Y'X^{e\delta}, Z) \in \mathbb{Z}[X, Y', Z]$. Note that F' is of positive Z-degree. So $f'(X, Y') \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y'^{-1})^*))$. We have written f' as in (8.5) with $p = \gamma q$ and we can therefore apply Conjecture 8.15 to see that there exists $R \in \mathbb{R}$ such that $f'(x, y') = f(x, y'x^{e\delta})$ converges in $\mathbb{C}$ for all $x, y' \in \mathbb{C}$ with $|x| > R$ and $|y'| > R$. We may assume $R \geq 0$. Now let $\lambda = e\delta + 1$. Let $x, y \in \mathbb{C}$ with $|x| > R$ and $|y| \geq |x|^\lambda = |x|^{e\delta+1}$. This means that we have $|yx^{-e\delta}| = |y| \cdot |x|^{-e\delta} \geq |x| \geq R$. Since $yx^{-e\delta} \in \mathbb{C}$, we have that $f'(x, yx^{-e\delta}) = f(x, y)$ converges, which was we wanted to prove. □

We require another generalization of Newton dots

Definition 8.17. *Let $f(X, Y) \in \mathbb{C}((X^{-1/q}))((Y^{-1/e}))$ for some $e, q \in \mathbb{N}$. We can write f as*

$$f(X, Y) = \sum_{n=-m}^{\infty} \sum_{l=-p_n}^{\infty} \omega_{-l/q,-m/e} X^{-l/q} Y^{-m/e}$$

with $n \in \mathbb{Z}$, $p_n \in \mathbb{Z}$ and with $\omega_{-l/q,-m/e} \in \mathbb{C}$. We take $\omega_{i,j} := 0$ for all $i, j \in \mathbb{Q}$ where $\omega_{i,j}$ has not been defined yet. We then define

$$D(f(X, Y)) := \{(i, j) \in Z^2 \mid \omega_{i,j} \neq 0\}$$

to be the set of Newton dots of f, when we view f as an expression in two variables.

Remark 8.18. *We can now write $f(X, Y)$ from previous definition as*

$$f(X, Y) = \sum_{(i,j) \in D(f(X,Y))} \omega_{i,j} X^i Y^j.$$

8.4 Generalization of Runge's Theorem

We will now try to generalize Runge's Theorem. We saw that the proof of Runge's Theorem 6.1 greatly depended on Theorem 6.9. It therefore seems logical to try to generalize this theorem. The next theorem will do this by trying to replace $\mathbb{Z}$ from Theorem 6.9 by $\mathbb{Z}[X]$ and view $f(X, Y) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ as Puiseux series in Y over $\overline{\mathbb{Q}}$. Because this theorem almost treats X as a constant, we will see that this theorem can only say something about the integral solutions $x, y, z \in \mathbb{Z}$ where $|x|$ is very small compared to $|y|$ and $|z|$. Note that this theorem requires an assumption that becomes redundant by Lemma 8.16 if we assume that Conjecture 8.15 holds.

Theorem 8.19. *Let $F(X, Y, Z) \in \mathbb{Z}[X, Y, Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X, Y]$. Let $f(X, Y) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ be a nonzero root of F. Let $K \subset \overline{\mathbb{Q}}(\{(X^{-1})^*\})$ be a finite field extension over $\mathbb{Q}(X^{-1})$ of degree $s := [K : \mathbb{Q}(X^{-1})]$ such that $f(X, Y) \in K(((Y^{-1})^*))$. Write f as in (4.8). So*

$$f(X, Y) = \sum_{n=-m}^{\infty} a_n Y^{-n/e},$$

with $m \in \mathbb{Z}$, $e \in \mathbb{N}$, and where all coefficients a_n lie in K such that $a_{-m} \neq 0$. Suppose that there exists $\mu', R' \in \mathbb{R}$ such that $f(x, y)$ converges in $\mathbb{C}$ for all $x, y \in \mathbb{C}$ with $|x| > R'$ and $|y| \geq |x|^{\mu'}$. Then there exists nonzero $P(X, Y, Z) \in \mathbb{Z}[X, Y, Z]$ and $\mu, R \in \mathbb{R}$ such that the following properties holds

- $\deg_Z P \le se$.
- *If $x, y \in \mathbb{C}$ with $|x| > R$ and $|y| \ge |x|^\mu$, then $f(x,y)$ converges in $\mathbb{C}$.*
- *If $x, y \in \mathbb{Z}$ with $|x| > R$ and $|y| \ge |x|^\mu$, and if $f(x,y) \in \mathbb{Z}$, then $P(x, y, f(x,y)) = 0$.*
- *If the set $\{(x,y) \in \mathbb{Z}^2 \mid |x| > R, |y| > |x|^\mu, f(x,y) \in \mathbb{Z}\}$ is Zariski-dense in $\mathbb{C}^2$, then we have $P(X, Y, f(X,Y)) = 0$ in $K((Y^{-1/e}))$.*
- *When viewed as a polynomial in Y and Z, P_λ is a monomial for every $\lambda \in \mathbb{R}_{>0}$ such that $\lambda \neq m/e$.*
- *If $\lambda = m/e > 0$, then P_λ is an element in $\mathbb{Z}[X]$ times a power of Y times a power of an irreducible polynomial in $\mathbb{Z}[X, Y, Z]$.*

Proof. This proof is very similar to the proof of Theorem 6.9.

Let $M \in \mathbb{N}_0$ be the cardinality of the set

$$B \coloneqq \{(j,k) \in \mathbb{Z}^2 \mid 1 \le k < se, -km/e \le j < 0\}. \tag{8.8}$$

Let $N \in \mathbb{N}$ satisfy

$$N \ge Me + (se-1)m.$$

Since K is a finite field extension over $\mathbb{Q}(X^{-1})$, there exists a monic irreducible polynomial $G(Y) \in \mathbb{Q}(X^{-1})[Y]$ of degree s which has a root $\theta_1 \in K$ that generates K. So $K = \mathbb{Q}(X^{-1})(\theta_1)$. Let L be the splitting field of $G(Y)$ over K. So there exists $\theta_2, \ldots, \theta_s \in L$ such that $G(Y) = (Y - \theta_1)\cdots(Y-\theta_s)$ in $L[Y]$. Any $c \in K$ can uniquely be written as $c_0 + c_1\theta_1 + \ldots + c_{s-1}\theta_1^{s-1}$ with $c_0, \ldots, c_{s-1} \in \mathbb{Q}(X^{-1})$. We denote the polynomial $c_0 + c_1 W + \ldots + c_{s-1}W^{s-1} \in \mathbb{Q}(X^{-1})[W]$ by $g_c(W)$. For $\sigma \in \{1, \ldots, s\}$, we call $c^{(\sigma)} \coloneqq g_c(\theta_\sigma) = c_0 + c_1\theta_\sigma + \ldots + c_{s-1}\theta_\sigma^{s-1} \in L$ the σ-th conjugate of c in K. In particular we have $c^{(1)} = c$.

Let $\zeta \in \mathbb{C}$ be a primitive e-th root of unity. This means that $\zeta^e = 1$ and that the elements $\zeta^0, \ldots, \zeta^{e-1}$ are all different solutions to $Z^e = 1$. It follows that the polynomial $Z^e - Y^{-1}$ in Z over $\mathbb{Q}[Y, Y^{-1}]$ factorizes over the field extension $\mathbb{Q}(\zeta)(Y^{-1/e})$ as $Z^e - Y^{-1} = (Z - \zeta^0 Y^{-1/e})\cdots(Z - \zeta^{e-1}Y^{-1/e})$. Notice that Y^{-1} is a unit in $\mathbb{Q}[Y, Y^{-1}]$. We now consider the following Laurent polynomial

$$Z - \sum_{n=-m}^{N} g_{a_n}(W) V^n \in \mathbb{Q}(X^{-1})[Y, Y^{-1}][Z][W][V, V^{-1}].$$

By Corollary 5.12 we find that

$$\prod_{\mathscr{E}=0}^{e-1}\left(Z - \sum_{n=-m}^{N} g_{a_n}(W)(\zeta^{\mathscr{E}} Y^{-1/e})^n\right) \in \mathbb{Q}(X^{-1})[Y, Y^{-1}][Z][W].$$

By Corollary 5.11 we find that

$$\prod_{\sigma=1}^{s}\prod_{\mathscr{E}=0}^{e-1}\left(Z - \sum_{n=-m}^{N} g_{a_n}(\theta_\sigma)(\zeta^{\mathscr{E}} Y^{-1/e})^n\right) \in \mathbb{Q}(X^{-1})[Y, Y^{-1}][Z].$$

We thus find for each $\beta = 0, \ldots, M$ that

$$F(X, Y, Z; \beta) \coloneqq Y^\beta \prod_{\sigma=1}^{s}\prod_{\mathscr{E}=0}^{e-1}\left(Z - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^n\right) \in \mathbb{Q}(X^{-1})[Y, Y^{-1}][Z]. \tag{8.9}$$

We view $F(X, Y, Z; \beta)$ as a Laurent polynomial in Y and Z over the field $\mathbb{Q}(X^{-1})$ and write $F(X, Y, Z; \beta) = \sum_{(j,k) \in D(F(X,Y,Z;\beta))} b_{\beta,j,k} Y^j Z^k$ with $b_{\beta,j,k} \in \mathbb{Q}(X^{-1})$, we see from (8.9) that any $(j,k) \in D(F(X, Y, Z; \beta))$ satisfies $0 \le k \le se$ and $j \in \mathbb{Z}$. The terms $b_{\beta,j,k} Y^j Z^k$ of $F(X, Y, Z; \beta)$ can thus be divided into the following three categories.

- Terms $b_{\beta,j,k} Y^j Z^k$ with $j \ge 0$.
- Terms $b_{\beta,j,k} Y^j Z^k$ with $j < 0$ and $j + km/e \le -1/e$.
- Terms $b_{\beta,j,k} Y^j Z^k$ with $j < 0$ and $j + km/e \ge 0$.

Note that the terms in the first category lie in $\mathbb{Q}(X^{-1})[Y,Z]$ and that the terms in the second category satisfy $\text{ord}_Y(b_{\beta,j,k}Y^j(f(X,Y))^k) \leq -1/e$. The terms in the third category satisfy $k \neq 0$, as otherwise $j + km/e = j \geq 0$ would contradict $j < 0$. They also satify $k \neq se$, as it would otherwise follow from (8.9) that $j = \beta \geq 0$, which again would contradict $j < 0$. We can conclude that the terms in the third category satisfy $(j,k) \in B$ as defined in (8.8). It can be seen that the terms in the first and second category do not satisfy $(j,k) \in B$. We can add terms of the same category together and get

$$F(X,Y,Z;\beta) = P(X,Y,Z;\beta) + S(X,Y,Z;\beta) + \sum_{(j,k)\in B} b_{\beta,j,k}Y^jZ^k,$$

with $P(X,Y,Z;\beta) \in \mathbb{Q}(X^{-1})[Y,Z]$ and $S(X,Y,Z;\beta) \in \mathbb{Q}(X^{-1})[Y,Y^{-1}][Z]$ with $\text{ord}_Y(S_\beta(X,Y,f(X,Y)) \leq -1/e$. We want to find $c_0,\ldots,c_M \in \mathbb{Q}(X^{-1})$, not all zero, such that $\sum_{\beta=0}^M c_\beta b_{\beta,j,k} = 0$ for all $(j,k) \in B$. This is a system of $|B| = M$ linear equations in $M+1$ variables over the field $\mathbb{Q}(X^{-1})$. It therefore indeed has a solution. We may scale these $c_0,\ldots,c_M$ by multiplying with any nonzero element in $\mathbb{Q}(X^{-1})$ and still find a suitable solution. We scale our solution such that $c_1,\ldots,c_M \in \mathbb{Z}[X]$ and such that $c_\beta \cdot P(X,Y,Z;\beta) \in \mathbb{Z}[X,Y,Z]$ for all $\beta \in \{0,\ldots,M\}$. This is possible since $\mathbb{Q}(X^{-1})$ is a fraction field of $\mathbb{Z}[X]$. This gives us

$$\sum_{\beta=0}^M c_\beta F(X,Y,Z;\beta) = \sum_{\beta=0}^M c_\beta P(X,Y,Z;\beta) + \sum_{\beta=0}^M c_\beta S(X,Y,Z;\beta),$$

hence

$$P = P(X,Y,Z) := \sum_{\beta=0}^M c_\beta P(X,Y,Z;\beta) = \sum_{\beta=0}^M c_\beta F(X,Y,Z;\beta) - \sum_{\beta=0}^M c_\beta S(X,Y,Z;\beta) \in \mathbb{Z}[X,Y,Z].$$

Note that it can be seen from this definition that $\deg_Z P \leq se$ holds. We are interested in the order of $P(X,Y,f(X,Y)) \in K(((Y^{-1})^*))$ when viewed as a formal Puiseux series at infinity in Y. First note that

$$f(X,Y) - \sum_{n=-m}^N a_n^{(\sigma)}(\zeta^{\mathscr{E}}Y^{-1/e})^n \in \mathbb{C}(((X^{-1})^*))((Y^{-1/e}))$$

holds for any $\mathscr{E} \in \{0,\ldots,e-1\}$ and $\sigma \in \{1,\ldots,s\}$. We see that

$$\begin{aligned}\text{ord}_Y\left(f(X,Y) - \sum_{n=-m}^N a_n^{(\sigma)}(\zeta^{\mathscr{E}}Y^{-1/e})^n\right) &\leq \max\left(\text{ord}_Y f(X,Y), \text{ord}_Y\left(\sum_{n=-m}^N a_n^{(\sigma)}(\zeta^{\mathscr{E}}Y^{-1/e})^n\right)\right)\\ &= \max(m/e, m/e)\\ &= m/e.\end{aligned}$$

In the case where $\mathscr{E} = 0$ and $\sigma = 1$ we find

$$\text{ord}_Y\left(f(X,Y) - \sum_{n=-m}^N a_n^{(1)}(\zeta^0Y^{-1/e})^n\right) = \text{ord}_Y\left(\sum_{n=N+1}^\infty a_nY^{-n/e}\right) \leq -(N+1)/e.$$

Together this gives us

$$\text{ord}_Y(F_\beta(X,Y,f(X,Y))) \leq \beta + (se-1)m/e - (N+1)/e \leq M + (se-1)m/e - N/e - 1/e \leq -1/e.$$

And we therefore have

$$\begin{aligned}\text{ord}_Y(P(X,Y,f(X,Y))) &= \text{ord}_Y\left(\sum_{\beta=0}^M c_\beta F_\beta(X,Y,f(X,Y)) - \sum_{\beta=0}^M c_\beta S_\beta(X,Y,f(X,Y))\right)\\ &\leq \max_\beta(\text{ord}_Y(F_\beta(X,Y,f(X,Y))), \text{ord}_Y(S_\beta(X,Y,f(X,Y))))\\ &\leq \max_\beta(-1/e, -1/e)\\ &= -1/e.\end{aligned} \tag{8.10}$$

There exists $\mu', R' \in \mathbb{R}$ such that $f(x,y)$ converges in $\mathbb{C}$ for all $x, y \in \mathbb{C}$ with $|x| > R'$ and $|y| \geq |x|^{\mu'}$. Since $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$, we have that $P(x,y,f(x,y))$ also converges for all such $x, y \in \mathbb{C}$. We can write

$$P(X,Y,f(X,Y)) = \sum_{(i,j) \in D(P(X,Y,f(X,Y)))} \omega_{i,j} X^i Y^j$$

for suitable $\omega_{i,j} \in \mathbb{C}$. We then have that

$$\sum_{(i,j) \in D(P(X,Y,f(X,Y)))} |\omega_{i,j} x^i y^j| < \infty$$

for all $x, y \in \mathbb{C}$ with $|x| > R'$ and $|y| \geq |x|^{\mu'}$. We can also write

$$P(X,Y,f(X,Y)) = \sum_{n=-m'}^{\infty} \omega_n(X) Y^{-n/e}$$

For suitable $m' \in \mathbb{Z}$ and $\omega_n(X) \in \mathbb{C}((X^{-1/q}))$. From (8.10) we get $m' \leq -1$. Since $f(X,Y)$ is integral over $\mathbb{Z}[X,Y]$, so is $P(X,Y,f(X,Y))$. So there exists nonzero $\varphi(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ such that

$$\varphi(X,Y,P(X,Y,f(X,Y))) = 0 \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*)). \tag{8.11}$$

We can apply Lemma 8.13 on $\varphi(X,Y,Z)$ and $P(X,Y,f(X,Y))$ to find that there exists $\gamma \in \mathbb{R}$ and $\delta \in \mathbb{R}_{\geq 0}$ such that $\operatorname{ord}_X \omega_n \leq \gamma + n\delta$ for all $n \in \{-m', -m'+1, \ldots\}$. So any $(i,j) \in D(P(X,Y,f(X,Y))$ satisfies $qi, ej \in \mathbb{Z}$ and $i \leq \gamma - je\delta$ and $j \leq -1/e < 0$. Now let $x_0, y_0 \in \mathbb{C}$ with $|x_0| > R'$ and $|y_0| \geq |x_0|^{\mu'}$. Let

$$A := \left(2 \sum_{(i,j) \in D(P(X,Y,f(X,Y)))} |\omega_{i,j} x_0^i y_0^j|\right)^e \in \mathbb{R}_{\geq 0}.$$

Let $R = \max(1, |x_0|, |Ay_0|) \in \mathbb{R}$ and $\mu = \max(\mu', e\delta + 1, e(\delta + \gamma) + 1, 1) \in \mathbb{R}$. Let $x, y \in \mathbb{C}$ with $|x| > R$ and $|y| > |x|^\mu$ and let $(i,j) \in D(P(X,Y,f(X,Y))$. If $i < 0$ we have

$$|x|^i |y|^j < |x_0|^i |x|^{j\mu} \leq |x_0|^i |x|^j < |x_0|^i |Ay_0|^j \leq A^{-1/e} |x_0|^i |y_0|^j.$$

If $i \geq 0$, we have

$$\begin{aligned}
|x|^i |y|^j &< |x|^{i+j\mu} \\
&\leq |x|^{\gamma - je\delta + j\mu} \\
&= |x|^{\gamma + j(\mu - e\delta - 1) + j} \\
&\leq |x|^{\gamma - (\mu/e - \delta - 1/e)} |x|^j \\
&< |x|^{\gamma - ((e(\delta+\gamma)+1)/e - \delta - 1/e)} |Ay_0|^j \\
&= |x|^0 |Ay_0|^j \\
&\leq |x_0|^i A^{-1/e} |y_0|^j \\
&= A^{-1/e} |x_0|^i |y_0|^j.
\end{aligned}$$

We therefore have found

$$\begin{aligned}
\sum_{(i,j) \in D(P(X,Y,f(X,Y)))} |\omega_{i,j} x^i y^j| &= \sum_{(i,j) \in D(P(X,Y,f(X,Y)))} |\omega_{i,j}| |x|^i |y|^j \\
&< \sum_{(i,j) \in D(P(X,Y,f(X,Y)))} |\omega_{i,j}| A^{-1/e} |x_0|^i |y_0|^j \\
&= A^{-1/e} \sum_{(i,j) \in D(P(X,Y,f(X,Y)))} |\omega_{i,j}| |x_0|^i |y_0|^j \\
&= A^{-1/e} \frac{1}{2} A^{1/e} \\
&= \frac{1}{2}
\end{aligned}$$

So this means that $|P(x,y,f(x,y))| < \frac{1}{2}$. If x, y and $f(x,y)$ are all integers, we then must have $P(x,y,f(x,y)) = 0$, since $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$. So we have found $R, \mu \in \mathbb{R}$ such that $P(x,y,f(x,y)) = 0$ holds for any pair of integers in the set

$$S_1 \coloneqq \{(x,y) \in \mathbb{Z}^2 \mid |x| > R, |y| > |x|^\mu, f(x,y) \in \mathbb{Z}\}.$$

Suppose that S_1 is Zariski-dense in C^2. We consider (8.11) again. We may assume that φ is irreducible in $\mathbb{Z}[X,Y,Z]$. We see that any $(x,y) \in S_1$ satisfies $0 = \varphi(x,y,P(x,y,f(x,y))) = \varphi(x,y,0)$. In particular this shows that $S_1 \subset T(\varphi(X,Y,0)) \subset \mathbb{C}^2$. Because S_1 is Zariski-dense in C^2, we see by definition of Zariski density that the Zariski-closed set $T(\varphi(X,Y,0))$ must be equal to C^2. So $\varphi(x,y,0) = 0$ for all $(x,y) \in \mathbb{C}^2$ and from this we see that $\varphi(X,Y,0)$ is the zero polynomial. So Z is a divisor of $\varphi(X,Y,Z)$. Since φ is irreducible, we therefore must have $\varphi(X,Y,Z) = \pm Z$, and this gives

$$0 = \varphi(X,Y,P(X,Y,f(X,Y))) = \pm P(X,Y,f(X,Y)) \in K((Y^{-1/e})).$$

We will now view $P(X,Y,Z)$ as a polynomial in Y and Z over the ring $\mathbb{Z}[X]$ and try to compute the λ-leading part of P for all $\lambda \in \mathbb{R}_{>0}$. Here we have $\deg_\lambda P = \max_{(j,k)\in D(P)}(j + \lambda k)$. For computing this we will apply the rules from Remark 4.60. Let γ be the largest element in $\{0,\dots,M\}$ such that $c_\gamma \neq 0$. First suppose that $\lambda > m/e$. This gives

$$\left(Z - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^n\right)_\lambda = Z,$$

for any $\sigma \in \{1,\dots,s\}$ and $\mathscr{E} \in \{0,\dots,e-1\}$. And therefore we have

$$F(X,Y,Z;\beta)_\lambda = (Y^\beta)_\lambda \prod_{\sigma=1}^{s} \prod_{\mathscr{E}=0}^{e-1} \left(Z - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^n\right)_\lambda = Y^\beta Z^{se}.$$

So $Y^\beta Z^{se} \in \mathbb{Z}[X,Y,Z]$ is the term of F_β that is of the highest λ-degree. This gives us

$$P(X,Y,Z;\beta)_\lambda = Y^\beta Z^{se},$$

and therefore we see that

$$P_\lambda = \left(\sum_{\beta=0}^{M} c_\beta P(X,Y,Z;\beta)\right)_\lambda = c_\gamma P(X,Y,Z;\gamma)_\lambda = c_\gamma Y^\gamma Z^{se},$$

which indeed is a monomial when viewed as a polynomial in Y and Z. Note that this additionally shows that P is nonzero since its λ-leading part is nonzero.

Now suppose that $\lambda < m/e$, from which can deduce that $m > 0$. This gives

$$\left(Z - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^n\right)_\lambda = -a_{-m}^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^{-m}).$$

for any $\sigma \in \{1,\dots,s\}$ and $\mathscr{E} \in \{0,\dots,e-1\}$. And therefore we have

$$\begin{aligned} F(X,Y,Z;\beta)_\lambda &= (Y^\beta)_\lambda \prod_{\sigma=1}^{s} \prod_{\mathscr{E}=0}^{e-1} \left(Z - \sum_{n=-m}^{N} a_n^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^n\right)_\lambda \\ &= Y^\beta \prod_{\sigma=1}^{s} \prod_{\mathscr{E}=0}^{e-1} -a_{-m}^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^{-m} \\ &= Y^\beta \prod_{\sigma=1}^{s} \pm(a_{-m}^{(\sigma)})^e Y^m \\ &= d_\beta Y^{\beta+ms}, \end{aligned}$$

for some nonzero $d_\beta \in \mathbb{Q}(X^{-1})$. We find in a similar way as previous case

$$P(X,Y,Z;\beta)_\lambda = d_\beta Y^{\beta+ms},$$

and therefore we see that

$$P_\lambda = \left(\sum_{\beta=0}^{M} c_\beta P(X,Y,Z;\beta)\right)_\lambda = c_\gamma d_\gamma Y^{\gamma+ms},$$

which is also a monomial.

Now suppose that $\lambda = m/e$, from which can deduce that $m > 0$. This gives

$$\left(Z - \sum_{n=-m}^{N} a_n^{(\sigma)} (\zeta^{\mathscr{E}} Y^{-1/e})^n\right)_\lambda = Z - a_{-m}^{(\sigma)} (\zeta^{\mathscr{E}} Y^{-1/e})^{-m}.$$

for any $\sigma \in \{1,\dots,s\}$ and $\mathscr{E} \in \{0,\dots,e-1\}$. So

$$P_\lambda = c_\gamma Y^\gamma \prod_{\sigma=1}^{s} \prod_{\mathscr{E}=0}^{e-1} (Z - a_{-m}^{(\sigma)} (\zeta^{\mathscr{E}} Y^{-1/e})^{-m}) \in \mathbb{Q}(X^{-1})[Y,Y^{-1}][Z] \subset \mathbb{Q}(X^{-1})(Y)[Z].$$

Let $m_1, e_1 \in \mathbb{N}$ be such that $\gcd(m_1,e_1) = 1$ and $m/e = m_1/e_1$. Let $h(X,Z) \in \mathbb{Z}[X,Z]$ be the minimum polynomial of ${a_{-m}}^{e_1}$, when viewed as a polynomial in Z over $\mathbb{Z}[X]$. So $h(X,Z)$ is irreducible in $Z[X,Z]$ and contains $Z = {a_{-m}}^{e_1}$ as a root. Let $b(X) \in \mathbb{Z}[X]$ be the leading coefficient of h and s_1 the Z-degree of h. Since h is irreducible, we have by Lemma 6.8 that

$$G(X,Y,Z) \coloneqq Y^{m_1 s_1} h(X, Z^{e_1} Y^{-m_1}) \in \mathbb{Z}[X,Y,Z]$$

is irreducible in $\mathbb{Z}[X,Y,Z]$, and therefore also irreducible in $\mathbb{Q}(X)(Y)[Z]$. For any $\sigma \in \{1,\dots,s\}$ and $\mathscr{E} \in \{0,\dots,e-1\}$ we have

$$\begin{aligned}
G(X,Y,{a_{-m}}^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^{-m}) &= Y^{m_1 s_1} h(X, ({a_{-m}}^{(\sigma)}(\zeta^{\mathscr{E}} Y^{-1/e})^{-m})^{e_1} Y^{-m_1}) \\
&= Y^{m_1 s_1} h(X, ({a_{-m}}^{(\sigma)})^{e_1} \zeta^{-me_1\mathscr{E}} Y^{me_1/e - m_1}) \\
&= Y^{m_1 s_1} h(X, ({a_{-m}}^{e_1})^{(\sigma)} \zeta^{-m_1 e \mathscr{E}}) \\
&= Y^{m_1 s_1} h(X, ({a_{-m}}^{e_1})^{(\sigma)} (\zeta^e)^{-m_1 \mathscr{E}}) \\
&= Y^{m_1 s_1} h(X, ({a_{-m}}^{e_1})^{(\sigma)}) \\
&= Y^{m_1 s_1} h(X, {a_{-m}}^{e_1})^{(\sigma)} \\
&= Y^{m_1 s_1} 0^{(\sigma)} \\
&= 0
\end{aligned}$$

Since all roots of P_λ, when viewed as a polynomial in Z, are roots of the irreducible polynomial G, we conclude that P_λ has no irreducible factors in $\mathbb{Q}(X)(Y)[Z]$ besides G. Since G is of Z-degree $s_1 e_1$ and P_λ of Z-degree se, it must follow that $s_1 e_1$ divides se and that $P_\lambda = cG^{se/(s_1 e_1)}$ for some $c \in \mathbb{Q}(X)(Y)$. By looking at the leading coefficient of P_λ and G when viewed as polynomials in Z, we see that $c = c_\gamma Y^\gamma b^{-se/(s_1 e_1)}$. So P_λ is $c_\gamma b^{-se/(s_1 e_1)}$ times a power of Y times a power of an irreducible polynomial in $\mathbb{Z}[X,Y,Z]$. From and $P_\lambda \in \mathbb{Z}[X,Y,Z]$ we conclude that P_λ is an nonzero element in $\mathbb{Z}[X]$ times a power of Y times a power of an irreducible polynomial in $\mathbb{Z}[X,Y,Z]$. This finishes our proof.

□

The proof of Runge's Theorem at some point switched the roles of the two variables and applied Theorem 6.9 again. Here we want to switch the roles of Y and Z. In order to still fully apply Theorem 8.19 after switching the roles of Y and Z, we need to know for any $(x,y,z) \in S(F)$ that if $|x|$ is very small compared to $|y|$, that $|x|$ is then also very small compared to $|z|$. For this we prove the following lemmas.

Lemma 8.20. *Let $F(X) = \sum_{i=1}^{n} a_i X^i \in \mathbb{Z}[X]$ be a nonconstant polynomial of degree $n \in \mathbb{N}$. Let*

$$h \coloneqq \max(|a_0|,\dots,|a_n|) \in \mathbb{N}.$$

Suppose that $x \in \mathbb{C}$ is a root of F. Then it follows that $|x| \le nh$.

Proof. If $|x| < 1$ we are immediately done, so suppose that $|x| \geq 1$. We then have the inequality

$$|x^n| \leq |a_n x^n| = |\sum_{i=0}^{n-1} a_i x^i| \leq \sum_{i=0}^{n-1} |a_i x^i| \leq \sum_{i=0}^{n-1} h|x^i| \leq nh|x^{n-1}|.$$

We divide on both sides by $|x^{n-1}|$ and this yields the result $|x| \leq nh$. $\square$

Lemma 8.21. *Let $F(X, Y, Z) \in \mathbb{Z}[X, Y, Z]$ be a nonconstant irreducible polynomial of positive Y-degree $d_2 \in \mathbb{N}$. Let $T(F)$ be the zero locus of F and $S(F)$ the intergral zero locus of F. Let S' be any subset of $S(F)$. Suppose that the set*

$$A_{R,\mu} \coloneqq \{(x, y, z) \in S' \mid |x| > R, |y| > |x|^{\mu}\}$$

is Zariski-dense in $T(F)$ for all $R, \mu \in \mathbb{R}$. Then it follows that the set

$$B_{R,\mu} \coloneqq \{(x, y, z) \in S' \mid |x| > R, |z| > |x|^{\mu}\}$$

is also Zariski-dense in $T(F)$ for all $R, \mu \in \mathbb{R}$.

Proof. We can write F as

$$\sum_{i=0}^{d_1} \sum_{j=0}^{d_2} \sum_{k=0}^{d_3} a_{i,j,k} X^i Y^j Z^k$$

for smallest possible $d_1, d_3 \in \mathbb{N}_0$ and with $a_{i,j,k} \in \mathbb{Z}$. Now let

$$h \coloneqq \max_{(i,j,k) \in D(F)} (|a_{i,j,k}|).$$

Let $R \in \mathbb{R}$ and $\mu \in \mathbb{R}$ be real numbers. Let $R' \coloneqq \max(R, d_2(d_1+1)(d_3+1)h)$ and let $\mu' \coloneqq \max(\mu, d_3\mu + d_1 + 1)$. We then have that $A_{R',\mu'}$ is Zariski-dense in $T(F)$. We define the three subsets

$$S_1 \coloneqq \{(x, y, z) \in A_{R',\mu'} \mid F(x, Y, z) = 0 \in \mathbb{Z}[Y]\},$$

$$S_2 \coloneqq \{(x, y, z) \in A_{R',\mu'} \mid z = 0\}$$

and

$$S_3 \coloneqq \{(x, y, z) \in A_{R',\mu'} \mid F(x, Y, z) \neq 0 \in \mathbb{Z}[Y], z \neq 0\}.$$

Note that $S_1 \cup S_2 \cup S_3 = A_{R',\mu'}$. We view F as a polynomial in Y and write

$$F(X, Y, Z) = \sum_{j=0}^{d_2} a_j(X, Z) Y^i$$

for suitable $a_j(X, Z) \in \mathbb{Z}[X, Z]$. For any $(x, y, z) \in S_1$ we then find that $a_j(x, z) = 0$ for all $j \in \{0, \ldots, d_2\}$. In particular this means that a_{d_2}, when viewed as a polynomial in $\mathbb{Z}[X, Y, Z]$, vanishes on S_1. Because a_{d_2} is not a multiple of F in $\mathbb{Z}[X, Y, Z]$, we can apply Corollary 7.17 to conclude that S_1 is not Zariski-dense in $T(F)$. We apply Corollary 7.17 on F and on the polynomial Z to see that S_2 is also not Zariski-dense in $T(F)$. Because $A_{R',\mu'}$ is Zariski-dense in $T(F)$ while S_1 and S_2 are not, we can apply Lemma 7.22 to see that S_3 then must be Zariski-dense in $T(F)$. For any $(x, y, z) \in S_3$ and for any $j \in \{0, \ldots, d_2\}$ we find that

$$\begin{aligned} |a_j(x, z)| &= |\sum_{i=0}^{d_1} \sum_{k=0}^{d_3} a_{i,j,k} x^i z^k| \\ &\leq \sum_{i=0}^{d_1} \sum_{k=0}^{d_3} |a_{i,j,k} x^i z^k| \\ &\leq \sum_{i=0}^{d_1} \sum_{k=0}^{d_3} h|x^i z^k| \\ &\leq (d_1 + 1)(d_3 + 1)h|x|^{d_1}|z|^{d_3}. \end{aligned}$$

By Lemma 8.20 we then see that

$$|x|^{\mu'} < |y| \leq d_2(d_1+1)(d_3+1)h|x|^{d_1}|z|^{d_3} \leq R'|x|^{d_1}|z|^{d_3} \leq |x|^{d_1+1}|z|^{d_3}.$$

This yields

$$|x|^{d_3\mu} \leq |x|^{\mu'-d_1-1} < |z|^{d_3}$$

and therefore

$$|x|^{\mu} < |z|.$$

From this we see that S_3 is a subset of $B_{R,\mu}$. We now apply Lemma 7.9 to conclude that $B_{R,\mu}$ is Zariski-dense in $T(F)$. □

Lemma 8.22. *Let $F(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ be a nonconstant irreducible polynomial. Let $(d_1,d_2,d_3) = (\deg_X F, \deg_Y F, \deg_Z F)$ and suppose that $d_2 \geq 1$. Let $T(F)$ be the zero locus of F and $S(F)$ the integral zero locus of F. Let S' be any subset of $S(F)$. Suppose that the set*

$$A_{R,\mu} := \{(x,y,z) \in S' \mid |x| > R, |y| > |x|^{\mu}\}$$

is Zariski-dense in $T(F)$ for all $R, \mu \in \mathbb{R}$. Then it follows that (d_2,d_3) is no Newton dot of F when we view F as a polynomial in Y and Z over $\mathbb{Z}[X]$.

Proof. We can write F as

$$F(X,Y,Z) = \sum_{i=0}^{d_1}\sum_{j=0}^{d_2}\sum_{k=0}^{d_3} a_{i,j,k}X^iY^jZ^k.$$

Let

$$h := \max_{(i,j,k)\in D(F)} |a_{i,j,k}|.$$

Suppose that (d_2,d_3) is a Newton dot of F when we view F as a polynomial in Y and Z over $\mathbb{Z}[X]$. Let $g(X) = \sum_{i=0}^{d_1} a_{i,d_2,d_3}X^i$ be its corresponding coefficient. Then $g(X)$ is a nonzero polynomial and therefore has at most d_1 roots in Z. Let $R \in \mathbb{R}$ such that R is larger than all these roots and such that

$$R > ((d_2+1)(d_3+1)-1)(d_1+1)h. \tag{8.12}$$

Let $\mu = d_1 + 1$. We will use results of the proof of the previous lemma. Note that F indeed satisfies the conditions of said lemma. We saw for a certain $R' \in \mathbb{R}$ and $\mu' \in \mathbb{R}$ with $\mu' \geq \mu$, that the set

$$S_3 := \{(x,y,z) \in A_{R',\mu'} \mid F(x,Y,z) \neq 0 \in \mathbb{Z}[Y], z \neq 0\}$$

is Zariski-dense in $T(F)$ and is also a subset of

$$B_{R,\mu} := \{(x,y,z) \in S' \mid |x| > R, |z| > |x|^{\mu}\}.$$

As a consequence we see that S_3 must be nonempty. So there exists $(x,y,z) \in A_{R',\mu'} \subset S(F)$ such that $|x| > R$, such that $|z| > |x|^{\mu}$ and such that $|y| > |x|^{\mu'} \geq |x|^{\mu}$. In particular we have $x,y,z \in \mathbb{Z}$ and $F(x,y,z) = 0$. For any $j \in \{0,\ldots,d_2\}$ and $k \in \{0,\ldots,d_3\}$ we have the inequality

$$\begin{aligned}|\sum_{i=0}^{d_1} a_{i,j,k}x^i| &\leq \sum_{i=0}^{d_1} |a_{i,j,k}x^i| \\ &\leq \sum_{i=0}^{d_1} h|x|^i \\ &\leq (d_1+1)h|x|^{d_1}.\end{aligned}$$

Now suppose that these j,k satisfy $(j,k) \neq (d_2,d_3)$. If $j \neq d_2$, we find

$$
\begin{aligned}
|\sum_{i=0}^{d_1} a_{i,j,k}x^iy^jz^k| &\le (d_1+1)h|x|^{d_1}|y|^j|z|^k \\
&\le (d_1+1)h|x|^{d_1}|y|^{d_2-1}|z|^{d_3} \\
&< (d_1+1)h|x|^{d_1}|x|^{-\mu}|y|^{d_2}|z|^{d_3} \\
&= (d_1+1)h|x|^{-1}|y|^{d_2}|z|^{d_3} \\
&< (d_1+1)hR^{-1}|y|^{d_2}|z|^{d_3}.
\end{aligned}
$$

If $k \neq d_3$, we also find

$$
|\sum_{i=0}^{d_1} a_{i,j,k}x^iy^jz^k| < (d_1+1)hR^{-1}|y|^{d_2}|z|^{d_3}
$$

by similar reasoning. By definition of R, we have $g(x) \neq 0$, so $|g(x)| \geq 1$. We now find

$$
\begin{aligned}
|y^{d_2}z^{d_3}| &\le |g(x)y^{d_2}z^{d_3}| \\
&= |\sum_{i=0}^{d_1} a_{i,d_2,d_3}x^iy^{d_2}z^{d_3}| \\
&= |\sum_{j=0}^{d_2-1}\sum_{k=0}^{d_3}\sum_{i=0}^{d_1} a_{i,j,k}x^iy^jz^k + \sum_{k=0}^{d_3}\sum_{i=0}^{d_1} a_{i,d_2,k}x^iy^{d_2}z^k| \\
&\le \sum_{j=0}^{d_2-1}\sum_{k=0}^{d_3}|\sum_{i=0}^{d_1} a_{i,j,k}x^iy^jz^k| + \sum_{k=0}^{d_3}|\sum_{i=0}^{d_1} a_{i,d_2,k}x^iy^{d_2}z^k| \\
&< ((d_2+1)(d_3+1)-1)(d_1+1)hR^{-1}|y|^{d_2}|z|^{d_3} \\
&< |y|^{d_2}|z|^{d_3} \\
&= |y^{d_2}z^{d_3}|.
\end{aligned}
$$

This gives a contradiction from which we may conclude that (d_2, d_3) is no Newton dot of F, when we view F as a polynomial in Y and Z over $\mathbb{Z}[X]$. □

We can now give our first attempt of generalizing Runge's Theorem.

Theorem 8.23. *Let $F(X,Y,Z) = \sum_{i=0}^{d_1}\sum_{j=0}^{d_2}\sum_{k=0}^{d_3} a_{i,j,k}X^iY^jZ^k \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive X-degree $d_1 \in \mathbb{N}$, of positive Y-degree $d_2 \in \mathbb{N}$ and of positive Z-degree $d_3 \in \mathbb{N}$. Suppose that $F(X,Y,Z)$ is irreducible in $\mathbb{Q}[X,Y,Z]$. Let $f_1(X,Y),\ldots,f_{d_3}(X,Y) \in \overline{Q}(\{(X^{-1})^*\})(((Y^{-1})^*))$ be the roots of F. Suppose for each $i \in \{1,\ldots,d_3\}$ that there exists $R_i \in \mathbb{R}$ and $\mu_i \in \mathbb{R}$ such that $f_i(x,y)$ converges for all $x, y \in \mathbb{C}$ with $|x| > R_i$ and $|y| \geq |x|^{\mu_i}$. Suppose that the set*

$$
A_{R,\mu} \coloneqq \{(x,y,z) \in S(F) \mid |x| > R, |y| > |x|^{\mu}\}
$$

is Zariski-dense in $T(F)$ for all $R, \mu \in \mathbb{R}$. View F as a polynomial in Y and Z. Then the following two properties hold:

1. *In the yz-plane, no point of $D(F)$ lies above the line connecting $(d_1, 0)$ and $(0, d_2)$.*
2. *For $\lambda = d_2/d_3$, the λ-leading part of F satisfies*

$$
\sum_{(j,k)\in D_\lambda(F)}\sum_{i=1}^{d_1} a_{i,j,k}X^iY^jZ^k = ap^k,
$$

where $a \in \mathbb{Z}[X], k \in \mathbb{N}$ and where $p = p(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ is an irreducible polynomial.

Proof. View F as polynomial in Y and Z over $\mathbb{Z}[X]$. If $\deg F = 1$, we have $d_2 = d_3 = 1$ and we can then write $F(X,Y,Z) = aY + bZ + c$ for some $a,b,c \in \mathbb{Z}[X]$ such that $a \neq 0$ and $b \neq 0$. The first property then immediately follows. We have $\lambda = d_2/d_3 = 1$, and the λ-leading part of F is $aX + bY$. Since this is irreducible in $\mathbb{Q}(X^{-1})[Y,Z]$, we have that the λ-leading part of F is indeed an element in $\mathbb{Z}[X]$ times (a power of) an irreducible polynomial in $\mathbb{Z}[X,Y,Z]$, which shows that the second property also holds. Now suppose that $\deg F > 1$. We in particular have $F \in \overline{Q}(\{(X^{-1})^*\})(((Y^{-1})^*))[Z]$. Since $\overline{Q}(\{(X^{-1})^*\})(((Y^{-1})^*))$ is algebraically closed, there exists $g(X,Y) \in \mathbb{Z}[X,Y]$ such that

$$F(X,Y,Z) = g(X,Y)\prod_{i=1}^{d_3}(Z - f_i(X,Y)). \tag{8.13}$$

We have here that $g(X,Y) \in \mathbb{Z}[X,Y]$ is the leading coefficient of F, when F is viewed as a polynomial in Z. If $f_i = 0$ for some $i \in \{1,\dots,d_3\}$, we see from (8.13) that Z would then be a factor of F, which leads to a contradiction since F is irreducible of positive Y-degree. So f_i is nonzero for all $i \in \{1,\dots,d_3\}$. Take the field $k = \mathbb{Q}(X^{-1})$. Then we have $F \in k[Y,Z]$, where F is of positive Z-degree. Since k is a subfield of the algebraically closed field $\overline{\mathbb{Q}}(((X^{-1})^*))$, there must exist $\overline{k} \subset \overline{\mathbb{Q}}(((X^{-1})^*))$ that is an algebraic closure of k. By Lemma 4.65, there then exists a subfield $K_i \subset \overline{k} \subset \overline{\mathbb{Q}}(((X^{-1})^*))$ for each $i \in \{1,\dots,d_3\}$ that is a finite field extension of k such that K_i contains all coefficients of f_i, when viewed as a Puiseux series at infinity in Y over $\overline{Q}(\{(X^{-1})^*\})$. Let $R := \max(R_1,\dots,R_{d_2})$ and let $\mu := \max(\mu_1,\dots,\mu_{d_2})$. We then see for each $i\{1,\dots,d_3\}$ that $f_i(x,y)$ converges in $\mathbb{C}$ for all $x,y \in \mathbb{C}$ with $|x| > R$ and $|y| \geq |x|^\mu$. We consider the sets

$$S_i = \{(x,y,z) \in A_{R,\mu} \mid f_i(x,y) = z\}$$

for each $i \in \{1,\dots,d_3\}$ and the set

$$S' = \{(x,y,z) \in A_{R,\mu} \mid g(x,y) = 0\}.$$

We see from (8.13) that $S' \cup S_1 \cup \ldots \cup S_{d_3} = A_{R,\mu}$. Because $g(X,Y)$ is no multiple of F in $\mathbb{Z}[X,Y,Z]$, we see from Corollary 7.17 that S' is not Zariski-dense in $T(F)$. Because $A_{R,\mu}$ is Zariski-dense in $T(F)$, it then follows from Lemma 7.22 that S_i is Zariski-dense in $T(F)$ for some $i \in \{1,\dots,d_3\}$. We fix this i. From Lemma 8.19 it then follows that there exists nonzero $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ and $e \in \mathbb{N}$ such that $P(X,Y,f_i(X,Y)) = 0$ in $K_i((Y^{-1/e}))$. When we view P and F as polynomials in Z over $\mathbb{Z}[X,Y]$, we see that F is irreducible and that one of its roots is also a root of P. Because of this we have that F is a factor of P in $\mathbb{Z}[X,Y,Z]$. So $P = FH$ for some $H \in \mathbb{Z}[X,Y,Z]$. We now view P, F and H as polynomials in Y and Z. By Remark 4.60 we have $P_\lambda = F_\lambda H_\lambda$ for all $\lambda \in \mathbb{R}_{>0}$. For $\lambda \neq d_2/d_3$ we have that P_λ is a monomial. Entire Subsection 3.4 stays true when we replace $\mathbb{Z}$ by any other unique factorization domain, in particular $\mathbb{Z}[X]$. Therefore it follows from Lemma 3.18 that F_λ is a monomial as well. Now let $\lambda = d_2/d_3$. Suppose that F_λ is also a monomial, we then have that $\tilde{F}$ is a monomial by Lemma 3.19. We then apply Lemma 3.20 to see that $(d_2,d_3) \in D(F)$. This however contradicts the result from Lemma 8.22. Therefore, F_λ is not a monomial. We have by Theorem 8.19 that P_λ is a element in $\mathbb{Z}[X]$ of a power of Y times a power of an irreducible polynomial $G(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$. Since F_λ is a factor of P_λ, we see that F_λ is also an element in $\mathbb{Z}[X]$ times a power of Y times a power of $G(X,Y,Z)$. Since F_λ is not a monomial, we have that $G(X,Y,Z)$ is not a monomial, when viewed as a polynomial in Y and Z. We now reverse the roles of Y and Z: let $F'(X,Y,Z) := F(X,Z,Y)$. We see that $F'(X,Y,Z)$ is of positive Y-degree d_3 and of positive Z-degree d_2. We also have $\deg F' = \deg F$. By Lemma 8.21 we have that the set

$$\{(x,y,z) \in S(F) \mid |x| > R', |z| > |x|^{\mu'}\}$$

is Zariski-dense in $T(F)$ for all $R', \mu' \in \mathbb{R}$. We have $(x,y,z) \in S(F)$ if and only if $(x,z,y) \in S(F')$. Therefore we have that the set

$$\{(x,y,z) \in S(F') \mid |x| > R', |y| > |x|^{\mu'}\}$$

is Zariski-dense in $T(F')$ for all $R', \mu' \in \mathbb{R}$. By applying the same reasoning as with F, we see that with $\lambda' = \deg_Y F'/\deg_Z F' = d_3/d_2 = \lambda^{-1}$ we have that $F'_{\lambda'} = F_\lambda$ is an element in $\mathbb{Z}[X]$ times a power of Z times a power of an irreducible polynomial $G'(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ that is not a monomial when viewed as a polynomial in Y and Z. Since $\mathbb{Z}[X,Y,Z]$ is a unique factorization domain we conclude that $G(X,Y,Z) = G'(X,Y,Z)$ and that F_λ is an element in $\mathbb{Z}[X]$ times a power of this irreducible $G(X,Y)$. This is the second property of this theorom. For the first property we notice that since $G(X,Y,Z)$ is irreducible and not a monomial when viewed as a polynomial in Y and Z, it is not divisible by Y and therefore there exists $a \in \mathbb{N}$ with $(a, 0) \in D_\lambda(G)$. Furthermore, because F_λ is an element in $\mathbb{Z}[X]$ times a power of G, we have that $(ka, 0) \in D_\lambda(F)$ for some $k \in \mathbb{N}$. We have the restriction $ka \leq d_2$ since F is of Y-degree d_2 and $D_\lambda(F) \subset D(F)$. For any $(j, k) \in D(F)$ we now have $j + \lambda k \leq \deg_\lambda(F) = ka + 0\lambda \leq d_1$. This can be rewritten as $d_3 j + d_2 k \leq d_2 d_3$, which shows that no point of $D(F)$ lies above the line connecting $(d_2, 0)$ and $(0, d_3)$. This is the first property of this theorem. □

We can use this theorem in the following example

Example 8.24. *Let $F(X, Y, Z) = Z^2 - X^2 - Y^2$. Then there exists $R \in \mathbb{R}$ and $\mu \in \mathbb{R}$ such that the set*

$$A_{R,\mu} := \{(x, y, z) \in S(F) \mid |x| > R, |y| > |x|^\mu\}$$

is not Zariski-dense in $T(F)$.

Proof. The roots of F are given by $f_\pm(X, Y) = \pm\sqrt{X^2 + Y^2}$. Let $g(V) \in \overline{\mathbb{Q}}(\{(V^{-1})^*\})$ be any of the two roots for W in the polynomial $W^2 - 1 - V^2 \in \mathbb{Z}[V, W]$. These two roots exist by Theorem 4.46. So $(g(V))^2 = 1 + V^2$ and there exists $R' \in \mathbb{R}$ such that $g(v)$ converges for all $v \in \mathbb{C}$ with $|v| > R'$. So $g(x^{-1}y)$ converges for all $x, y \in \mathbb{C}$ with $x \neq 0$ and $|x^{-1}y| > R'$. This equation translates to $|y| > R'|x|$. We let $R_1 = \max(1, R')$ and $\mu_1 = 2$. Any $x, y \in \mathbb{C}$ with $|x| > R_1$ and $|y| \geq |x|^{\mu_1}$ then satisfies $|y| \geq |x|^2 \geq R_1|x|$. We now have $f_\pm(X, Y) = \pm\sqrt{X^2 + Y^2} = \pm X\sqrt{1 + X^{-2}Y^2} = \pm X g(X^{-1}Y)$. So $f_\pm(x, y)$ converges for all $x, y \in \mathbb{C}$ with $|x| > R_1$ and $|y| \geq |x|^{\mu_1}$. F is irreducible in $\mathbb{Q}[X, Y, Z]$ because $X^2 + Y^2$ is not a square in $\mathbb{Q}[X, Y]$. We have $d_2 = \deg_Y F = 2$ and $d_3 = \deg_Z F = 2$. We view F as polynomial in Y and Z over $\mathbb{Z}[X]$ and find $F_\lambda = Z^2 - Y^2 = (Z - Y)(Z + Y)$. Since this is not an element in $\mathbb{Z}[X]$ times an irreducible polynomial in $\mathbb{Z}[X, Y, Z]$, We then find the desired result by applying Theorem 8.23. □

In the example above we find by symmetry of X and Y that there also exist $R' \in \mathbb{R}$ and $\mu' \in \mathbb{R}$ such that the set

$$B_{R',\mu'} := \{(x, y, z) \in S(F) \mid |y| > R', |x| > |y|^{\mu'}\}$$

is not Zariski-dense in $T(F)$. This does not necessarily mean that $S(F)$ is also not Zariski-dense in $T(F)$. We saw in Example 8.5 that this is not the case.

Let $F \in \mathbb{Z}[X, Y, Z]$ be an irreducible polynomial of positive X-degree $d_1 \in \mathbb{N}$, of positive Y-degree $d_2 \in \mathbb{N}$ and of positive Z-degree $d_3 \in \mathbb{N}$. Also suppose that there exists $R, R', \mu, \mu' \in \mathbb{R}$ such that the sets

$$A_{R,\mu} := \{(x, y, z) \in S(F) \mid |x| > R, |y| > |x|^\mu\}$$

and

$$B_{R',\mu'} := \{(x, y, z) \in S(F) \mid |y| > R', |x| > |y|^{\mu'}\}$$

are both not Zariski-dense in $T(F)$. Let $G(X, Y, Z) = \prod_{a=-R}^{R} (X - a) \in \mathbb{Z}[X, Y, Z]$. We see that any $(x, y, z) \in S(F)$ with $|x| \leq R$ satisfies $g(x, y, z) = 0$. We therefore see by Corollary 7.17 that the set

$$S_1 = \{(x, y, z) \in S(F) \mid |x| \leq R\}$$

is not Zariski-dense in $T(F)$. By similar reasoning we see that the set

$$S_2 = \{(x, y, z) \in S(F) \mid |y| \leq R'\}$$

is also not Zariski-dense in $T(F)$. We now consider the set

$$S_3 = \{(x, y, z) \in S(F) \mid |x|^{\mu'^{-1}} \leq |y| \leq |x|^\mu\}.$$

We note that

$$S(F) = A_{R,\mu} \cup B_{R',\mu'} \cup S_1 \cup S_2 \cup S_3.$$

We use Lemma 7.22 to conclude that $S(F)$ is Zariski-dense in $T(F)$ if and only if S_3 is Zariski-dense in $T(F)$. We can divide S_3 in $n \in \mathbb{N}$ smaller sets. We let $\alpha = (\mu - \mu'^{-1})/n$ and take

$$S'_t = \{(x, y, z) \in S(F) \mid |x|^{\mu'^{-1} + t\alpha} \leq |y| \leq |x|^{\mu'^{-1} + (t+1)\alpha}\}.$$

We have

$$\bigcup_{t=0}^{n-1} S'_t = S_3.$$

We again apply Lemma 7.22 to see that $S(F)$ is Zariski-dense in $T(F)$ if and only if S'_t is Zariski-dense in $T(F)$ for some $t \in \{0, \dots, n-1\}$. Because $n \in \mathbb{N}$ can be as large as we want, we see that $S(F)$ is Zariski-dense in $T(F)$ if and only if there exists $\varepsilon > 0$ and $\lambda \in \mathbb{R}$ such that the set

$$\{(x, y, z) \in S(F) \mid |x|^{\lambda-\varepsilon} \le |y| \le |x|^{\lambda+\varepsilon}\}$$

is Zariski-dense in $T(F)$. So the elements in this last set satisfy $|y| \approx |x|^{\lambda}$. Because of this we only need to focus on integral solutions where $|y|$ and $|x|$ are roughly the same. In our second attempt to generalize Runge's Theorem 6.1, we do not want to have restrictions on the ratio between $|x|$ and $|y|$. If $f(X, Y)$ is a root of F, we then want $f(x, y)$ to converge in $\mathbb{C}$ for all $x, y \in \mathbb{C}$ such that $|x|$ and $|y|$ are bigger then some constant. For our second attempt we need to generalize Newton dots and degrees to a case with three variables. They are intuitively the same as the case with two variables and also satisfy similar rules.

Definitions 8.25. *Let*

$$F(X, Y, Z) := \sum_{k=0}^{d_3} \sum_{n=-m}^{\infty} \sum_{l=-p}^{\infty} a_{n,l,k} X^{-l/q} Y^{-n/e} Z^k,$$

for some $e, q \in \mathbb{N}$*,* $d_3 \in \mathbb{N}_0$ *and* $m, p \in \mathbb{Z}$ *and with* $a_{n,l,k} \in \mathbb{C}$*. Let* $\mu, \lambda \in \mathbb{R}_{>0}$*. We then define*

$$D(F) := \{(-l/q, -n/e, k) \mid a_{n,l,k} \neq 0\},$$

and

$$\deg_{\mu,\lambda}(F) := \max_{(i,j,k)\in D(F)} (i + j\mu + k\lambda)$$

and

$$D_{\mu,\lambda}(F) := \{(i, j, k) \in D(F) \mid i + j\mu + k\lambda = \deg_{\mu,\lambda}(F)\}$$

and

$$F_{\mu,\lambda} := \sum_{(i,j,k)\in D_{\mu,\lambda}(F)} a_{-ej,-qi,k} X^i Y^j Z^k.$$

Remark 8.26. *Let*

$$G(X, Y, Z) := \sum_{k=0}^{d_3} \sum_{n=-m}^{\infty} \sum_{l=-p}^{\infty} a_{n,l,k} X^{-l/q} Y^{-n/e} Z^k,$$

for some $e, q \in \mathbb{N}$*,* $d_3 \in \mathbb{N}_0$ *and* $m, p \in \mathbb{Z}$ *and with* $a_{n,l,k} \in \mathbb{C}$ *and let*

$$G(X, Y, Z) := \sum_{k=0}^{d_3} \sum_{n=-m}^{\infty} \sum_{l=-p}^{\infty} b_{n,l,k} X^{-l/q} Y^{-n/e} Z^k,$$

with $b_{n,l,k} \in \mathbb{C}$*. We then have the following for all* $\mu, \lambda \in \mathbb{R}_{>0}$*:*

- $\deg_{\mu,\lambda}(G + H) \le \max(\deg_{\mu,\lambda} G, \deg_{\mu,\lambda} H)$*.*
- *If* $\deg_{\mu,\lambda} G > \deg_{\mu,\lambda} H$*, then* $(G + H)_{\mu,\lambda} = G_{\mu,\lambda}$*.*
- $\deg_{\mu,\lambda}(GH) \le \deg_{\mu,\lambda} G + \deg_{\mu,\lambda} H$*.*
- $(GH)_{\mu,\lambda} = G_{\mu,\lambda} H_{\mu,\lambda}$*.*

We now start a second attempt to generalize Theorem 6.9. Note that one requirement of this theorem becomes redundant if Conjecture 8.15 holds.

Theorem 8.27. *Let $F(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ be a polynomial of positive Z-degree. We view F as a polynomial in Z over $\mathbb{Z}[X,Y]$. Let $f(X,Y) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ be a nonzero root of F. Suppose that we can write f as*

$$f(X,Y) = \sum_{n=-m}^{\infty} \sum_{l=-p}^{\infty} a_{n,l} X^{-l/q} Y^{-n/e}, \tag{8.14}$$

for some $m,p \in \mathbb{Z}$, and for some $e,q \in \mathbb{N}$ and where all coefficients $a_{n,l}$ lie in $\overline{\mathbb{Q}}$. Suppose that there exists $R' \in \mathbb{R}$ such that $f(x,y)$ converges for all $x,y \in \mathbb{C}$ with $|x| > R'$ and $|y| > R'$. Suppose that there exists $\varphi \in \mathbb{R}_{>0}$ and $\pi_1, \pi_2 \in \mathbb{R}$ such that

$$\pi_1 < -l/q + (n/e)\varphi < \pi_2 \qquad \forall n,l \in \mathbb{Z}, a_{n,l} \neq 0. \tag{8.15}$$

Then their exists nonzero $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ and $u \in \mathbb{C}$ such that the following properties hold for all $R \in \mathbb{R}$ with $R > |u|$.

- *$f(x,y)$ converges in $\mathbb{C}$ for all $x,y \in \mathbb{C}$ with $|x| > R$ and $|y| > R$.*
- *If $x,y \in \mathbb{Z}$ with $|x| > R$ and $|y| > R$, and if $f(x,y) \in \mathbb{Z}$, then we have $P(x,y,f(x,y)) = 0$.*
- *If the set $\{(x,y) \in \mathbb{Z}^2 \mid |x| > R, |y| > R, f(x,y) \in \mathbb{Z}\}$ is Zariski-dense in $\mathbb{C}^2$, then we have $P(X,Y,f(X,Y)) = 0$ in $\overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$.*
- *Let $\mu \in \mathbb{R}_{>0}$ and $\lambda \in \mathbb{R}_{>0}$ and $H = \deg_{\mu,\lambda}(f(X,Y))$. We consider three cases*
  - *If $H < \lambda$, we have that $P_{\mu,\lambda}(X,Y,Z)$ is a monomial in $\mathbb{Z}[X,Y,Z]$.*
  - *If $H > \lambda$, we have $P_{\mu,\lambda}(X,Y,Z) \in \mathbb{Z}[X,Y]$.*
  - *If $H = \lambda$, we have that $P_{\mu,\lambda}(X,Y,Z)$ is a monomial in $\mathbb{Z}[X,Y]$ times a power of an irreducible factor in $\mathbb{Z}[X,Y,Z]$.*

Proof. By Lemma 8.11 we see that all $a_{n,l}$ lie in some algebraic number field K. Let $s \in \mathbb{N}$ be the degree of this number field. Let $E \in \mathbb{N}$ be large enough, such that the linear inequalities

$$-1 + seq\pi_1 \leq i - j\varphi \leq 1 + seq\pi_2$$

and

$$i + j < -E$$

have no common solution $(i,j) \in Z^2$ with either $i > 0$ or $j > 0$. Because f is nonzero, we may assume m and p to be as small as possible, so there exists smallest possible $n', l' \in \mathbb{Z}$ with $n' \geq -m$ and $l' \geq -p$ such that $a_{n',-p} \neq 0$ and $a_{-m,l'} \neq 0$ both hold. Define the set

$$\begin{aligned} B := \{(i,j,k) \in \mathbb{Z}^3 \mid\ & 0 \leq k \leq seq, \\ & \min(i,j) < 0, \\ & i + j + (m/e + p/q)k \geq -E, \\ & -1 + (seq - k)\pi_1 < i - j\varphi < 1 + (seq - k)\pi_2\}. \end{aligned} \tag{8.16}$$

by combining these inequalities, one can show that B is a finite set. Let $M \in \mathbb{N}$ be the cardinality of B. For each $\beta \in \mathbb{N}_0$, we let

$$\overline{\beta} := \lfloor \beta\varphi \rfloor \in \mathbb{N}_0,$$

where $\lfloor . \rfloor$ is the floor function. As a consequence we have

$$|\overline{\beta} - \beta\varphi| < 1. \tag{8.17}$$

Let $T \in \mathbb{N}_0$ be large enough, such that

$$T \geq n'sq \tag{8.18}$$

and such that

$$\overline{T} = \lfloor T\varphi \rfloor \geq l'se \tag{8.19}$$

both hold. Let $N \in \mathbb{N}$ be large enough, such that

$$\lfloor (T+M)\varphi \rfloor + T + M + (seq-1)(p/q + m/e) + \max(m/e, p/q) - N < -E \tag{8.20}$$

and such that

$$N \geq \max(l'/q, n'/e) \tag{8.21}$$

both hold.

Since K is a finite field extension over $\mathbb{Q}$ of degree s, there exists a monic irreducible polynomial $G(X) \in \mathbb{Q}[X]$ of degree s which has a root $\theta_1 \in K$ that generates K. So $K = \mathbb{Q}(\theta_1)$. Let L be the splitting field of $G(X)$ over K. So there exists $\theta_2, \ldots, \theta_s \in L$ such that $G(X) = (X - \theta_1) \cdots (X - \theta_s)$ in $L[X]$. Any $c \in K$ can uniquely be written as $c_0 + c_1\theta_1 + \ldots + c_{s-1}\theta_1^{s-1}$ with $c_0, \ldots, c_{s-1} \in \mathbb{Q}$. We denote the polynomial $c_0 + c_1 W + \ldots + c_{s-1}W^{s-1}$ by $g_c(W)$. For $\sigma \in \{1, \ldots, s\}$, we call $c^{(\sigma)} := g_c(\theta_\sigma) = c_0 + c_1\theta_\sigma + \ldots + c_{s-1}\theta_\sigma^{s-1} \in L$ the σ-th conjugate of c in K. In particular we have $c^{(1)} = c$ and $0^{(\sigma)} = 0$.

Let $\zeta_1 \in \mathbb{C}$ be a primitive q-th root of unity. This means that $\zeta_1^q = 1$ and that the elements $\zeta_1^0, \ldots, \zeta_1^{q-1}$ are all different solutions to $Y^q = 1$. It follows that the polynomial $Y^q - X^{-1}$ in Y over $\mathbb{Q}[X, X^{-1}]$ factorizes over the field extension $\mathbb{Q}(\zeta_1)(X^{-1/q})$ as $Y^q - X^{-1} = (Y - \zeta_1^0 X^{-1/q}) \cdots (Y - \zeta_1^{q-1} X^{-1/q})$. Notice that X^{-1} is a unit in $\mathbb{Q}[X, X^{-1}]$. In a similar way we let ζ_2 be a primitive e-th root of unity. We now consider the following Laurent polynomial in V

$$Z - \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} g_{a_{n,l}}(W) V^l U^n \in \mathbb{Q}[X, X^{-1}][Y, Y^{-1}][Z][W][U, U^{-1}][V, V^{-1}].$$

By Corollary 5.12 we find that

$$\prod_{\mathscr{E}_1=0}^{q-1} \left(Z - \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} g_{a_{n,l}}(W) (\zeta_1^{\mathscr{E}_1} X^{-1/q})^l U^n \right) \in \mathbb{Q}[X, X^{-1}][Y, Y^{-1}][Z][W][U, U^{-1}].$$

We apply Corollary 5.12 again and find that

$$\prod_{\mathscr{E}_2=0}^{e-1} \prod_{\mathscr{E}_1=0}^{q-1} \left(Z - \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} g_{a_{n,l}}(W) (\zeta_1^{\mathscr{E}_1} X^{-1/q})^l (\zeta_2^{\mathscr{E}_2} Y^{-1/e})^n \right) \in \mathbb{Q}[X, X^{-1}][Y, Y^{-1}][Z][W].$$

We now apply Corollary 5.11 we find that

$$\prod_{\sigma=1}^{s} \prod_{\mathscr{E}_2=0}^{e-1} \prod_{\mathscr{E}_1=0}^{q-1} \left(Z - \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} g_{a_{n,l}}(\theta_\sigma) (\zeta_1^{\mathscr{E}_1} X^{-1/q})^l (\zeta_2^{\mathscr{E}_2} Y^{-1/e})^n \right) \in \mathbb{Q}[X, X^{-1}][Y, Y^{-1}][Z]. \tag{8.22}$$

For each $\beta \in \mathbb{N}_0$ we then define

$$F(X, Y, Z; \beta) := X^{\overline{\beta}} Y^{\beta} \prod_{\sigma=1}^{s} \prod_{\mathscr{E}_2=0}^{e-1} \prod_{\mathscr{E}_1=0}^{q-1} \left(Z - \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} a_{n,l}^{(\sigma)} (\zeta_1^{\mathscr{E}_1} X^{-1/q})^l (\zeta_2^{\mathscr{E}_2} Y^{-1/e})^n \right). \tag{8.23}$$

From (8.22) we then find

$$F(X, Y, Z; \beta) \in \mathbb{Q}[X, X^{-1}][Y, Y^{-1}][Z]. \tag{8.24}$$

We can write

$$F(X, Y, Z; \beta) = \sum_{(i,j,k) \in D(F(X,Y,Z;\beta))} b_{\beta,i,j,k} X^i Y^j Z^k,$$

for suitable $b_{\beta,i,j,k} \in \mathbb{Q}$. We see from (8.23) and (8.24) that if $(i, j, k) \in D(F(X, Y, Z; \beta))$, that $i, j, k \in \mathbb{Z}$ and $0 \leq k \leq seq$ must hold. Note that we may assume that $\pi_1 \leq 0$ as well as $\pi_2 \geq 0$ hold. For the case where $\beta = 0$, we can use (8.15) and (8.23) to find that

$$(seq - k)\pi_1 \leq i - j\varphi \leq (seq - k)\pi_2 \qquad \forall (i, j, k) \in D(F(X, Y, Z; 0)). \tag{8.25}$$

For any $\beta \in \mathbb{N}_0$ we have

$$\begin{aligned} F(X, Y, Z; \beta) &= X^{\overline{\beta}} Y^{\beta} F(X, Y, Z; 0) \\ &= X^{\overline{\beta}} Y^{\beta} \sum_{(i,j,k) \in D(F(X,Y,Z;0))} b_{0,i,j,k} X^i Y^j Z^k \\ &= \sum_{(i,j,k) \in D(F(X,Y,Z;0))} b_{0,i,j,k} X^{i+\overline{\beta}} Y^{j+\beta} Z^k. \end{aligned}$$

From this we can see that $(i,j,k) \in D(F(X,Y,Z;\beta))$ holds if and only if $(i-\overline{\beta}, j-\beta, k) \in D(F(X,Y,Z;0))$. We combine this with (8.25) and (8.17) and then find for every $\beta \in \mathbb{N}_0$ and every $(i,j,k) \in D(F(X,Y,Z;\beta))$ that

$$\begin{aligned} -1 + (seq - k)\pi_1 &< \overline{\beta} - \beta\varphi + (i - \overline{\beta}) - (j - \beta)\varphi \\ &= i - j\varphi \\ &= \overline{\beta} - \beta\varphi + (i - \overline{\beta}) - (j - \beta)\varphi \\ &< 1 + (seq - k)\pi_2. \end{aligned} \tag{8.26}$$

So $i - j\varphi$ is bounded above and below for all $(i,j,k) \in D(F(X,Y,Z;\beta))$. For any $\beta \in \{T, \dots, T+M\}$ we will divide the terms $b_{\beta,i,j,k}X^iY^jZ^k$ of $F(X,Y,Z;\beta)$ into the following three categories.

- Terms $b_{\beta,i,j,k}X^iY^jZ^k$ with $i, j \geq 0$.
- Terms $b_{\beta,i,j,k}X^iY^jZ^k$ with $i < 0$, or $j < 0$ and with $i + j + (m/e + p/q)k < -E$.
- Terms $b_{\beta,i,j,k}X^iY^jZ^k$ with $i < 0$, or $j < 0$ and with $i + j + (m/e + p/q)k \geq -E$.

Note that the terms in the first category lie in $\mathbb{Q}[X,Y,Z]$. We see from (8.26) that the terms in the third category satisfy $(i,j,k) \in B$. The terms from the first and second category does not satisfy $(i,j,k) \in B$. For each $\beta \in \{T, \dots, T+M\}$ we can add terms of the same category together and get

$$F(X,Y,Z;\beta) = P(X,Y,Z;\beta) + S(X,Y,Z;\beta) + \sum_{(i,j,k)\in B} b_{\beta,i,j,k}X^iY^jZ^k,$$

with $P(X,Y,Z;\beta) \in \mathbb{Q}[X,Y,Z]$ and $S(X,Y,Z;\beta) \in \mathbb{Q}[X,X^{-1}][Y,Y^{-1}][Z]$. For the elements $(i,j,k) \in B$ for which $(i,j,k) \notin D(F(X,Y,Z;\beta))$ we have defined $b_{\beta,i,j,k} = 0$.

We want to find $c_T, \dots, c_{M+T} \in \mathbb{Z}$, not all zero, such that $\sum_{\beta=T}^{M+T} c_\beta b_{\beta,i,j,k} = 0$ for all $(i,j,k) \in B$. This is a system of $|B| = M$ linear equations in $|\{T, \dots, M+T\}| = M+1$ variables. It therefore indeed has a solution. We may scale these $c_T, \dots, c_{T+M}$ by any nonzero integer and still find a suitable solution. We scale by a nonzero integer such that $c_\beta \cdot P(X,Y,Z;\beta) \in \mathbb{Z}[X,Y,Z]$ for all $\beta \in \{T, \dots, T+M\}$. This gives us

$$\sum_{\beta=T}^{T+M} c_\beta F(X,Y,Z;\beta) = \sum_{\beta=T}^{T+M} c_\beta P(X,Y,Z;\beta) + \sum_{\beta=T}^{T+M} c_\beta S(X,Y,Z;\beta),$$

We define

$$P = P(X,Y,Z) := \sum_{\beta=T}^{T+M} c_\beta P(X,Y,Z;\beta) = \sum_{\beta=T}^{T+M} c_\beta F(X,Y,Z;\beta) - \sum_{\beta=T}^{T+M} c_\beta S(X,Y,Z;\beta) \tag{8.27}$$

and see that $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ holds.

We are interested in the terms of $P(X,Y,f(X,Y)) \in K(\{(X^{-1})^*\})(((Y^{-1})^*))$. First we look at $D(f(X,Y))$. If $(i,j) \in D(f(X,Y))$, we see by (8.14) that $i \leq p/q$ and $j \leq m/e$, hence $i + j \leq p/q + m/e$. We have the same boundaries for any

$$(i,j) \in D\left(-\sum_{n=-m}^{eN}\sum_{l=-p}^{qN} a_{n,l}^{(\sigma)}(\zeta_1^{\mathscr{E}_1}X^{-1/q})^l(\zeta_2^{\mathscr{E}_2}Y^{-1/e})^n\right)$$

for all $\mathscr{E}_1 \in \{0, \dots, q-1\}$, $\mathscr{E}_2 \in \{0, \dots, e-1\}$ and $\sigma \in \{1, \dots, s\}$. We can add these Puiseux series together and thus keep the same boundaries for any

$$(i,j) \in D\left(f(X,Y) - \sum_{n=-m}^{eN}\sum_{l=-p}^{qN} a_{n,l}^{(\sigma)}(\zeta_1^{\mathscr{E}_1}X^{-1/q})^l(\zeta_2^{\mathscr{E}_2}Y^{-1/e})^n\right)$$

for all $\mathscr{E}_1 \in \{0, \ldots, q-1\}$, $\mathscr{E}_2 \in \{0, \ldots, e-1\}$ and $\sigma \in \{1, \ldots, s\}$. In the case with $\mathscr{E}_1 = \mathscr{E}_2 = 0$ and $\sigma = 1$ we find

$$\begin{aligned} f(X,Y) - \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} a_{n,l}^{(\sigma)} (\zeta_1^{\mathscr{E}_1} X^{-1/q})^l (\zeta_2^{\mathscr{E}_2} Y^{-1/e})^n \\ = \sum_{n=-m}^{\infty} \sum_{l=-p}^{\infty} a_{n,l} X^{-l/q} Y^{-n/e} - \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} a_{n,l} X^{-l/q} Y^{-n/e}. \end{aligned}$$

From this we can observe that its Newton dots (i,j) all satisfy either $i < -N$ or $j < -N$. Together with $i \le p/q$ and $j \le m/e$, this tells us that $i + j < \max(m/e, p/q) - N$. So for all $\beta \in \{T, \ldots, T+M\}$ we see from (8.23) that the Newton dots (i,j) in $D(F(X,Y,f(X,Y);\beta))$ satisfy the following inequality:

$$\begin{aligned} i + j &< \overline{\beta} + \beta + (seq - 1)(p/q + m/e) + \max(m/e, p/q) - N \\ &\le \lfloor (T+M)\varphi \rfloor + T + M + (seq-1)(p/q + m/e) + \max(m/e, p/q) - N \\ &< -E. \end{aligned}$$

So we have

$$i + j < -E \qquad \forall (i,j) \in D(F(X,Y,f(X,Y);\beta)) \tag{8.28}$$

for every $\beta \in \{T, \ldots, T+M\}$.

We will now look at $S(X,Y,f(X,Y);\beta)$ for each $\beta \in \{T, \ldots, T+M\}$. We saw that $S(X,Y,Z;\beta)$ only consists of terms of the form $b_{\beta,i,j,k} X^i Y^j Z^k$, where $i + j + (m/e + p/q)k < -E$. We also saw that any $(i,j) \in D(f(X,Y))$ satisfies $i + j \le p/q + m/e$. Let $(i,j,k) \in D(S(X,Y,Z;\beta))$. We then find that every $(i', j') \in D(b_{\beta,i,j,k} X^i Y^j (f(X,Y))^k)$ also satisfies $i' + j' \le i + j + (m/e + p/q)k < -E$. This combined with (8.27) and (8.28) gives

$$i + j < -E \qquad \forall (i,j) \in D(P(X,Y,f(X,Y))). \tag{8.29}$$

Now let $(i,j,k) \in D(P(X,Y,Z;\beta)) \subset D(F(X,Y,Z;\beta))$. From (8.26) we have

$$-1 + (seq - k)\pi_1 \le i - j\varphi \le 1 + (seq - k)\pi_2.$$

This combined with (8.15) shows that

$$-1 + seq\pi_1 \le i - j\varphi \le 1 + seq\pi_2 \qquad \forall (i,j) \in D(P(X,Y,f(X,Y);\beta))$$

holds for any $\beta \in \{T, \ldots, T+M\}$. And therefore we have

$$-1 + seq\pi_1 \le i - j\varphi \le 1 + seq\pi_2 \qquad \forall (i,j) \in D(P(X,Y,f(X,Y))). \tag{8.30}$$

We compare (8.29) and (8.30) to the definition of E to find

$$i \le 0 \text{ and } j \le 0 \qquad \forall (i,j) \in D(P(X,Y,f(X,Y))). \tag{8.31}$$

There exists $R' \in \mathbb{R}$ such that $P(x,y,f(x,y))$ converges for all $x, y \in \mathbb{C}$ with $|x| > R'$ and $|y| > R'$. Because $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$, we have that $P(X,Y,f(X,Y))$ also converges for such x and y. We may assume that $R' > 1$ holds. We can write

$$P(X,Y,f(X,Y)) = \sum_{(i,j) \in D(P(X,Y,f(X,Y)))} \omega_{i,j} X^i Y^j$$

with $\omega_{i,j} \in \mathbb{C}$. Let $u' \in \mathbb{C}$ satisfy $|u'| > R'$. We then have

$$\sum_{(i,j) \in D(P(X,Y,f(X,Y)))} |\omega_{i,j}| |u'|^i |u'|^j < \psi$$

for some $\psi \in \mathbb{R}$. Let $k \in \mathbb{N}$. We then find

$$\begin{aligned}\sum_{(i,j)\in D(P(X,Y,f(X,Y)))} |\omega_{i,j}||ku'|^i|ku'|^j &= \sum_{(i,j)\in D(P(X,Y,f(X,Y)))} |\omega_{i,j}||u'|^i|u'|^j|k|^{i+j} \\ &\leq \sum_{(i,j)\in D(P(X,Y,f(X,Y)))} |\omega_{i,j}||u'|^i|u'|^j|k|^{-E} \\ &= k^{-E} \sum_{(i,j)\in D(P(X,Y,f(X,Y)))} |\omega_{i,j}||u'|^i|u'|^j \\ &< k^{-E}\psi.\end{aligned}$$

We will now choose $k \in \mathbb{N}$ large enough such that $k^{-E}\psi < 1/2$. So for $u := ku'$ we have

$$\sum_{(i,j)\in D(P(X,Y,f(X,Y)))} |\omega_{i,j}||u|^{i+j} < 1/2.$$

Now let $R \in \mathbb{R}$ such that $R > |u| > R'$. Let $x, y \in \mathbb{C}$ satisfy $|x|, |y| > R$. Let $(i,j) \in D(P(X,Y,f(X,Y)))$. We then have $i \leq 0$ and $j \leq 0$ by (8.31). This shows us that

$$\begin{aligned}\sum_{(i,j)\in D(P(X,Y,f(X,Y)))} |\omega_{i,j}||x|^i|y|^j &\leq \sum_{(i,j)\in D(P(X,Y,f(X,Y)))} |\omega_{i,j}||u|^{i+j} \\ &< 1/2.\end{aligned}$$

Therefore we have $|P(x,y,f(x,y))| < 1/2$ for all $x, y \in \mathbb{C}$ with $|x| > R$ and $|y| > R$. If additionally we have $x, y \in \mathbb{Z}$ and $f(x,y) \in \mathbb{Z}$ we must have $P(x,y,f(x,y)) \in \mathbb{Z}$ and therefore $P(x,y,f(x,y)) = 0$. If the set

$$\{(x,y) \in \mathbb{Z}^2 \mid |x| > R, |y| > R, f(x,y) \in \mathbb{Z}\}$$

is Zariski-dense in $\mathbb{C}^2$, then $P(X,Y,f(X,Y)) = 0$ in $\overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$. The proof of this is similar to the corresponding proof in Theorem 8.19.

Now let $\mu, \lambda \in \mathbb{R}_{>0}$. Let

$$h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2) := \sum_{n=-m}^{eN} \sum_{l=-p}^{qN} a_{n,l}^{(\sigma)} (\zeta_1^{\mathscr{E}_1} X^{-1/q})^l (\zeta_2^{\mathscr{E}_2} Y^{-1/e})^n \tag{8.32}$$

for any $\sigma \in \{1,\ldots,s\}$, $\mathscr{E}_2 \in \{0,\ldots,e-1\}$ and $\mathscr{E}_1 \in \{0,\ldots,q-1\}$. Note that $\deg_{\mu,\lambda}(h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))$ and $D_{\mu,\lambda}(h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))$ are independent on σ, $\mathscr{E}_2$ and $\mathscr{E}_1$. We have $a_{n',-p} \neq 0$. We then see from (8.21) that $(p/q, -n'/e, 0) \in D(h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))$. This Newton dot has the property that its first entry p/q is the largest possible entry among all Newton dots in $D(h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))$, and that its second entry $-n'/e$ is the largest among all these Newton dots where the first entry is p/q. Also, its third entry is always zero, because $h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2)$ does not depend on Z. We use (8.23) and then deduce for any $k \in \{0,\ldots,seq\}$ and $\beta \in \{T,\ldots,T+M\}$ that

$$\begin{aligned}(\overline{\beta} + (seq-k)p/q, \beta + (seq-k)(-n'/e), k) &\in D(X^{\overline{\beta}} Y^{\beta} \prod_{\sigma=1}^{s} \prod_{\mathscr{E}_2=0}^{e-1} \prod_{\mathscr{E}_1=0}^{q-1} (Z - h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))) \\ &= D(F(X,Y,Z;\beta))).\end{aligned}$$

If $-n'/e \geq 0$, we directly have $\beta + (seq-k)(-n'/e) \geq 0$. Otherwise we use (8.18) and find

$$\beta + (seq-k)(-n'/e) \geq T + (seq)(-n'/e) \geq 0 \tag{8.33}$$

for all $k \in \{0,\ldots,seq\}$ and $\beta \in \{T,\ldots,T+M\}$. Now let $(i,j,k) \in D_{\mu,\lambda}(F(X,Y,Z;\beta))$. Suppose that $j < 0$. We then use (8.33) to find

$$\begin{aligned}\deg_{\mu,\lambda}(F(X,Y,Z;\beta)) &= i + j\mu + k\lambda \\ &< \overline{\beta} + (seq-k)p/q + (\beta + (seq-k)(-n'/e))\mu + k\lambda \\ &\leq \deg_{\mu,\lambda}(F(X,Y,Z;\beta)).\end{aligned}$$

This leads to a contradiction from which we may deduce that $j \geq 0$. Similar reasoning, where we use $a_{-l'/q,m/e} \neq 0$ and (8.19), yields $i \geq 0$. As a consequence, we find by definition of $P(X,Y,Z;\beta)$ that $(i,j,k) \in D_{\mu,\lambda}(P(X,Y,Z;\beta))$. Because $D(P(X,Y,Z;\beta)) \subset D(F(X,Y,Z;\beta))$ holds, we then find that

$$F_{\mu,\lambda}(X,Y,Z;\beta) = P_{\mu,\lambda}(X,Y,Z;\beta) \tag{8.34}$$

for all $\beta \in \{T,\ldots,T+M\}$.

We saw that $\deg_{\mu,\lambda}(h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))$ was independent on σ, $\mathscr{E}_1$ and $\mathscr{E}_2$. Let $(n,l) \in \mathbb{Z}^2$ such that $(-l/q,-n/e) \in D(f(X,Y))$. We then have $a_{n,l} \neq 0$. Suppose that $(-l/q,-n/e) \notin D(h(X,Y;1;0;0))$. We then see from (8.14) and (8.32) that $n > eN$ or $l > qN$ must hold. We also have $n \geq -m$ and $l \geq -p$. Suppose that $n > eN$. From (8.21) and $a_{n',-p} \neq 0$ we then have

$$\begin{aligned}-l/q - n/e\mu &< p/q - N\mu \\ &\leq p/q - n'/e\mu \\ &\leq \deg_{\mu,\lambda}(f(X,Y)).\end{aligned}$$

From this we see that $(-l/q - n/e) \notin D_{\mu,\lambda}(f(X,Y))$. By similar reasoning we reach the same conclusion if $l > qN$. As a consequence we find

$$H = \deg_{\mu,\lambda}(f(X,Y)) = \deg_{\mu,\lambda}(h(X,Y;1;0;0)) = \deg_{\mu,\lambda}(h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))$$

for all $\sigma \in \{1,\ldots,s\}$, $\mathscr{E}_2 \in \{0,\ldots,e-1\}$ and $\mathscr{E}_1 \in \{0,\ldots,q-1\}$.

Now let $\gamma \in \{T,\ldots,T+M\}$ be the largest such that $c_\gamma \neq 0$. We consider three cases: First suppose that

$$H < \lambda.$$

We then have

$$(Z - h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))_{\mu,\lambda} = Z$$

for any $\sigma \in \{1,\ldots,s\}$, $\mathscr{E}_2 \in \{0,\ldots,e-1\}$ and $\mathscr{E}_1 \in \{0,\ldots,q-1\}$. We therefore see from (8.23) and (8.34) that

$$P_{\mu,\lambda}(X,Y,Z;\beta) = F_{\mu,\lambda}(X,Y,Z;\beta) = X^{\overline{\beta}}Y^{\beta}Z^{seq},$$

hence

$$P_{\mu,\lambda}(X,Y,Z) = c_\gamma X^{\overline{\gamma}}Y^{\gamma}Z^{seq}.$$

From this result we can see that $P_{\mu,\lambda}(X,Y,Z)$ is nonzero. Because the case $H < \lambda$ is always possible by taking λ large enough, we see from this that P is nonzero. Now suppose that

$$H > \lambda.$$

We then have

$$(Z - h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))_{\mu,\lambda} = -h_{\mu,\lambda}(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2)$$

for any $\sigma \in \{1,\ldots,s\}$, $\mathscr{E}_2 \in \{0,\ldots,e-1\}$ and $\mathscr{E}_1 \in \{0,\ldots,q-1\}$. We therefore see from (8.23) and (8.34) that

$$P_{\mu,\lambda}(X,Y,Z;\beta) = F_{\mu,\lambda}(X,Y,Z;\beta) = X^{\overline{\beta}}Y^{\beta}\prod_{\sigma=1}^{s}\prod_{\mathscr{E}_2=0}^{e-1}\prod_{\mathscr{E}_1=0}^{q-1} -h_{\mu,\lambda}(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2).$$

This combined with $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ gives

$$P_{\mu,\lambda}(X,Y,Z) \in \mathbb{Z}[X,Y].$$

Now suppose that

$$H = \lambda.$$

We then have

$$(Z - h(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2))_{\mu,\lambda} = Z - h_{\mu,\lambda}(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2)$$

for any $\sigma \in \{1, \ldots, s\}$, $\mathscr{E}_2 \in \{0, \ldots, e-1\}$ and $\mathscr{E}_1 \in \{0, \ldots, q-1\}$. We therefore see from (8.23) and (8.34) that

$$P_{\mu,\lambda}(X,Y,Z;\beta) = F_{\mu,\lambda}(X,Y,Z;\beta) = X^{\overline{\beta}}Y^{\beta} \prod_{\sigma=1}^{s} \prod_{\mathscr{E}_2=0}^{e-1} \prod_{\mathscr{E}_1=0}^{q-1} (Z - h_{\mu,\lambda}(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2)),$$

hence

$$P_{\mu,\lambda}(X,Y,Z) = c_{\gamma}X^{\overline{\gamma}}Y^{\gamma} \prod_{\sigma=1}^{s} \prod_{\mathscr{E}_2=0}^{e-1} \prod_{\mathscr{E}_1=0}^{q-1} (Z - h_{\mu,\lambda}(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2)) \in \mathbb{Z}[X,Y,Z]. \tag{8.35}$$

Because P is nonzero, so is $P_{\mu,\lambda}(X,Y,Z)$. We see that $h_{\mu,\lambda}(X,Y;1;0;0)$ is a root of $P_{\mu,\lambda}(X,Y,Z)$, a nonzero polynomial in $\mathbb{Z}[X,Y,Z]$. It therefore is also a root of an irreducible factor $G(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ of $P_{\mu,\lambda}(X,Y,Z)$. So

$$\begin{aligned} G(X,Y,h_{\mu,\lambda}(X,Y;1;0;0)) &= G(X,Y, \sum_{(-l/q,-n/e)\in D_{\mu,\lambda}(h(X,Y;1;0;0))} a_{n,l}^{(1)}(\zeta_1^0 X^{-1/q})^l(\zeta_2^0 Y^{-1/e})^n) \\ &= 0. \end{aligned}$$

We substitute $X^{-1/q}$ by X' and $Y^{-1/e}$ by Y'. This yields

$$G(X'^{-q}, Y'^{-e}, \sum_{(-l/q,-n/e)\in D_{\mu,\lambda}(h(X,Y;1;0;0))} a_{n,l}^{(1)}(\zeta_1^0 X')^l(\zeta_2^0 Y')^n) = 0. \tag{8.36}$$

When we view

$$G(X'^{-q}, Y'^{-e}, \sum_{(-l/q,-n/e)\in D_{\mu,\lambda}(h(X,Y;1;0;0))} g_{a_{n,l}}(W)(\zeta_1^0 X')^l(\zeta_2^0 Y')^n) \in \mathbb{Q}[W,X',Y'].$$

as polynomial in X', Y' with coefficients in W, we see from (8.36) that each coefficient, as a polynomial in W over $\mathbb{Q}$, has θ_1 as a root. As a consequence these coefficient must also have θ_σ as root for all $\sigma \in \{1, \ldots, s\}$. So we have

$$G(X'^{-q}, Y'^{-e}, \sum_{(-l/q,-n/e)\in D_{\mu,\lambda}(h(X,Y;1;0;0))} a_{n,l}^{(\sigma)}(\zeta_1^0 X')^l(\zeta_2^0 Y')^n) = 0. \tag{8.37}$$

for all $\sigma \in \{1, \ldots, s\}$. We now let $E_1 \in \{0, \ldots, q-1\}$ and $E_2\{0, \ldots, e-1\}$ and substitute X' by $\zeta_1^{\mathscr{E}_1}X''$ and Y' by $\zeta_2^{\mathscr{E}_2}Y''$. We then find the following from (8.36):

$$\begin{aligned} &G(X''^{-q}, Y''^{-e}, \sum_{(-l/q,-n/e)\in D_{\mu,\lambda}(h(X,Y;1;0;0))} a_{n,l}^{(\sigma)}(\zeta_1^{\mathscr{E}_1}X'')^l(\zeta_2^{\mathscr{E}_2}Y'')^n) \\ =&G((\zeta_1^q)^{-\mathscr{E}_1}X''^{-q}, (\zeta_2^e)^{-\mathscr{E}_2}Y''^{-e}, \sum_{(-l/q,-n/e)\in D_{\mu,\lambda}(h(X,Y;1;0;0))} a_{n,l}^{(\sigma)}(\zeta_1^{\mathscr{E}_1}X'')^l(\zeta_2^{\mathscr{E}_2}Y'')^n) \\ =&G((\zeta_1^{\mathscr{E}_1}X'')^{-q}, (\zeta_2^{\mathscr{E}_2}Y'')^{-e}, \sum_{(-l/q,-n/e)\in D_{\mu,\lambda}(h(X,Y;1;0;0))} a_{n,l}^{(\sigma)}(\zeta_1^{\mathscr{E}_1}X'')^l(\zeta_2^{\mathscr{E}_2}Y'')^n) \\ =&0. \end{aligned}$$

From this we find that

$$G(X,Y,h_{\mu,\lambda}(X,Y;\sigma;\mathscr{E}_1;\mathscr{E}_2)) = 0$$

for all $\sigma \in \{1, \ldots, s\}$, $\mathscr{E}_2 \in \{0, \ldots, e-1\}$ and $\mathscr{E}_1 \in \{0, \ldots, q-1\}$. So every root of $P_{\mu,\lambda}(X,Y,Z)$, when viewed as a polynomial in Z, is also a root of G. We therefore find that $P_{\mu,\lambda}(X,Y,Z)$ is an element in $\mathbb{Z}[X,Y]$ times a power of the irreducible polynomial $G(X,Y,Z)$. To find this element in $\mathbb{Z}[X,Y]$ we can compare the leading coefficients of $P_{\mu,\lambda}$ and $G(X,Y,Z)$, when we view them as polynomials in $\mathbb{Z}$. The leading coefficient of $P_{\mu,\lambda}$ is $c_\gamma X^{\overline{\gamma}}Y^\gamma$, as we can see by (8.35). Let $b(X,Y) \in \mathbb{Z}[X,Y]$ be the leading coefficient of G. Since G is a factor of $P_{\mu,\lambda}$, we also have that $b(X,Y)$ is a factor of $c_\gamma X^{\overline{\gamma}}Y^\gamma$. From this we conclude that $P_{\mu,\lambda}$ is indeed a monomial in $\mathbb{Z}[X,Y]$ times a power of an irreducible element in $\mathbb{Z}[X,Y,Z]$. □

We can now start with our second attempt to generalize Runge's Theorem 6.1.

Theorem 8.28. *Let $F(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ be an irreducible polynomial of positive Z-degree d_3. Let*

$$f_1(X,Y),\ldots,f_{d_3}(X,Y) \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$$

be the roots of F. Suppose that for each $t \in \{1,\ldots,d_3\}$, there exists $\varphi_t \in \mathbb{R}_{>0}$ and $\pi_{t,1},\pi_{t,2} \in \mathbb{R}$ such that

$$\pi_{t,1} < i - j\varphi_t < \pi_{t,2} \qquad \forall (i,j) \in D(f_t(X,Y)).$$

Suppose for each $t \in \{1,\ldots,d_3\}$ that there exists $R_t \in \mathbb{R}$ such that $f_t(x,y)$ converges for all $x,y \in \mathbb{C}$ with $|x| > R_t$ and $|y| > R_t$. If $S(F)$ is Zariski-dense in $T(F)$, then there exists $H_\mu \in \mathbb{R}$ for any $\mu \in \mathbb{R}_{>0}$ such that the following holds for all $\lambda \in \mathbb{R}_{>0}$:

- *If $H_\mu < \lambda$, we have that $F_{\mu,\lambda}(X,Y,Z)$ is a monomial in $\mathbb{Z}[X,Y,Z]$.*
- *If $H_\mu > \lambda$, we have $F_{\mu,\lambda}(X,Y,Z) \in \mathbb{Z}[X,Y]$.*
- *If $H_\mu = \lambda$, we have that $F_{\mu,\lambda}(X,Y,Z)$ is a monomial in $\mathbb{Z}[X,Y]$ times a power of an irreducible factor in $\mathbb{Z}[X,Y,Z]$.*

Proof. If $F(X,Y,Z) = \pm Z$, we can take $H_\mu = 0$ for each μ and this problem then becomes trivial. We therefore will ignore this case and as a consequence find that the roots $f_t(X,Y)$ are all nonzero. We take $R' \in \mathbb{R}$ such that $R' \geq \max(R_1,\ldots,R_{d_3},1)$. We can write

$$F(X,Y,Z) = G(X,Y)\prod_{t=1}^{d_3}(Z - f_t(X,Y)) \tag{8.38}$$

where $G(X,Y) \in \mathbb{Z}[X,Y]$ is the leading coefficient of F, when viewed as a polynomial in Z over $\mathbb{Z}[X,Y,Z]$. For each $t \in \{1,\ldots,d_2\}$ we define

$$S_t \coloneqq \{(x,y,z) \in \mathbb{Z}^3 \mid |x| > R', |y| > R', z = f_t(x,y)\}.$$

We also define the sets

$$S' \coloneqq \{(x,y,z) \in S(F) \mid G(x,y) = 0\}$$

and

$$S_X \coloneqq \{(x,y,z) \in S(F) \mid |x| \leq R'\}$$

and

$$S_Y \coloneqq \{(x,y,z) \in S(F) \mid |y| \leq R'\}.$$

We see from (8.38) that we have

$$S(F) = S_1 \cup \ldots \cup S_{d_3} \cup S' \cup S_X \cup S_Y.$$

From Lemma 7.22 we see that one of these subsets must be Zariski-dense in $T(F)$. We will show that S', S_X and S_Y are not Zariski-dense in $T(F)$. Because $G(X,Y)$ is a nonzero polynomial in $\mathbb{Z}[X,Y,Z]$ of zero Z-degree it is certainly no multiple of F. By definition of S', we see that G vanishes on all points in S'. We then use Corollary 7.17 to see that S' is not Zariski-dense in $T(F)$. Now let

$$g(X,Y,Z) \coloneqq \prod_{a \in \{-R',\ldots,R'-1,R'\}} (X-a) \in \mathbb{Z}[X,Y,Z].$$

We directly see that $g(x,y,z) = 0$ for all $|x| \leq R'$. We apply Corollary 7.17 again and see that S_X is not Zariski-dense in $T(F)$. The same holds for S_Y by similar reasoning. So S_t must be Zariski-dense for some $t \in \{1,\ldots,d_3\}$. We fix this t and take $f \coloneqq f_t$. We can now apply Theorem 8.27 and find the existence of nonzero $P(X,Y,Z) \in \mathbb{Z}[X,Y,Z]$ and $u \in \mathbb{C}$ such the properties from Theorem 8.27 hold for all $R \in \mathbb{R}$ with $R > |u|$. Because R and R' are independent of eachother and because both are only bounded from below, we may assume that $R = R'$. We now want to show that the set $\overline{S} \coloneqq \{(x,y) \in \mathbb{Z}^2 \mid |x| > R, |y| > R, f(x,y) \in \mathbb{Z}\}$ is Zariski-dense in $C^2 = T(0)$. Suppose that this is not the case. By Lemma 7.13 we then see that there must exist $G'(X,Y) \in \mathbb{C}[X,Y]$ such that G' vanishes on $\overline{S}$, but not on C^2. Now define

$$\overline{G}(X,Y,Z) = G'(X,Y) \in \mathbb{Z}[X,Y,Z].$$

Let $(x,y,z) \in S_t$. Because $f(x,y) = z \in \mathbb{Z}$, we get $(x,y) \in \overline{S}$. This shows that $\overline{G}(x,y,z) = G'(x,y) = 0$. So $\overline{G}$ vanishes on S_t. Because S_t is Zariski-dense in $T(F)$, we see by Lemma 7.13 that $\overline{G}$ must also vanish on $T(F)$. Hilbert's Nullstellensatz 7.15 then tells us that some power of G must have F as a factor. This is not possible as F is of positive Z-degree while $\overline{G}$ is of zero Z-degree. We therefore see that $\overline{S}$ is indeed Zariski-dense in C^2. We therefore see from Theorem 8.27 that $P(X,Y,f(X,Y)) = 0$. So $f(X,Y)$ is a root of P, when viewed as a polynomial in Z. Because $f(X,Y)$ is also a root of the irreducible polynomial $F(X,Y,Z)$, we may conclude that F is a factor of P. By Remark 8.26 we then see for any $\mu, \lambda \in \mathbb{R}_{>0}$ that $F_{\mu,\lambda}$ is also a factor of $P_{\mu,\lambda}$. Because $\mathbb{Z}[X,Y,Z]$ is a unique factorization domain, we can use Theorem 8.27 and take $H_\mu := \deg_{\mu,\lambda}(f(X,Y))$ to find the desired results. Note that H_μ does not depend on λ since $f(X,Y)$ is of zero Z-degree. □

Remark 8.29. *In the previous theorem we required the existence of $R_t \in \mathbb{R}$ for each $t \in \{1,\ldots,d_3\}$, such that $f_t(x,y)$ converges for all $x,y \in \mathbb{C}$ with $|x| > R_t$ and $|y| > R_t$. This requirement becomes redundant if we assume that Conjecture 8.15 holds. This is because of the following: Let $t \in \{1,\ldots,d_3\}$. Let $(i,j) \in D(f_t(X,Y))$. From (8.1) we see that j is bounded from above. The inequality $i < \pi_{t,2} + j\varphi_t$ then shows that i is also bounded from above. We therefore can write $f_t(X,Y)$ as in (8.2). We then can apply Conjecture 8.15 to see that there exists $R_t \in \mathbb{R}$ such that $f_t(x,y)$ converges in $\mathbb{C}$ for all $x,y \in \mathbb{C}$ with $|x| > R_t$ and $|y| > R_t$.*

We can apply Theorem 8.28 to the following examples:

Example 8.30. *Let*

$$F(X,Y,Z) = Z^2 - (X^2Y + 1)(X+Y)^2Y.$$

Then $S(F)$ is not Zariski-dense in $T(F)$.

Proof. We assume that $S(F)$ is Zariski-dense in $T(F)$ and will reach a contradiction. We view F as a polynomial in Z and see that F is a quadratic polynomial whose coefficients 1, 0 and $-(X^2Y+1)(X+Y)^2Y$ do not share a common irreducible factor in $\mathbb{Z}[X,Y]$. Because $(X^2Y+1)(X+Y)^2Y$ is not a square in $\mathbb{Z}[X,Y]$, we see that F is irreducible in $\mathbb{Z}[X,Y,Z]$. Let $g(V) \in \overline{\mathbb{Q}}(\{(V^{-1})^*\})$ be any of the two roots for W in the polynomial $W^2 - 1 - V \in \mathbb{Z}[V,W]$. So $(g(V))^2 = 1 + V$. Because $g(V)$ is a convergent Puiseux series, there exists $R \in \mathbb{R}$ such that $g(v)$ converges for all $v \in \mathbb{C}$ with $|v| > R$. We can write

$$g(V) = \sum_{t \in D(g(V))} a_t V^t$$

for suitable $a_t \in \overline{\mathbb{Q}}$. We substitute $V = X^2Y$ and find

$$g(X^2Y)(X+Y)Y^{1/2} = \sum_{t \in D(g(V))} a_t X^{2t+1}Y^{t+1/2} + \sum_{t \in D(g(V))} a_t X^{2t}Y^{t+3/2}.$$

We also find

$$(g(X^2Y)(X+Y)Y^{1/2})^2 = g^2(X^2Y)(X+Y)^2Y = (1+X^2Y)(X+Y)^2Y,$$

which shows that $f(X,Y) := g(X^2Y)(X+Y)Y^{1/2} \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ is a root of F. If $x,y \in \mathbb{C}$ satisfy $|x| > \max(R,1)$ and $|y| > \max(R,1)$ they also satisfy $|x^2y| > R$. This shows that $g(x^2y)$ hence $f(x,y)$ converges for such x,y. We see that F has at most two roots since it is of positive Z-degree 2. We had two choices for g and therefore have found both roots of F. If $(i,j) \in D(f(X,Y))$, there must exist $t \in D(g(V))$ such that $(i,j) = (2t+1, t+1/2)$ or such that $(i,j) = (2t, t+3/2)$. We take $\varphi = 2$ and find either

$$i - j\varphi = 2t + 1 - (t+1/2)\cdot 2 = 0.$$

or

$$i - j\varphi = 2t - (t+3/2)\cdot 2 = -2.$$

So we take $\pi_1 = -3$ and $\pi_2 = 1$ and find that

$$\pi_1 < i - j\varphi < \pi_2$$

for all $(i,j) \in D(f(X,Y))$. We can therefore apply Theorem 8.28. We let $\mu = 1$ and $\lambda = 2$. We then find

$$F_{\mu,\lambda}(X,Y,Z) = Z^2 - X^2Y^2(X+Y)^2 = (Z + XY(X+Y))(Z - XY(X+Y)).$$

This is not a monomial in $\mathbb{Z}[X,Y,Z]$, not an element in $\mathbb{Z}[X,Y]$ and not a monomial in $\mathbb{Z}[X,Y]$ times a power of an irreducible factor in $\mathbb{Z}[X,Y,Z]$. Theorem 8.28 then tells us that we have reached a contradiction, from which we may conclude that $S(F)$ is not Zariski-dense in $T(F)$. □

Example 8.31. *Let*

$$F(X,Y,Z) = Z^3 - X^2Y(XY^5+2).$$

Then $S(F)$ is not Zariski-dense in $T(F)$.

Proof. We assume that $S(F)$ is Zariski-dense in $T(F)$ and will reach a contradiction. We view F as a polynomial in Z and see that F is a cubic polynomial whose coefficients 1, 0, 0, and $-X^2Y(XY^5+2)$ do not share a common irreducible factor in $\mathbb{Z}[X,Y]$. Because $X^2Y(XY^5+2)$ is not a cube in $\mathbb{Z}[X,Y]$, we see that F is irreducible in $\mathbb{Z}[X,Y,Z]$. Let $g(V) \in \overline{\mathbb{Q}}(\{(V^{-1})^*\})$ be any of the three roots for W in the polynomial $W^3 - 2 - V \in \mathbb{Z}[V,W]$. So $(g(V))^3 = 2 + V$ and $g(v)$ converge for all $v \in \mathbb{C}$ with $|v| > R$ for some R. We can write

$$g(V) = \sum_{t \in D(g(V))} a_t V^t$$

for suitable $a_t \in \overline{\mathbb{Q}}$. We substitute $V = XY^5$ and find

$$g(XY^5)X^{2/3}Y^{1/3} = \sum_{t \in D(g(V))} a_t X^{t+2/3} Y^{5t+1/3}.$$

We also find

$$(g(XY^5)X^{2/3}Y^{1/3})^3 = g^3(XY^5)X^2Y = X^2Y(XY^5+2),$$

which shows that $f(X,Y) \coloneqq g(XY^5)X^{2/3}Y^{1/3} \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ is a root of F. Any $x,y \in \mathbb{C}$ with $|x| > \max(R,1)$ and $|y| > \max(R,1)$ also satisfy $|xy^5| > R$. So $g(xy^5)$ hence $f(x,y)$ converge for such x,y. We see that F has at most three roots since it is of positive Z-degree 3. We had three choices for g and therefore have found all three roots of F. If $(i,j) \in D(f(X,Y))$, there must exist $t \in D(g(V))$ such that $(i,j) = (t + 2/3, 5t + 1/3)$. We take $\varphi = 1/5$ and find

$$i - j\varphi = t + 2/3 - (5t + 1/3)/5 = 2/3 - 1/15 = 3/5.$$

So we take $\pi_1 = 0$ and $\pi_2 = 1$ and find that

$$\pi_1 < i - j\varphi < \pi_2$$

for all $(i,j) \in D(f(X,Y))$. We can therefore apply Theorem 8.28. We let $\mu = 1$ and $\lambda = 3$. We then find

$$F_{\mu,\lambda}(X,Y,Z) = Z^3 - X^3Y^6 = (Z - XY^2)(Z^2 + XY^2Z + X^2Y^4).$$

This is not a monomial in $\mathbb{Z}[X,Y,Z]$, not an element in $\mathbb{Z}[X,Y]$ and not a monomial in $\mathbb{Z}[X,Y]$ times a power of an irreducible factor in $\mathbb{Z}[X,Y,Z]$. Theorem 8.28 then tells us that we have reached a contradiction, from which we may conclude that $S(F)$ is not Zariski-dense in $T(F)$. □

It can be interesting to check whether all conditions on Theorem 8.28 are really necessary. The previous examples show by contradiction that being Zariski-dense is a necessary condition. The following example shows that being irreducible is also a necessary condition on F.

Example 8.32. *Let $F \coloneqq (X+Z)(Y+Z)$. We see by Example 8.1 that $S(X+Z)$ is Zariski-dense in $T(X+Z)$ and that $S(Y+Z)$ is Zariski-dense in $T(Y+Z)$. We then apply Lemma 7.18 to see that $S(F)$ is Zariski-dense in $T(F)$. As a polynomial in Z, we see that F has two roots, namely $-X$ and $-Y$. They both converge for all x,y. We directly find the existence of $\varphi \in \mathbb{R}_{>0}$ and $\pi_1, \pi_2 \in \mathbb{R}$ such that*

$$\pi_1 < i - j\varphi < \pi_2$$

holds for any $(i,j) \in D(-X) = \{(1,0)\}$ and for any $(i,j) \in D(-Y) = \{(0,1)\}$. But if we take $\mu = \lambda = 1$, we see that $F_{\mu,\lambda} = F = (X+Z)(Y+Z)$. This is not a monomial in $\mathbb{Z}[X,Y,Z]$, not an element in $\mathbb{Z}[X,Y]$ and also not a monomial in $\mathbb{Z}[X,Y]$ times a power of an irreducible factor in $\mathbb{Z}[X,Y,Z]$.

The next example shows that the boundaries on the Newton dots are a necessary condition on F.

Example 8.33. *Let $F := Y - (2X + Z)Z$. After swapping the roles of Y and Z we can use Example 8.1 to see that $S(F)$ is Zariski-dense in $T(F)$. When we view F as a polynomial in Z, we use the quadratic formula to see that its roots are given by*

$$\frac{-2X \pm \sqrt{4X^2 - 4Y}}{2} = -X \pm X\sqrt{X^2 + Y}.$$

We let $g(V) \in \overline{\mathbb{Q}}(\{(V^{-1})^*\})$ be any of the two roots for W in the polynomial $W^2 - 1 - V \in \mathbb{Z}[V, W]$. So $(g(V))^2 = 1 + V$. We can write

$$g(V) = \sum_{t \in D(g(V))} a_t V^t$$

for suitable $a_t \in \overline{\mathbb{Q}}$. We substitute $V = X^{-2}Y$ and find

$$g(X^{-2}Y)X = \sum_{t \in D(g(V))} a_t X^{-2t+1} Y^t.$$

We also find

$$(g(X^{-2}Y)X)^2 = g^2(X^{-2}Y)X^2 = X^2 + Y,$$

which shows that $f(X, Y) := g(X^{-2}Y)X - X \in \overline{\mathbb{Q}}(\{(X^{-1})^*\})(((Y^{-1})^*))$ is a root of F. We see that F has at most two roots since it is of positive Z-degree. We had two choices for g and therefore have found both roots of F. If $(i, j) \in D(f(X, Y))$, we either have $(i, j) = (1, 0)$ or there must exist $t \in D(g(V))$ such that $(i, j) = (1 - 2t, t)$. Because the elements $t \in D(g(V))$ are not bounded from below we see that $1 - 2t$ is not bounded from above. So for any $\varphi \in \mathbb{R}_{>0}$ see see that there exists no upper boundary for $i - j\varphi = 1 - 2t - t\varphi$. So π_2 from Theorem 8.28 does not exist. We let $\mu = 1/2$ and $\lambda = 1$. We then find

$$F_{\mu,\lambda}(X, Y, Z) = -2XZ - Z^2 = -(2X + Z)Z.$$

This is not a monomial in $\mathbb{Z}[X, Y, Z]$, not an element in $\mathbb{Z}[X, Y]$ and not a monomial in $\mathbb{Z}[X, Y]$ times a power of an irreducible factor in $\mathbb{Z}[X, Y, Z]$. We therefore see that the boundary from Theorem 8.28 is a necessary condition.

One might wonder if we can allow less strong boundaries for Theorem 8.28. Is it for example good enough if the roots $f_i(X, Y)$ of $F(X, Y, Z)$ are of the form in (8.2)? This unfortunately is not the case as the following example shows.

Example 8.34. *Let $F := Z^2 - (X^2 - 1)(Y^2 - 1)$. We saw in Example 8.6 that $S(F)$ is Zariski-dense in $T(F)$ and that F is irreducible. When we view F as a polynomial in Z, we use the quadratic formula to see that its roots are given by*

$$f_{\pm}(X, Y) = \pm\sqrt{(X^2 - 1)(Y^2 - 1)}.$$

We let $g(V) \in \overline{\mathbb{Q}}(\{(V^{-1})^*\})$ be any of the two roots for W in the polynomial $W^2 - V^2 + 1 \in \mathbb{Z}[V, W]$. So $(g(V))^2 = V^2 - 1$ and there exists $R \in \mathbb{R}$ such that $g(v)$ converges for all $v \in \mathbb{C}$ with $|v| > R$. We can write

$$g(V) = \sum_{t \in D(g(V))} a_t V^t$$

for suitable nonzero $a_t \in \overline{\mathbb{Q}}$. By doing the algorithm to find the Taylor expansion of $g(V)$, we find

$$g(V) = \pm\sqrt{V^2 - 1} = \pm(V - \frac{1}{2}V^{-1} - \frac{1}{8}V^{-3} - \frac{1}{16}V^{-5} - \frac{5}{128}V^{-7} - \dots)$$

and we see that $\mathrm{ord}_V g(V) = 1$. We substitute V by X and by Y and find

$$\sqrt{X^2 - 1} = \sum_{i \in D(g(V))} a_i X^i$$

and

$$\sqrt{Y^2 - 1} = \sum_{j \in D(g(V))} a_j y^j.$$

We therefore have that the two roots of $F(X,Y,Z)$ are given by

$$\begin{aligned} f_{\pm}(X,Y) &= \pm\sqrt{(X^2-1)(Y^2-1)} \\ &= \pm\sqrt{(X^2-1)}\sqrt{(Y^2-1)} \\ &= \pm g(X)g(Y) \\ &= \pm \sum_{i\in D(g(V))} a_i X^i \sum_{j\in D(g(V))} a_j y^j \\ &= \pm \sum_{i\in D(g(V))} \sum_{j\in D(g(V))} a_i a_j X^i Y^j. \end{aligned}$$

We see that $f_{\pm}(x,y) = g(x)g(y)$ converge for all $x,y \in \mathbb{C}$ with $|x| > R$ and $|y| > R$. Also, for any $(i,j) \in D(f_{\pm}(X,Y))$ we must have $i \in D(g(V))$ and $j \in D(g(V))$. So both i and j are bounded from above. We therefore see that the roots $f_{\pm}(X,Y)$ are of the form in (8.2). *Now let $\mu = 1$ and $\lambda = 2$. We then find*

$$F_{\mu,\lambda}(X,Y,Z) = Z^2 - X^2Y^2 = (Z+XY)(Z-XY).$$

This is not a monomial in $\mathbb{Z}[X,Y,Z]$, not an element in $\mathbb{Z}[X,Y]$ and not a monomial in $\mathbb{Z}[X,Y]$ times a power of an irreducible factor in $\mathbb{Z}[X,Y,Z]$. We therefore see that Theorem 8.28 does not work if we weaken the boundary condition such that roots of the form in (8.2) *get accepted.*

The first drawback of Theorem 8.28 is that in order to apply this theorem we do need to know what the form is of the roots of the polynomial $F(X,Y,Z)$. We did not need to do such thing in Runge's Theorem 6.1. The second drawback is that the conditions are very strong. We therefore can not use this theorem often. Remark 8.29 tells us that we can remove the condition that demands $f(x,y)$ to converge if Conjecture 8.15 is true.

9 Further Reading

Runge's Theorem helps us to prove that certain binary Diophantine equations have only finitely many integral solutions. In these cases, we might wonder whether one could find boundaries on the size of such solutions. Hilliker and Straus [7] obtained such bounds. Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a polynomial such that we can prove by Runge's Theorem 6.1 that $F(x,y) = 0$ holds for only finitely many $x,y \in \mathbb{Z}$. Let $d = \max_{\deg_X F, \deg_Y F}$ and $h = \mathrm{ht}(F)$. Hilliker and Straus have showed that

$$\max|x|,|y| \leq \begin{cases} 4(h+1)^2 & \text{if } d = 1. \\ (8dh)^{d^{2d^3}} & \text{if } d > 1. \end{cases}$$

An improvement on these boundaries has been made by P. Walsh [14]. He found the following boundary on these $x,y \in \mathbb{Z}$

$$\max|x|,|y| \leq (2d)^{18d^7} h^{12d^6}.$$

In this thesis we have made an attempt to generalize Runge's Theorem to Diophantine equations in three dimensions. Enrico Bombieri [2] has given a theorem that, as pointed out by John Coates, may be viewed as a generalization to Runge's Theorem. This generalization works over all higher dimensions and requires a fast knowledge of algebraic geometry. His generalization has been described by Aaron Levin [9].

References

[1] M.J. Barwick et al. The Newton polygon and the Puiseux characteristic. Master's thesis, University of Sydney., 2011.

[2] E. Bombieri. On Weil's "Théorème de Décomposition". *American Journal of Mathematics*, 105(2):295–308, 1983.

[3] K. Conrad. Symmetric polynomials. `http://www.math.uconn.edu/~kconrad/blurbs/galoistheory/symmfunction.pdf`. Accessed: 2019-01-10.

[4] D. Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*, volume 150. Springer Science & Business Media, 2013.

[5] H. Grauert and R. Remmert. *Analytische Stellenalgebren*, volume 176. Springer-Verlag, 2013.

[6] A Grytczuk and A Schinzel. On Runge's Theorem about Diophantine equations. *Colloq. Math. Soc. J. Bolyai*, 60:329–356, 1992.

[7] D.L. Hilliker and E.G. Straus. Determination of bounds for the solutions to those binary Diophantine equations that satisfy the hypotheses of Runge's Theorem. *Transactions of the American Mathematical Society*, 280(2):637–657, 1983.

[8] M. Klazar. Runge's Theorem on Diophantine equations. `https://kam.mff.cuni.cz/~klazar/ostrava2.pdf`, 2013. Accessed: 2018-04-05.

[9] A. Levin. Variations on a theme of Runge: effective determination of integral points on certain varieties. *J. Théor. Nombres Bordeaux*, 20(2):385–417, 2008.

[10] B. Mannaa and T. Coquand. Dynamic Newton–Puiseux Theorem. *Journal of Logic & Analysis*, 5, 2013.

[11] K.J. Nowak. Some elementary proofs of Puiseux's theorems. *Univ. Iagel. Acta Math*, 38:279–282, 2000.

[12] C. Runge. Ueber ganzzahlige Lösungen von Gleichungen zwischen zwei Veränderlichen. *Journal für die reine und angewandte Mathematik*, 100:425–435, 1887.

[13] P. Stevenhagen. Number rings. `http://websites.math.leidenuniv.nl/algebra/ant.pdf`, 2017. Accessed: 2018-04-05.

[14] P.G. Walsh. A quantitative version of Runge's Theorem on Diophantine equations. *Acta Arith*, 62(2):157–172, 1992.