HAL open science

Algebraic certificates for Budan's theorem

Daniel Bembé

▶ To cite this version:

Daniel Bembé. Algebraic certificates for Budan's theorem. General Mathematics [math.GM]. Université de Franche-Comté, 2011. English. NNT : 2011BESA2022 . tel-00839189

HAL Id: tel-00839189

https://theses.hal.science/tel-00839189v1

Submitted on 27 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Algebraic certificates for Budan's theorem

Daniel Bembé

Besançon/München 2011

Algebraic certificates for Budan's theorem

Daniel Bembé

Thesis
prepared at
l'U.F.R. des sciences et techniques de l'Université de Franche-Comté
and
der Fakultät für Mathematik, Informatik und Statistik
der Universität München
to achieve
le grade de docteur de l'Université de Franche-Comté
spécialité : Mathématiques et applications
and
den Grad des Doktors der Naturwissenschaften
der Universität München.

Besançon/München, August 2, 2011

The defence will take place at the Université de Franche-Comté on August 2, 2011.

Members of the jury will be

Werner Bley (Professor an der Universität München)
[examinateur alternatif/Ersatzprüfer],

André Galligo (Professeur à l'Université Nice Sophia Antipolis)
[rapporteur/auswärtiger Gutachter],

Laureano González-Vega (Profesor titular de Universidad de Cantabria)
[présidente du jury/Zweitberichterstatter],

Henri Lombardi (Maître de Conférences à l'Université de Franche-Comté, HDR)
[directeur de thèse/dritter Prüfer],

Hervé Perdry (Maître de Conférences à l'Université Paris-Sud)
[examinateur/Ersatzprüfer],

Peter Schuster (Privatdozent an der Universität München)
[directeur de thèse/Erstberichterstatter],

Helmut Schwichtenberg (Professor an der Universität München)
[examinateur/vierter Prüfer].

Contents

1 Budan's theorem and virtual roots **17**
1.1 Budan's theorem 17
1.2 Virtual Roots 23

2 An algebraic certificate for Budan's theorem **33**
2.1 What is an algebraic certificate for Budan's theorem? 33
2.2 An algebraic certificate for Budan's theorem 34
2.3 An algorithm to calculate all linear certificates 40

3 A certificate for Budan's theorem constructed in polynomial time **49**
3.1 A certificate for Budan's theorem constructed in polynomial time 49
3.2 About the real roots of certain linear combinations of the polynomials p_i . . . 55
3.2.1 Motivation 55
3.2.2 Example and theorem 57
3.2.3 Generalization; example and theorem 64

4 Conclusion and further work **77**

Thanks

When I look around at the university, I see two types of supervisors. Those who care a little more about their students and those who care a little less about their students. And then there are those who do everything for their students: I had two of them. Dear Henry, dear Peter, you inspired me mathematically, you taught me philosophically, and you supported me in times of crisis. Thanks a million for this fantastic job.
In all the years I felt very much at home at the mathematical institute of the Universität München. Not only that it equipped me with bread for biting and a desk, no, all his friendly staff always supported me by all means. Here Helmut Schwichtenberg admitted me to the logics group although we had not met before, here Otto Forster assisted me with mathematical and Helmut Zöschinger with all-embracing wisdom. The two years where I had the fortune to be employed at the Universität der Bundeswehr München were not less convenient; dear Cornelius, thank you very much for your unconditional support. During all my visits to Besançon the Univerité de Franche-Comté always provided an excellent workplace. These trips were enabled by the kind financial aid of the Bayersich-Französische Hochschulzentrum/Centre de Coopération Universitaire Franco-Baravois and the Deutsch-Französische Hochschule/Université franco-allemande.
During this work I often used the computer algebra system PARI/GP to calculate with examples; the algorithm in section 2.3 is implemented in PARI. I thank the developers for this fantastic tool.
Dear André, I thank you for your cooperation. Dear Nathan, Matthew, Freiric, Amelie, without your language skills the English and French would have thrown me off track sometimes.
Let it be a stimulating discussion, a good popularity, which helps recover the working mood, a coffee break: You all know why you appear here: Thanks, Mama, Papa, Mamai, Lukas, Hannah, Denis, Michi, Flo, Simon, Josef, Basil, Freiric, Andreas, Claudia, Sebastian, Kathrin, Kuniko, Samuel, Vroni, Miriam, Natalia, Fabienne with Noah and Nathan, Karol, Spatzlmauslbusslbärle and Karin Beck.
If someone is looking for places that are particularly suitable in order to insert examples in dissertations: I recommend the Café Altschwabing and the Café Batty Baristas in Munich as well as the Café du Théâtre in Besançon.

Introduction and motivation

By the use of Zorn's lemma and the law of the excluded middle one can show that every field has an algebraic closure and every ordered field has a real closure [BCR, BPR, CR].
From a computational point of view, to deal with a field means to have (i) for every field element at least one representative implemented in the machine and certain representatives for 0 and 1; (ii) algorithms for addition and multiplication; (iii) a test if a representative equals 0, and – if not – an algorithm for computing the inverse. (A field providing (iii) is called "discrete".) If we speak about calculating in the algebraic closure of such an implemented field F (for example $\mathrm{F} = \mathbb{Q}$) this means that for every polynomial $f \in \mathrm{F}[X]$, we can calculate in a field extension where f admits a root. Thereby the items (i), (ii), (iii) of the field extension shall be reduced to those of the base field. From a computational point of view, this is hard since factorization of f is hard.
To deal with an ordered field means to have in addition (iv) a sign test. If we speak about calculating in the real closure of an ordered field R (again for example $\mathrm{R} = \mathbb{Q}$) this means that for every polynomial $g \in \mathrm{R}[X]$ which admits both a negative and a positive value, we can calculate in a field extension where g admits a root. Hollkott [Hol] showed that the order enables us to avoid factorizing g. (An example for this: Consider $g_1 := X^2 - 2$ and $g_2 := X^2 - 3 \in \mathbb{Q}[X]$. To calculate in an extension of $\mathbb{Q}$ where $g_1 g_2$ admits a root we must choose one of the factors g_1 or g_2. If we can not factorize $g_1 g_2$ we can use the fact that in the real closure of $\mathbb{Q}$ we have

$$-\sqrt{3} < -\sqrt{2} < \sqrt{2} < \sqrt{3}, \qquad (*)$$

i.e., the order enables us to distinguish the roots. And by deciding for one of the four positions in $(*)$ we can decide for one of the factors g_1 or g_2 without knowing it, i.e., without factorizing $g_1 g_2$.) Hollkott showed for an arbitrary (even non-archimedian) ordered field R and a polynomial g as above how to reduce a calculation in an ordered field extension where g admits a root to finitely many calculations in R. Furthermore, he showed without using Zorn's lemma that the object we calculate in fulfills the axioms of an ordered field. This means he presented a constructive existence proof for the real closure. In his proof he deduced the assertions of Rolle's and Sturm's theorems by an induction on the degree of g, which leads to an algorithm of very high complexity in the degree of g.
A constructive existence proof of conceptual nature for the real closure was presented by Lombardi and Roy [LR1]. Here, in the induction step, the algebraic theorem about finite growth is used instead of Rolle's theorem to show that a polynomial with positive derivative in an interval increases there. Implementation of the algorithms hidden in [LR1] is related to Hörmander's algorithm [BCR].
Furthermore, the arguments in [CR] give reason to an algorithm for reducing a calculation in an ordered field extension of R where g as above admits a root to calculations in R. This

algorithm which uses specialized Sturm series [GLRR] is presented in [BPR]. But here the correctness – i.e., the fact that the object we calculate in fulfills the axioms of an ordered field – is proved by the argument that this object is embedded in the real closure. Hence this algorithm alone can not be seen as constructive existence proof for the real closure.
In this work we present two algebraic certificates for Budan's theorem. Budan's theorem claims the following. Let R be an ordered field, $f \in \mathrm{R}[X]$ of degree n and $a, b \in \mathrm{R}$ with $a < b$. Then the number of sign changes in the sequence $(f(b), f'(b), \ldots, f^{(n)}(b))$ is not greater than the number of sign changes in the sequence $(f(a), f'(a), \ldots, f^{(n)}(a))$. This enables us to count real roots in a similar way to the real root counting by Sturm's theorem. (Budan's count of real roots is today known as "Budan-Fourier count" which, indeed, counts so called virtual roots which comprehend the real roots [CLLR, GLM].) An algebraic certificate for Budan's theorem is a certain kind of proof which leads from the negation of the assumption to the contradictory algebraic identity $0 > 0$. In particular, what we present are linear certificates; compare [CLR, Schr]. The existence of such a certificate already follows from the Baby Positivstellensatz [CLR, Schr] together with any proof of Budan's theorem. The algorithm for our first certificate (chapter 2) is based on the historical proof by Budan which uses only combinatorial arguments [Bud, Bor]. It has a complexity exponential in the degree of f. The algorithm for the second certificate (chapter 3) is based on mixed Taylor series [Lom2] and polynomials $\prod_{k=0}^{i-1}(X-k) \in \mathbb{R}[X]$ and shows a smaller complexity: The main calculation is solving a linear system; this is polynomial in the degree of f.
Compared to Budan's theorem, all known algebraic certificates for Sturm's theorem are much more complicated. This could be in connection with the fact that all known proofs of Sturm's theorem use heavily all roots of the polynomials in the Sturm sequence.
The motivation for the present work lies in the objective to find new algorithms based on Budan's theorem to reduce calculations in ordered field extensions to calculations in the base field. On the one hand we would like to reduce the complexities of the calculations, on the other hand we are interested in new correctness proofs for these algorithms. Here the presented certificates can be helpful. If we succeed with this idea it would mean a new constructive existence proof for the real closure. Furthermore, we are interested in a construction of the real closure of a non-discrete field (for example, many subfields of $\mathbb{R}$) where our considerations can also be helpful.
The constructive papers [Lom1, LR2] about Stengle's Positivstellensatz [CLR, Kri, Ste] show a connection to the construction of the real closure. A special case of Stengle's Positivstellensatz is known as Hilbert's 17th problem. Furthermore, Schmüdgen's Positivstellensatz is of interest [Kri, Schm]; for a constructive version see [Schw]. For general references about Positivstellensätze consider [BCR, PD, Sche].

In chapter 1 we present the historical proof of Budan's theorem, define virtual roots and explain some of their properties. This should be helpful for motivating and understanding the certificates. Besides adaption of the historical arguments, everything up to corollary 1.2.6 comes from [Bud, CLLR, GLM]. The last statements are original.
Chapters 2 and 3 contain the two certificates. All of them is invented by the Ph.D. student inspired by the ideas of his supervisors Henri Lombardi and Peter Schuster and the cited literature. Chapter 2 is mainly published in [Bem].

Introduction et motivation

Par l'utilisation du lemme de Zorn et le principe du tiers exclu on peut montrer que chaque corps a une clôture algébrique et chaque corps ordonné a une clôture réelle [BCR, BPR, CR]. D'un point de vue informatique, si nous parlons d'un corps, ça veut dire que nous avons (i) pour chaque élément du corps au moins une représentation dans la machine et certaines représentations pour les 0 et les 1 ; (ii) des algorithmes pour l'addition et la multiplication ; (iii) un test si une représentation correspond à 0, et – sinon – un algorithme pour calculer l'inverse. (Un corps fournissant (iii) est appelé « discret ».) Si nous parlons de calculer dans la clôture algébrique d'un tel corps F (par exemple $\mathrm{F} = \mathbb{Q}$), ça veut dire que pour tout polynôme $f \in \mathrm{F}[X]$, nous pouvons calculer dans une extension du corps de base dans laquelle f admet une racine. Ainsi les éléments (i), (ii), (iii) de l'extension doivent être ramenés au corps de base. D'un point de vue informatique, cela est difficile car la factorisation de f est difficile.

Si nous parlons d'un corps ordonné, ça veut dire que nous avons en plus (iv) un test du signe. Si nous parlons de calculer dans la clôture réelle d'un corps ordonné R (prenons, par exemple, à nouveau $\mathrm{R} = \mathbb{Q}$), ça veut dire que pour tout polynôme $g \in \mathrm{R}[X]$ qui admet aussi bien une valeur négative qu'une valeur positive, nous pouvons calculer dans une extension du corps de base où g admet une racine. Hollkott [Hol] a montré que l'ordre nous permet d'éviter de factoriser g. (Un exemple : Nous considèrons $g_1 := X^2 - 2$ et $g_2 := X^2 - 3 \in \mathbb{Q}[X]$. Pour calculer dans une extension de $\mathbb{Q}$ où $g_1 g_2$ admet une racine nous devons choisir l'un des facteurs g_1 ou g_2. Si nous ne pouvons pas factoriser $g_1 g_2$, nous pouvons utiliser le fait que dans la clôture réelle de $\mathbb{Q}$ nous avons

$$-\sqrt{3} < -\sqrt{2} < \sqrt{2} < \sqrt{3}, \tag{$*$}$$

et l'ordre nous permet de distinguer les racines. En choisissant l'une des quatre positions de $(*)$ nous pouvons nous décider pour l'un des facteurs g_1 ou g_2 sans le connaître, c'est à dire, sans factoriser $g_1 g_2$.) Hollkott a montré pour un corps ordonné arbitraire (même non-archimèdien) R et un polynôme g comme ci-dessus comment réduire un calcul dans une extension ordonnée où g admet une racine à un nombre fini des calculs dans R. Par ailleurs, il a montré, sans utiliser le lemme de Zorn que l'objet, dans lequel nous calculons, satisfait les axiomes d'un corps ordonné. Cela signifie qu'il a présenté une preuve constructive de l'existence de la clôture réelle. Dans sa preuve, il a déduit les assertions des théorèmes de Rolle et de Sturm par une récurrence sur le degré de g, qui conduit à un algorithme de complexité très élevée dans le degré de g.

Une preuve constructive de nature conceptuelle de l'existence de la clôture réelle a été présentée par Lombardi et Roy [LR1]. Ici, dans l'hérédité de la recurrence, le théorème algébrique sur la croissance finie est utilisé au lieu du théorème de Rolle pour montrer qu'un polynôme avec dérivée positive dans un intervalle y augmente. Implémenter les algorithmes

cachés dans [LR1] est lié à l'Algorithme de Hörmander [BCR].
Par ailleurs, les arguments dans [CR] conduisent à un algorithme pour réduire un calcul dans une extension ordonnée de R où g comme ci-dessus admet une racine aux calculs dans R. Cet algorithme qui utilise des suites de Sturm spécialisées [GLRR] est présenté dans [BPR]. Mais ici, l'exactitude – à savoir, le fait que l'objet, dans lequel nous calculons, satisfait les axiomes d'un corps ordonné – est prouvée par l'argument que cet objet a un plongement dans la clôture réelle. Ainsi, cet algorithme seul ne peut pas être considéré comme une preuve constructive d'existence de la clôture réelle.
Dans ce travail, nous présentons deux certificats algébriques pour le théorème de Budan. Le théorème de Budan s'énonce comme suit : Soit R un corps ordonné, $f \in \mathrm{R}[X]$ de degré n et $a, b \in \mathrm{R}$ avec $a < b$. Alors, le nombre de variations de signe dans la suite $(f(b), f'(b), \ldots, f^{(n)}(b))$ n'est pas supérieur au nombre de variations de signe dans la séquence $(f(a), f'(a), \ldots, f^{(n)}(a))$. Cela nous permet de compter des racines réelles d'une manière similaire au comptage des racines réelles par le théorème de Sturm. (Compter des racines réelles à la Budan est aujourd'hui connu comme « Budan-Fourier count ». En effet, il compte des racines dites virtuelles qui comprennent les racines réelles [CLLR, GLM].) Un certificat algébrique pour le théorème de Budan est un certain type de preuve qui mène de la négation de l'hypothèse à l'identité algébrique contradictionelle $0 > 0$. En particulier, ce que nous présentons sont des certificats linéaires ; on pourra comparer avec [CLR, Schr]. L'existence d'un tel certificat résulte déjà de la Baby Positivstellensatz [CLR, Schr] en prenant une preuve du théorème de Budan. L'algorithme pour notre premier certificat (chapitre 2) est basé sur la preuve historique par Budan, qui utilise uniquement des arguments combinatoires [Bud, Bor]. Il a une complexité exponentielle dans le degré de f. L'algorithme pour le deuxième certificat (chapitre 3) est basé sur des suites de Taylor mixtes [Lom2] et des polynômes $\prod_{k=0}^{i-1}(X-k) \in \mathbb{R}[X]$ et exhibe une plus petite complexité : Le calcul principal est la résolution d'un système linéaire, ce qui est polynomiale dans le degré de f.
Comparé au théorème de Budan, tous les certificats algébriques connus pour le théorème de Sturm sont beaucoup plus compliqués. Cela pourrait être en relation avec le fait que toutes les preuves connues du théorème de Sturm utilisent copieusement toutes les racines de tous les polynômes dans la suite de Sturm.
La motivation pour ce travail réside dans l'objectif de trouver des nouveaux algorithmes basés sur le théorème de Budan pour réduire des calculs en extensions ordonnées aux calculs dans le corps de base. D'une part, nous aimerions réduire la complexité des calculs, d'autre part nous nous intéressons à des nouvelles preuves de l'exactitude de ces algorithmes. Ici, les certificats présentés peuvent être utiles. Si nous réussissons avec cette idée, cela signifierait une nouvelle preuve constructive de l'existence de la clôture réelle. Par ailleurs, nous sommes intéressés à une construction de la clôture réelle d'un corps non-discret (par exemple, de nombreux sous-corps de $\mathbb{R}$) où nos considérations peuvent également être utiles.
Les travaux constructifs [Lom1, LR2] concernant le Positivstellensatz de Stengle [CLR, Kri, Ste] montrent une connexion avec la construction de la clôture réelle. Un cas special du Positivstellensatz de Stengle est connu comme le dix-septième problème de Hilbert. Par ailleurs, il faut citer le Positivstellensatz de Schmüdgen [Kri, Schm] ; pour une version constructive voir [Schw]. Pour des références générales sur les Positivstellensätze, on peut lire [BCR, PD, Sche].

Dans le chapitre 1 nous présentons la preuve historique du théorème de Budan, définissons des racines virtuelles et expliquons certaines de leurs propriétés. Cela sera utile pour motiver et comprendre les certificats. Outre l'adaptation des arguments historiques, tout jusqu'au

corollaire 1.2.6 vient de [Bud, CLLR, GLM]. Les dernières assertions sont originales.
Les chapitres 2 et 3 contiennent les deux certificats. Ils ont été tous les deux inventés par le doctorant inspiré par les idées de ses superviseurs Henri Lombardi et Peter Schuster et la littérature citée. Le chapitre 2 est principalement publié dans [Bem].

Einleitung und Motivation

Mit Hilfe des Zornschen Lemmas und des Satzes vom ausgeschlossenen Dritten lässt sich zeigen, dass jeder Körper einen algebraischen Abschluss und jeder geordnete Körper einen reellen Abschluss besitzt [BCR, BPR, CR].
Wenn wir im Rahmen der Computeralgebra von einem Körper sprechen, meinen wir damit (i) die Implementierung jedes Körperelements durch mindestens einen Repräsentanten in der Maschine und festgelegte Repräsentanten für 0 und 1; (ii) Algorithmen für Addition und Multiplikation; (iii) einen Test, ob ein Repräsentant gleich 0 ist, und – falls nicht – einen Algorithmus zum Berechnen des Inversen. (Ein Körper, der über (iii) verfügt, wird als „diskret“ bezeichnet.) Wenn wir davon sprechen, im algebraischen Abschluss eines so implementierten Körpers F zu rechnen (z.B. $\mathrm{F} = \mathbb{Q}$), bedeutet das, dass wir für jedes Polynom $f \in \mathrm{F}[X]$ in einer Körpererweiterung rechnen können, in der f eine Nullstelle annimmt. Dabei sollen die Punkte (i), (ii), (iii) der Körpererweiterung auf die des Grundkörpers zurückgeführt werden. Dies ist nicht leichter als das Faktorisieren von f.
Wenn wir von einem geordneten Körper sprechen, meinen wir damit, dass wir zusätzlich über (iv) einen Vorzeichentest verfügen. Wenn wir davon sprechen, im reellen Abschluss eines geordneten Körpers R zu rechnen (auch hier z.B. $\mathrm{R} = \mathbb{Q}$), bedeutet das, dass wir für jedes Polynom $g \in \mathrm{R}[X]$, welches sowohl negative als auch positive Werte annimmt, in einer Körpererweiterung rechnen können, in der g eine Nullstelle annimmt. Hollkott [Hol] hat gezeigt, wie die Ordnung uns in die Lage versetzt, dabei auf das Faktorisieren von g zu verzichten. (Ein Beispiel hierfür: Betrachte $g_1 := X^2 - 2$ und $g_2 := X^2 - 3 \in \mathbb{Q}[X]$. Um in einer Körpererweiterung zu rechnen, in der $g_1 g_2$ eine Nullstelle annimmt, müssen wir uns für einen der Faktoren g_1 oder g_2 entscheiden. Wenn wir aber $g_1 g_2$ nicht faktorisieren können, können wir ausnutzen, dass im rellen Abschluss von $\mathbb{Q}$

$$-\sqrt{3} < -\sqrt{2} < \sqrt{2} < \sqrt{3} \qquad (*)$$

gilt. Die Ordnung ermöglicht es uns also, die Nullstellen zu unterscheiden. Und durch Entscheiden für eine der vier Positionen in $(*)$ können wir uns für einen der Faktoren g_1 oder g_2 entscheiden, ohne ihn zu kennen, d.h. ohne $g_1 g_2$ zu faktorisieren.) Hollkott zeigte für einen beliebigen (nicht notwendig archimedisch) geordneten Körper R und ein Polynom g wie oben, wie sich eine Berechnung in einer geordneten Körpererweiterung, in der g eine Nullstelle annimmt, auf endlich viele Berechnungen in R zurückführen lässt. Darüberhinaus zeigte er, ohne das Zornsche Lemma zu verwenden, dass das Objekt, in dem wir rechnen, die Axiome eines geordneten Körpers erfüllt. Dies bedeutet, dass er einen konstruktiven Existenzbeweis für den reellen Abschluss angab. In seinem Beweis leitete er die Aussagen der Sätze von Rolle und Sturm induktiv über den Grad von g her, was zu einem Algorithmus mit sehr hoher Komplexität im Grad von g führt.
Einen konstruktiven Existenzbeweis von konzeptioneller Natur gaben Lombardi and Roy an

[LR1]. Hier wird im Induktionsschritt statt auf den Satz von Rolle auf den algebraischen Satz zur endlichen Zunahme zurückgegriffen, um zu zeigen, dass ein Polynom in einem Intervall steigt, in dem seine Ableitung positive Werte annimmt. Der aus [LR1] extrahierte Algorithmus zeigt Verwandtschaft zum Algorithmus von Hörmander [BCR].
Desweiteren führen die in [CR] vorgestellten Methoden zu einem Algorithmus, um die Berechnungen in einer geordneten Körpererweiterung von R, in der g wie oben eine Nullstelle annimmt, auf Berechungen im Grundkörper R zurückzuführen. Dieser Algorithmus basiert auf speziellen Sturmschen Ketten [GLRR] und ist in [BPR] ausgeführt. Allerdings wird hier seine Korrektheit – d.h. dass das Objekt, in dem wir rechnen, die Axiome eines geordneten Körpers erfüllt – auf seine Einbettung in den reellen Abschluss zurückgeführt. Insofern kann dieser Algorithmus nicht als konstruktiver Existenzbeweis für den reellen Abschluss gesehen werden.
In der vorliegenden Arbeit stellen wir zwei algebraische Zertifikate für den Satz von Budan vor. Die Aussage des Satzes von Budan ist folgende: Seien R ein geordneter Körper, $f \in \mathrm{R}[X]$ vom Grad n und $a, b \in \mathrm{R}$ mit $a < b$. Dann ist die Anzahl der Vorzeichenwechsel in der Folge $(f(b), f'(b), \ldots, f^{(n)}(b))$ nicht größer als die Anzahl der Vorzeichenwechsel in der Folge $(f(a), f'(a), \ldots, f^{(n)}(a))$. Dies ermöglicht uns ähnlich den Sturmschen Ketten das Zählen reeller Nullstellen. (Das Zählen von Nullstellen mit Hilfe des Satzes von Budan wird heute als „Budan-Fourier count“ bezeichnet. Gezählt werden dabei mehr als die reellen Nullstellen – sogenannte virtuelle Nullstellen [CLLR, GLM].) Ein algebraisches Zertifikat für den Satz von Budan ist ein Beweis, in welchem aus der Negation der Aussage des Satzes von Budan die widersprüchliche algebraische Identität $0 > 0$ gefolgert wird. Die hier vorgestellten Zertifikate werden als lineare Zertifikate bezeichnet (vlg. [CLR, Schr]). Die Existenz eines solchen Zertifkates wird bereits vom Baby Positivstellensatz [CLR, Schr] zusammen mit einem beliebigen Beweis für den Satz von Budan garantiert. Der Algorithmus für unser erstes Zertifikat (Kapitel 2) orientiert sich am historischen Beweis von Budan, der mit ausschließlich kombinatorischen Mitteln auskommt [Bud, Bor]. Seine Komplexität ist exponentiell im Grad von f. Der Algorithmus für das zweite Zertifkat (Kapitel 3) basiert auf gemischten Taylorfolgen [Lom2] und Polynomen $\prod_{k=0}^{i-1}(X-k) \in \mathbb{R}[X]$ und zeigt eine geringere Komplexität: Diese ergibt sich hauptsächlich aus der Lösung eines linearen Gleichungssystems, und die Komplexität hierfür ist polynomial im Grad von f.
Verglichen mit diesen algebraischen Zertifikaten für den Satz von Budan sind alle bekannten algebraischen Zertifikate für den Satz von Sturm wesentlich komplizierter. Dies steht möglicherweise damit im Zusammenhang, dass in allen bekannten Beweisen für den Satz von Sturm auf alle Nullstellen aller in der Sturmschen Kette stehender Polynome zurückgegriffen wird.
Die vorliegende Arbeit ist motiviert von der Suche nach neuen, auf dem Satz von Budan basierenden, Algorithmen zum Zurückführen der Berechnungen in geordneten Körpererweiterung auf solche im Grundkörper. Dadurch erhoffen wir uns einerseits, die Komplexität derartiger Berechnungen zu senken, und andererseits, Beweise für die Korrektheit dieser Algorithmen zu finden, die auf neuen Argumenten beruhen. Dabei können die hier vorgestellten Zertifikate eine Rolle spielen. Neue derartige Algorithmen wären neue konstruktive Existenzbeweise für den reellen Abschluss. Weiterhin interessiert uns die Konstruktion des reellen Abschlusses eines nicht diskreten Körpers (wie beispielsweise viele Unterkörper von $\mathbb{R}$); auch hier können unsere Argumente hilfreich sein.
Im Zusammenhang mit der Konstruktion des reellen Abschlusses sind die konstruktiven Arbeiten [Lom1, LR2] über den Stengleschen Positivstellensatz [CLR, Kri, Ste] zu erwähnen,

dessen Spezialfall als Hilberts 17. Problem bekannt ist. Ebenso der Schmüdgensche Positivstellensatz [Kri, Schm] und seine konstruktive Version [Schw]. Allgemeinere Arbeiten über Positivstellensätze finden sich in [BCR, PD, Sche].

In Kapitel 1 stellen wir den historischen Beweis des Satzes von Budan vor, definieren virtuelle Nullstellen und zeigen einige ihrer Eigenschaften. Dies soll zu Verständnis und Motivation der Zertifikate beitragen. Abgesehen von der Überarbeitung der historischen Beweise findet sich alles bis zum Korollar 1.2.6 in [Bud, CLLR, GLM]. Die letzten Aussagen dieses Kapitels sind original.
Kapitel 2 und 3 bringen die beiden Zertifikate. Abgesehen von der zitierten Literatur ist hier der Doktorand – inspiriert von den Ideen seiner Doktorväter Henri Lombardi and Peter Schuster – als Urheber anzusehen. Kapitel 2 is teilweise veröffentlicht ([Bem]).

Chapter 1

Budan's theorem and virtual roots

In this chapter we give an overview about Budan's theorem and virtual roots. Everything up to corollary 1.2.6 can be found in [Bud, CLLR, GLM].
Budan's theorem is presented in the appendix of the historic paper [Bud] which contains only sketches of proofs. Using exactly their arguments we precisely formulate these proofs.
The Budan-Fourier count of virtual roots is similar to the Sturm count of real roots. From a constructive point of view, the latter seems complicated since all known proofs use all real roots of the polynomials in the Sturm sequence. To prove that the Budan-Fourier count of virtual roots results to a non-negative number only combinatoric calculations in a real field are needed. This gives a special interest to Budan's theorem from a constructive point of view.
Our definition of virtual roots comes from [GLM]. The correspondence to Budan's theorem is already shown in [CLLR] in the more general context of **f**-derivatives. The continuity of the virtual root functions gives them a special interest from a constructive point of view. Moreover, we describe some properties of virtual roots and virtual multiplicities in the end of the chapter.

1.1 Budan's theorem

In this section let $\mathbf{Q}$ denote a real field (i.e., ordered, for example $\mathbb{Q}$).

Definition 1.1.1.

i) For α, $\gamma \in \mathbf{Q}$, let the *sign* of α

$$\operatorname{sign}(\alpha) := \begin{cases} -1 & \text{if } \alpha < 0, \\ 0 & \text{if } \alpha = 0, \\ +1 & \text{if } \alpha > 0 \end{cases}$$

and $[\alpha, \gamma] := \{\beta \in \mathbf{Q} | \alpha \leq \beta \leq \gamma\}$ ($[\alpha, \gamma[$, $]\alpha, \gamma]$ and $]\alpha, \gamma[$ equivalently).

ii) For a sequence $(a_0, \ldots, a_n) \in (\mathbf{Q} \setminus \{0\})^{n+1}$, the *number of sign changes* $\mathbf{V}(a_0, \ldots, a_n)$ is defined inductively in the following way:

$$\mathbf{V}(a_0) := 0, \qquad \mathbf{V}(a_0, \ldots, a_i) := \begin{cases} \mathbf{V}(a_0, \ldots, a_{i-1}) & \text{if } a_{i-1}a_i > 0, \\ \mathbf{V}(a_0, \ldots, a_{i-1}) + 1 & \text{if } a_{i-1}a_i < 0; \end{cases}$$

to determine the number of sign changes of a sequence $(b_0, \dots, b_n) \in \mathbf{Q}^{n+1}$, delete the zeros in $(b_0, \dots, b_n)$ and apply the preceding case ($\mathbf{V}$ of the empty sequence equals 0).

iii) For a polynomial $f = a_0 + a_1 X + \cdots + a_n X^n \in \mathbf{Q}[X]$ of degree n, the *sequence of coefficients* is defined to be $(a_0, \dots, a_n)$. The number of sign changes of its coefficients is denoted by

$$\mathbf{V}(f) := \mathbf{V}(a_0, \dots, a_n).$$

iv) Let $\mathbf{S} : \mathbf{Q}^{n+1} \to \mathbf{Q}^{n+1}$ be the map that maps a sequence to the sequence summed in the following way:

$$\mathbf{S}(a_0, \dots, a_n) := \Big(\sum_{j=0}^{n} a_j, \sum_{j=1}^{n} a_j, \dots, a_{n-1} + a_n, a_n\Big).$$

v) Let $f \in \mathbf{Q}[X]$, $\zeta \in \mathbf{Q}$. The *real multiplicity* $\mathrm{rmult}_f(\zeta)$ is, by definition, the number $m \geq 0$ for which $(X - \zeta)^m$ divides f and $(X - \zeta)^{m+1}$ does not. If $rmult_f(\zeta) \geq 1$, we say that ζ is a real root of f.

Lemma 1.1.2. *For a sequence $(a_0, \dots, a_n) \in \mathbf{Q}^{n+1}$, a polynomial $f \in \mathbf{Q}[X]$, $f \neq 0$, and $s \geq 1$, we have*

i) $\mathbf{V}(a_0, \dots, a_n) \geq \mathbf{V}(\mathbf{S}(a_0, \dots, a_n))$,

ii) $\mathbf{V}(f \cdot (X + 1)) \leq \mathbf{V}(f) < \mathbf{V}(f \cdot (X - 1))$,

iii)

$$\begin{aligned}\mathbf{S}^s(a_0, \dots, a_n) = &\Big(\sum_{j=0}^{n} \binom{s-1+j}{s-1} a_j, \sum_{j=1}^{n} \binom{s-2+j}{s-1} a_j, \dots \\ &\dots, \sum_{j=n-1}^{n} \binom{s-n+j}{s-1} a_j, \sum_{j=n}^{n} \binom{s-(n+1)+j}{s-1} a_j\Big).\end{aligned}$$

Proof.

i) This is easily seen by induction on the length of the sequence. As the summation is done from the right to the left the induction is backwards: For a sequence of length 1 we have $\mathbf{S}(a_0) = (a_0)$. Let the assumption be true for a sequence $(a_1, \dots, a_n) \neq (0, \dots, 0)$ of length $n \geq 1$. To prove the assumption for the sequence $(a_0, a_1, \dots, a_n)$, it suffices to consider the case where the original sequence does not gain a sign change while the summed sequence already has the maximal number of sign changes, i.e., $\mathbf{V}(a_0, a_1, \dots, a_n) = \mathbf{V}(a_1, \dots, a_n)$ and $\mathbf{V}(\mathbf{S}(a_1, \dots, a_n)) = \mathbf{V}(a_1, \dots, a_n)$. Choosing $1 \leq k \leq n$ minimal with $a_k \neq 0$ leads to

$$a_0 a_k \geq 0 \qquad \text{and} \qquad a_k\Big(\sum_{j=1}^{n} a_j\Big) \geq 0$$

which shows that

$$a_0\Big(\sum_{j=1}^{n} a_j\Big) \geq 0.$$

ii) This is similarly seen by induction. Let $(a_0, \dots, a_n)$ be the sequence of coefficients of f. Let the claim be shown for $(a_0, \dots, a_{n-1}) \neq (0, \dots, 0)$.
The sequence of coefficients of $f \cdot (X+1)$ is $(a_0, a_0 + a_1, \dots, a_{n-1} + a_n, a_n)$. It suffices to consider the case where the original sequence does not gain a sign change, i.e., $a_k a_n \geq 0$, where $0 \leq k \leq n-1$ is maximal with $a_k \neq 0$. Then the right-most non-zero element of $(a_0, a_0 + a_1, \dots, a_{n-1})$ also is a_k, which shows the claim.
The sequence of coefficients of $f \cdot (X-1)$ is $(-a_0, a_0 - a_1, \dots, a_{n-1} - a_n, a_n)$. It suffices to consider the case where the original sequence gains a sign change, i.e., $a_n \neq 0$ and $a_{n-1} a_n \leq 0$. Then $(a_{n-1} - a_n) a_n < 0$, which shows the claim.

iii) Induction on s. $s = 1$:

$$\Big(\sum_{j=0}^{n} a_j, \dots, a_{n-1} + a_n, a_n\Big) = \Big(\sum_{j=0}^{n} \binom{j}{0} a_j, \dots, \binom{0}{0} a_{n-1} + \binom{1}{0} a_n, \binom{0}{0} a_n\Big)$$

$s - 1 \to s$: We write the sequence as

$$\Big(\sum_{j=i}^{n} \binom{s-(i+1)+j}{s-1} a_j\Big)_{0 \leq i \leq n}$$

and present an induction on i backwards. $i = n$:

$$\sum_{j=n}^{n} \binom{s-(n+1)+j}{s-1} a_j = \sum_{j=n}^{n} \binom{(s-1)-(n+1)+j}{(s-1)-1} a_j$$

$i + 1 \to i$:

$$\begin{aligned}
&\sum_{j=i+1}^{n} \binom{s-(i+2)+j}{s-1} a_j + \sum_{j=i}^{n} \binom{(s-1)-(i+1)+j}{(s-1)-1} a_j \\
&= \sum_{j=i+1}^{n} \left[\binom{s-(i+2)+j}{s-1} + \binom{s-(i+2)+j}{(s-1)-1}\right] a_j + \binom{s-2}{s-2} a_i \\
&= \sum_{j=i+1}^{n} \binom{s-(i+1)+j}{s-1} a_j + \binom{s-1}{s-1} a_i = \sum_{j=i}^{n} \binom{s-(i+1)+j}{s-1} a_j
\end{aligned}$$

□

If ζ is a positive root of f, then executing the coordinate transformation $X \mapsto \zeta X$, the multiplication $f \cdot (X-1)$ and the back-transformation $X \mapsto \frac{1}{\zeta} X$ generalizes lemma 1.1.2 ii) to arbitrary positive roots. This leads to the following assertion, called Descartes' rule: If f has degree n, k negative and l positive roots, then

$$l \leq \mathbf{V}(f) \leq n - k. \tag{1.1.1}$$

In case $k + l = n$, (1.1.1) fixes $\mathbf{V}(f)$. Budan's theorem provides some information about $\mathbf{V}(f)$ in general.

Theorem 1.1.3 (Budan)**.** *Let* $\mathbf{Q}$ *be a real field,* $f \in \mathbf{Q}[X]$ *and* $\delta \in \mathbf{Q}$, $\delta > 0$. *For the polynomial* $f \circ (X + \delta)$, *we get*

$$\mathbf{V}(f \circ (X + \delta)) \leq \mathbf{V}(f).$$

Proof. First, let $\delta := 1$. If $f = a_0 + \cdots + a_n X^n$, then $f \circ (X+1) = b_0 + \cdots + b_n X^n$ with

$$b_i = \sum_{j=i}^{n} \binom{j}{i} a_j.$$

According to lemma 1.1.2 iii) the sequence $(b_0, \ldots, b_n)$ appears in the diagonal of the matrix $(a_{s,i})$ whose rows consist of $\mathbf{S}^s(a_0, \ldots, a_n)$:

$$\begin{pmatrix} a_{0,0} & \cdots & a_{0,n} \\ a_{1,0} & \cdots & a_{1,n} \\ a_{2,0} & \cdots & a_{2,n} \\ \vdots & & \vdots \\ a_{n+1,0} & \cdots & a_{n+1,n} \end{pmatrix} = \begin{pmatrix} & (a_0, \ldots, a_n) \\ \mathbf{S} & (a_0, \ldots, a_n) \\ \mathbf{S}^2 & (a_0, \ldots, a_n) \\ & \vdots \\ \mathbf{S}^{n+1} & (a_0, \ldots, a_n) \end{pmatrix}$$

$$= \begin{pmatrix} a_0 & \cdots & a_{n-2} & a_{n-1} & a_n \\ \sum_{j=0}^{n} \binom{j}{0} a_j & \cdots & \sum_{j=n-2}^{n} \binom{2-n+j}{0} a_j & \sum_{j=n-1}^{n} \binom{1-n+j}{0} a_j & \sum_{j=n}^{n} \binom{0-n+j}{0} a_j \\ \vdots & \sum_{j=1}^{n} \binom{j}{1} a_j & \iddots & \sum_{j=n-1}^{n} \binom{2-n+j}{1} a_j & \sum_{j=n}^{n} \binom{1-n+j}{1} a_j \\ & & \ddots & \iddots & \sum_{j=n}^{n} \binom{2-n+j}{2} a_j \\ & & & \sum_{j=n-1}^{n} \binom{j}{n-1} a_j & \vdots \\ & & & \cdots & \sum_{j=n}^{n} \binom{j}{n} a_j \end{pmatrix}$$

The following line-by-line comparison leads to $\mathbf{V}(b_0, \ldots, b_n) \leq \mathbf{V}(a_0, \ldots, a_n)$:

$$\mathbf{V}(a_{n+1,n}) = \mathbf{V}(a_{n,n}), \tag{1.1.2}$$

$$\mathbf{V}(a_{n,n-1}, a_{n+1,n}) = \mathbf{V}(a_{n,n-1}, a_{n,n}) \leq \mathbf{V}(a_{n-1,n-1}, a_{n-1,n}), \tag{1.1.3}$$

where in (1.1.3) the $=$ follows from (1.1.2) and $a_{n+1,n} = a_{n,n} = a_n \neq 0$, the $\leq$ is Lemma 1.1.2 i). Equivalently

$$\begin{aligned} \mathbf{V}(a_{n-1,n-2}, a_{n,n-1}, a_{n+1,n}) &\leq \mathbf{V}(a_{n-1,n-2}, a_{n-1,n-1}, a_{n-1,n}) \\ &\leq \mathbf{V}(a_{n-2,n-2}, a_{n-2,n-1}, a_{n-2,n}), \\ &\vdots \\ \mathbf{V}(b_0, \ldots, b_n) = \mathbf{V}(a_{1,0}, a_{2,1}, \ldots, a_{n,n-1}, a_{n+1,n}) &\leq \mathbf{V}(a_{1,0}, \ldots, a_{1,n}) \\ &\leq \mathbf{V}(a_{0,0}, \ldots, a_{0,n}) = \mathbf{V}(a_0, \ldots, a_n). \end{aligned}$$

We have proved the claim for $\delta = 1$. For arbitrary $\delta > 0$, it follows with a coordinate transformation:

$$\begin{aligned}\mathbf{V}(f) &= \mathbf{V}(a_0, a_1, \ldots, a_n) = \mathbf{V}(a_0, a_1\delta, \ldots, a_n\delta^n) = \mathbf{V}(f \circ (\delta X)) \\ &\geq \mathbf{V}(f \circ (\delta X) \circ (X+1)) = \mathbf{V}(f \circ (\delta X + \delta)) = \mathbf{V}(f \circ (\delta X + \delta) \circ (\frac{1}{\delta} X)) \\ &= \mathbf{V}(f \circ (X + \delta)).\end{aligned}$$

□

The next corollary will enable us to "real root counting", i.e., to give an upper bound for the number or real roots in an interval $]\alpha, \beta]$.

Corollary 1.1.4. *Let $f \in \mathbf{Q}[X]$, $\zeta \in \mathbf{Q}$ and $m := \mathrm{rmult}_f(\zeta)$. For every $\alpha, \beta \in \mathbf{Q}$ with $\alpha < \zeta \leq \beta$, we have*

$$\mathbf{V}(f \circ (X + \beta)) \leq \mathbf{V}(f \circ (X + \alpha)) - m.$$

Proof. Lemma 1.1.2 ii). claims

$$\mathbf{V}(g \cdot (X+1)) \leq \mathbf{V}(g) < \mathbf{V}(g \cdot (X-1)).$$

Let $\delta_1 > 0$ and $\delta_2 \geq 0$. With coordinate transformations we get

$$\mathbf{V}(g \cdot (X + \delta_2)) \leq \mathbf{V}(g) \leq \mathbf{V}(g \cdot (X - \delta_1)) - 1.$$

Applying this formula m times leads to

$$\mathbf{V}(g \cdot (X + \delta_2)^m) \leq \mathbf{V}(g) \leq \mathbf{V}(g \cdot (X - \delta_1)^m) - m. \tag{1.1.4}$$

Defining now

$$g := \frac{f}{(X - \zeta)^m},$$

$\delta_1 := \zeta - \alpha > 0$ and $\delta_2 := \beta - \zeta \geq 0$ shows the desired:

$$\begin{aligned}\mathbf{V}(f \circ (X + \beta)) &= \mathbf{V}((g \circ (X + \beta)) \cdot (X + \beta - \zeta)^m) = \mathbf{V}((g \circ (X + \beta)) \cdot (X + \delta_2)^m) \\ &\leq \mathbf{V}(g \circ (X + \beta)) && (1.1.5) \\ &\leq \mathbf{V}(g \circ (X + \alpha)) && (1.1.6) \\ &\leq \mathbf{V}((g \circ (X + \alpha)) \cdot (X - \delta_1)^m) - m && (1.1.7) \\ &= \mathbf{V}((g \circ (X + \alpha)) \cdot (X - (\zeta - \alpha))^m) - m = \mathbf{V}(f \circ (X + \alpha)) - m,\end{aligned}$$

where (1.1.5) and (1.1.7) follow from (1.1.4) while (1.1.6) is Budan's theorem as $\alpha < \beta$. □

Budan's theorem (Ferdinand François Désiré Budan de Boislaurent (1761—1840), [Bor]) is presented in the appendix of [Bud]. According to [Akr], it was communicated for the first time in 1807. Fourier knew the result a few years before Budan. Budan's communication to the Académie des Sciences was examined by Lagrange and Legendre (rapporteur). They considered the paper as essentially correct, but the Académie did not yet publish the result; they waited for Fourier's paper which appeared in 1820 containing a very complicated proof [Fou1]. In 1822 [Bud] was finally published.

Budan's counting of roots is today known as "Budan-Fourier count". While Budan only uses the sequence of coefficients of f, according to [Vin1], Fourier introduces in [Fou2] the sequence of derivatives $(f(a), f'(a), \ldots, f^{(n)}(a))$ which is today known as "Fourier's sequence". Furthermore, Fourier gives a proof for the claim of lemma 1.2.6 iii), which is not mentioned in [Bud] but easily deducible with the method used in the proof of corollary 1.1.4.
[Bud] does not present any definition. Lemma 1.1.2 i) is proved in the same way while the proof of ii) is omitted with a reference to [Seg]. Lemma 1.1.2 iii) and theorem 1.1.3 are proved in the same way, where the inductions are introduced by the cases "$i = n$, $i = n-1$, $i = n-2$ etc.". The statement of corollary 1.1.4 is presented twice. The case $\operatorname{rmult}_f(\zeta) = 1$ is proved using the fact that f changes sign at ζ. Therefore the constant coefficients of $f(X+\beta)$ and $f(X+\alpha)$ differ in signs (choosing α and β near to ζ). This leads to a greater number of sign changes as the highest coefficients match. For even $\operatorname{rmult}_f(\zeta)$ this argument does not suffice. Therefore [Bud] also presents our corollary 1.1.4.
The main part of [Bud] deals with the numerical approximation of real roots using Budan's theorem which is enhanced in [Vin1, Vin2]. We want to present one of their examples.

Example 1.1.5. Budan's method for approximating real roots of a polynomial $f_0(X_0)$: If $f_0(X_0)$ has a real root in $]0,1[$, use the transformation $X_0 = \frac{1}{X_1}$. Then the function

$$\begin{aligned} f_1(X_1) :=& a_0 X_1^n + a_1 X_1^{n-1} + \cdots + a_n = X_1^n (a_0 + a_1 X_1^{-1} + \cdots + a_n X_1^{-n}) \\ =& X_1^n (a_0 + a_1 X_0 + \cdots + a_n X_0^n) = X_1^n f_0(X_0), \end{aligned}$$

which is the polynomial in X_1 with coefficients in reverse direction, has a real root ≥ 1, which can be approximated by substituting integer values for X_1: If $f_1(X_1)$ has a real root in $]\alpha_1, \beta_1[$ with $1 \geq \alpha_1 \geq \beta_1$, then $f_0(X_0)$ has a real root in $]\frac{1}{\beta_1}, \frac{1}{\alpha_1}[$. Using this again with $X_i = \frac{1}{X_{i+1}} + \alpha_i$ and the corresponding real roots in the intervals with integer boundaries $]\alpha_{i+1}, \beta_{i+1}[$ we get the real root of $f_0(X_0)$ as the chain fraction

$$\lim_{n\to\infty} \alpha_1 + \cfrac{1}{\alpha_2 + \ddots + \frac{1}{\alpha_n}}.$$

We consider the polynomial $f_0(X_0) := X_0^2 - 2 \in \mathbb{R}[X]$ and look for a positive real root:

		V
$f_0(X_0) =$	$X_0^2 - 2$	1
$f_0(X_0+1) =$	$X_0^2 + 2X_0 - 1$	1
$f_0(X_0+2) =$	$X_0^2 + 4X_0 + 2$	0

Therefore $f_0(X_0)$ has a real root in $]1,2]$. Let be $\alpha_0 := 1$ and $X_0 = \frac{1}{X_1} + 1$. Then

$$f_1(X_1) = -X_1^2 + 2X_1 + 1 = X_1^2 (X_1^{-2} + 2X_1^{-1} - 1) = X_1^2 f_0(X_0).$$

Now we look for a positive real root of $f_1(X_1)$:

		V
$f_1(X_1) =$	$-X_1^2 + 2X_1 + 1$	1
$f_1(X_1+1) =$	$-X_1^2 + 2$	1
$f_1(X_1+2) =$	$-X_1^2 - 2X_1 + 1$	1
$f_1(X_1+3) =$	$-X_1^2 - 4X_1 - 2$	0

Therefore $f_1(X_1)$ has a real root in $]2,3]$. Let be $\alpha_1 := 2$ and $X_1 = \frac{1}{X_2} + 2$. Then

$$f_2(X_2) = X_2^2 - 2X_2 - 1 = X_2^2(-X_2^{-2} - 2X_2^{-1} + 1) = X_2^2 f_1(X_1).$$

And since $f_2(X_2) = -f_1(X_2)$ it also has a real root in $]2,3]$ and so on. We get

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2+\ddots}}.$$

1.2 Virtual Roots

In this section let $\mathbf{R}$ denote a real closed field (for example $\mathbb{R}$) and $f^{(i)}$ the i-th derivative of a polynomial f. The idea of virtual roots comes from the desire to define for every $n \geq 1$ and $1 \leq k \leq n$ a function

$$\rho_{n,k} : \{f \in \mathbf{R}[X] \text{ monic of degree } n\} = \mathbf{R}^n \to \mathbf{R}$$

such that

$$\rho_{n,k} : \mathbf{R}^n \to \mathbf{R} \text{ is continuous,}$$

for fixed f and for every real root ζ of f, we have for at least one k

$$\rho_{n,k}(f) = \zeta$$

and for $n \geq 2$

$$\rho_{n,1}(f) \leq \rho_{n-1,1}(f') \leq \rho_{n,2}(f) \leq \rho_{n-1,2}(f') \leq \cdots \leq \rho_{n-1,n-1}(f') \leq \rho_{n,n}(f).$$

If f is fixed, we write only $\rho_{n,k}$:

Definition 1.2.1. Let $f \in \mathbf{R}[X]$ be of degree n and $\zeta \in \mathbf{R}$.

i) Let $\rho_{j,0} := -\infty$ and $\rho_{j,j+1} := \infty$ for $0 \leq j \leq n$. For $1 \leq j \leq n$, the j virtual roots of $f^{(n-j)}$,

$$\rho_{j,1} \leq \cdots \leq \rho_{j,j},$$

are defined inductively:

a) For for $1 \leq k \leq j-1$, let be defined the $\rho_{j-1,k}$ in such a way that for $1 \leq k \leq j$,

$$f^{(n-j+1)}(x) f^{(n-j+1)}(y) \geq 0$$

for all $x, y \in \mathbf{R}_{j-1,k} := [\rho_{j-1,k-1}, \rho_{j-1,k}]$ (resp. the half-open interval in case $k \in \{1, j\}$).

b) Then for $1 \leq k \leq j$, the $\rho_{j,k} \in \mathbf{R}_{j-1,k}$ is defined by the inequality

$$|f^{(n-j)}(\rho_{j,k})| \leq |f^{(n-j)}(x)|$$

for all $x \in \mathbf{R}_{j-1,k}$. This is well-defined since $f^{(n-j)}$ is strictly monotone on $\mathbf{R}_{j-1,k}$. Three cases can appear (figure 1.2.1):

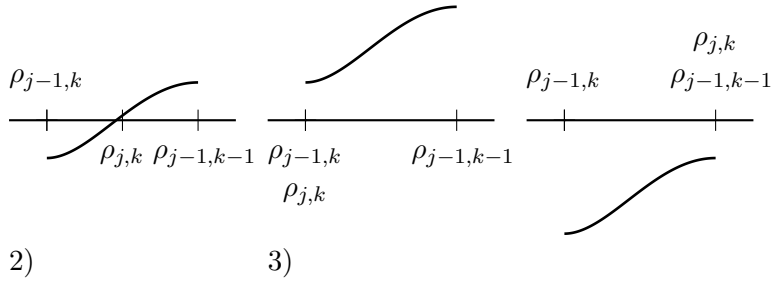


Figure 1.2.1: Definition 1.2.1 i)b)

1) $\rho_{j-1,k-1} = \rho_{j,k} = \rho_{j-1,k}$.
2) $f^{(n-j)}$ admits a real root in $]\rho_{j-1,k-1}, \rho_{j-1,k}[$ then $\rho_{j,k}$ equals this real root.
3) $f^{(n-j)}$ does not admit a real root in $]\rho_{j-1,k-1}, \rho_{j-1,k}[$ then $\rho_{j,k}$ is the point with the least absolute value under $f^{(n-j)}$. Hence it equals either $\rho_{j-1,k-1}$ or $\rho_{j-1,k}$.

c) We get for $1 \leq k \leq j+1$,

$$f^{(n-j)}(x) f^{(n-j)}(y) \geq 0$$

for all $x, y \in \mathbf{R}_{j,k} = [\rho_{j,k-1}, \rho_{j,k}]$ (resp. the half-open interval in case $k \in \{1, j+1\}$).

ii) If ζ is a virtual root of f, the *virtual multiplicity* is, by definition,

$$\mathrm{vmult}_f(\zeta) := l - k + 1$$

choosing k minimal and l maximal such that $\rho_{n,k} = \zeta = \rho_{n,l}$. Otherwise, $\mathrm{vmult}_f(\zeta) := 0$.

iii) If $\mathrm{vmult}_f(\zeta) > \mathrm{rmult}_f(\zeta) = 0$, we call ζ a virtual non-real root of f.

Example 1.2.2. We want to consider the virtual roots of the polynomial $f := X^2 + a \in \mathbb{R}[X]$ with $a \in \mathbb{R}$. For $a \leq 0$, we have $\rho_{2,1}, \rho_{2,2} = \mp\sqrt{a}$ and for $a > 0$, $\rho_{2,1}, \rho_{2,2} = 0$ which is the real root of f'. The question about the relation between virtual and complex roots is evident. Of course, from all complex roots of a polynomial its virtual roots can be deduced. The impossibility of the other direction is shown by this example as different polynomials have the same virtual roots.

Lemma 1.2.3. *If $f \in \mathbf{R}[X]$ of degree n, $\delta \in \mathbf{R}$ and $(a_0, \ldots, a_n)$ the coefficients of $f \circ (X + \delta)$, then we get for all i*

$$\mathrm{sign}(a_i) = \mathrm{sign}(f^{(i)}(\delta)).$$

Proof. This is easily seen using Taylor's formula:

$$f \circ (X + \delta) = \frac{f(\delta)}{0!} + \frac{f'(\delta)}{1!} X + \cdots + \frac{f^{(n)}(\delta)}{n!} X^n \qquad \square$$

With this notation Budan's theorem reads in the following way:

Theorem 1.2.4 (Budan)**.** *Let $\mathbf{Q}$ be a real field, $f \in \mathbf{Q}[X]$ and $x, y \in \mathbf{Q}$, with $x < y$. Then*

$$\mathbf{V}(f(x) \ldots, f^{(n)}(x)) \geq \mathbf{V}(f(y) \ldots, f^{(n)}(y)).$$

We next show that Budan-Fourier counts the virtual roots (which is a different proof for Budan's theorem, by the way):

Theorem 1.2.5. *Let $f \in \mathbf{R}[X]$ be of degree n, $\rho_{n,1} \leq \cdots \leq \rho_{n,n}$ its virtual roots and $\rho_{n,0} = -\infty$, $\rho_{n,n+1} = \infty$. Then we have for $1 \leq k \leq n+1$ with $\rho_{n,k-1} \neq \rho_{n,k}$*

$$x \in [\rho_{n,k-1}, \rho_{n,k}[\qquad \Longleftrightarrow \qquad \mathbf{V}(f(x), f'(x), \ldots, f^{(n)}(x)) = n+1-k$$

(resp. $x \in]-\infty, \rho_{n,1}[$) in case $k = 1$).

Proof. By induction on the degree j of $f^{(n-j)}$. Let $\rho_{j,1} \leq \cdots \leq \rho_{j,j}$ denote the virtual roots of $f^{(n-j)}$ and $\rho_{j,0} = -\infty$, $\rho_{j,j+1} = \infty$.
Let $j = 0$. Then $]\rho_{0,0}, \rho_{0,1}[= \mathbf{R}$ and $\mathbf{V}(f^{(n)}(x)) = 0$ for all $x \in \mathbf{R}$.
Let $j > 0$ and the statement be true for $j-1$. Let $1 \leq k \leq j+1$ with $\rho_{j-1,k-1} \neq \rho_{j-1,k}$ and consider $x \in [\rho_{j-1,k-1}, \rho_{j-1,k}[$. In case i)b)2) of definition 1.2.1, we get

$$\begin{aligned}
f^{(n-j+i)}(x) f^{(n-j)}(x) &< 0 \qquad \text{for} \qquad \rho_{j-1,k-1} = x, \\
f^{(n-j+1)}(x) f^{(n-j)}(x) &< 0 \qquad \text{for} \qquad \rho_{j-1,k-1} < x < \rho_{j,k}, \\
f^{(n-j)}(x) &= 0 \qquad \text{for} \qquad \rho_{j,k} = x, \\
f^{(n-j+1)}(x) f^{(n-j)}(x) &> 0 \qquad \text{for} \qquad \rho_{j,k} < x < \rho_{j-1,k},
\end{aligned}$$

for the smallest $i \geq 1$ with $f^{(n-j+i)}(\rho_{j-1,k-1}) \neq 0$. In case i)b)3), the same argument holds. □

From theorem 1.2.5 we can easily deduce the continuity of the virtual root-functions. Strong means like the mean value theorem show again corollary 1.1.4:

Corollary 1.2.6.

i) Let $f = a_0 + \cdots + a_n X^n \in \mathbf{R}[X]$ be of degree n and $\rho_{n,k}$ its virtual roots. For every k the function $\rho_{n,k} : \mathbf{R}^n \to \mathbf{R}$, considered as function of $(a_0, \ldots, a_{n-1}) \in \mathbf{R}^n$, is continuous w.r.t. the euclidian topology (i.e., the topology generated by the open balls).

ii) For every $a \in \mathbf{R}$

$$\operatorname{rmult}_f(a) \leq \operatorname{vmult}_f(a).$$

iii) For every $a \in \mathbf{R}$

$$\operatorname{vmult}_f(a) - \operatorname{rmult}_f(a) \text{ is even.}$$

iv) (Corollary 1.1.4.) For $x, y \in \mathbf{R}$ with $x < y$

$$0 \leq \sum_{a \in]x,y]} \operatorname{rmult}_f(a) \leq \mathbf{V}(f(x) \ldots, f^{(n)}(x)) - \mathbf{V}(f(y) \ldots, f^{(n)}(y)).$$

v) Let $\zeta < \eta$ be two successive real roots of f. Then

$$\sum_{a \in]\zeta,\eta[} \operatorname{rmult}_{f'}(a) \text{ is odd.}$$

Proof. i) Let $a := \rho_{n,k}(f)$ be the k-th virtual root of f and $\epsilon \in \mathbf{R}$, $\epsilon > 0$ such that $f^{(i)}(a-\epsilon)f^{(i)}(a+\epsilon) \neq 0$ for $0 \leq i \leq n$. Now change the coefficients of f in such a minimal way that (1.2.1) holds and denote the resulting polynomial by $\tilde{f}$.

$$f^{(i)}(a-\epsilon)\tilde{f}^{(i)}(a-\epsilon) > 0 \qquad \text{and} \qquad f^{(i)}(a+\epsilon)\tilde{f}^{(i)}(a+\epsilon) > 0 \tag{1.2.1}$$

for $0 \leq i \leq n$. From theorem 1.2.5 we get $\rho_{n,k}(\tilde{f}) \in]a-\epsilon, a+\epsilon]$.

ii) This follows from the following fact, which can be derived from the mean value theorem, applied inductively on f and its derivatives: Let $g \in \mathbf{R}[X]$ of degree ≥ 1. For every $a \in \mathbf{R}$ exists an $\epsilon > 0$ such that

$$\begin{aligned}(-1)^{\mathrm{rmult}_g(a)} g(x)g(y) &> 0, \\ g(y)g'(y) &> 0\end{aligned} \tag{1.2.2}$$

for every $x \in]a-\epsilon, a[$ and $y \in]a, a+\epsilon[$.

iii) This follows from (1.2.2) and $f^{(n)}(x)f^{(n)}(y) > 0$.

iv) This follows from ii) as

$$\mathbf{V}(f(x)\ldots, f^{(n)}(x)) - \mathbf{V}(f(y)\ldots, f^{(n)}(y)) = \sum_{a \in]x,y]} \mathrm{vmult}_f(a).$$

v) This is also a consequence of the mean value theorem: As $f(a) \neq 0$ for $a \in]\zeta, \eta[$, we have

$$f(x)f(y) > 0 \qquad \text{and} \qquad f'(x)f'(y) < 0$$

for every $x \in]\zeta, \zeta+\epsilon[$ and $y \in]\eta-\epsilon, \eta[$, ϵ sufficiently small. Hence

$$\mathbf{V}(f'(x)\ldots, f^{(n)}(x)) - \mathbf{V}(f'(y)\ldots, f^{(n)}(y)) = \sum_{a \in]x,y]} \mathrm{vmult}_{f'}(a)$$

is odd. And, according to iii), so is also

$$\sum_{a \in]x,y]} \mathrm{rmult}_{f'}(a).$$

□

The next lemma provides some information about what happens to the virtual multiplicity when integrating the polynomial. While the real multiplicity of a real root a of f can become zero when integrating f, the virtual multiplicity can only decrease about one:

Lemma 1.2.7. *Let $f \in \mathbf{R}[X]$, $a \in \mathbf{R}$. The following cases and only them can appear:*

i) $\mathrm{rmult}_f(a) = 0 = \mathrm{rmult}_{f'}(a)$ *and* $\mathrm{vmult}_f(a) = \mathrm{vmult}_{f'}(a)$*;*

ii) $\mathrm{rmult}_f(a) = \mathrm{rmult}_{f'}(a) + 1$ *and* $\mathrm{vmult}_f(a) = \mathrm{vmult}_{f'}(a) + 1$*;*

iii) $\mathrm{rmult}_f(a) = 0 < \mathrm{rmult}_{f'}(a)$ *and*
$\mathrm{vmult}_f(a) - \mathrm{vmult}_{f'}(a) \in \{-1, 0, 1\}$,

where the additional condition that as well $\mathrm{vmult}_f(a) - \mathrm{rmult}_f(a)$ *as* $\mathrm{vmult}_{f'}(a) - \mathrm{rmult}_{f'}(a)$ *is even has to be fulfilled in every case.*

Proof. This follows from the definitons and corresponding examples. □

Obviously, every virtual root of f is a real root of f or one of its derivatives. The question if a real root of a derivative of f is a virtual root of f is answered by the following lemma:

Lemma 1.2.8. *Let* $f \in \mathbf{R}[X]$ *be of degree* n *and* $a \in \mathbf{R}$. *Let* m *be the number of* $i \in \{0, \dots, n\}$ *for which the following holds:* $f^{(i)}(a) = 0$ *and it exists an* $\epsilon > 0$ *such that*

$$\begin{aligned} f^{(i-1)}(y)f^{(i)}(y) &> 0 \\ f^{(i)}(x)f^{(i)}(y) &< 0 \\ f^{(i+1)}(y)f^{(i)}(y) &> 0 \end{aligned}$$

for every $x \in]a - \epsilon, a[$ *and* $y \in]a, a + \epsilon[$. *Then*

$$\mathrm{vmult}_f(a) = \begin{cases} 2m & \text{if } f(a) \neq 0, \\ 2m + 1 & \text{if } f(a) = 0. \end{cases}$$

Proof. This follows by induction on the degree of f and lemma 1.2.7. □

Example 1.2.9. We want to mention random polynomials as in [BG]. I.e., polynomials whose coefficients are distributed according to a certain continuous probability distribution. For a random polynomial f, we have that two different derivatives of f (f itself included) do not have a real root in common. Hence the virtual multiplicity is always ≤ 2.

The next lemmata give easy bounds for virtual roots and a characterization:

Lemma 1.2.10. *Let* $f \in \mathbf{R}[X]$ *be a monic polynomial of degree* n.

i) If $f(X) = a_0 + a_1 X + \cdots + X^n$, *then*

$$-\max_i |a_i| - 1 < \eta < \max_i |a_i| + 1$$

for every virtual root η *of* f.

ii) If $f(X) = \prod_{1 \leq i \leq n_1}(X - \zeta_i) \prod_{1 \leq j \leq n_2}((X - d_j)^2 + e_j^2)$, *then*

$$\min_{i,j}(\zeta_i, d_j) \leq \eta \leq \max_{i,j}(\zeta_i, d_j)$$

for every virtual root η *of* f.

Proof. i) Let $M \geq \max_i |a_i|$.

$$\begin{aligned} f(M+1) =& a_0 + \cdots + a_{n-1}(M+1)^{n-1} + (M+1)^n \\ =& a_0 + \cdots + (a_{n-1} + M + 1)(M+1)^{n-1} \\ \geq& a_0 + \cdots + a_{n-2}(M+1)^{n-2} + (M+1)^{n-1} \\ & \vdots \\ \geq& a_0 + (M+1) \geq 1 \end{aligned}$$

Therefore $f(X)$ has no real root for $X \geq \max_i |a_i| + 1$ or $X \leq -\max_i |a_i| - 1$ (for the left boundary consider $f(-X)$ resp. $-f(-X)$, if n is odd). We still have to consider the derivatives of f; they also have no real roots outside the boundaries, since the absolute values of the coefficients of $f'(X) = \frac{1}{n}a_1 + \frac{2}{n}a_2X + \cdots + \frac{n-1}{n}a_{n-1}X^{n-2} + X^{n-1}$ are smaller, than those of f. The claim follows, since at every virtual root, one of the derivatives has a real root.

ii) As $\mathbf{R}$ is real closed f is the product of linear and quadratic factors. Denote them by $g_i(X) := X - \zeta_i$ and $h_j(X) := (X - d_j)^2 + e_j^2$ with ζ_i, d_j, $e_j \in \mathbf{R}$. We consider the g_i, h_j and its first/second derivatives: $g_i = X - \zeta_i$, $g_i' = 1$ and $h_j = (X - d_j)^2 + e_j^2$, $h_j' = 2(X - d_j)$, $h_j'' = 2$. All of them are > 0 for $X > \max_{i,j}(\zeta_i, d_j)$. The k-th derivative of f can be written as

$$f^{(k)} = \sum_{\substack{\lambda=(l_1,\ldots,l_{n_1},m_1,\ldots,m_{n_2}) \\ l_1+\cdots+l_{n_1}+m_1+\cdots+m_{n_2}=k}} \left(c_\lambda \prod_i g_i^{(l_i)} \prod_j h_j^{(m_j)} \right) \tag{1.2.3}$$

for certain $c_\lambda \geq 1$. Therefore for $k \in [0 \ldots n]$ we get $f^{(k)}(X) > 0$ for $X > \max_{i,j}(\zeta_i, d_j)$ (as at least one summand of (1.2.3) is > 0). This gives the boundary on the right: for the boundary on the left we consider $(-1)^n f(-X)$. It follows from the preceding that $((-1)^n f(-X))^{(k)}(X) > 0$ for $X > -\min_{i,j}(\zeta_i, d_j)$, and therefore $(f(X))^{(k)}(X)$ has constant sign for $X < \min_{i,j}(\zeta_i, d_j)$. The claim follows. □

Lemma 1.2.11. *For every $n \geq 1$ and $1 \leq k \leq n$, let $\tau_{n,k} : \{f \in \mathbf{R}[X]$ of degree $n\} \to \mathbf{R}$ be functions. The following assertions are equivalent:*

i) For fixed f and for every real root ζ of f, we have for at least one k

$$\tau_{n,k}(f) = \zeta, \tag{1.2.4}$$

for the number $m_f(\zeta)$, defined in correspondance to $\mathrm{vmult}_f(\zeta)$ in definition 1.2.1 ii), we have

$$m_f(\zeta) - \mathrm{rmult}_f(\zeta) \text{ is even} \tag{1.2.5}$$

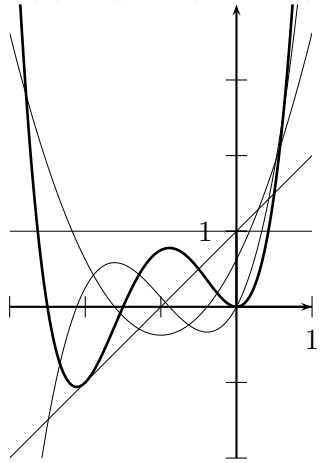
and for $n \geq 2$

$$\tau_{n,1}(f) \leq \tau_{n-1,1}(f') \leq \tau_{n,2}(f) \leq \tau_{n-1,2}(f') \leq \cdots \leq \tau_{n-1,n-1}(f') \leq \tau_{n,n}(f). \tag{1.2.6}$$

ii) $\tau_{n,k}(f) = \rho_{n,k}(f)$.

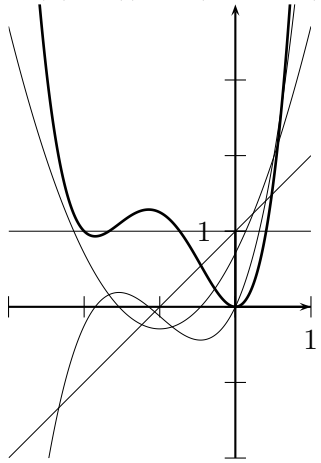
Proof. ii) $\Rightarrow$ i) is clear. i) $\Rightarrow$ ii) is shown by induction on the degree n. In case $n = 1$ (1.2.4) fixes $\tau_{1,1}(f) = \rho_{1,1}(f)$. $n - 1 \to n$: For $j \leq n - 1$ we already have $\tau_{j,k}(f) = \rho_{j,k}(f)$. Consider the intervals $\mathbf{R}_{n-1,k}(f) := [\tau_{n-1,k-1}(f), \tau_{n-1,k}(f)]$ (resp. the half-open interval in case $k \in \{1, n\}$) as in definition 1.2.1 i). (1.2.6) is $\tau_{n,k}(f) \in \mathbf{R}_{n-1,k}(f)$. This fixes the $\tau_{n,k}(f)$ for which $\tau_{n-1,k-1}(f) = \tau_{n-1,k}(f)$. (1.2.4) fixes the $\tau_{n,k}(f)$ for which f admits a real root in the inner of $\mathbf{R}_{n-1,k}(f)$. And from (1.2.5) we get a unique way to put the remaining $\tau_{n,k}(f)$ at one of the boundaries of $\mathbf{R}_{n-1,k}(f)$. □

i) $f_i(x) = ((x+2)^2 - 0.25)x^2$



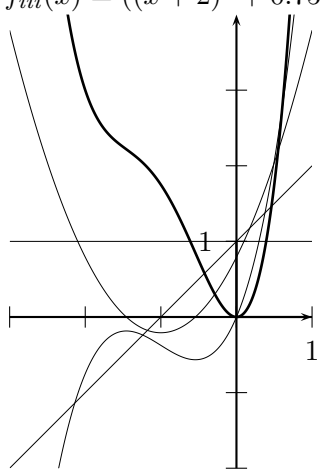
$$\left(\begin{array}{c|ccc|cccc|c}
+ & - & - & - & + & + & + & + & + \\
- & - & + & + & + & + & - & - & + \\
+ & + & + & - & - & - & - & + & + \\
- & - & - & - & - & + & + & + & + \\
+ & + & + & + & + & + & + & + & + \\
\hline
4 & 3 & 3 & 3 & 2 & 2 & 2 & 2 & 0
\end{array}\right)$$

ii) $f_{ii}(x) = ((x+2)^2 + 0.25)x^2$



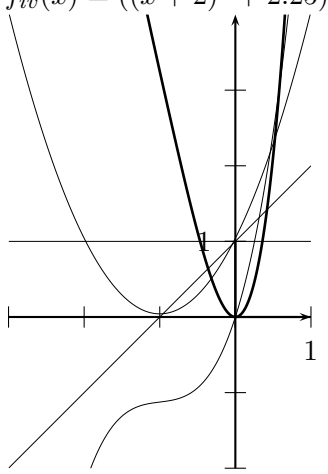
$$\left(\begin{array}{c|cc|ccc|c}
+ & + & + & + & + & + & + \\
- & + & + & - & - & - & + \\
+ & + & - & - & - & + & + \\
- & - & - & - & + & + & + \\
+ & + & + & + & + & + & + \\
\hline
4 & 2 & 2 & 2 & 2 & 2 & 0
\end{array}\right)$$

iii) $f_{iii}(x) = ((x+2)^2 + 0.75)x^2$



$$\left(\begin{array}{c|ccc|c}
+ & + & + & + & + \\
- & - & - & - & + \\
+ & - & - & + & + \\
- & - & + & + & + \\
+ & + & + & + & + \\
\hline
4 & 2 & 2 & 2 & 0
\end{array}\right)$$

iv) $f_{iv}(x) = ((x+2)^2 + 2.25)x^2$



$$\left(\begin{array}{c|c|c}
+ & + & + \\
- & - & + \\
+ & + & + \\
- & + & + \\
+ & + & + \\
\hline
4 & 2 & 0
\end{array}\right)$$

Figure 1.2.2: Example 1.2.12

Example 1.2.12. Figure 1.2.2 shows the graphs of four polynomials $\mathbb{R} \to \mathbb{R}$ of degree 4 (thick lines) and their monic (i.e., $f_i'/1!$, $f_i^{(2)}/2!$...) derivatives (thin lines). We call the table below the graph of f_i the *Budan table* of f_i. In the first (top) row the vertical lines represent the real roots of f_i. The $-$ resp. $+$ signs display the intervals where f_i admits negative resp. positive values. In the second row the real roots of f_i' are displayed, and so on. The bottom row shows

$$\mathbf{V}(f_i(x), f_i'(x), f_i^{(2)}(x), f_i^{(3)}(x), f_i^{(4)}(x)).$$

The virtual roots of f_i are located at the positions where these numbers $\mathbf{V}$ differ and are represented by vertical lines.

i) f_i admits 4 real roots:

x	$\mathrm{vmult}_{f_i}(x)$	$\mathrm{rmult}_{f_i}(x)$	virtual roots of f_i located at x
-2.50	1	1	$\rho_{4,1}$
-1.50	1	1	$\rho_{4,2}$
0.00	2	2	$\rho_{4,3}$, $\rho_{4,4}$

We want to consider the virtual roots $\rho_{4,1}$ and $\rho_{4,2}$ of the polynomial $((x+2)^2+c)x^2$ as functions of $-0.25 \le c \le 0.25$. For $-0.25 \le c \le 0$, we have the two real roots $-2.5 \le \rho_{4,1} \le 2$ and $-1.5 \ge \rho_{4,2} \ge 2$. For $c = 0$, we have the double real root $\rho_{4,1} = 0 = \rho_{4,2}$. And for $0 < c \le 0.25$, we have a constant pair of virtual roots which are no longer real roots at $\rho_{4,1} = 0 = \rho_{4,2}$.

ii) f_{ii} admits 2 real roots:

x	$\mathrm{vmult}_{f_{ii}}(x)$	$\mathrm{rmult}_{f_{ii}}(x)$	virtual roots of f_{ii} located at x
≈ -1.58	2	0	$\rho_{4,1}$, $\rho_{4,2}$
0.00	2	2	$\rho_{4,3}$, $\rho_{4,4}$

According to lemma 1.2.8, $\rho_{4,1} = \rho_{4,2}$ are located at the smallest real root of f_{ii}'. In this case, f_{ii} has a minimum with a positive value.

iii) f_{iii} admits 2 real roots:

x	$\mathrm{vmult}_{f_{iii}}(x)$	$\mathrm{rmult}_{f_{iii}}(x)$	virtual roots of f_{iii} located at x
≈ -1.46	2	0	$\rho_{4,1}$, $\rho_{4,2}$
0.00	2	2	$\rho_{4,3}$, $\rho_{4,4}$

Here, $\rho_{4,1} = \rho_{4,2}$ are located at the smallest real root of $f_{iii}^{(2)}$. No minimum is visible anymore.

iv) f_{iv} admits 2 real roots:

x	$\mathrm{vmult}_{f_{iv}}(x)$	$\mathrm{rmult}_{f_{iv}}(x)$	virtual roots of f_{iv} located at x
-1.00	2	0	$\rho_{4,1}$, $\rho_{4,2}$
0.00	2	2	$\rho_{4,3}$, $\rho_{4,4}$

Here, $\rho_{4,1} = \rho_{4,2}$ are located at the real root of $f_{iv}^{(3)}$.

At the end, let us consider the connected components of a Budan table. I.e., the subsets of places in the Budan table for which every two of them can be connected by horizontal or vertical steps visiting only places with the same sign. It is easily seen that every connected component is unbounded to the left. And every one which is bounded to the right causes a virtual root at its rightmost point. (Compare [BG].)

Example 1.2.13. We want to consider the product of two polynomials f and g. In general, the set of virtual roots of fg does not contain that of f. Furthermore, multiplication can change the order of the real roots in the following sense: Consider

$$f := 0.1X^4 + 0.6X^3 - X^2 + 0.6X + 0.1, \qquad g := X^2 + X + 1 \in \mathbb{R}[X]$$

(figure 1.2.3; solid lines) and their product fg (dotted line).

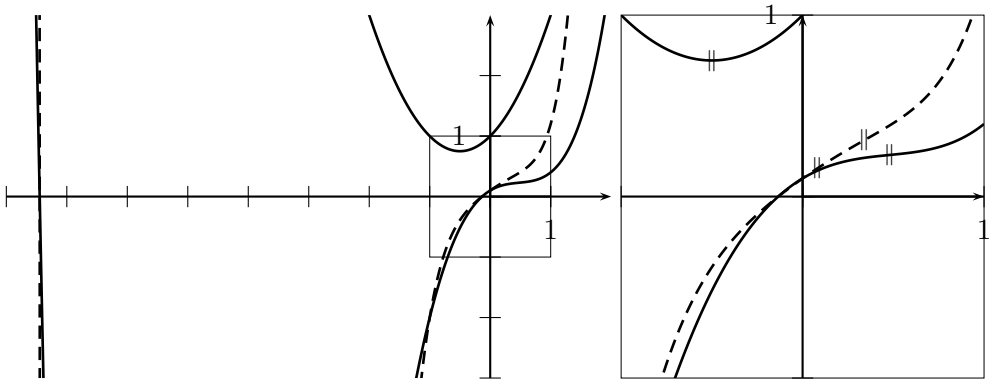


Figure 1.2.3: Example 1.2.13

x	virtual roots of f at x	virtual roots of g at x	virtual roots of g at x
≈ -7.45	$\rho_{4,1}(f)$		$\rho_{6,1}(fg)$
-0.50		$\rho_{2,1}(g)$, $\rho_{2,2}(g)$	
≈ -0.13	$\rho_{4,2}(f)$		$\rho_{6,2}(fg)$
≈ 0.08			$\rho_{6,3}(fg)$, $\rho_{6,4}(fg)$
≈ 0.34			$\rho_{6,5}(fg)$, $\rho_{6,6}(fg)$
≈ 0.48	$\rho_{4,3}(f)$, $\rho_{4,4}(f)$		

We see that in the ordered union of the virtual roots of f and g,

$$\rho_{4,1}(f) < \rho_{2,1}(g) = \rho_{2,2}(g) < \rho_{4,2}(f) < \rho_{4,3}(f) = \rho_{4,4}(f),$$

the real roots appear at position 1 and 4 while $\rho_{6,1}(fg)$, $\rho_{6,2}(fg)$ are the real roots of fg.

Chapter 2

An algebraic certificate for Budan's theorem

2.1 What is an algebraic certificate for Budan's theorem?

According to [CLR], an algebraic certificate (certificate for short) is a proof of a certain claim by precisely algebraic identities (equalities, inequalities).
In chapters 2 and 3 we present two algorithms which calculate algebraic certificates for Budan's theorem. Both of them receive as input data:
$n \in \mathbb{N}$ and two sequences of sign conditions

$$(\sigma_0, \ldots, \sigma_n),\ (\tilde{\sigma}_0, \ldots, \tilde{\sigma}_n) \in \{-1, 0, +1\}^{n+1} \tag{2.1.1}$$

such that

$$\mathbf{V}(\sigma_0, \ldots, \sigma_n) < \mathbf{V}(\tilde{\sigma}_0, \ldots, \tilde{\sigma}_n).$$

They calculate:
For $0 \leq i \leq n$ some coefficients z_i, $\tilde{z}_i \in \mathbb{Z}$ such that

$$\sigma_i z_i \geq 0, \qquad \tilde{\sigma}_i \tilde{z}_i \geq 0, \tag{2.1.2}$$

where at least one of the inequalities in (2.1.2) is a strict one, and

$$\sum_i z_i \frac{f^{(i)}(0)}{i!} + \sum_i \tilde{z}_i \frac{f^{(i)}(1)}{i!} = 0 \tag{2.1.3}$$

for every polynomial $f \in \mathbf{Q}[X]$ of degree n over an arbitrary ordered field $\mathbf{Q}$.
This result means the following: Let $a < b \in \mathbf{Q}$. After making the coordinate transformation $g(X) := f((X-a)/(b-a))$, the assumptions (2.1.2), (2.1.3) and

$$\operatorname{sign}(g^{(i)}(a)) = \sigma_i, \qquad \operatorname{sign}(g^{(i)}(b)) = \tilde{\sigma}_i \tag{2.1.4}$$

for all i lead to the contradiction $0 < 0$. This contradiction proves the claim of Budan's theorem in the form of theorem 1.2.4 for the special sign conditions (2.1.1) and every polynomial $g \in \mathbf{Q}[X]$ of degree n.

We speak of a linear certificate since (2.1.3) is linear combination of the $f^{(i)}(0)$ and $f^{(i)}(1)$. For linear incompatibilities and linear certificates consider [CLR, Schr].
Our situation is related to the Baby Positivstellensatz ([CLR], theorem 5.7) which claims in a more general context: The impossibility to find a polynomial $g \in \mathbf{Q}[X]$, such that (2.1.4) holds, implies the existence of an equality like (2.1.3). And, according to Budan's theorem, such a g is impossible. Therefore the existence of a linear certificate follows from the Baby Positivstellensatz. Furthermore, the means to calculate the certificate are provided by linear programming in an ordered group [Schr]. Nevertheless, we consider our certificates interesting as in a constructive context it is interesting in which way things are proved resp. in which way algorithms calculate.
Our algorithm in chapter 3 is based on mixed Taylor series [Lom2] and polynomials $\prod_{k=0}^{i-1}(X-k) \in \mathbb{R}[X]$ and calculates in polynomial time in the degree of g.
The algorithm in section 2.2 is based on the historical proof by Budan and has exponential complexity. This shows, furthermore, the general difference of our two algorithms.
In section 2.3 we present a further algorithm to calculate certificates. But this algorithm itself does not deliver arguments for a non-empty result and therefore it can not be considered as proof for Budan's theorem. In the naive version presented in section 2.3 this algorithm also has a bad complexity.

2.2 An algebraic certificate for Budan's theorem

Definition 2.2.1. Let be $2 \leq n \in \mathbb{N}$.

i) Let $\mathrm{Ab}(a_1, \ldots, a_n)$ denote the free abelian group generated by $a_1, \ldots, a_n$.

ii) Let
$I_n := \{0, \ldots, n\} \times \{1, \ldots, n\}$ denote the index set of the Budan matrix;
$I_{n\triangledown} := \{(i,j) \in I_n | i + j \leq n+1\}$ the index set of the upper left triangle;
$I_{n}- := \{(i,j) \in I_n | i = 0\}$ the index set of the top row;
$I_{n/} := \{(i,j) \in I_n | i + j = n+1\}$ the index set of the diagonal and
$I_{n\nearrow} := I_{n}- \cup I_{n/}$.

iii) Let the *Budan matrix* of dimension n be the $(n+1)\times n$-matrix

$$\begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_n \\ \sum\limits_{k=1}^{1}\binom{1-k}{0}a_k & \sum\limits_{k=1}^{2}\binom{2-k}{0}a_k & \sum\limits_{k=1}^{3}\binom{3-k}{0}a_k & \dots & \sum\limits_{k=1}^{n}\binom{n-k}{0}a_k \\ \sum\limits_{k=1}^{1}\binom{2-k}{1}a_k & \sum\limits_{k=1}^{2}\binom{3-k}{1}a_k & \ddots & \sum\limits_{k=1}^{n-1}\binom{n-k}{1}a_k & 0 \\ \sum\limits_{k=1}^{1}\binom{3-k}{2}a_k & \ddots & \iddots & \iddots & \vdots \\ \vdots & \sum\limits_{k=1}^{2}\binom{n-k}{n-2}a_k & \iddots & & \vdots \\ \sum\limits_{k=1}^{1}\binom{n-k}{n-1}a_k & 0 & \dots & \dots & 0 \end{pmatrix} \tag{2.2.1}$$

$$=: \big(\alpha_{i,j}\big)_{(i,j)\in I_n} \in \mathrm{Ab}(a_1,\dots,a_n)^{(n+1)\times n}.$$

iv) For $A = I_{n\triangledown}$ (resp. $I_{n\triangleright}$), let a *sign condition* σ on A be a map

$$\sigma : A \to \{-1,0,1\} \subset \mathbb{Z}, \quad (i,j) \mapsto \sigma_{i,j}.$$

v) Two pairs (i,j), $(\tilde{\imath},\tilde{\jmath}) \in I_{n\triangledown}$ are said to lie in the same *connected component* w.r.t. a sign condition σ on $I_{n\triangledown}$ (in symbols: $(i,j) \sim_\sigma (\tilde{\imath},\tilde{\jmath})$) if there exists a *path*

$$\big((i,j) = (i_0,j_0),\ (i_1,j_1),\dots,(i_m,j_m) = (\tilde{\imath},\tilde{\jmath})\big)$$

in $I_{n\triangledown}$ with $\sigma_{i_k,j_k} = \sigma_{i,j}$ for all k and

$$(i_{k+1},j_{k+1}) \in \{(i_k+1,j_k),\ (i_k-1,j_k),\ (i_k,j_k+1),\ (i_k,j_k-1),\ (i_k,j_k)\}$$

for all $k < m$.

vi) For a sign condition σ on $I_{n\triangledown}$, let the *connected component* of $(i,j) \in I_{n\triangledown}$ be

$$C(i,j) := \{(\tilde{\imath},\tilde{\jmath}) \in I_{n\triangledown} | (\tilde{\imath},\tilde{\jmath}) \sim_\sigma (i,j)\}.$$

vii) For $A = I_{n\triangledown}$ (resp. $I_{n\triangleright}$) and a sign condition σ on A with $\sigma_{i,j} \neq 0$ for at least one $(i,j) \in I_{n/}$, let a *linear incompatibility* (resp. a *linear certificate*) z w.r.t. σ be a map

$$z : A \to \mathbb{Z}, \quad (i,j) \mapsto z_{i,j}$$

such that $\sigma_{i,j}z_{i,j} \geq 0$ for all $(i,j) \in A$, $\sigma_{i,j}z_{i,j} > 0$ for at least one $(i,j) \in I_{n/}$ and

$$\sum_{(i,j)\in A} z_{i,j}\alpha_{i,j} = 0$$

in $\mathrm{Ab}(a_1,\dots,a_n)$.

Remark 2.2.2.

i) Notice that the top row of the Budan matrix is subscripted $(0,1),\dots,(0,n)$.

ii) For all following considerations exclusively elements of the upper left triangle of the Budan matrix $I_{n\triangledown}$ play a role.

iii) For example the Budan matrix for $n := 5$.

$$\begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ a_1 & a_1+a_2 & a_1+a_2+a_3 & a_1+a_2+a_3+a_4 & a_1+a_2+a_3+a_4+a_5 \\ a_1 & 2a_1+a_2 & 3a_1+2a_2+a_3 & 4a_1+3a_2+2a_3+a_4 & 0 \\ a_1 & 3a_1+a_2 & 6a_1+3a_2+a_3 & 0 & 0 \\ a_1 & 4a_1+a_2 & 0 & 0 & 0 \\ a_1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

iv) $\alpha_{(i,j)} = \alpha_{(i,j-1)} + \alpha_{(i-1,j)}$ for all $(i,j) \in I_{n\triangledown}$ with $i \geq 1$ and $j \geq 2$.

v) $(i,j) \sim_\sigma (\tilde{\imath},\tilde{\jmath})$ means that there is a path through $I_{n\triangledown}$ from (i,j) to $(\tilde{\imath},\tilde{\jmath})$ for which every move is either up, down, to the left or to the right (not diagonal) and every visited element has the same sign (0 is regarded as sign distinct from ± 1).

vi) $\sim_\sigma$ is an equivalence relation; therefore the

$$\{C(i,j) \in \mathcal{P}(I_{n\triangledown}) | (i,j) \in I_{n\triangledown}\}$$

define a partition of $I_{n\triangledown}$.

Let K be an ordered field and h a homomorphism from $\mathrm{Ab}(a_1,\dots,a_n)$ to the additive group of K.

vii) For every polynomial $f(x) := h(a_1)x^{n-1} + \cdots + h(a_{n-1})x + h(a_n) \in K[X]$, the sequence

$$\Big(\frac{f(0)}{0!}, \frac{f'(0)}{1!}, \dots, \frac{f^{(n-1)}(0)}{(n-1)!}\Big) = (h(a_n), h(a_{n-1}), \dots, h(a_1))$$

appears in the top row and the sequence

$$\Big(\frac{f(1)}{0!}, \frac{f'(1)}{1!}, \dots, \frac{f^{(n-1)}(1)}{(n-1)!}\Big) = (h(\alpha_{1,n}), h(\alpha_{2,n-1}), \dots, h(\alpha_{n,1}))$$

in the diagonal of the Budan matrix.

viii) A linear incompatibility (resp. linear certificate) expresses a contradiction to the hypothesis

$$\mathrm{sign}(h(\alpha_{i,j})) = \sigma_{i,j}$$

for all $(i,j) \in I_{n\triangledown}$ (resp. $I_{n\triangleright}$) as it claims $0 < \sum z_{i,j} h(\alpha_{i,j}) = 0$.

Lemma 2.2.3. *Let σ be a sign condition on $I_{n\triangledown}$ such that the sequence $(\sigma_{n,1}, \sigma_{n-1,2}, \ldots, \sigma_{1,n})$ has more sign changes than $(\sigma_{0,1}, \ldots, \sigma_{0,n})$ (in particular, $\sigma_{i,j} \neq 0$ for at least one $(i,j) \in I_{n/}$.) Then there exists a linear incompatibility*

$$z : I_{n\triangledown} \to \mathbb{Z}$$

w.r.t. σ.

Proof. First, we want to prove the following claim: There exists at least one $(\tilde{\imath}, \tilde{\jmath}) \in I_{n/}$ with $\sigma_{\tilde{\imath},\tilde{\jmath}} \neq 0$ and $\min_{(i,j)\in C(\tilde{\imath},\tilde{\jmath})}(i) > 0$, i.e., there is at least one nonzero connected component which touches the diagonal and does not touch the top row.
Proof by contradiction. Supposed $\min(i) = 0$ for all nonzero $(\tilde{\imath}, \tilde{\jmath}) \in I_{n/}$. Then for arbitrary $(\tilde{\imath}, \tilde{\jmath})$, $(\hat{\imath}, \hat{\jmath}) \in I_{n/}$ with opposite nonzero signs and $\tilde{\jmath} < \hat{\jmath}$ and arbitrary $(0, \tilde{\jmath}') \in C(\tilde{\imath}, \tilde{\jmath})$, $(0, \hat{\jmath}') \in C(\hat{\imath}, \hat{\jmath})$ we get $\tilde{\jmath}' < \hat{\jmath}'$.
(Otherwise let be $(\tilde{\imath}_k, \tilde{\jmath}_k)_k \in C(\tilde{\imath}, \tilde{\jmath})$ resp. $(\hat{\imath}_k, \hat{\jmath}_k)_k \in C(\hat{\imath}, \hat{\jmath})$ two pathes which connect $(\tilde{\imath}, \tilde{\jmath})$ with $(0, \tilde{\jmath}')$ resp. $(\hat{\imath}, \hat{\jmath})$ with $(0, \hat{\jmath}')$. W.l.o.g. $\tilde{\imath}_k = \hat{\imath}_k$ for all k with $\tilde{\imath}_k \leq \hat{\imath}$ and $(\hat{\imath}_k, \hat{\jmath}_k) = (\hat{\imath}, \hat{\jmath})$ for all k with $\tilde{\imath}_k > \hat{\imath}$ — otherwise this can be achieved by inserting additional steps since both connected components contain some (i,j) for every $i \leq \hat{\imath}$ (here we use that they both touch the top row). Now let k_0 be the smallest k for which $\tilde{\jmath}_k > \hat{\jmath}_k$. Then $\tilde{\jmath}_{k_0-1} = \tilde{\jmath}_{k_0} - 1 = \hat{\jmath}_{k_0}$ which is a contradiction since the two connected components have different signs. (Figure 2.2.1.)) This means that two connected components cannot cross each other. It follows that

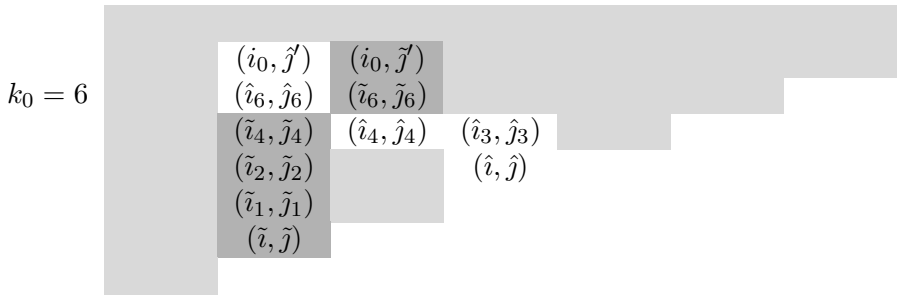


Figure 2.2.1: Proof of lemma 2.2.3; connected components cannot cross.

we get for every pair $(\tilde{\imath}, \tilde{\jmath})$, $(\hat{\imath}, \hat{\jmath})$ in the diagonal as above, a pair $(0, \tilde{\jmath}')$, $(0, \hat{\jmath}')$ in the top row as above which means that the top row has at least as many sign changes as the diagonal which was excluded by the assumption, and the first claim is proven.
Next, we take such an $(\tilde{\imath}, \tilde{\jmath}) \in I_{n/}$ with $\sigma_{\tilde{\imath},\tilde{\jmath}} \neq 0$ and $\min_{(i,j)\in C(\tilde{\imath},\tilde{\jmath})}(i) > 0$. Every $\alpha_{i,j}$ with $(i,j) \in C(\tilde{\imath}, \tilde{\jmath})$ and $j \geq 2$ can be written as the sum of the element one column to the left and the element one row up: $\alpha_{i,j} = \alpha_{i,j-1} + \alpha_{i-1,j}$; and every $\alpha_{i,1}$ equals the element one row above: $\alpha_{i,1} = \alpha_{i-1,1}$. It follows by induction that every $\alpha_{i,j}$ with $(i,j) \in C(\tilde{\imath}, \tilde{\jmath})$ — in particular $\alpha_{\tilde{\imath},\tilde{\jmath}}$ himself — can be written as sum of direct neighbors of $C(\tilde{\imath}, \tilde{\jmath})$:

$$\alpha_{\tilde{\imath},\tilde{\jmath}} = \sum_{(i,j)\in D(\tilde{\imath},\tilde{\jmath})} n_{i,j}\alpha_{i,j} \tag{2.2.2}$$

with $n_{i,j} \in \mathbb{N}$ (including 0) and

$$D(\tilde{\imath}, \tilde{\jmath}) := \{(i,j) \in I_{n\triangledown} \setminus C(\tilde{\imath}, \tilde{\jmath}) | (i, j+1) \in C(\tilde{\imath}, \tilde{\jmath}) \vee (i+1, j) \in C(\tilde{\imath}, \tilde{\jmath})\}.$$

And from the connectedness of $C(\tilde{\imath},\tilde{\jmath})$ follows that $\sigma_{i,j}\sigma_{\tilde{\imath},\tilde{\jmath}} \leq 0$ for $(i,j) \in D(\tilde{\imath},\tilde{\jmath})$. Therefore (2.2.2) leads with

$$\begin{aligned} z_{\tilde{\imath},\tilde{\jmath}} &:= \sigma_{\tilde{\imath},\tilde{\jmath}}, & \\ z_{i,j} &:= \sigma_{i,j} n_{i,j} & \text{for } (i,j) \in D(\tilde{\imath},\tilde{\jmath}), \\ z_{i,j} &:= 0 & \text{otherwise} \end{aligned}$$

to

$$\sum_{(i,j)\in I_{n\triangledown}} z_{i,j}\alpha_{i,j} = 0.$$

Since $\sigma_{i,j} z_{i,j} \geq 0$ for all $(i,j) \in I_{n\triangledown}$, $\sigma_{\tilde{\imath},\tilde{\jmath}} z_{\tilde{\imath},\tilde{\jmath}} = 1$ and $(\tilde{\imath},\tilde{\jmath}) \in I_{n\diagup}$ we are done. □

Example 2.2.4. For $n := 7$, figure 2.2.2 shows a sign condition σ on $I_{7\triangledown}$ with four sign changes in the top row and five in the diagonal. The colors mean:

$\sigma_{i,j} = -1$	$\sigma_{i,j} = 0$	$\sigma_{i,j} = +1$

a	b	c	d	e	f	g
a	$a+b$	$a+b+c$	$a+b+c+d$	$a+b+c+d$ $+e$	$a+b+c+d$ $+e+f$	$a+b+c+d$ $+e+f+g$
a	$2a+b$	$3a+2b+c$	$4a+3b+2c$ $+d$	$5a+4b+3c$ $+2d+e$	$6a+5b+4c$ $+3d+2e+f$	
a	$3a+b$	$6a+3b+c$	$10a+6b+3c$ $+d$	$15a+10b$ $+6c+3d+e$		
a	$4a+b$	$10a+4b+c$	$20a+10b$ $+4c+d$			
a	$5a+b$	$15a+5b+c$				
a	$6a+b$					
a						

Figure 2.2.2: Example 2.2.4.

We have for $(\tilde{\imath},\tilde{\jmath}) := (4,4)$ that $\sigma_{4,4} = 1$ and $\min_{(i,j)\in C(4,4)}(i) = 3 > 0$. Therefore — starting at $(4,4)$ — we can construct the linear incompatibility

$$(20a+10b+4c+d) - (10a+6b+3c+d) - (4a+b) - (3a+b) - (3a+2b+c) = 0,$$

which leads to the contradiction $0 < 0$ since $0 < (20a+10b+4c+d)$ and the other summands are nonnegative.

Lemma 2.2.5. *Let σ^* be a sign condition on $I_{n\diagup}$, and for every sign condition σ on $I_{n\triangledown}$ with $\sigma|_{I_{n\diagup}} = \sigma^*$ let be given a linear incompatibility $z^\sigma : I_{n\triangledown} \to \mathbb{Z}$. Then there exists a linear certificate*

$$z^* : I_{n\diagup} \to \mathbb{Z}$$

w.r.t. σ^*.

Proof. Let Σ denote the sign conditions σ on $I_{n\triangledown}$ with $\sigma|_{I_{n\nearrow}} = \sigma^*$, and let

$$o : \{1, \dots, (n-1)n/2\} \to I_{n\triangledown} \setminus I_{n\nearrow}$$

be an arbitrary order on $I_{n\triangledown} \setminus I_{n\nearrow}$. By induction on k we will show for every $\sigma \in \Sigma$, the existence of a linear incompatibility $z^{\sigma,k}$ with

$$z^{\sigma,k}_{o(l)} = 0$$

for all $1 \le l \le k$.
In the base case ($k := 0$) let be $z^{\sigma,0} := z^{\sigma}$.
In the inductive step ($k-1 \to k$), for every $\sigma \in \Sigma$, a linear incompatibility $z^{\sigma,k-1}$ with $z^{\sigma,k-1}_{o(l)} = 0$ for all $1 \le l < k$ is provided.
Now, let $\sigma \in \Sigma$ be arbitrary and define σ^- (resp. σ^+) by

$$\sigma^-_{i,j} \text{ (resp. } \sigma^+_{i,j}) := \begin{cases} -1 \text{ (resp. } +1) & \text{if } (i,j) = o(k), \\ \sigma_{i,j} & \text{otherwise.} \end{cases}$$

Now, consider $z^{\sigma^-,k-1}$ and $z^{\sigma^+,k-1}$. If $z^{\sigma^-,k-1}_{o(l)} z^{\sigma^+,k-1}_{o(l)} = 0$ we are done. Otherwise we define

$$z^{\sigma,k}_{i,j} := z^{\sigma^+,k-1}_{o(k)} z^{\sigma^-,k-1}_{i,j} - z^{\sigma^-,k-1}_{o(k)} z^{\sigma^+,k-1}_{i,j} \tag{2.2.3}$$

for all $(i,j) \in I_{n\triangledown}$. Since as well $z^{\sigma^+,k-1}_{o(l)}$ as $-z^{\sigma^-,k-1}_{o(l)}$ are positive, we have

$$\begin{aligned} z^{\sigma,k}_{o(k)} &= 0 && \text{by (2.2.3),} \\ z^{\sigma,k}_{o(l)} &= 0 && \text{for } l < k \text{ by induction hypothesis,} \\ \sigma_{i,j} z^{\sigma,k}_{i,j} &\ge 0 && \text{for all } (i,j) \in I_{n\triangledown} \text{ by induction hypothesis,} \\ \sigma_{i,j} z^{\sigma,k}_{i,j} &> 0 && \text{for at least one } (i,j) \in I_{n\diagup} \text{ by induction hypothesis.} \end{aligned}$$

End of inductive step.
Finally we define

$$z^*_{i,j} := z^{\sigma,(n-1)n/2}_{i,j}$$

for $(i,j) \in I_{n\nearrow}$ and an arbitrary sign condition $\sigma \in \Sigma$. Take in mind that $z^{\sigma,(n-1)n/2}_{i,j} = 0$ for $(i,j) \notin I_{n\nearrow}$. Therefore z^* is a linear certificate as desired. □

Lemma 2.2.3 and lemma 2.2.5 together lead to

Theorem 2.2.6. *Let σ^* be a sign condition on $I_{n\nearrow}$ such that the sequence $(\sigma_{n,1}, \sigma_{n-1,2}, \dots, \sigma_{1,n})$ has more sign changes than $(\sigma_{0,1}, \dots, \sigma_{0,n})$. Then there exists a linear certificate $z^* : I_{n\nearrow} \to \mathbb{Z}$ w.r.t. σ^*.*

Example 2.2.7. For the sign condition σ^* on $I_{7\nearrow}$

$$\begin{aligned} (\sigma_{0,1}, \dots, \sigma_{0,7}) &:= (+1, \quad 0, -1, +1, +1, -1, \quad 0), \\ (\sigma_{7,1}, \dots, \sigma_{1,7}) &:= (+1, +1, -1, +1, -1, +1, -1) \end{aligned}$$

which is displayed in figure 2.2.2, is

$$(z_{0,1}, \dots, z_{0,7}, z_{7,1}, \dots, z_{1,7}) := (15, 0, -2, 0, 1, 0, 0, 0, 0, -4, 3, -1, 0, 0)$$

a linear certificate since $\sigma_{i,j} z_{i,j} \geq 0$ for all (i, j), $\sigma_{5,3} z_{5,3} > 0$ and

$$\begin{aligned} 0 <& 15(a) - 2(c) + (e) - 4(15a + 5b + c) + 3(20a + 10b + 4c + d) \\ & - (15a + 10b + 6c + 3d + e) = 0. \end{aligned}$$

2.3 An algorithm to calculate all linear certificates

In this section we present an algorithm of real linear programming to calculate for a fixed sign condition (2.1.1) all possible certificates, i.e., all possible rational $2(n+1)$-tuples $(z_i, \tilde{z}_i)$ for which (2.1.2) and (2.1.3) hold. To do so, we consider (2.1.3) as a linear system whose vector space of solutions is restricted by the sign conditions (2.1.2). Geometrically, this means the intersection of half spaces defined by (2.1.2). The set of solutions is described by the convex hull of finitely many points. Its non-emptiness is not clear from the following algorithm but from the preceding arguments.
We used the computer algebra system PARI/GP [`http://pari.math.u-bordeaux.fr/`] to implement this algorithm. We would like to thank the authors of this helpful tool.

Definition 2.3.1. Let $2 \leq n \in \mathbb{N}$.

i) For $x \in \mathbb{Q}$, let $|x|$ denote its absolute value;

ii) for a matrix a, let a^t denote its transposed;

iii) for $a = (a_1, \dots, a_n)^t$, $b = (b_1, \dots, b_n)^t \in \mathbb{Q}^n$, let $\langle a, b \rangle := \sum_{i=1}^n a_i b_i$ denote the scalar product;

iv) for $0 \neq a \in \mathbb{Q}^n$, let

$$\{\langle a, x \rangle = 0\} := \{x \in \mathbb{Q}^n | \langle a, x \rangle = 0\}$$

denote the hyperplane with normal a and

$$\begin{aligned} \{-\langle a, x \rangle \geq 0\} &:= \{\langle a, x \rangle \leq 0\} := \{x \in \mathbb{Q}^n | \langle a, x \rangle \leq 0\}, \\ \{-\langle a, x \rangle \leq 0\} &:= \{\langle a, x \rangle \geq 0\} := \{x \in \mathbb{Q}^n | \langle a, x \rangle \geq 0\} \end{aligned}$$

the corresponding half spaces (resp. "<", ">");

v) for $X \subset \mathbb{Q}^n$, let

$$\operatorname{conv}(X) := \Big\{ \sum_{k=1}^{l} a_k x_k \Big| x_k \in X,\, 0 \leq a_k \in \mathbb{Q},\, \sum_{k=1}^{l} a_k = 1 \Big\}$$

denote the convex hull of X;

vi) for b, $c \in \mathbb{Q}^n$, let $\overline{bc} := \operatorname{conv}(b, c)$ denote the line between a and b;

vii) for B, $C \subset \mathbb{Q}^n$, let $\overline{BC} := \bigcup_{(b,c)\in B\times C} \overline{bc}$;

viii) let $B : \{1,\dots,2n-1\} \to I_{n\nearrow}\backslash\{(0,1)\}$ denote the following correspondence to the indices of the Budan matrix (2.2.1)

$$\begin{aligned}&\big(B(1)\ ,B(2)\dots B(n)\ ,B(n+1)\dots B(2n-1)\big)\\ :=&\big((n,1),(0,2)\dots(0,n),(n-1,2)\dots(1,n)\qquad\big).\end{aligned}$$

Lemma 2.3.2. *Let σ be a sign condition on $I_{n\nearrow}$ with $\sigma_{0,1} = \sigma_{n,1}$,*

$$\mathbf{U}^n := \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \binom{n-1}{n-2} & \dots & \binom{n-1}{1} & \binom{n-1}{0} \\ 0 & 1 & \ddots & \vdots & \vdots & \binom{n-2}{n-2} & \dots & \binom{n-2}{1} & \binom{n-2}{0} \\ \vdots & 0 & \ddots & 0 & \vdots & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 1 & 0 & \vdots & \ddots & \binom{1}{1} & \binom{1}{0} \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 & \binom{0}{0} \end{pmatrix} \in \mathbb{Q}^{n\times(2n-1)},$$

and let V be the set of $v := (v_1,\dots,v_{2n-1})^t \in \mathbb{Z}^{2n-1}$ for which

$$\mathbf{U}^n v = 0$$

and which fulfill the sign condition σ, i.e.,

$$\sigma_{B(k)} v_k \geq 0 \tag{2.3.1}$$

for all k.
Then we have for every linear certificate z w.r.t. σ that

$$(z_{B(1)} + z_{0,1}, z_{B(2)}, \dots, z_{B(2n-1)})^t \in V, \tag{2.3.2}$$

and for every $v \in V$ with $\sigma_{B(i)} v_k > 0$ for at least one $k \in \{1, n+1, \dots, 2n-1\}$ is

$$z_{i,j} := \begin{cases} 0 & \text{for } (i,j) = (0,1) \text{ and} \\ v_{B^{-1}(i,j)} & \text{otherwise} \end{cases}$$

a linear certificate w.r.t. σ.

Proof. Consider the monomorphism h from $\mathrm{Ab}(a_1,\dots,a_n)$ to the additive group of $\mathbb{Q}^n$ which throws

$$a_1 \mapsto (1,0,\dots,0)^t, \dots, a_n \mapsto (0,\dots,0,1)^t.$$

Then the columns of $\mathbf{U}^n$ correspond to the elements of the top row and diagonal of the Budan matrix, i.e.,

$$\begin{aligned}\mathbf{U}^n =&\big(h(\alpha_{B(1)}),\dots,h(\alpha_{B(2n-1)})\big)\\ =&\big(h(\alpha_{0,1}) = h(\alpha_{n,1}), h(\alpha_{0,2}),\dots,h(\alpha_{0,n}), h(\alpha_{n-1,2}),\dots,h(\alpha_{1,n})\big).\end{aligned}$$

Now the assertions follow directly from the definition of linear certificate. □

Lemma 2.3.3. *Under the conditions of lemma 2.3.2 let be*

$$\mathbf{V}^n := \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \vdots \\ \mathbf{v}_{n-1} \\ \mathbf{v}_n \\ \mathbf{v}_{n+1} \\ \vdots \\ \vdots \\ \mathbf{v}_{2n-1} \end{pmatrix} := \begin{pmatrix} -\binom{n-1}{n-2} & \dots & -\binom{n-1}{1} & -\binom{n-1}{0} \\ -\binom{n-2}{n-2} & \dots & -\binom{n-2}{1} & -\binom{n-2}{0} \\ 0 & \ddots & \vdots & \vdots \\ \vdots & \ddots & -\binom{1}{1} & -\binom{1}{0} \\ 0 & \dots & 0 & -\binom{0}{0} \\ 1 & \ddots & \vdots & 0 \\ 0 & \ddots & 0 & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in \mathbb{Q}^{(2n-1)\times(n-1)}$$

and

$$W := \{x \in \mathbb{Z}^{n-1} | \sigma_{B(k)} \langle \mathbf{v}_k^t, x \rangle \geq 0 \text{ for all } k\} \tag{2.3.3}$$

the intersection of all $(n-1)$-dimensional half spaces, whose normals are the rows $\mathbf{v}_k$, for which $\sigma_{B(k)} \neq 0$, and which have the orientation demanded by $\sigma_{B(k)}$.
Then we get for the set V of lemma 2.3.2

$$V = \{\mathbf{V}^n w | w \in W\} \subset \mathbb{Z}^{2n-1}. \tag{2.3.4}$$

Proof. We have

$$\{v \in \mathbb{Z}^{2n-1} | \mathbf{U}^n v = 0\} = \{\mathbf{V}^n w | w \in \mathbb{Z}^{n-1}\}$$

(for the necessity of w to be an integer consider the lower half of $\mathbf{V}^n$). It remains to prove the equivalence between the conditions (2.3.1) and (2.3.3), which is clear since $v_k = \langle \mathbf{v}_k^t, x \rangle$ for every

$$(v_1, \dots, v_{2n-1})^t = v = \mathbf{V}^n x \in \{v \in \mathbb{Z}^{2n-1} | \mathbf{U}^n v = 0\} \supset V$$

and every k. □

After these terrible definitions, the situation shall be explained by an example:

Example 2.3.4. Let $n := 3$ and σ according to (2.3.5).

$$\mathbf{U}^3 = \begin{pmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix}; \quad \mathbf{V}^3 = \begin{pmatrix} -2 & -1 \\ -1 & -1 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \sigma(0,1) = 1 \\ \sigma(0,2) = -1 \\ \sigma(0,3) = -1 \\ \sigma(2,2) = -1 \\ \sigma(1,3) = 1 \end{pmatrix}. \tag{2.3.5}$$

With $w := (-1, 2)^t$, we get $v = \mathbf{V}^3 w = (0, -1, -2, -1, 2)^t$ which fulfills σ. How can such a $w \in W$ be found systematically? The next theorem describes how to find all $w \in W$.

Theorem 2.3.5. *Let σ be a sign condition on $I_{n\nearrow}$ with $\sigma_{i,j} \neq 0$ for all (i,j) and $\sigma_{0,1} = \sigma_{n,1}$. With $\mathbf{v}_k$ as in lemma 2.3.3 define the sets $W_0, \dots, W_n \in \mathbb{Q}^{n-1}$ inductively.*

$$
\begin{aligned}
W_0 &:= \{ \begin{pmatrix} \sigma_{B(n+1)} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sigma_{B(2n-1)} \end{pmatrix} \}; \\
W_1 &:= \big(W_0 \cap \{\sigma_{B(1)} \langle \mathbf{v}_1^t, x\rangle \geq 0\}\big) \cup \big(\overline{B_0 C_0} \cap \{\langle \mathbf{v}_1^t, x\rangle = 0\}\big); \\
&\vdots \\
W_n &:= \big(W_{n-1} \cap \{\sigma_{B(n)} \langle \mathbf{v}_n^t, x\rangle \geq 0\}\big) \cup \big(\overline{B_{n-1} C_{n-1}} \cap \{\langle \mathbf{v}_n^t, x\rangle = 0\}\big)
\end{aligned}
$$

with $B_k := W_k \cap \{\langle \mathbf{v}_{k+1}^t, x\rangle > 0\}$ and $C_k := W_k \cap \{\langle \mathbf{v}_{k+1}^t, x\rangle < 0\}$ for all k. Then with

$$\widetilde{W} = \{r \cdot w \in \mathbb{Q}^{n-1} | r > 0,\ w \in \operatorname{conv}(W_n)\} \cap \mathbb{Z}^{n-1},$$

(2.3.4) and (2.3.2) define a one-to-one relation between $\tilde{w} \in \widetilde{W}$ and the linear certificates z w.r.t. σ with $z_{0,1} = 0$.

Remark 2.3.6. The general case with $\sigma_{B(k)} = 0$ for some k can be done in the following way: For $1 \leq k \leq n$ skip the calculation of W_k and let be $W_k := W_{k-1}$ and go on with calculating W_{k+1}. For $k > n$ calculate both W_n for $\sigma_{B(k)} = -1$ and $\sigma_{B(k)} = 1$ and take the union of the solutions.

Proof. For $k > n$, (2.3.3) defines a restriction to a certain octant of the $\mathbb{Z}^{n-1}$. Since the length of the considered vectors does not matter, we take only the vectors with sum of absolute coordinates one. With

$$W_{\text{norm}} := \Big\{(w_1, \dots, w_{n-1})^t \in \mathbb{Q}^{n-1} \Big| \sum_k |w_k| = 1\Big\}$$

we get

$$
\begin{aligned}
\operatorname{conv}(W_0) =& \{(w_1, \dots, w_{n-1})^t \in W_{\text{norm}} | \sigma_{B(k)} w_{k-n} \geq 0 \text{ for } n+1 \leq k \leq 2n-1\} \\
=& W_{\text{norm}} \cap \{x \in \mathbb{Q}^{n-1} | \sigma_{B(k)} \langle \mathbf{v}_k^t, x\rangle \geq 0 \text{ for } n+1 \leq k \leq 2n-1\}.
\end{aligned}
$$

The next lemma 2.3.7 shows for $1 \leq k \leq n$,

$$\operatorname{conv}(W_k) = \operatorname{conv}(W_{k-1}) \cap \{\sigma_{B(k)} \langle \mathbf{v}_k^t, x\rangle \geq 0\},$$

which shows that $\operatorname{conv}(W_n)$ fulfills all sign conditions of (2.3.3).
To verify the one-to-one take a linear certificate z w.r.t. σ with $z_{0,1} = 0$. (2.3.4) leads to a unique $v \in V$ and (2.3.2) to a unique $w \in W$. To see that $w \in \widetilde{W}$ write it as

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_{n-1} \end{pmatrix} = \sum_k |w_k| \begin{pmatrix} \frac{w_1}{\sum_k |w_k|} \\ \vdots \\ \frac{w_{n-1}}{\sum_k |w_k|} \end{pmatrix}.$$

For the other direction take a $\tilde{w} \in \widetilde{W}$ which leads to a unique $v \in V$ and to a unique linear certificate z with $z_{0,1} = 0$; since $\tilde{w} \neq 0 \Rightarrow$ at least for one $n+1 \leq k \leq 2n-1$ is $z_{B(k)} = v_k \neq 0$ and therefore $\sigma_{B(k)} z_{B(k)} > 0$.
It remains to prove

Lemma 2.3.7. *Let be $Y \subset \mathbb{Q}^n$, $0 \neq a \in \mathbb{Q}^n$, $B := Y \cap \{\langle a, x\rangle > 0\}$ and $C := Y \cap \{\langle a, x\rangle < 0\}$. Then we have*

$$\operatorname{conv}(Y) \cap \{\langle a, x\rangle \geq 0\} = \operatorname{conv}\left((Y \cap \{\langle a, x\rangle \geq 0\}) \cup (\overline{BC} \cap \{\langle a, x\rangle = 0\})\right).$$

Proof. "$\supset$" is clear since

$$\operatorname{conv}(Y) \cap \{\langle a, x\rangle \geq 0\} \supset (Y \cap \{\langle a, x\rangle \geq 0\}) \cup (\overline{BC} \cap \{\langle a, x\rangle = 0\})$$

and u, $v \in \{\langle a, x\rangle \geq 0\} \Rightarrow (\lambda u + (1-\lambda)v) \in \{\langle a, x\rangle \geq 0\}$ for $0 \leq \lambda \leq 1$ since $\langle a, (\lambda u + (1-\lambda)v)\rangle = \lambda\langle a, u\rangle + (1-\lambda)\langle a, v\rangle \geq 0$.
For "$\subset$" let d be an arbitrary point of $\operatorname{conv}(B \cup C) \cap \{\langle a, x\rangle \geq 0\}$, i.e.,

$$d := \sum \beta_k b_k + \sum \gamma_l c_l$$

with $b_k \in B$, $c_l \in C$, β_k, $\gamma_l \geq 0$ with $\sum \beta_k + \sum \gamma_l = 1$ and $\langle a, d\rangle \geq 0$.
Furthermore with $\beta^* := \sum \beta_k$ and $\gamma^* := \sum \gamma_l$ let

$$b^* := \sum \frac{\beta_k}{\beta^*} b_k, \qquad c^* := \sum \frac{\gamma_l}{\gamma^*} c_l, \qquad \tilde{d} := \overline{b^* c^*} \cap \{\langle a, x\rangle = 0\}.$$

Then $d \in \overline{b^* c^*}$, and with $\tilde{d} = \lambda_0 b^* + (1-\lambda_0)c^*$ and $d = \lambda_1 b^* + (1-\lambda_1)c^*$ we get $\lambda_0 \leq \lambda_1$, since $\langle a, \lambda b^* + (1-\lambda)c^*\rangle = \lambda\langle a, b^*\rangle + (1-\lambda)\langle a, c^*\rangle$ and $\langle a, \tilde{d}\rangle = 0 \leq \langle a, d\rangle$. I.e.,

$$d = \frac{\lambda_1 - \lambda_0}{1-\lambda_0} b^* + \left(1 - \frac{\lambda_1 - \lambda_0}{1-\lambda_0}\right)\tilde{d} \in \overline{b^* \tilde{d}}.$$

(Figure 2.3.1.)

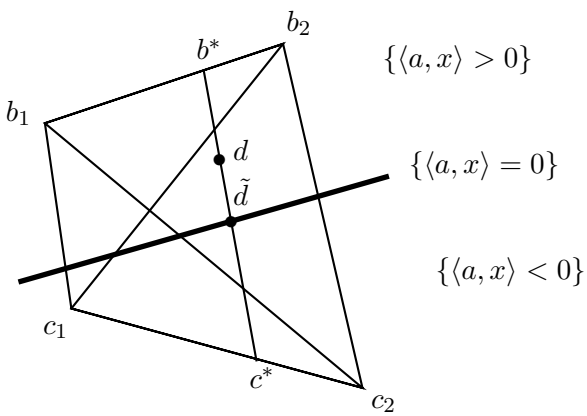


Figure 2.3.1: Proof of lemma 2.3.7.

The next argument shows that

$$\tilde{d} \in \operatorname{conv}\big((Y \cap \{\langle a, x\rangle \geq 0\}) \cup (\overline{BC} \cap \{\langle a, x\rangle = 0\})\big) =: RHS;$$

according to lemma 2.3.8, $\overline{b_k c^*} \cap \{\langle a, x\rangle = 0\} \in RHS$ for every k and therefore — again with 2.3.8 — $\tilde{d} = \overline{b^* c^*} \cap \{\langle a, x\rangle = 0\} \in RHS$.
It follows from $d \in \overline{b^* \tilde{d}}$ that $d \in RHS$.
And arbitrary $d' \in \operatorname{conv}(Y) \cap \{\langle a, x\rangle \geq 0\}$ can be written as

$$d' = \sum \alpha'_j a'_j + \sum \beta'_k b'_k + \sum \gamma'_l c'_l$$

with $a'_j \in Y \cap \{\langle a, x\rangle = 0\} \subset RHS$, $b'_k \in B$, $c'_l \in C$ and therefore also $d' \in RHS$.
It remains to prove

Lemma 2.3.8. *With the notions of lemma 2.3.7, let be $b \in B$, $c_l \in C$ and $\gamma'_l \geq 0$ with $\sum \gamma'_l = 1$ for $1 \leq l \leq m$ and $c^* := \sum \gamma'_l c_l$. Then we have*

$$\overline{bc^*} \cap \{\langle a, x\rangle = 0\} \in \operatorname{conv}\left(\overline{\{b\}\{c_1, \ldots, c_m\}} \cap \{\langle a, x\rangle = 0\}\right).$$

Proof. First let be $m = 2$. Let $(\lambda_1 b + (1-\lambda_1)c_1) = \overline{bc_1} \cap \{\langle a, x\rangle = 0\}$ and $(\lambda_2 b + (1-\lambda_2)c_2) = \overline{bc_2} \cap \{\langle a, x\rangle = 0\}$. With

$$\tilde{\gamma'_1} := \frac{\gamma'_1(1-\lambda_2)}{\gamma'_1(1-\lambda_2) + \gamma'_2(1-\lambda_1)}$$

and

$$\tilde{\gamma'_2} := \frac{\gamma'_2(1-\lambda_1)}{\gamma'_1(1-\lambda_2) + \gamma'_2(1-\lambda_1)}$$

we get

$$\begin{aligned}&\tilde{\gamma'_1}(\lambda_1 b + (1-\lambda_1)c_1) + \tilde{\gamma'_2}(\lambda_2 b + (1-\lambda_2)c_2) = \overline{bc^*} \cap \{\langle a, x\rangle = 0\}\\ &\in \operatorname{conv}\left(\overline{\{b\}\{c_1, c_2\}} \cap \{\langle a, x\rangle = 0\}\right)\end{aligned}$$

as $0 \leq \tilde{\gamma'_1}, \tilde{\gamma'_2} \leq 1$. (Consider Figure 2.3.2; the grey triangles are similar.)
For greater m the claim follows by induction since

$$c^* = \sum_{l=1}^{m} \gamma'_l c_l = \Big(\sum_{l=1}^{m-1} \gamma'_l\Big) \frac{\sum_{l=1}^{m-1} \gamma'_l c_l}{\sum_{l=1}^{m-1} \gamma'_l} + \gamma'_m c_m$$

and $(\sum_{l=1}^{m-1} \gamma'_l c_l)/(\sum_{l=1}^{m-1} \gamma'_l) \in \operatorname{conv}(\{c_1, \ldots, c_{m-1}\})$. □ □ □

Example 2.3.9. continues example 2.3.4. Now W resp. W_3 as in theorem 2.3.5 shall be calculated.

$$\begin{aligned}W = \mathbb{Z}^2 &\cap \{\langle(-2,-1)^t, x\rangle \geq 0\}\\ &\cap \{\langle(-1,-1)^t, x\rangle \leq 0\}\\ &\cap \{\langle(\ \ 0,-1)^t, x\rangle \leq 0\}\\ &\cap \{\langle(\ \ 1,\ \ 0)^t, x\rangle \leq 0\}\\ &\cap \{\langle(\ \ 0,\ \ 1)^t, x\rangle \geq 0\}\end{aligned}$$

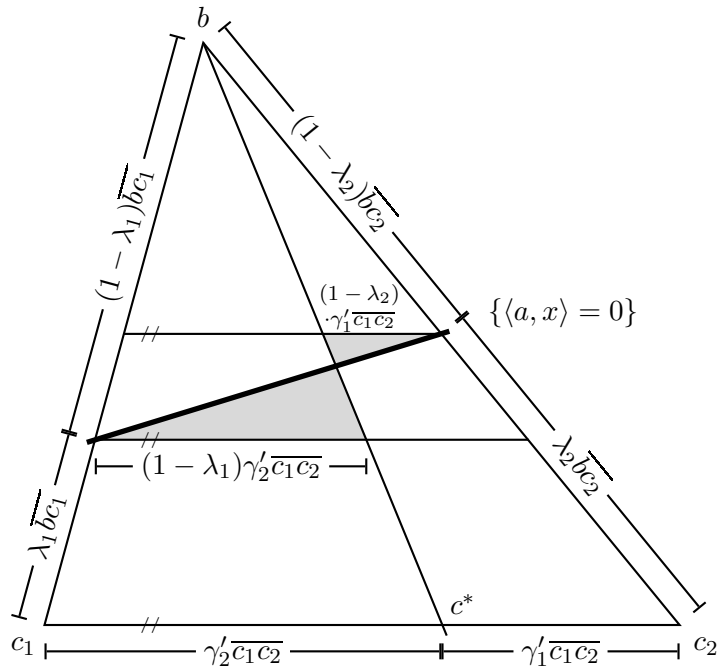


Figure 2.3.2: Proof of lemma 2.3.8.

$$W_0 = \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$$

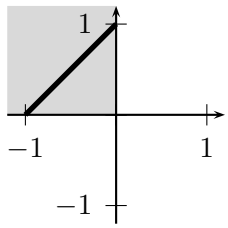


Figure 2.3.3: $\operatorname{conv}(W_0)$.

thus the line $\operatorname{conv}(W_0)$ represents $\{rw | r \geq 0,\ w \in \operatorname{conv}(W_0)\}$, the quadrant determined by the lower two conditions. According to theorem 2.3.5, we get

$$\begin{aligned} W_1 =& \left(\left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \cap \{\langle (-2,-1)^t, x\rangle \geq 0\} \right) \\ &\cup \left(\overline{\begin{pmatrix} -1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}} \cap \{\langle (-2,-1)^t, x\rangle = 0\} \right) \\ =& \left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix} \right\}; \end{aligned}$$

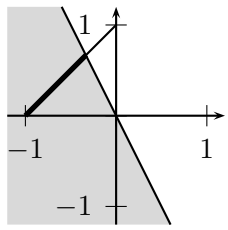


Figure 2.3.4: $\operatorname{conv}(W_1)$.

$$
\begin{aligned}
W_2 =& \left(\left\{\begin{pmatrix}-1\\0\end{pmatrix}, \begin{pmatrix}-\frac{1}{3}\\\frac{2}{3}\end{pmatrix}\right\} \cap \{\langle(-1,-1)^t, x\rangle \leq 0\}\right) \\
&\cup \left(\overline{\begin{pmatrix}-1\\0\end{pmatrix}\begin{pmatrix}-\frac{1}{3}\\\frac{2}{3}\end{pmatrix}} \cap \{\langle(-1,-1)^t, x\rangle = 0\}\right) \\
=& \left\{\begin{pmatrix}-\frac{1}{3}\\\frac{2}{3}\end{pmatrix}\right\} \cup \left\{\begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\end{pmatrix}\right\};
\end{aligned}
$$

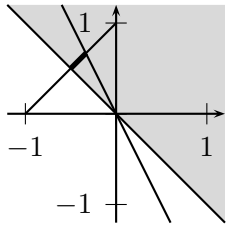


Figure 2.3.5: $\operatorname{conv}(W_2)$.

$$
\begin{aligned}
W_3 =& \left(\left\{\begin{pmatrix}-\frac{1}{3}\\\frac{2}{3}\end{pmatrix}, \begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\end{pmatrix}\right\} \cap \{\langle(0,-1)^t, x\rangle \leq 0\}\right) \\
&\cup \left(\emptyset \cap \{\langle(0,-1)^t, x\rangle = 0\}\right) \\
=& \left\{\begin{pmatrix}-\frac{1}{3}\\\frac{2}{3}\end{pmatrix}, \begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\end{pmatrix}\right\};
\end{aligned}
$$

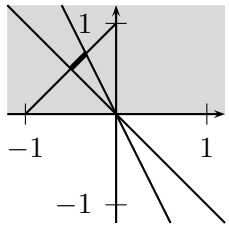


Figure 2.3.6: $\operatorname{conv}(W_3)$.

and, finally,

$$
W = \left\{ r \cdot w \in \mathbb{Q}^2 | r \in \mathbb{N},\, w \in \overline{\begin{pmatrix}-\frac{1}{3}\\\frac{2}{3}\end{pmatrix}\begin{pmatrix}-\frac{1}{2}\\\frac{1}{2}\end{pmatrix}} \right\} \cap \mathbb{Z}^2.
$$

Chapter 3

A certificate for Budan's theorem constructed in polynomial time

3.1 A certificate for Budan's theorem constructed in polynomial time

In this section we present an algorithm to calculate a certificate for Budan's theorem in polynomial time in the degree of the polynomial. Before we state the main theorem 3.1.5, we present a detailed example for a better understanding. The proof of theorem 3.1.5 is not too long but refers to theorem 3.2.7 which is the main result of section 3.2.

Definition 3.1.1.

i) Let the naturals $\mathbb{N} = \{0, 1, \dots\}$ include 0;

ii) let $\mathbb{Q}_{\geq 0} := \{x \in \mathbb{Q} | x \geq 0\}$ and $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} | x \geq 0\}$.

Let be $1 \leq r \leq n$ and $v \in \mathbb{Q}^n$ (resp. $\mathbb{R}^n$ in section 3.2.2);

iii) for $0 \leq i \leq n-1$ let

$$p_i(X) := \prod_{k=0}^{i-1} (X - k) \in \mathbb{R}[X]$$

denote such a real polynomial of degree i ($p_0(X) := 1$);

iv) for $1 \leq i \leq n$ and $v =: (v_1, \dots, v_n)$ let

$$\pi_i(v) := v_i$$

denote the projection of v on the *coordinate* at *position* i;

v) let

$$q_v(X) := \sum_{i=1}^{n} \pi_i(v) p_{i-1}(X)$$

denote the linear combination of the p_i with coefficients v;

vi) let the *length* of v

$$L(v) := \sum_{i=1}^{n} \operatorname{sign}(|\pi_i(v)|)$$

be the number of its nonzero coordinates;

vii) let

$$\{1,\dots,n\}^{r<} := \{(c_1,\dots,c_r) \in \{1,\dots,n\}^r | c_1 < \dots < c_r\}$$

denote the r-tuples of {1,. . . ,n} ordered in a strictly ascending way;

viii) for $L(v) = r$ and $1 \le j \le r$ let denote $c_j(v)$ the position of the j-th nonzero coordinate of v, i.e.,

$$1 \le c_1(v) < \dots < c_r(v) \le n \qquad \text{with} \qquad \pi_{c_j(v)}(v) \neq 0$$

for all j, and

$$C(v) := \big(c_1(v),\dots,c_r(v)\big) \in \{1,\dots,n\}^{r<};$$

ix) for naturals $0 \le z_1 < \dots < z_{r-1}$ let denote

$$\mathbf{C}_{(z_1,\dots,z_{r-1})} \subset \bigcup_{\rho=1}^{r} \mathbb{N}^\rho$$

the union of the sets of ρ-tuples of naturals $(c_1,\dots,c_\rho)$ for which the following holds

a) $1 \le c_1 < \dots < c_\rho$;
b) $z_{r-\rho} < c_1 - 1$;
c) $c_j - 1 \le z_{r-\rho+j-1}$ for $2 \le j \le \rho$.

(In case $\rho = r$ consider $z_{r-\rho}$ as -1 and in case $\rho = 1$ consider $z_{r-\rho+1}$ as $+\infty$; note that only for $\rho = 1$ c_1 can be an arbitrary natural $> z_{r-\rho} + 1$ while for $\rho > 1$ holds $c_\rho - 1 \le z_{r-1}$.)

Remark 3.1.2 (about the definitions). Since the polynomials p_i and q_v are used frequently consider the example $v := (12, -6, 0, 1) \in \mathbb{Q}^4$. $L(v) = 3$; $C(V) = (c_1(v), c_2(v), c_3(v)) = (1, 2, 4) \in \{1,\dots,4\}^{3<}$; $q_v(X) = 12p_0 - 6p_1 + p_3$ is shown in figure 3.1.1.
Note that for a v with $C(v) = (c_1)$ the corresponding $q_v = \pi_{c_1}(v)p_{c_1-1}(X)$ has degree $c_1 - 1$ and its greatest root at $c_1 - 2$.
For an explanation of $\mathbf{C}_{(z_1,\dots,z_{r-1})}$, consider remark 3.2.4.

Example 3.1.3 (for the certificate). Consider a polynomial $f \in \mathbb{R}[X]$ of degree $n := 4$ and $a < b \in \mathbb{R}$. Budan's theorem claims that the number of sing changes in the sequence $(f(b), f'(b), \dots, f^{(n)}(b))$ is not greater than the number of sign changes in the sequence $(f(a), \dots, f^{(n)}(a))$.

For the certificate, we take two sequences of signs such that at b we have more sign changes than at a. And from this we will derive the contradiction $0 < 0$. For example

$$\begin{aligned}\alpha &:= (\alpha_0, \dots, \alpha_4) := (+1, +1, +1, -1, +1) \qquad \text{and}\\ \beta &:= (\beta_0, \dots, \beta_4) := (-1, +1, -1, -1, +1).\end{aligned}$$

α shall correspond to $(f(a), \dots, f''''(a))$, β to $(f(b), \dots, f''''(b))$. Now the certificate will give coefficients $a_0, \dots, a_4, b_0, \dots, b_4 \in \mathbb{Q}[(b-a)]$ with

$$\alpha_i a_i \geq 0, \qquad \beta_i b_i \geq 0 \tag{3.1.1}$$

for all i (at least one of the inequalities being a strict one) and

$$\sum_{i=0}^{4} a_i f^{(i)}(a) + \sum_{i=0}^{4} b_i f^{(i)}(b) = 0; \tag{3.1.2}$$

which leads to the contradiction $0 < 0$.
To find such a_i, b_i we consider the Taylor expansions of f and its derivatives at a in the variable b, multiplied by some powers of $(b-a)$.

$$\begin{aligned}
f(b) &= f(a)+f'(a)(b-a)+\tfrac{f''(a)}{2!}(b-a)^2+\tfrac{f'''(a)}{3!}(b-a)^3+\tfrac{f''''(a)}{4!}(b-a)^4\\
f'(b)(b-a) &= \qquad f'(a)(b-a)+f''(a)(b-a)^2+\tfrac{f'''(a)}{2!}(b-a)^3+\tfrac{f''''(a)}{3!}(b-a)^4\\
f''(b)(b-a)^2 &= \qquad\qquad f''(a)(b-a)^2+f'''(a)(b-a)^3+\tfrac{f''''(a)}{2!}(b-a)^4\\
f'''(b)(b-a)^3 &= \qquad\qquad\qquad f'''(a)(b-a)^3+f''''(a)(b-a)^4
\end{aligned} \tag{3.1.3}$$

We next take the coefficient vector $v = (12, -6, 0, 1)$ — whose calculation will be explained later — multiply the equations of (3.1.3) by these coefficients and add them.

$$\begin{aligned}
12f(b) &= 12f(a)+12f'(a)(b-a)+12\tfrac{f''(a)}{2!}(b-a)^2+12\tfrac{f'''(a)}{3!}(b-a)^3+12\tfrac{f''''(a)}{4!}(b-a)^4\\
-6f'(b)(b-a) &= \qquad -6f'(a)(b-a)-6f''(a)(b-a)^2-6\tfrac{f'''(a)}{2!}(b-a)^3-6\tfrac{f''''(a)}{3!}(b-a)^4\\
f'''(b)(b-a)^3 &= \qquad\qquad\qquad f'''(a)(b-a)^3+f''''(a)(b-a)^4\\
\hline
12f(b) - 6f'(b)(b-a) &+ f'''(b)(b-a)^3\\
&= 12f(a)+6f'(a)(b-a)+0\tfrac{f''(a)}{2!}(b-a)^2+0\tfrac{f'''(a)}{3!}(b-a)^3+\tfrac{1}{2}\tfrac{f''''(a)}{4!}(b-a)^4
\end{aligned} \tag{3.1.4}$$

(3.1.4) defines the

$$\begin{aligned}
(a_0, \dots, a_4) &:= (12, 6(b-a), 0, 0, \tfrac{1}{2}(b-a)^4) && \text{from the right side of (3.1.4) and}\\
(b_0, \dots, b_4) &:= (-12, 6(b-a), 0, -(b-a)^3, 0) && \text{from the left side with reversed signs.}
\end{aligned}$$

(3.1.1) holds and (3.1.4) becomes to (3.1.2), which shows Budan's claim for the particular signs α and β and every degree-4 polynomial. The question is if such coefficients v exist in general (depending on α, β) and how to calculate them?
Calculation of $v = (v_1, v_2, v_3, v_4)$. Obviously, we need

$$-v_i \beta_{i-1} \geq 0 \qquad \text{for } 1 \leq i \leq 4; \tag{3.1.5}$$

$$\alpha_j \sum_{i=0}^{j} \frac{v_{i+1}}{(j-i)!} \geq 0 \qquad \text{for } 0 \leq j \leq 4. \tag{3.1.6}$$

($(b-a)$ is positive. In the case $j=4$ take $v_5 := 0$.) After multiplication with $j!$ (3.1.6) reads as

$$\alpha_j q_v(j) = \alpha_j \sum_{i=0}^{3} v_{i+1} p_i(j) = \alpha_j \sum_{i=0}^{j} v_{i+1} \prod_{k=0}^{i-1} (j-k) = \alpha_j \sum_{i=0}^{j} v_{i+1} \frac{j!}{(j-i)!} \geq 0. \tag{3.1.7}$$

This is where the polynomials p_i and q_v come into play. In particular, we look for a q_v with

$$q_v(0) \geq 0,\ q_v(1) \geq 0,\ q_v(2) \geq 0,\ q_v(3) \leq 0,\ q_v(4) \geq 0,$$

i.e., q_v changes sign simultaneously with α. Therefore we mark in the sequences α and β the positions left to the sign changes and define $(z_1, z_2) := (2, 3)$ corresponding to α (e.g., $z_1 = 2$ comes from $\alpha_2\alpha_3 = -1$) and $(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3) := (0, 1, 3)$ corresponding to β. We now enter theorem 3.2.7 with (z_1, z_2) and $(c_1, c_2, c_3) := (\tilde{c}_1 + 1, \tilde{c}_2 + 1, \tilde{c}_3 + 1)$ and get back $v = (12, -6, 0, 1)$ such that

$$q_v(X) = 12p_0(X) - 6p_1(X) + p_3(X) = 12(1) - 6(X) + (X(X-1)(X-2))$$

(Figure 3.1.1) fulfills (3.1.7).

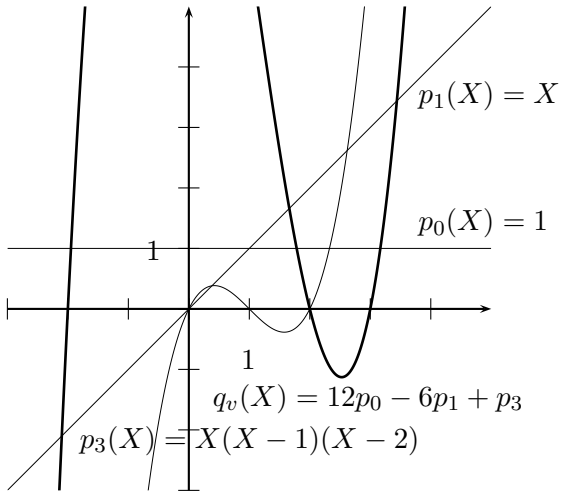


Figure 3.1.1: Example; q_v.

Furthermore, the nonzero coordinates of v alternate in signs which leads with the choice of $(\tilde{c}_1, \tilde{c}_2, \tilde{c}_3)$ to (3.1.5).
In general, theorem 3.2.7 provides such a q if $(c_1, \ldots, c_r) \in \mathbf{C}_{(z_1, \ldots, z_{r-1})}$.
Before we state the general assertion it follows

Remark 3.1.4 (about zeros in the sign sequences). Theorem 3.1.5 only deals with sign sequences $\{+1, -1\}^{n+1}$. Sign sequences $\{+1, -1, 0\}^{n+1}$ can be reduced to strict ones by assigning the zeros arbitrarily negative or positive. We only must take care that at least one of the inequalities (3.1.1) is strict (as it is $a_n\alpha_n$ as long as $\alpha_n = 1$).

Theorem 3.1.5. *Let be R an ordered field, $a, b \in R$ with $a < b$, $n \geq 1$ and $(\alpha_0, \ldots, \alpha_n)$, $(\beta_0, \ldots, \beta_n) \in \{+1, -1\}^{n+1}$ with $\alpha_n = \beta_n = +1$ such that $(\beta_0, \ldots, \beta_n)$ has at least one sign change more than $(\alpha_0, \ldots, \alpha_n)$.*

There is an algorithm $\mathbf{Z}$ *which receives* $(\alpha_0, \dots, \alpha_n)$ *and* $(\beta_0, \dots, \beta_n)$ *and provides coefficients* a_i, $b_i \in \mathbb{Q}[(b-a)]$ $(0 \le i \le n)$ *with* $\alpha_i a_i \ge 0$, $\beta_i b_i \ge 0$ *for all* i *such that*

$$\sum_{i=0}^{n} a_i f^{(i)}(a) + \sum_{i=0}^{n} b_i f^{(i)}(b) = 0 \tag{3.1.8}$$

for every polynomial $f \in R[X]$ *of degree* $\le n$. *The complexity of* $\mathbf{Z}$ *is the complexity to solve a linear system in* n *variables with coefficients of bit length* $n \log n$.

Proof. Let A denote the number of sign changes in $(\alpha_0, \dots, \alpha_n)$ and $0 \le z_1 < z_2 < \dots < z_A < n$ the indices left to sign changes, i.e., $\alpha_{z_j}\alpha_{z_j+1} = -1$ for $j \in \{1, \dots, A\}$. In the same way — concerning $(\beta_0, \dots, \beta_n)$ — define B and $0 \le \tilde{c}_1 < \tilde{c}_2 \dots < \tilde{c}_B < n$.
W.l.o.g., let be $B = A + 1$, $\tilde{c}_1 = 0$ and

$$\tilde{c}_{j+1} \le z_j \tag{3.1.9}$$

for $j \in \{1, \dots, A\}$. We can make this assumption since there is exactly one $i_0 \in \{0, \dots, n-1\}$ for which the subsequences $(\alpha_{i_0}, \dots, \alpha_n)$, $(\beta_{i_0}, \dots, \beta_n)$ fulfill it; if $i_0 \ne 0$ consider the theorem with these subsequences and $f^{(i_0)}$ instead of f.
With $c_j := \tilde{c}_j + 1$ we get

$$(c_1, \dots, c_B) \in \mathbf{C}_{(z_1, \dots, z_A)}$$

since in definition 3.1.1 ix) a) is clear; b) is nothing to show as $\rho = B = r$; c) is (3.1.9), i.e., $c_j - 1 = \tilde{c}_j \le z_{j-1}$ for $2 \le j \le B$.
With $\tau := t := B \le n$, $z_1, \dots, z_A$ and $c_1, \dots, c_B$ enter now theorem 3.2.7. It provides a vector $v \in \mathbb{Q}^n$ with

$$C(v) = (c_1, \dots, c_B) \qquad \text{and} \qquad \pi_{c_B}(v) = 1;$$

$$q_v(z_1) = \dots = q_v(z_A) = 0; \tag{3.1.10}$$

$$\operatorname{sign}(\pi_{c_j}(v)) = (-1)^{B-j} \tag{3.1.11}$$

for $1 \le j \le B$ and for $x \in \mathbb{N}$

$$\operatorname{sign}(q_v(x)) = \begin{cases} (-1)^{B-1} & \text{for} \quad x < z_1; \\ (-1)^{B-j} & \text{for} \quad z_{j-1} < x < z_j \quad \text{and} \quad 1 < j \le A; \\ 1 & \text{for} \quad z_A < x. \end{cases} \tag{3.1.12}$$

Now define $b_i := 0$ for $i \notin \{\tilde{c}_1, \dots, \tilde{c}_B\}$ and for $1 \le j \le B$

$$b_{\tilde{c}_j} := -\pi_{c_j}(v)(b-a)^{\tilde{c}_j}.$$

From $\beta_n = 1$ follows $\beta_{\tilde{c}_j} = (-1)^{B-j+1}$. And with $b - a > 0$ and (3.1.11)

$$\operatorname{sign}(\beta_{\tilde{c}_j} b_{\tilde{c}_j}) = (-1)^{B-j+1}\operatorname{sign}(-\pi_{c_j}(v)) = -(-1)^{B-j+1}(-1)^{B-j} = 1,$$

as desired.
Define next for $0 \le i \le n$,

$$a_i := (b-a)^i \sum_{\{j \mid \tilde{c}_j \le i\}} \frac{\pi_{c_j}(v)}{(i - \tilde{c}_j)!}.$$

By the help of (3.1.12), we can determine $\operatorname{sign}(a_i)$; we get

$$\begin{aligned} i! \sum_{\{j|\tilde{c}_j \le i\}} \frac{\pi_{c_j}(v)}{(i-\tilde{c}_j)!} &= \sum_{\{j|\tilde{c}_j \le i\}} \pi_{c_j}(v) \frac{i!}{(i-\tilde{c}_j)!} = \sum_{\{j|\tilde{c}_j \le i\}} \pi_{c_j}(v) \prod_{k=0}^{\tilde{c}_j-1} (i-k) \\ &= \sum_{\{j|\tilde{c}_j \le i\}} \pi_{c_j}(v) p_{\tilde{c}_j}(i) = \sum_{\{j|\tilde{c}_j \le i\}} \pi_{c_j}(v) p_{c_j-1}(i) \\ &= \sum_{k=1}^{n} \pi_k(v) p_{k-1}(i) = q_v(i), \end{aligned}$$

where the last line follows from $\pi_i(v) = 0$ if $i \notin \{\tilde{c}_1, \dots, \tilde{c}_B\}$ and $p_k(i) = 0$ if $k > i$. For $z_{j-1} < i < z_j$ holds $\alpha_i = (-1)^{A-j+1}$ and with (3.1.12)

$$\operatorname{sign}(\alpha_i a_i) = (-1)^{A-j+1} q_v(i) = (-1)^{A-j+1}(-1)^{B-j} = 1$$

(equally for $i < z_1$ or $> z_A$). With (3.1.10) we have $\alpha_i a_i \ge 0$ for all i as desired.
Last (3.1.8) follows from Taylor's theorem — applied to the $\tilde{c}_j$th derivative of f — $f^{(\tilde{c}_j)}(b) = \sum_{i=\tilde{c}_j}^{n} (b-a)^{i-\tilde{c}_j} \frac{f^{(i)}(a)}{(i-\tilde{c}_j)!}$.

$$\begin{aligned} &\sum_{i=0}^{n} b_i f^{(i)}(b) && + \sum_{i=0}^{n} a_i f^{(i)}(a) \\ =&\sum_{j=1}^{B} -\pi_{c_j}(v)(b-a)^{\tilde{c}_j} f^{(\tilde{c}_j)}(b) && + \sum_{i=0}^{n} (b-a)^i \sum_{\{j|\tilde{c}_j \le i\}} \frac{\pi_{c_j}(v)}{(i-\tilde{c}_j)!} f^{(i)}(a) \\ =&\sum_{j=1}^{B} -\pi_{c_j}(v)(b-a)^{\tilde{c}_j} f^{(\tilde{c}_j)}(b) && + \sum_{j=1}^{B} \sum_{i=\tilde{c}_j}^{n} (b-a)^i \frac{\pi_{c_j}(v)}{(i-\tilde{c}_j)!} f^{(i)}(a) \\ =&\sum_{j=1}^{B} \left[-\pi_{c_j}(v)(b-a)^{\tilde{c}_j} f^{(\tilde{c}_j)}(b) + \sum_{i=\tilde{c}_j}^{n} (b-a)^i \frac{\pi_{c_j}(v)}{(i-\tilde{c}_j)!} f^{(i)}(a) \right] \\ =&\sum_{j=1}^{B} \pi_{c_j}(v)(b-a)^{\tilde{c}_j} \left[-f^{(\tilde{c}_j)}(b) + \sum_{i=\tilde{c}_j}^{n} (b-a)^{i-\tilde{c}_j} \frac{f^{(i)}(a)}{(i-\tilde{c}_j)!} \right] = 0 \end{aligned}$$

To determine the complexity to calculate the a_i, $b_i \in \mathbb{Q}[(b-a)]$ form the $(\alpha_0, \dots, \alpha_n)$ and $(\beta_0, \dots, \beta_n)$, we follow the proof above and theorem 3.2.7 which contributes most of the effort. We give an upper bound for the complexity.

i) We neglect the effort to find i_0, $A < n$ and the $z_1 < z_2 < \dots < z_A$ and $\tilde{c}_1 < \tilde{c}_2 \dots < \tilde{c}_B$.

ii) We alculate the vector $v \in \mathbb{Q}^n$ in a way unlike the proof of theorem 3.2.7. We first calculate the naturals $p_{c_j-1}(z_i)$ for $1 \le i \le A$, $1 \le j \le B$; for fixed i this can be done successively since

$$p_{c_{j+1}-1}(z_i) = p_{c_j-1}(z_i) \prod_{c_j-1}^{c_{j+1}-2} (z_i - k).$$

For every i, this means at most $c_B - 3$ multiplications of naturals $\leq (c_B - 2)!$. For a multiplication of two l-bit numbers we calculate a complexity of $O(l^2)$. With $A < B \leq n$ this leads (added up for all i) to a complexity $O(n^2(n \log n)^2)$.

iii) From corollary 3.2.10 we have the invertibility of the matrix

$$\mathbf{M} := \begin{pmatrix} p_{c_1-1}(z_1) & p_{c_2-1}(z_1) & \dots & p_{c_A-1}(z_1) \\ p_{c_1-1}(z_2) & p_{c_2-1}(z_2) & \dots & p_{c_A-1}(z_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_{c_1-1}(z_A) & p_{c_2-1}(z_A) & \dots & p_{c_A-1}(z_A) \end{pmatrix} \in \mathbb{N}^{A,A}. \qquad \mathbf{m} := \begin{pmatrix} -p_{c_B-1}(z_1) \\ -p_{c_B-1}(z_2) \\ \vdots \\ -p_{c_B-1}(z_A) \end{pmatrix} \in \mathbb{N}^A.$$

We next solve the system of linear equations $\mathbf{M}x = \mathbf{m}$.
This is a linear system in n variables with coefficients of bit length $n \log n$.

iv) It follows that v is defined by

$$\pi_i(v) = \begin{cases} x_j & \text{if } i = c_j \text{ for } 1 \leq j \leq A; \\ 1 & \text{if } i = c_B; \\ 0 & \text{otherwise} \end{cases}$$

with $(x_1, \dots, x_A) := x$ form above. The calculation of a_i, b_i from the given v is negligible again.

All together, ii) has a lower complexity than iii). For further complexity considerations consider [BP, GG]. □

3.2 About the real roots of certain linear combinations of the polynomials $p_i(X) := \prod_{k=0}^{i-1}(X - k)$

3.2.1 Motivation

When we developed the certificate for Budan's theorem — theorem 3.1.5 — the question arose if for

$$0 \leq c_1 < \dots < c_r \leq z_1 < \dots < z_{r-1} \in \mathbb{N}$$

the polynomials $p_{c_1}, \dots, p_{c_r}$ can be combined linearly in such a way that for the linear combination $q \neq 0$ holds $q(z_j) = 0$ for $1 \leq j \leq r-1$? In particular, our requests on q were even stronger; q should have constant sign between every pair (z_j, z_{j+1}), i.e., it should admit the roots z_j but no other roots $\geq c_r$. As the

$$\big(p_{c_1}(z_1), \dots, p_{c_1}(z_{r-1})\big), \dots, \big(p_{c_r}(z_1), \dots, p_{c_r}(z_{r-1})\big)$$

are r vectors of dimension $r-1$ the question of the existence of a nonzero linear combination which provides enough roots is answered. The assumption that this linear combination is unique up to multiplicity and that it does not admit any other real roots $\geq c_r$ besides the z_j seems to be more difficult to prove.
The following situation is comparatively easier. If we took as basic elements the polynomials $\tilde{p}_i(X) := X^i \in \mathbb{R}[X]$ instead of the above-defined p_i and asked the same question it would

follow from Budan's theorem that the corresponding linear combination $\tilde{q}$ admits only the positive roots $z_1, \ldots, z_{r-1}$. This is because $\tilde{q}$ would be a polynomial with at most r nonzero coefficients and therefore at most $r-1$ sign changes between its coefficients. Then Budan's theorem would claim at most $r-1$ positive roots. But we could not find such an easy argument for the polynomials p_i.
Our first approach to the question was the conjecture that the following would hold: "*Let p and q be two polynomials of different degree and $Z_p := \{x \in \mathbb{R} | p'(x) = 0 \vee p''(x) = 0 \vee \ldots\}$ the set of all roots of all derivatives of p; similarly Z_q. Then $(p-q)$ can only admit 2 roots $> \max(Z_p, Z_q)$.*" Which turned out to be wrong (figure 3.2.1).

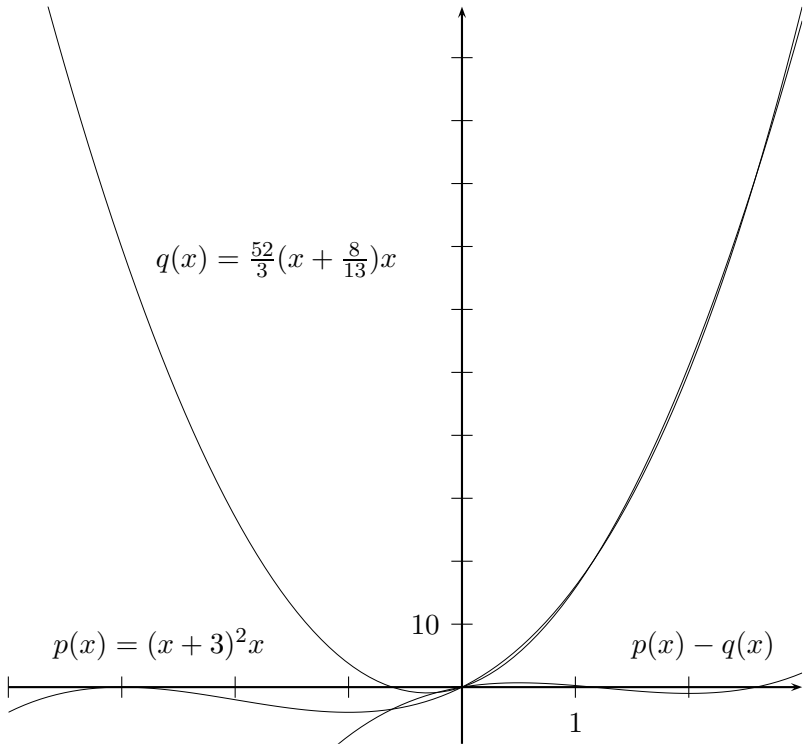


Figure 3.2.1: Counterexample; $(p-q)$ admits 3 roots $> \max(Z_p, Z_q) = -4/13$.

We finally succeeded in proving the uniqueness of q up to multiplicity and that it does not admit more real roots $\geq c_r$ than the z_j by the following argument. If the c_j are consecutive — i.e., $c_{j+1} = c_j + 1$ for all j — the claim follows from the following degree argument. We can write

$$p_{c_j}(X) = p_{c_1}(X) \prod_{k=c_1}^{c_j - 1} (X - k)$$

and therefore every linear combination $q = p_{c_1} q_1$ with a polynomial q_1. And since $\deg(q) = \deg(p_{c_r}) = \deg(p_{c_1+r-1}) = c_1 + r - 1$ we get $\deg(q_1) = r - 1$ and from this the claim. The general case with non-consecutive c_j can be reduced to these special cases.
The whole proof is presented in section 3.2.2. In section 3.2.3 We prove a more general case which is needed for the proof of theorem 3.1.5. We consider it useful to keep section 3.2.2 as on

the one hand it helps understanding section 3.2.3 and on the other hand it is not completely comprehended by section 3.2.3.

3.2.2 Example and theorem

Example 3.2.1 (for theorem 3.2.2). Since the proof of theorem 3.2.2 is a bit technical we give an example which explains the idea. Let be $(c_1, c_2, c_3, z_1, z_2) := (2, 5, 6, 5, 8)$ (i.e., we consider the polynomials p_1, p_4, p_5). We will first compute coefficients v_2 and v_5 such that with the coefficient vector $v := (v_1, \ldots, v_6) := (0, v_2, 0, 0, v_5, 1)$ and

$$q_v = \sum_{i=1}^{6} v_i p_{i-1}$$

$q_v(5) = q_v(8) = 0$. We will then show that $q_v(x) > 0$ for $x \in]4, 5[\cup]8, \infty[$ and $q_v(x) < 0$ for $x \in]5, 8[$.
The example is illustrated in figure 3.2.2.

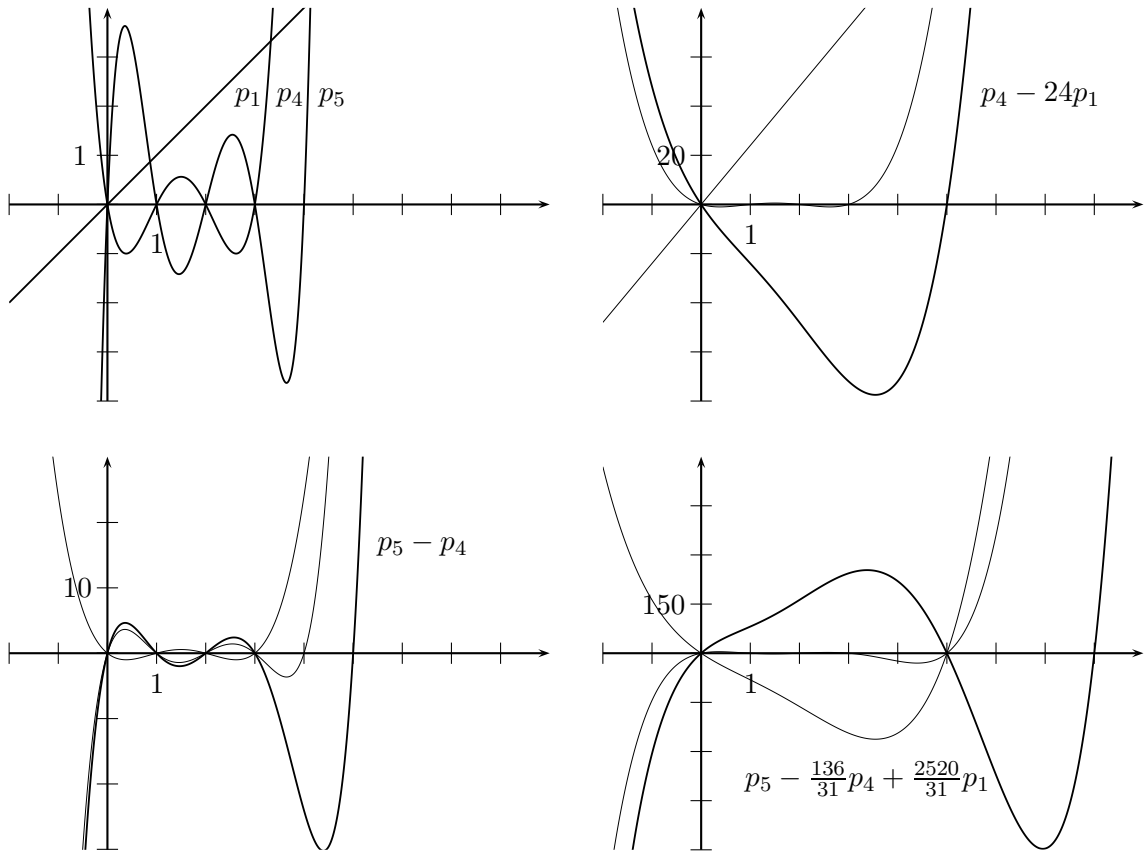


Figure 3.2.2: Example.

In the first step, we compute linear combinations which admit a root at 5. We consider q_α with $\alpha := (0, 0, 0, 0, -1, 1)$ for which $q_\alpha(5) = 0$. As also $q_\alpha(0) = q_\alpha(1) = q_\alpha(2) = q_\alpha(3) = 0$ and $\deg(q_\alpha) = 5$ we get $q_\alpha(x) < 0$ for $x \in]4, 5[$ and $q_\alpha(x) > 0$ for $x \in]5, \infty[$.
We consider q_β with $\beta := (0, -24, 0, 0, 1, 0)$ for which $q_\beta(5) = 0$. Here the degree argument

does not work directly but on the following way. We consider $\gamma := (0,-4,1,0,0,0)$, $\delta := (0,0,-3,1,0,0)$, $\varepsilon := (0,0,0,-2,1,0)$ which all admit a root at 5. And we know about the q_γ, q_δ, q_ε from their roots and degrees that they are negative in $]4,5[$ and positive in $]5,\infty[$. We now write

$$\beta = \begin{pmatrix} 0 \\ -24 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = 6 \begin{pmatrix} 0 \\ -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} = 6\gamma + 2\delta + \varepsilon. \tag{3.2.1}$$

It follows form the positivity of the coefficients $6, 2, 10$ that $q_\beta(x) < 0$ for $x \in]4,5[$ and $q_\beta(x) > 0$ for $x \in]5,\infty[$. The choice of nonnegative coefficients in (3.2.1) was possible because of the alternating signs of the nonzero entries in γ, δ, ε.

In the second step, we use the linear combinations from the first step to compute a linear combination which admits roots at 5 and 8. We consider q_v with $v := \alpha - (105/31)\beta = (0, 2520/31, 0, 0, -136/31, 1)$ for which $q_v(5) = q_v(8) = 0$. Here we consider $\eta := (0,28,-10,1,0,0)$, $\theta := (0,0,18,-8,1,0)$, $\iota := (0,0,0,10,-6,1)$ which also admit roots at 5 and 8. And we know about the q_η, q_θ, q_ι that they are positive in $]4,5[\cup]8,\infty[$ and negative in $]5,8[$. We now write

$$\begin{aligned} v = \begin{pmatrix} 0 \\ \frac{2520}{31} \\ 0 \\ 0 \\ -\frac{136}{31} \\ 1 \end{pmatrix} &= \frac{50}{31} \underbrace{\begin{pmatrix} 0 \\ \frac{252}{5} \\ 0 \\ -\frac{31}{5} \\ 1 \\ 0 \end{pmatrix}}_{=\kappa} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 10 \\ -6 \\ 1 \end{pmatrix} = \frac{50}{31} \underbrace{\left[\frac{9}{5} \begin{pmatrix} 0 \\ 28 \\ -10 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 18 \\ -8 \\ 1 \\ 0 \end{pmatrix} \right]}_{=:\kappa} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 10 \\ -6 \\ 1 \end{pmatrix} \\ &= \frac{90}{31} \begin{pmatrix} 0 \\ 28 \\ -10 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{50}{31} \begin{pmatrix} 0 \\ 0 \\ 18 \\ -8 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 10 \\ -6 \\ 1 \end{pmatrix} = \frac{90}{31}\eta + \frac{50}{31}\theta + \iota. \end{aligned} \tag{3.2.2}$$

Form the positivity of the coefficients $90/31, 50/31, 1$ we get $q_v(x) > 0$ for $x \in]4,5[\cup]8,\infty[$ and $q_v(x) < 0$ for $x \in]5,8[$, as desired.

The question is if it is always possible to find nonnegative coefficients as in (3.2.2). Yes, it is. This can be seen in the following way. Consider $\kappa = (5/9)\eta + \theta$. The alternating signs of the nonzero entries in η resp. θ make it possible to write a linear combination κ of η and θ with nonnegative coefficients such that the 3rd coordinate is eliminated. And as again — the nonzero entries of κ have alternating signs — it is possible to write a linear combination v of κ and ι with nonnegative coefficients such that even the 4th coordinate is eliminated. The alternation in the signs of the nonzero entries of the occurring linear combinations follows from a uniqueness argument which is particularly explained in the proof of

Theorem 3.2.2. *Let be* $1 \le t \le n$,

$$C := (c_1, \dots, c_t) \in \{1, \dots, n\}^{t<} \qquad \textit{and} \qquad n-2 < z_1 < \dots < z_{t-1} \in \mathbb{R}.$$

Then there exists a unique vector $v \in \mathbb{R}^n$ with

$$C(v) = C \qquad \text{and} \qquad \pi_{c_t}(v) = 1 \tag{3.2.3}$$

such that

$$q_v(z_1) = \cdots = q_v(z_{t-1}) = 0. \tag{3.2.4}$$

Furthermore, q_v does not admit any other real root in $]n-2, \infty[$, i.e.,

$$n-2 < x \in \mathbb{R} \qquad \text{and} \qquad q_v(x) = 0 \qquad \Longrightarrow \qquad x \in \{z_1, \ldots, z_{t-1}\}, \tag{3.2.5}$$

and the algebraic multiplicities of the roots $z_1, \ldots, z_{t-1}$ are 1.
Furthermore, the nonzero coordinates of v have alternating signs, i.e.,

$$\operatorname{sign}(\pi_{c_j}(v)) = (-1)^{t-j}$$

for $1 \leq j \leq t$.

Proof. In the case that $C = (k+1, \ldots, k+r)$ for some $0 \leq k \leq n-r$ and (3.2.3) and (3.2.4) are fulfilled, (3.2.5) follows from a degree argument. We want to show that for arbitrary C v can be written as linear combination with positive coefficients of some v_k with $C(v_k) = (k+1, \ldots, k+r)$ with $0 \leq k \leq n-r$. Then (3.2.5) follows, as for $x > n-2$ all the $q_{v_k}(x)$ are simultaneously negative resp. zero resp. positive.
We want to show the claim by an induction over $1 \leq r \leq t$.
Base case; $r := 1$.
Let $c_1 \leq n$ be arbitrary and define $v \in \mathbb{R}^n$ by $\pi_{c_1}(v) = 1$ and $\pi_i(v) = 0$ for $i \neq c_1$. Then

i) the conditions $C(v) = (c_1)$ and $\pi_{c_1}(v) = 1$ determine v in a unique way; therefore it shall be denoted by $v_{()}^{(c_1)} := v$;

ii) $\operatorname{sign}(\pi_{c_1}(v)) = 1 = (-1)^{r-1}$;

iii) $q_v(x) = 1 \cdot p_{c_1-1}(x) > 0$ for all $x > n-2 \geq c_1 - 2$.

Inductive step for $1 < r \leq t$; $r-1 \to r$.
In this step is provided that for every $1 \leq s < r$ and every arbitrary $(c_1, \ldots, c_s) \in \{1, \ldots, n\}^{s<}$ we have a $v_{(z_1,\ldots,z_{s-1})}^{(c_1,\ldots,c_s)} \in \mathbb{R}^n$ which is uniquely determined by the following conditions 1. and 2..

1. $C(v_{(z_1,\ldots,z_{s-1})}^{(c_1,\ldots,c_s)}) = (c_1, \ldots, c_s) \qquad \text{and} \qquad \pi_{c_s}(v_{(z_1,\ldots,z_{s-1})}^{(c_1,\ldots,c_s)}) = 1$;

2. $q_{v_{(z_1,\ldots,z_{s-1})}^{(c_1,\ldots,c_s)}}(z_j) = 0$ for $1 \leq j < s$.

Additionally is provided that

3. $\operatorname{sign}(\pi_{c_j}(v_{(z_1,\ldots,z_{s-1})}^{(c_1,\ldots,c_s)})) = (-1)^{s-j}$ for $1 \leq j \leq s$;

4. $q_{v_{(z_1,\ldots,z_{s-1})}^{(c_1,\ldots,c_s)}}(x) = 0$ and $x > n-2 \Longrightarrow x \in \{z_1, \ldots, z_{s-1}\}$, and the algebraic multiplicities of the roots $z_1, \ldots, z_{s-1}$ are 1.

In the first part of the inductive step we want to prove for all $C \in \{1,\dots,n\}^{r<}$ the existence of a $v^C_{(z_1,\dots,z_{r-1})} \in \mathbb{R}^n$ which is uniquely determined by the conditions 1., 2. and fulfills property 3..
Let $C := (c_1,\dots,c_r) \in \{1,\dots,n\}^{r<}$ be arbitrary. We want to write $v^{(c_1,\dots,c_r)}_{(z_1,\dots,z_{r-1})}$ as a difference of

$$v_2 := v^{(c_2,\dots,c_r)}_{(z_1,\dots,z_{r-2})} \qquad \text{and} \qquad v_1 := v^{(c_1,\dots,c_{r-1})}_{(z_1,\dots,z_{r-2})}.$$

Both of them have $L(v_2) = L(v_1) = r-1$, i.e., for them the induction hypothesis holds. And since $\pi_{c_r}(v_2) = \pi_{c_{r-1}}(v_1) = 1$ we have $\lim_{z\to\infty} q_{v_2}(z) = \lim_{z\to\infty} q_{v_1}(z) = \infty$. As $z_{r-2} < z_{r-1}$ we get from 4.

$$q_{v_2}(z_{r-1}) > 0 \qquad \text{and} \qquad q_{v_1}(z_{r-1}) > 0.$$

Define

$$y := \frac{q_{v_2}(z_{r-1})}{q_{v_1}(z_{r-1})} > 0 \qquad \text{and} \qquad v := v_2 - yv_1.$$

For v we have $q_v(z_j) = 0$ for $1 \le j \le r-1$ which is condition 2.. 1. and 3. can be seen in the following way. Consider $\pi_{c_j}(v_2) - y\pi_{c_j}(v_1) = \pi_{c_j}(v)$ for $1 \le j \le r$ (compare figure 3.2.6; remember that $C(v_2) = (c_2,\dots,c_r)$, $C(v_1) = (c_1,\dots,c_{r-1})$).

$$\begin{aligned}
&\operatorname{sign}(\pi_{c_{r-1}(v_2)}(v_2)) = (-1)^{(r-1)-(r-1)} && \Longrightarrow \quad \operatorname{sign}(\pi_{c_r(v)}(v)) = (-1)^{r-r};\\
&\operatorname{sign}(\pi_{c_{j-1}(v_2)}(v_2)) = (-1)^{(r-1)-(j-1)} && \text{and}\\
&\operatorname{sign}(\pi_{c_j(v_1)}(v_1)) \quad = (-1)^{(r-1)-j} && \Longrightarrow \quad \operatorname{sign}(\pi_{c_j(v)}(v)) = (-1)^{r-j} \qquad (1 < j < r);\\
&\operatorname{sign}(\pi_{c_1(v_1)}(v_1)) \quad = (-1)^{(r-1)-1} && \Longrightarrow \quad \operatorname{sign}(\pi_{c_1(v)}(v)) = (-1)^{r-1}.
\end{aligned}$$

$$\begin{aligned}
v_2 &= (\qquad\qquad\qquad\qquad \dots, \pi_{c_{r-3}(v_2)}(v_2) > 0, \dots, \pi_{c_{r-2}(v_2)}(v_2) < 0, \dots, \pi_{c_{r-1}(v_2)}(v_2) > 0, \dots)\\
v_1 &= (\dots, \pi_{c_{r-3}(v_1)}(v_1) > 0, \dots, \pi_{c_{r-2}(v_1)}(v_1) < 0, \dots, \pi_{c_{r-1}(v_1)}(v_1) > 0, \dots \qquad\qquad\qquad)
\end{aligned}$$

Figure 3.2.3: Alternating signs of the nonzero coordinates of $v = v_2 - yv_1$.

I.e., the coordinates of v at positions C have alternating signs and are nonzero, in particular, thus $C(v) = C$.
Till here, we have shown that v fulfills 1., 2. and 3.; now the uniqueness comes. Let $\tilde{v} \in \mathbb{R}^n$ also fulfill 1. and 2. and $w := v - \tilde{v}$. Since $\pi_{c_r}(w) = 1 - 1 = 0$ holds $L(w) < r$. But the cases $1 \le L(w) < r$ are impossible, because in these cases from the induction hypothesis follows that w, divided by its highest nonzero coordinate,

$$\frac{1}{\pi_{c_{L(w)}(w)}(w)} w,$$

fulfills the conditions 1. w.r.t. $C(w)$, 2. for $1 \le j < L(w)$ and thus even 4. which claims $q_w(z_{r-1}) \neq 0$. Therefore $L(w) = 0$ and thus $\tilde{v} = v =: v^C_{(z_1,\dots,z_{r-1})}$.

We have now proven for all $C \in \{1, \ldots, n\}^{r<}$ the existence of a $v^C_{(z_1,\ldots,z_{r-1})}$ which is uniquely determined by the conditions 1., 2. and fulfills property 3..
In the second part of the inductive step we want to prove with the help of the first part and lemma 3.2.3 that these $v^C_{(z_1,\ldots,z_{r-1})}$ also fulfill property 4.
Define

$$\begin{aligned} u_1 \quad &:= v^{(1,\ldots,r)}_{(z_1,\ldots,z_{r-1})}; \\ &\vdots \\ u_{n-r+1} &:= v^{(n-r+1,\ldots,n)}_{(z_1,\ldots,z_{r-1})}. \end{aligned}$$

Evidently, $u_1, \ldots, u_{n-r+1}$ fulfill property I. of lemma 3.2.3.. Property II. can be seen in the following way. Let be $u := \sum_{k=1}^{n-r+1} x_k u_k$ with $x_k \in \mathbb{R}$ an arbitrary linear combination of the u_k with $0 < L(u) \leq r$ and highest nonzero coordinate $\pi_{c_{L(u)}(u)}(u) = 1$. Then we have

$$q_u(z_j) = \sum_{k=1}^{n-r+1} x_k q_{u_k}(z_j) = 0$$

for $1 \leq j \leq r-1$. This again makes the cases $1 \leq L(u) < r$ impossible, because in these cases u fulfills the conditions 1. w.r.t. $C(u)$, 2. for $1 \leq j < L(u)$, thus 4. by the induction hypothesis which leads to $q_u(z_{r-1}) > 0$. Thus $L(u) = r$, and since u fulfills the conditions 1. and 2., from the first part follows that u is uniquely determined by these conditions, thus

$$u = v^{C(u)}_{(z_1,\ldots,z_{r-1})}.$$

And, particularly,

$$\operatorname{sign}(\pi_{c_j(u)}(u)) = (-1)^{r-j}$$

for $1 \leq j \leq r$ which is property II. of lemma 3.2.3. Now we enter lemma 3.2.3 with the $u_1, \ldots, u_{n-r+1}$, and it provides for every $C := (c_1, \ldots, c_r) \in \{1, \ldots, n\}^{r<}$ some $0 \leq y_k \in \mathbb{R}$ $(1 \leq k \leq n-r+1)$ with

$$C\Big(\sum_{k=1}^{n-r+1} y_k u_k\Big) = C \qquad \text{and} \qquad \pi_{c_r}\Big(\sum_{k=1}^{n-r+1} y_k u_k\Big) = 1,$$

thus from the uniqueness

$$\sum_{k=1}^{n-r+1} y_k u_k = v^C_{(z_1,\ldots,z_{r-1})} \qquad \text{and} \qquad q_{v^C_{(z_1,\ldots,z_{r-1})}}(X) = \sum_{k=1}^{n-r+1} y_k q_{u_k}(X).$$

The reduction of $q_{v^C_{(z_1,\ldots,z_{r-1})}}$ to the q_{u_k} enables us to apply a degree argument to determine

the roots $> n-2$. From the definitions we get

$$\begin{aligned} q_{u_k}(X) =& q_{v^{(k,\dots,k+r-1)}_{(z_1,\dots,z_{r-1})}}(X) = \sum_{i=k}^{k+r-1} \pi_i(v^{(k,\dots,k+r-1)}_{(z_1,\dots,z_{r-1})}) p_{i-1}(X) \\ =& \sum_{i=k}^{k+r-1} \pi_i(v^{(k,\dots,k+r-1)}_{(z_1,\dots,z_{r-1})}) \prod_{j=0}^{i-2}(X-j) \\ =& \prod_{j=0}^{k-2}(X-j) \left(\sum_{i=k}^{k+r-1} \pi_i(v^{(k,\dots,k+r-1)}_{(z_1,\dots,z_{r-1})}) \prod_{j=k-1}^{i-2}(X-j) \right) = \prod_{j=0}^{k-2}(X-j) \prod_{j=1}^{r-1}(X-z_j) \end{aligned}$$

since $q_{u_k}(z_j) = 0$ for $1 \le j < r$, $\deg(q_{u_k}) = k+r-2$, $\pi_{k+r-1}(v^{(k,\dots,k+r-1)}_{(z_1,\dots,z_{r-1})}) = 1$ and p_{k+r-2} is monic.

$$\begin{aligned} q_{v^C_{(z_1,\dots,z_{r-1})}}(X) &= \sum_{k=1}^{n-r+1} y_k q_{u_k}(X) = \sum_{k=1}^{n-r+1} y_k \prod_{j=0}^{k-2}(X-j) \prod_{j=1}^{r-1}(X-z_j) \\ &= \prod_{j=1}^{r-1}(X-z_j) \underbrace{\left(\sum_{k=1}^{n-r+1} y_k \prod_{j=0}^{k-2}(X-j) \right)}_{=:\tilde{q}(X)} \end{aligned}$$

And since $j \le n-r-1 \le n-2$ in the definition of $\tilde{q}(X)$ and $y_k \ge 0$, we get $\tilde{q}(z) > 0$ for $z > n-2$. Thus from $q_{v^C_{(z_1,\dots,z_{r-1})}}(x) = 0$ and $x > n-2$ follows $x \in \{z_1, \dots, z_{r-1}\}$, and the algebraic multiplicities of the roots z_j are 1; which is property 4.. □

Lemma 3.2.3. *Let $2 \le r \le n$ and $u_1, \dots, u_{n-r+1} \in \mathbb{R}^n$ be n-dimensional vectors for which*

I. $L(u_k) = r$ and highest nonzero coordinate $\pi_{c_{L(u_k)}(u_k)}(u_k) = 1$ for $1 \le k \le n-r+1$ and

$$\begin{aligned} C(u_1) &= (1, \dots, r); \\ &\vdots \\ C(u_{n-r+1}) &= (n-r+1, \dots, n). \end{aligned}$$

II. For every linear combination with nonnegative coefficients $v \in \{\sum_{k=1}^{n-r+1} \mathbb{R}_{\ge 0} u_k\}$ with $0 < L(v) \le r$ and $\pi_{c_{L(v)}(v)}(v) = 1$ we have

$$L(v) = r \qquad \text{and} \qquad \operatorname{sign}(\pi_{c_j(v)}(v)) = (-1)^{r-j} \qquad \text{for } 1 \le j \le r,$$

i.e., the signs of its nonzero coordinates alternate.

Let be

$$\mathbf{U}^{L=r}_{\ge 0} := \left\{ v \in \left\{ \sum_{k=1}^{n-r+1} \mathbb{R}_{\ge 0} u_k \right\} \middle| L(v) = r \text{ and } \pi_{c_{L(v)}(v)}(v) = 1 \right\}$$

the linear combinations of the u_k with nonnegative coefficients, length r and highest nonzero coordinate 1. Then

$$C(\mathbf{U}^{L=r}_{\ge 0}) = \{1, \dots, n\}^{r<},$$

i.e., for every choice of r positions of $\{1,\dots,n\}$ exists a linear combination of the u_k with nonnegative coefficients whose nonzero coordinates are exactly at the chosen positions.

Proof. Let r, n and an arbitrary $C := (c_1,\dots,c_r) \in \{1,\dots,n\}^{r<}$ be fixed.
We want to show inductively that for every $0 \le j < r$ exist certain $v \in \mathbf{U}_{\ge 0}^{L=r}$ with $(c_1(v),\dots,c_j(v)) = (c_1,\dots,c_j)$. In particular, for every $0 \le j < r$ and every $0 \le d \le n - c_j - (r-j)$ exists such a v with

$$\begin{aligned} C(v) &= \big(c_1(v),\dots,c_j(v),c_{j+1}(v) \quad ,c_{j+2}(v) \quad ,\dots,c_r(v) \qquad\qquad \big) \\ &= \big(c_1 \quad ,\dots,c_j \quad ,c_j+d+1,c_j+d+2,\dots,c_j+d+(r-j)\big), \end{aligned} \tag{3.2.6}$$

i.e., the first j nonzero coordinates of v correspond to C and the other nonzero coordinates are consecutive, starting at a position $> c_j$.
Then for the v with $j := r-1$ and $d := c_r - c_{r-1} - 1$ we will have $C(v) = C$ as desired.
Base case; $j = 0$.
Here it must be shown that for every $0 \le e \le n-r$ exists a $w_e \in \mathbf{U}_{\ge 0}^{L=r}$ with $C(w_e) = (e+1,\dots,e+r)$. We can take $w_e := u_{e+1}$ itself, which of course is $\in \mathbf{U}_{\ge 0}^{L=r}$.
Inductive step; $j \to j+1$.
In this step for $0 \le d \le n - c_j - (r-j)$ some $v_d \in \mathbf{U}_{\ge 0}^{L=r}$ are provided with

$$\begin{aligned} C(v_d) &= \big(c_1(v_d),\dots,c_j(v_d),c_{j+1}(v_d) \quad ,\dots,c_r(v_d) \qquad\qquad \big) \\ &= \big(c_1 \quad ,\dots,c_j \quad ,c_j+d+1,\dots,c_j+d+(r-j)\big). \end{aligned}$$

And it must be shown that for every $0 \le e \le n - c_{j+1} - (r-(j+1))$ exists a $w_e \in \mathbf{U}_{\ge 0}^{L=r}$ with

$$\begin{aligned} C(w_e) &= \big(c_1(w_e),\dots,c_{j+1}(w_e),c_{j+2}(w_e) \quad ,\dots,c_r(w_e) \qquad\qquad \big) \\ &= \big(c_1 \quad ,\dots,c_{j+1} \quad ,c_{j+1}+e+1,\dots,c_{j+1}+e+(r-(j+1))\big). \end{aligned}$$

We want to show this by induction on $0 \le e \le n - c_{j+1} - (r-(j+1))$.
Base case; $e = 0$.
For $e = 0$ we have

$$\begin{aligned} &\big(c_1,\dots,c_j,c_{j+1} \qquad\qquad ,c_{j+1}+e+1 \qquad\qquad ,\dots,c_{j+1}+e+(r-(j+1))\big) = \\ &\big(c_1,\dots,c_j,c_j+(c_{j+1}-c_j-1)+1,c_j+(c_{j+1}-c_j-1)+2,\dots,c_j+(c_{j+1}-c_j-1)+(r-j)\big). \end{aligned}$$

Therefore $w_0 := v_{c_{j+1}-c_j-1} \in \mathbf{U}_{\ge 0}^{L=r}$ does the desired.
Inductive step; $e \to e+1$.
In this step a $w_e \in \mathbf{U}_{\ge 0}^{L=r}$ is provided with

$$C(w_e) = \big(c_1,\dots,c_{j+1},c_{j+1}+e+1,\dots,c_{j+1}+e+(r-(j+1))\big)$$

and a $w_{e+1} \in \mathbf{U}_{\ge 0}^{L=r}$ must be found with

$$C(w_{e+1}) = \big(c_1,\dots,c_{j+1},c_{j+1}+e+2,\dots,c_{j+1}+e+(r-j)\big).$$

This w_{e+1} will be expressed as a linear combination of w_e and $v_{c_{j+1}-c_j+e}$. We have

$$\begin{aligned} &C(w_e) \\ &= \big(c_1,\dots,c_j,c_{j+1},c_{j+1}+e+1 \qquad\qquad ,\dots,c_{j+1}+e+(r-(j+1))\big) \qquad \text{and} \\ &C(v_{c_{j+1}-c_j+e}) \\ &= \big(c_1,\dots,c_j, \qquad c_j+(c_{j+1}-c_j+e)+1,\dots,c_{j+1}+e+(r-(j+1)),c_{j+1}+e+(r-j)\big), \end{aligned} \tag{3.2.7}$$

in particular, $c_{j+2}(w_e) = c_{j+1} + e + 1 = c_{j+1}(v_{c_{j+1}-c_j+e})$. Furthermore, from property II. follows

$$\begin{aligned} \operatorname{sign}(\pi_{c_{j+2}(w_e)}(w_e)) &= (-1)^{r-(j+2)} \qquad \text{and} \\ \operatorname{sign}(\pi_{c_{j+1}(v_{c_{j+1}-c_j+e})}(v_{c_{j+1}-c_j+e})) &= (-1)^{r-(j+1)}. \end{aligned}$$

Therefore

$$y := -\frac{\pi_{c_{j+1}(v_{c_{j+1}-c_j+e})}(v_{c_{j+1}-c_j+e})}{\pi_{c_{j+2}(w_e)}(w_e)} > 0$$

and

$$w_{e+1} := v_{c_{j+1}-c_j+e} + y \cdot w_e \in \Big\{ \sum_{k=1}^{n-r+1} \mathbb{R}_{\geq 0} u_k \Big\}.$$

Moreover, we have

$$\pi_{c_{j+1}+e+1}(w_{e+1}) = \pi_{c_{j+1}+e+1}(v_{c_{j+1}-c_j+e}) + y \cdot \pi_{c_{j+1}+e+1}(w_e) = 0. \tag{3.2.8}$$

How do the other coordinates of w_{e+1} look like? From (3.2.7) and (3.2.8) follows that at least all of them besides them at the r positions

$$\big(c_1, \ldots, c_{j+1}, c_{j+1} + e + 2, \ldots, c_{j+1} + e + (r-j)\big)$$

are 0. Since $\pi_{c_{j+1}+e+(r-j)}(w_{e+1}) = \pi_{c_{j+1}+e+(r-j)}(v_{c_{j+1}-c_j+e}) = 1$ follows $0 < L(w_{e+1}) \leq r$ thus

$$L(w_{e+1}) = r$$

with property II. (which even postulates that the nonzero coordinates of w_{e+1} have alternating signs, which will be needed in the next inductive step). Therefore $w_{e+1} \in \mathbf{U}_{\geq 0}^{L=r}$ does the desired.
End of inductive step.
End of inductive step.
As announced, the inductions provide for every $0 \leq j < r$ and $0 \leq d \leq n - c_j - (r-j)$ a $v \in \mathbf{U}_{\geq 0}^{L=r}$ for which (3.2.6) holds. For $j := r-1$ and $d := c_r - c_{r-1} - 1$, we get

$$\begin{aligned} C(v) &= \big(c_1(v), \ldots, c_{r-1}(v), c_r(v) \qquad\qquad\qquad\qquad \big) \\ &= \big(c_1 \quad, \ldots, c_{r-1} \quad, c_{r-1} + (c_r - c_{r-1} - 1) + 1\big) = C \end{aligned}$$

as desired. □

3.2.3 Generalization; example and theorem

In this section we will generalize theorem 3.2.2. Theorem 3.2.2 provides for some

$$(c_1, \ldots, c_r) \in \{1, \ldots, n\}^{r<} \qquad \text{and} \qquad n - 2 < z_1 < \cdots < z_{r-1} \in \mathbb{R}$$

a linear combination

$$q := \sum_{j=1}^{r} v_{c_j} p_{c_j-1}$$

for which $q(z_j) = 0$ for all j. From $c_r \leq n < z_1 + 2$ follows that all demanded roots z_j are greater than the roots of all the used polynomials p_{c_j-1}. But it is also possible to allow the c_j to be greater under certain conditions which are specialized in remark 3.2.4. If $z_1 \in \mathbb{N}$ and $z_1 + 1 < c_1 < c_2 < \dots$ then

$$p_{c_j-1}(z_1) = \prod_{k=0}^{c_j-2} (z_1 - k) = 0$$

for all j; therefore the demand $q(z_1) = 0$ is fulfilled anyway by every linear combination q of the p_{c_j-1} and can be omitted. This leads to the situation that no longer r but only $\rho \leq r$ polynomials p_{c_j-1} are needed to achieve the demanded roots $z_1, \dots, z_{r-1}$.
The proof of the general theorem 3.2.7 equals the proof of theorem 3.2.2 where $\{1, \dots, n\}^{r<}$ is replaced by $\mathbf{C}_{(z_1,\dots,z_{r-1})}$. A little restriction to theorem 3.2.2 (where $z_1, \dots, z_{r-1} \in \mathbb{R}$) appears since here only natural roots $z_1, \dots, z_{r-1}$ are allowed. On the other hand this enables us to use in the whole section coordinate vectors $v \in \mathbb{Q}^n$ instead of $\mathbb{R}^n$.

Remark 3.2.4 (about definition 3.1.1 ix) of $\mathbf{C}_{(z_1,\dots,z_{r-1})}$). In section 3.2.2 we used for $r-1$ roots $z_1 < \dots < z_{r-1}$ r numbers $c_1 < \dots < c_r$ with $c_r - 1 \leq z_1$. Here we use only $\rho \leq r$ numbers $c_1 < \dots < c_\rho$ whose order is mixed with the z_j in the following way. Place c_1 arbitrarily such that $c_1 - 1 \neq z_j$ for all j and then place for every $z_j > c_1 - 1$ a $c_{j+1-(r-\rho)}$ such that $c_{j+1-(r-\rho)} - 1 \leq z_j$. Then $(c_1, \dots, c_\rho) \in \mathbf{C}_{(z_1,\dots,z_{r-1})}$; compare figure 3.2.4.

$$\begin{array}{llll}
z_1 & < z_2 & < z_3 & < z_4 < c_1 - 1 \\
z_1 < c_1 - 1 < c_2 - 1 < c_3 - 1 < c_4 - 1 & \leq z_2 & < z_3 & < z_4 \\
z_1 < c_1 - 1 < c_2 - 1 < c_3 - 1 & \leq z_2 \leq c_4 - 1 & \leq z_3 & < z_4 \\
z_1 < c_1 - 1 < c_2 - 1 & \leq z_2 \leq c_3 - 1 & \leq z_3 \leq c_4 - 1 & \leq z_4
\end{array}$$

Figure 3.2.4: Examples for $(c_1, \dots, c_\rho) \in \mathbf{C}_{(z_1,\dots,z_4)}$.

Definition 3.2.5.

i) Let $1 \leq r \leq s \leq n$. An ordered r-tuple $C := (c_1, \dots, c_r) \in \{1, \dots, n\}^{r<}$ is said to be a subset of an ordered s-tuple $\tilde{C} := (\tilde{c}_1, \dots, \tilde{c}_s) \in \{1, \dots, n\}^{s<}$ (written $C \subset \tilde{C}$) if $\{c_1, \dots, c_r\} \subset \{\tilde{c}_1, \dots, \tilde{c}_s\}$.

ii) Let $1 \leq t \leq n$ and $0 \leq z_1 < \dots < z_{t-1} \leq n \in \mathbb{N}$. Let the *restriction* w.r.t. $(z_1, \dots, z_{t-1})$ be the map

$$R_{(z_1,\dots,z_{t-1})} : \{1, \dots, n\}^{t<} \to \mathbf{C}_{(z_1,\dots,z_{t-1})}, \quad (c_1, \dots, c_t) \mapsto (c_{1+t-\tau}, \dots, c_t)$$

which maps an ordered t-tuple $(c_1, \dots, c_t)$ to the uniquely defined τ-tuple $(1 \leq \tau \leq t)$ $(c_{1+t-\tau}, \dots, c_t)$ such that $(c_{1+t-\tau}, \dots, c_t) \in \mathbf{C}_{(z_1,\dots,z_{t-1})}$.
(For existence and uniqueness consider lemma 3.2.8 iv).)

Example 3.2.6 (for theorem 3.2.7). Let $r := 4$ and $(c_1, c_2, c_3, z_1, z_2, z_3) := (3, 5, 7, 1, 4, 6)$. Since

$$z_1 < c_1 - 1 < c_2 - 1 \leq z_2 \quad \text{and} \quad c_3 - 1 \leq z_3 \qquad \Longrightarrow \qquad (c_1, c_2, c_3) \in \mathbf{C}_{(z_1,z_2,z_3)}$$

with $\rho = 3$ (this will turn out as sufficiently that the following works). We will first calculate v_3, v_5 such that with $v := (0, 0, v_3, 0, v_5, 0, 1)$ and

$$q_v = \sum_{i=1}^{7} \pi_i(v) p_{i-1}$$

$q_v(1) = q_v(4) = q_v(6) = 0$. We will then show that q_v is positive at $\{2, 3, 7, 8, \dots\}$ and negative at $\{5\}$.
The example is illustrated in figure 3.2.5.

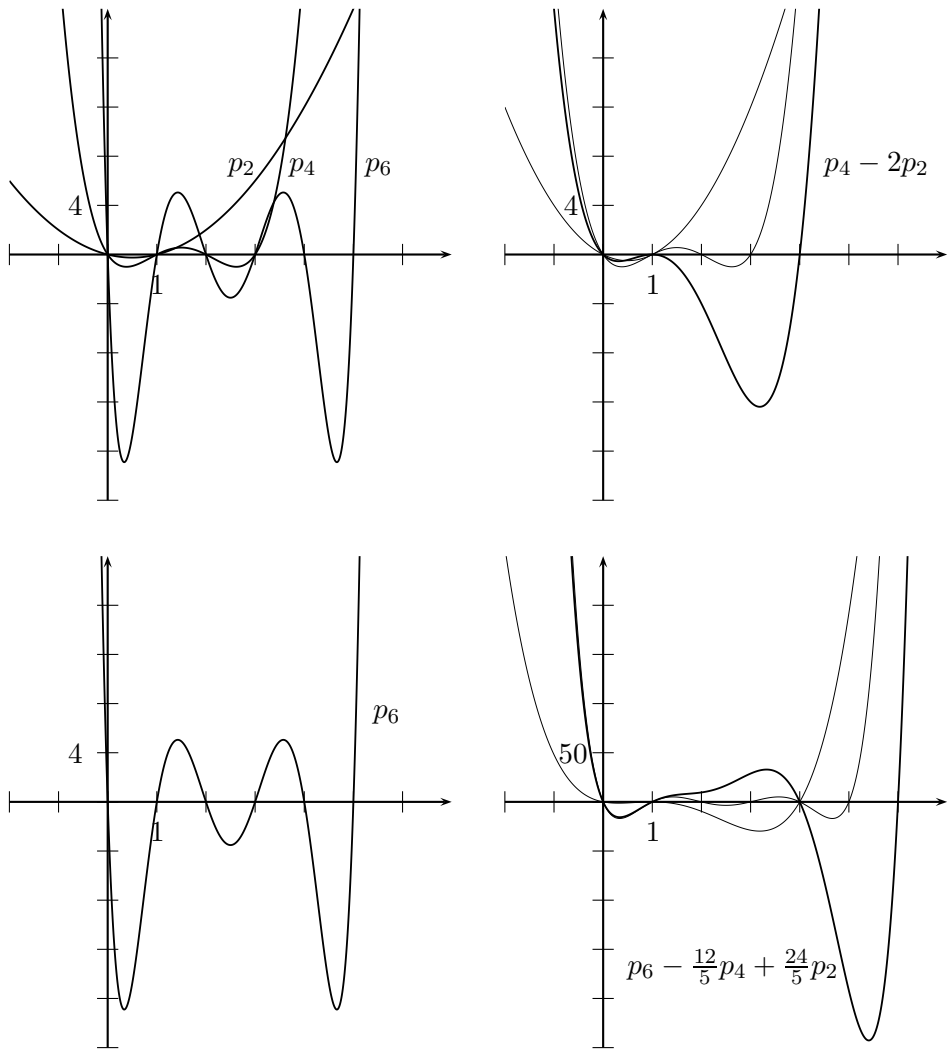


Figure 3.2.5: Example.

In the first step we consider linear combinations which admit a root at 4. We take q_α with

$\alpha := (0,0,-2,0,1,0,0)$ and q_β with $\beta := (0,0,0,0,0,0,1)$.
In the second step we compute q_v with $v := \beta - (12/5)\alpha = (0,0,24/5,0,-12/5,0,1)$ for which $q_v(4) = q_v(6) = 0$. To determine the signs of q_v at the naturals we consider $\gamma := (0,0,8,-5,1,0,0)$, $\delta := (0,0,0,3,-3,1,0)$ and $\varepsilon := (0,0,0,0,0,-1,1)$ which also admit roots at 4 and 6. But for the q_γ, q_δ, q_ε follows from their degrees that they are nonnegative at $\{2,3\}$, negative at $\{5\}$ and positive at $\{7,8,\dots\}$. We now write

$$v = \begin{pmatrix} 0 \\ 0 \\ \frac{24}{5} \\ 0 \\ -\frac{12}{5} \\ 0 \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 \\ 0 \\ \frac{24}{5} \\ 0 \\ -\frac{12}{5} \\ 1 \\ 0 \end{pmatrix}}_{=\zeta} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} = \underbrace{\left[\frac{3}{5} \begin{pmatrix} 0 \\ 0 \\ 8 \\ -5 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 3 \\ -3 \\ 1 \\ 0 \end{pmatrix} \right]}_{=:\zeta} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} = \tfrac{3}{5}\gamma + \delta + \varepsilon. \tag{3.2.9}$$

Because of the positive coefficients $3/5, 1, 1$ we are done.
Why can we find these positive coefficients? Consider ζ in (3.2.9) where the 4th coordinate is eliminated; here the coefficients are positive since γ and δ have alternating signs at their nonzero coordinates. The question is why does this also hold for ζ? And again, the answer is given by the uniqueness of ζ (later, the uniqueness will be shown to be implied by $C(\zeta) = (3,5,6) \in \mathbf{C}_{(1,4,6)}$). The details come in

Theorem 3.2.7. *Let $1 \le \tau \le t \le n$,*

$$0 \le z_1 < \cdots < z_{t-1} \le n \in \mathbb{N} \qquad \textit{and} \qquad C := (c_1, \dots, c_\tau) \in \mathbf{C}_{(z_1,\dots,z_{t-1})}$$

with $c_\tau \le n$.
Then there exists a unique vector $v \in \mathbb{Q}^n$ with

$$C(v) = C \qquad \textit{and} \qquad \pi_{c_\tau}(v) = 1$$

such that

$$q_v(z_1) = \cdots = q_v(z_{t-1}) = 0.$$

Furthermore, the nonzero coordinates of v have alternating signs, i.e.,

$$\operatorname{sign}(\pi_{c_j}(v)) = (-1)^{\tau - j}$$

for $1 \le j \le \tau$;
and for natural x with $0 \le x < c_1 - 1$ we have $q_v(x) = 0$ and with $c_1 - 1 \le x$

$$\operatorname{sign}(q_v(x)) = \begin{cases} (-1)^{t-1} & \textit{for} \quad x < z_1; \\ (-1)^{t-j} & \textit{for} \quad z_{j-1} < x < z_j \quad \textit{and} \quad 1 < j < t; \\ 1 & \textit{for} \quad z_{t-1} < x. \end{cases} \tag{3.2.10}$$

Proof. We want to show the claim by induction over $1 \leq \rho \leq \tau$.
Base case; $\rho := 1$.
In this case from $(c_1) \in \mathbf{C}_{(z_1,\dots,z_{t-1})}$ follows $z_{t-1} < c_1 - 1$. Thus let be $c_1 \geq z_{t-1} + 2$ arbitrary (arbitrary if $t = 1$) and define $v \in \mathbb{Q}^n$ by $\pi_{c_1}(v) = 1$ and $\pi_i(v) = 0$ for $i \neq c_1$. Then

i) the conditions $C(v) = (c_1)$ and $\pi_{c_1}(v) = 1$ determine v in a unique way; therefore it shall be denoted by $v^{(c_1)}_{(z_1,\dots,z_{t-1})} := v$;

ii) $\operatorname{sign}(\pi_{c_1}(v)) = 1 = (-1)^{\rho-1}$;

iii) $q_v(x) = 1 \cdot p_{c_1-1}(x) \begin{cases} = 0 & \text{for natural } x < c_1 - 1 \text{ (particularly, for } x \in \{z_1, \dots, z_{t-1}\}); \\ > 0 & \text{for } c_1 - 1 \leq x. \end{cases}$

Inductive step for $1 < \rho \leq \tau$; $\rho - 1 \to \rho$.
In this step is provided that for every $1 \leq \sigma < \rho$, $\sigma \leq s \leq n$, $0 \leq z_1 < \dots < z_{s-1} \leq n \in \mathbb{N}$ and $C := (c_1, \dots, c_\sigma) \in \mathbf{C}_{(z_1,\dots,z_{s-1})}$ with $c_\sigma \leq n$ exists a $v^{(c_1,\dots,c_\sigma)}_{(z_1,\dots,z_{s-1})} \in \mathbb{Q}^n$ which is uniquely determined by the conditions 1. and 2.

1. $C(v^{(c_1,\dots,c_\sigma)}_{(z_1,\dots,z_{s-1})}) = C \qquad \text{and} \qquad \pi_{c_\sigma}(v^{(c_1,\dots,c_\sigma)}_{(z_1,\dots,z_{s-1})}) = 1$;

2. $q_{v^{(c_1,\dots,c_\sigma)}_{(z_1,\dots,z_{s-1})}}(z_j) = 0$ for $1 \leq j < s$.

Additionally is provided that

3. $\operatorname{sign}(\pi_{c_j}(v^{(c_1,\dots,c_\sigma)}_{(z_1,\dots,z_{s-1})})) = (-1)^{\sigma-j}$ for $1 \leq j \leq \sigma$;

4. $q_{v^{(c_1,\dots,c_\sigma)}_{(z_1,\dots,z_{s-1})}}(x)$ admits the signs as in (3.2.10) where v is replaced by $v^{(c_1,\dots,c_\sigma)}_{(z_1,\dots,z_{s-1})}$ and t by s.

In the first part of the inductive step we want to prove for all $(c_1, \dots, c_\rho) \in \mathbf{C}_{(z_1,\dots,z_{t-1})}$ the existence of a $v^{(c_1,\dots,c_\rho)}_{(z_1,\dots,z_{t-1})} \in \mathbb{Q}^n$ which is uniquely determined by the conditions 1., 2. and fulfills property 3..
Let $0 \leq z_1 < \dots < z_{t-1} \leq n \in \mathbb{N}$ be as in the assumption and fixed and $C := (c_1, \dots, c_\rho) \in \mathbf{C}_{(z_1,\dots,z_{t-1})}$ be arbitrary and fixed. We want to write $v^{(c_1,\dots,c_\rho)}_{(z_1,\dots,z_{t-1})}$ as a difference of

$$v_2 := v^{(c_{1+\rho-\tilde{\rho}},\dots,c_\rho)}_{(z_1,\dots,z_{t-2})} \qquad \text{and} \qquad v_1 := v^{(c_1,\dots,c_{\rho-1})}_{(z_1,\dots,z_{t-2})}$$

with a $1 \leq \tilde{\rho} < \rho$ which is provided by lemma 3.2.8 ii). Lemma 3.2.8 also claims $(c_{1+\rho-\tilde{\rho}}, \dots, c_\rho)$, $(c_1, \dots, c_{\rho-1}) \in \mathbf{C}_{(z_1,\dots,z_{t-2})}$. And since $L(v_2) = \tilde{\rho} < \rho$ and $L(v_1) = \rho - 1$, for both vectors the induction hypothesis holds. Furthermore, since

$$c_1 - 1 < c_{1+\rho-\tilde{\rho}} - 1 \leq c_\rho - 1 \leq z_{t-1} \qquad \text{and} \qquad z_{t-2} < z_{t-1}$$

we get from 4. that $q_{v_2}(z_{t-1}) > 0$ and $q_{v_1}(z_{t-1}) > 0$. Define

$$y := \frac{q_{v_2}(z_{t-1})}{q_{v_1}(z_{t-1})} > 0 \qquad \text{and} \qquad v := v_2 - y v_1.$$

For v holds $q_v(z_j) = 0$ for $1 \leq j \leq t-1$ which is condition 2.. 1. and 3. can be seen in the following way. Consider $\pi_{c_j}(v_2) - y\pi_{c_j}(v_1) = \pi_{c_j}(v)$ for $1 \leq j \leq \rho$ (compare figure 3.2.6; remember that $C(v_2) = (c_{1+\rho-\tilde{\rho}}, \ldots, c_\rho)$, $C(v_1) = (c_1, \ldots, c_{\rho-1})$).

$$\begin{aligned}
&\operatorname{sign}(\pi_{c_{\tilde{\rho}}(v_2)}(v_2)) &&= (-1)^{\tilde{\rho}-\tilde{\rho}} && \Rightarrow \operatorname{sign}(\pi_{c_r(v)}(v)) = (-1)^{\rho-\rho}; \\
&\operatorname{sign}(\pi_{c_{j-(\rho-\tilde{\rho})}(v_2)}(v_2)) &&= (-1)^{\tilde{\rho}-(j-(\rho-\tilde{\rho}))} && \text{and} \\
&\operatorname{sign}(\pi_{c_j(v_1)}(v_1)) &&= (-1)^{(\rho-1)-j} && \Rightarrow \operatorname{sign}(\pi_{c_j(v)}(v)) = (-1)^{\rho-j} \quad (\rho-\tilde{\rho} < j < \rho); \\
&\operatorname{sign}(\pi_{c_j(v_1)}(v_1)) &&= (-1)^{(\rho-1)-j} && \Rightarrow \operatorname{sign}(\pi_{c_j(v)}(v)) = (-1)^{\rho-j} \quad (1 \leq j \leq \rho-\tilde{\rho}).
\end{aligned}$$

$$\begin{aligned}
v_2 &= (\qquad\qquad\qquad\qquad ..., \pi_{c_{\tilde{\rho}-2}(v_2)}(v_2) > 0, ..., \pi_{c_{\tilde{\rho}-1}(v_2)}(v_2) < 0, ..., \pi_{c_{\tilde{\rho}}(v_2)}(v_2) > 0, ...) \\
v_1 &= (..., \pi_{c_{\rho-3}(v_1)}(v_1) > 0, ..., \pi_{c_{\rho-2}(v_1)}(v_1) < 0, ..., \pi_{c_{\rho-1}(v_1)}(v_1) > 0, ...\qquad\qquad)
\end{aligned}$$

Figure 3.2.6: Alternating signs of the nonzero coordinates of $v = v_2 - yv_1$.

I.e., the coordinates of v at positions C have alternating signs and are nonzero, in particular, thus $C(v) = C$.
Till here we have shown that v fulfills 1., 2. and 3.; now the uniqueness comes. Let $\tilde{v} \in \mathbb{Q}^n$ also fulfill 1. and 2. and $w := v - \tilde{v}$. Since $\pi_{c_\rho}(w) = 1 - 1 = 0$ we have $L(w) < \rho$. Assume $1 \leq \sigma := L(w) < \rho$ then lemma 3.2.8 iii) claims $C(w) \in \mathbf{C}_{(z_{t-\sigma+1},\ldots,z_{t-1})}$. And since $q_w(z_j) = 0$ for $1 \leq j \leq t-1$, the induction hypothesis holds for w, divided by its highest nonzero coordinate,

$$\frac{1}{\pi_{c_\sigma(w)}(w)} w = v^{C(w)}_{(z_{t-\sigma+1},\ldots,z_{t-1})}.$$

Therefore property 4. and

$$c_1(w) - 1 = c_{1+\rho-\sigma} - 1 \leq z_{t-\rho+(1+\rho-\sigma)-1} = z_{t-\sigma}$$

lead to $q_w(z_{t-\sigma}) \neq 0$. Which is a contradiction, and therefore $L(w) = 0$, i.e., $\tilde{v} = v =: v^C_{(z_1,\ldots,z_{t-1})}$.
We have now proven for all $(c_1, \ldots, c_\rho) \in \mathbf{C}_{(z_1,\ldots,z_{t-1})}$ the existence of a $v^{(c_1,\ldots,c_\rho)}_{(z_1,\ldots,z_{t-1})} \in \mathbb{Q}^n$ which is uniquely determined by the conditions 1., 2. and fulfills property 3..
In the second part of the inductive step we want to prove with the help of the first part and lemma 3.2.9 that these $v^{(c_1,\ldots,c_\rho)}_{(z_1,\ldots,z_{t-1})}$ also fulfill property 4.
Lemma 3.2.9 deals with only $\rho - 1$ roots which means no loss of generality since we can omit the roots z_j with $z_j < c_1 - 1$, which is fulfilled for $1 \leq j \leq t - \rho$. Therefore in the following, we take only the roots $(z_{t-\rho+1}, \ldots, z_{t-1})$; for short let denote $R := R_{(z_{t-\rho+1},\ldots,z_{t-1})}$ the restriction w.r.t. $(z_{t-\rho+1}, \ldots, z_{t-1})$. Define

$$\begin{aligned}
u_1 &:= v^{R(1,\ldots,\rho)}_{(z_{t-\rho+1},\ldots,z_{t-1})}; \\
&\vdots \\
u_{n-r+1} &:= v^{R(n-\rho+1,\ldots,n)}_{(z_{t-\rho+1},\ldots,z_{t-1})},
\end{aligned}$$

which all are yet defined since $R(j,\dots,\rho+j-1)$ $(1\le j\le n-\rho+1)$ contains at most ρ elements. Evidently, $u_1,\dots,u_{n-r+1}$ fulfill property I. of lemma 3.2.9..
Property II. can be seen in the following way. Let $\tilde{C}\in\mathbf{C}_{(z_{t-\rho+1},\dots,z_{t-1})}$ be arbitrary and $u:=\sum_{k=1}^{n-\rho+1}y_ku_k$ with $y_k\in\mathbb{Q}$ be an arbitrary linear combination of the u_k such that

$$C(u)\subset\tilde{C}$$

and highest nonzero coordinate $\pi_{c_{L(u)}(u)}(u)=1$. Then we have

$$q_u(z_j)=\sum_{k=1}^{n-\rho+1}y_kq_{u_k}(z_j)=0 \tag{3.2.11}$$

for $t-\rho+1\le j\le\ t-1$. Assumed

$$C(u)=:(c_1',\dots,c_{\sigma'}')\neq\tilde{C}=:(\tilde{c}_1,\dots,\tilde{c}_\sigma),$$

i.e., $\sigma'<\sigma$, we get from lemma 3.2.8 iii) that $C(u)\in\mathbf{C}_{(z_{t-\sigma'+1},\dots,z_{t-1})}$. And as

$$c_1'-1\le\tilde{c}_{1+\sigma-\sigma'}-1\le z_{t-\sigma'}$$

follows $q_u(z_{t-\sigma'})\neq 0$ which is a contradiction to (3.2.11). Thus $C(u)=\tilde{C}\in\mathbf{C}_{(z_{t-\rho+1},\dots,z_{t-1})}$. Furthermore, since $L(u)\le\rho$ we get that u is uniquely determined by the induction hypothesis (in case $L(u)<\rho$) or the first part of the inductive step (in case $L(u)=\rho$), i.e.,

$$u=v^{C(u)}_{(z_{t-\rho+1},\dots,z_{t-1})}.$$

And particularly

$$\operatorname{sign}(\pi_{c_j(u)}(u))=(-1)^{\sigma'-j}$$

for $1\le j\le\sigma'$ which is property II. of lemma 3.2.9.
Now we enter lemma 3.2.9 with the $u_1,\dots,u_{n-\rho+1}$; note that the roots called $(z_{t-\rho+1},\dots,z_{t-1})$ here, are called $(z_1,\dots,z_{\rho-1})$ in lemma 3.2.9. The Lemma provides for every $C:=(c_1,\dots,c_\rho)\in\mathbf{C}_{(z_{t-\rho+1},\dots,z_{t-1})}$ some $0\le y_k\in\mathbb{Q}$ $(1\le k\le n-\rho+1)$ with

$$C\Big(\sum_{k=1}^{n-\rho+1}y_ku_k\Big)=C\qquad\text{and}\qquad\pi_{c_\rho}\Big(\sum_{k=1}^{n-\rho+1}y_ku_k\Big)=1,$$

thus from the uniqueness

$$\sum_{k=1}^{n-\rho+1}y_ku_k=v^C_{(z_{t-\rho+1},\dots,z_{t-1})}\qquad\text{and}\qquad q_{v^C_{(z_{t-\rho+1},\dots,z_{t-1})}}(X)=\sum_{k=1}^{n-\rho+1}y_kq_{u_k}(X).$$

To use the degree argument for the q_{u_k}, we must consider $R(k,\dots,k+\rho-1)$ for $1\le k\le n-\rho+1$. Let $0\le r_k<\rho$ be such that

$$R(k,\dots,k+\rho-1)=R_{(z_{t-\rho+1},\dots,z_{t-1})}(k,\dots,k+\rho-1)=(k+r_k,\dots,k+\rho-1),$$

i.e., the restriction deletes the r_k leftmost elements; then from the definition of R follows that

$$z_{t-\rho+r_k} < k + r_k - 1 < z_{t-\rho+1+r_k}. \tag{3.2.12}$$

Furthermore,

$$\begin{aligned}
q_{u_k}(X) =& q_{v^{(k+r_k,\dots,k+\rho-1)}_{(z_{t-\rho+1},\dots,z_{t-1})}}(X) = \sum_{i=k+r_k}^{k+\rho-1} \pi_i(v^{(k+r_k,\dots,k+\rho-1)}_{(z_{t-\rho+1},\dots,z_{t-1})})p_{i-1}(X) \\
&= \sum_{i=k+r_k}^{k+\rho-1} \pi_i(v^{(k+r_k,\dots,k+\rho-1)}_{(z_{t-\rho+1},\dots,z_{t-1})}) \prod_{j=0}^{i-2}(X-j) \\
&= \prod_{j=0}^{k+r_k-2}(X-j) \left(\sum_{i=k+r_k}^{k+\rho-1} \pi_i(v^{(k+r_k,\dots,k+\rho-1)}_{(z_{t-\rho+1},\dots,z_{t-1})}) \prod_{j=k+r_k-1}^{i-2}(X-j) \right) \\
&= \prod_{j=0}^{k+r_k-2}(X-j) \prod_{j=t-\rho+1+r_k}^{t-1}(X-z_j)
\end{aligned}$$

since q_{u_k} admits the $(\rho-1-r_k)$ roots $z_{t-\rho+1+r_k},\dots,z_{t-1}$, which are not yet contained in the set of roots $\{0,1,\dots,k+r_k-2\}$, (3.2.12) and $\deg(q_{u_k}) = k+\rho-2$. (Furthermore, $\pi_{k+\rho-1}(v^{(k+r_k,\dots,k+\rho-1)}_{(z_{t-\rho+1},\dots,z_{t-1})})) = 1$ and $p_{k+\rho-2}$ is monic.)
Thus for every natural x we get $q_{u_k}(x) = 0$ for $x \le k+r_k-2$ and for $x > k+r_k-2$

$$\operatorname{sign}(q_{u_k}(x)) = \begin{cases} (-1)^{\rho-1-r_k} & \text{for} \quad x < z_{t-\rho+1+r_k}; \\ (-1)^{t-j} & \text{for} \quad z_{j-1} < x < z_j \quad \text{and} \quad (t-\rho+1+r_k) < j < t; \\ 1 & \text{for} \quad z_{t-1} < x. \end{cases}$$

Since $k = c_1$ is the smallest k for which $y_k \neq 0$, and $r_{c_1} = 0$, we get by considering

$$q_{v^C_{(z_{t-\rho},\dots,z_{t-1})}}(X) = \sum_{k=1}^{n-\rho+1} y_k q_{u_k}(X)$$

that for every natural x we get $q_{v^C_{(z_{t-\rho+1},\dots,z_{t-1})}}(X) = 0$ for $x \le c_1 - 2$ and for $x > c_1 - 2$

$$\operatorname{sign}(q_{v^C_{(z_{t-\rho+1},\dots,z_{t-1})}}(X)) = \begin{cases} (-1)^{t-1} & \text{for} \quad x < z_1; \\ (-1)^{t-j} & \text{for} \quad z_{j-1} < x < z_j \quad \text{and} \quad 1 < j < t; \\ 1 & \text{for} \quad z_{t-1} < x. \end{cases}$$

Finally, since $z_{t-\rho} \le c_1 - 2$, $q_{v^C_{(z_{t-\rho+1},\dots,z_{t-1})}}$ admits the roots $z_1,\dots,z_{t-\rho}$, anyway, thus $v^C_{(z_1,\dots,z_{t-1})} = v^C_{(z_{t-\rho+1},\dots,z_{t-1})}$. □

Lemma 3.2.8. *Let $2 \le \rho \le t$, $0 \le z_1 < \dots < z_{t-1} \in \mathbb{N}$ and $(c_1,\dots,c_\rho) \in \mathbf{C}_{(z_1,\dots,z_{t-1})}$. Then*

i) $(c_1,\dots,c_{\rho-1}) \in \mathbf{C}_{(z_1,\dots,z_{t-2})}$;

ii) it exists a $1 \le \tilde{\rho} < \rho$ such that $(c_{1+\rho-\tilde{\rho}},\dots,c_\rho) \in \mathbf{C}_{(z_1,\dots,z_{t-2})}$;

iii) for $1 \leq \sigma < \rho$ and an arbitrary subset $(\tilde{c}_1, \ldots, \tilde{c}_\sigma) \in \{c_1, \ldots, c_\rho\}^\sigma$ with $\tilde{c}_1 < \cdots < \tilde{c}_\sigma$, we have $(\tilde{c}_1, \ldots, \tilde{c}_\sigma) \in \mathbf{C}_{(z_{t-\sigma+1}, \ldots, z_{t-1})}$.

Let $1 \leq \rho$, $1 \leq t$, $0 \leq z_1 < \cdots < z_{t-1} \in \mathbb{N}$ and $1 \leq c_1 < \cdots < c_\rho \in \mathbb{N}$ be arbitrary such that the number of $j \in \{1, \ldots, t-1\}$ for which $c_1 - 1 \leq z_j$ is less than ρ.

iv) Then there exists one and only one $1 \leq \sigma \leq \rho$ such that $(c_\sigma, \ldots, c_\rho) \in \mathbf{C}_{(z_1, \ldots, z_{t-1})}$; σ is the maximal number $\leq \rho$ for which $c_\sigma - 1 > z_{(t-1)-(\rho-\sigma)}$.

Proof. i) follows directly from the definition 3.1.1 ix) of $\mathbf{C}_{(z_1, \ldots, z_{t-1})}$.
ii) Let be $\tilde{\rho} < \rho$ the smallest number for which $z_{(t-1)-\tilde{\rho}} < c_{1+\rho-\tilde{\rho}} - 1$ holds — which exists since $z_{(t-1)-(\rho-1)} < c_1 - 1 < c_2 - 1 = c_{1+\rho-(\rho-1)} - 1$; b) of definition 3.1.1 ix) is fulfilled. Then we get for $2 \leq j \leq \tilde{\rho}$ that $z_{(t-1)-(\tilde{\rho}-j+1)} \geq c_{1+\rho-(\tilde{\rho}-j+1)} - 1$, i.e., $c_{j+\rho-\tilde{\rho}} - 1 \leq z_{(t-1)-\tilde{\rho}+j-1}$, as desired in c).
iii) As the number of concerned roots equals $\sigma - 1$ the demand 3.1.1 ix) b) is empty, and we must only show c), i.e., $\tilde{c}_j - 1 \leq z_{t-1-(\sigma-j)}$ for $2 \leq j \leq \sigma$. From the assumption follows $c_j - 1 \leq z_{t-\rho+j-1}$ for $2 \leq j \leq \rho$ and, furthermore, we have $\tilde{c}_j \leq c_{j+(\rho-\sigma)}$ for $1 \leq j \leq \sigma$ which leads to

$$\tilde{c}_j - 1 \leq c_{j+(\rho-\sigma)} - 1 \leq z_{t-\rho+j+(\rho-\sigma)-1} = z_{t-1-(\sigma-j)}$$

for $2 \leq j \leq \sigma$.
iv) Existence. Let σ be the maximal number $\leq \rho$ for which $c_\sigma - 1 > z_{(t-1)-(\rho-\sigma)}$. Which exists since from the assumption follows that at least $c_1 - 1 > z_{t-\rho} = z_{(t-1)-(\rho-1)}$. Then definition 3.1.1 ix) b) demands $z_{t-(\rho-\sigma+1)} < c_\sigma - 1$ which holds; 3.1.1 ix) c) demands $c_j - 1 \leq z_{t-\rho+j-1}$ for $\sigma + 1 \leq j \leq \rho$ which is true since

$$c_j - 1 > z_{t-\rho+j-1} \tag{3.2.13}$$

contradicts the maximality of σ. Thus $(c_\sigma, \ldots, c_\rho) \in \mathbf{C}_{(z_1, \ldots, z_{t-1})}$.
Uniqueness. If σ is not maximal, the contradiction follows from (3.2.13). □

It finally remains to formulate a correspondence to lemma 3.2.3. It turns out that lemma 3.2.3 can be adapted to lemma 3.2.9 by replacing systematically — as well in the assumption as in the proof — every $(c_1, \ldots, c_t) \in \{1, \ldots, n\}^{t<}$ by its restriction $R(c_1, \ldots, c_t) = (c_{1+t-\tau}, \ldots, c_t)$ which arises from $(c_1, \ldots, c_t)$ by deleting some of its leftmost entries such that $(c_{1+t-\tau}, \ldots, c_t) \in \mathbf{C}_{(z_1, \ldots, z_{t-1})}$.

Lemma 3.2.9. *Let $2 \leq \rho \leq n$, $0 \leq z_1 < \cdots < z_{\rho-1} \leq n \in \mathbb{N}$ and $u_1, \ldots, u_{n-\rho+1} \in \mathbb{Q}^n$ be n-dimensional vectors for which*

I. $C(u_k) \in \mathbf{C}_{(z_1, \ldots, z_{\rho-1})}$ and highest nonzero coordinate $\pi_{c_{L(u_k)}(u_k)}(u_k) = 1$ for $1 \leq k \leq n - \rho + 1$ and

$$\begin{aligned} C(u_1) &= R_{(z_1, \ldots, z_{\rho-1})}(1, \ldots, \rho); \\ &\vdots \\ C(u_{n-\rho+1}) &= R_{(z_1, \ldots, z_{\rho-1})}(n - \rho + 1, \ldots, n). \end{aligned}$$

II. For every $C \in \mathbf{C}_{(z_1,\dots,z_{\rho-1})}$ and every linear combination with nonnegative coefficients $0 \neq v \in \{\sum_{k=1}^{n-\rho+1} \mathbb{Q}_{\geq 0} u_k\}$ with $C(v) \subset C$ and $\pi_{c_{L(v)}(v)}(v) = 1$ we have

$$C(v) = C \qquad \text{and} \qquad \operatorname{sign}(\pi_{c_j(v)}(v)) = (-1)^{L(v)-j} \qquad \text{for } 1 \leq j \leq L(v).$$

Let be

$$\mathbf{U}_{\geq 0}^{C \in \mathbf{C}} := \Big\{ v \in \Big\{ \sum_{k=1}^{n-\rho+1} \mathbb{Q}_{\geq 0} u_k \Big\} \Big| C(v) \in \mathbf{C}_{(z_1,\dots,z_{\rho-1})} \text{ and } \pi_{c_{L(v)}(v)}(v) = 1 \Big\}$$

the linear combinations of the u_k with nonnegative coefficients whose nonzero coordinates lie in $\mathbf{C}_{(z_1,\dots,z_{\rho-1})}$ and the highest of them equals 1. Then

$$C(\mathbf{U}_{\geq 0}^{C \in \mathbf{C}}) = \mathbf{C}_{(z_1,\dots,z_{\rho-1})},$$

i.e., for every $C \in \mathbf{C}_{(z_1,\dots,z_{\rho-1})}$ exists a linear combination of the u_k with nonnegative coefficients whose nonzero coordinates are exactly at the positions C.

Proof. To prove lemma 3.2.9 we want to adapt the proof of 3.2.3 and explain only where the proof of 3.2.3 must be transformed. Note that in this section it suffices to work over $\mathbb{Q}$.
Let first r in 3.2.3 be $r := \rho$. Since $z_1, \dots, z_{\rho-1}$ are fixed let R denote $R_{(z_1,\dots,z_{\rho-1})}$. We use some vectors w_e, v, w_{e+1} of 3.2.3, while the new vectors for 3.2.9 are denoted by $\tilde{w}_e$, $\tilde{v}$, $\tilde{w}_{e+1}$. What we will do is the following. In 3.2.3 in an outer induction over $0 \leq j < \rho$ and an inner induction over $0 \leq d \leq n - c_j - (\rho - j)$ the existence of some $v_{j,d} \in \mathbf{U}_{\geq 0}^{L=\rho}$ is proven with

$$\begin{aligned} C(v_{j,d}) &= \big(c_1(v_{j,d}), \dots, c_j(v_{j,d}), c_{j+1}(v_{j,d}), c_{j+2}(v_{j,d}), \dots, c_\rho(v_{j,d}) \qquad \big) \\ &= \big(c_1 \qquad , \dots, c_j \qquad , c_j + d + 1, c_j + d + 2, \dots, c_j + d + (\rho - j)\big). \end{aligned}$$

Such that $v_{r-1,c_\rho - c_{\rho-1} - 1}$ does the desired.
In 3.2.9 the $\rho - 1$ roots lie in such a way that $z_1 < \dots < z_{\rho-1} \leq n$; while in 3.2.3 holds $n - 1 \leq z_1$. Therefore in general, for an arbitrary $v \in \mathbb{Q}^n$ of length ρ is $C(v) \notin \mathbf{C}_{(z_1,\dots,z_{\rho-1})}$. We will use the restriction and define some $\tilde{v}$ with $C(\tilde{v}) = R(C(v)) \in \mathbf{C}_{(z_1,\dots,z_{\rho-1})}$. To do this, we follow the same outer and inner induction and define some $\tilde{v}_{j,d}$ for which holds

- $\tilde{v}_{j,d} \in \mathbf{U}_{\geq 0}^{C \in \mathbf{C}}$;
- $C(\tilde{v}_{j,d}) = R(C(v_{j,d}))$,

which is defined, since $L(v_{j,d}) = \rho$ while the number of roots is only $\rho - 1$. As this works for arbitrary $(c_1, \dots, c_\rho) \in \{1, \dots, n\}^{\rho <}$, we get for every $(\tilde{c}_1, \dots, \tilde{c}_\sigma) \in \mathbf{C}_{(z_1,\dots,z_{\rho-1})}$ with $\sigma \leq \rho$ (at least) one $(\tilde{c}_1, \dots, \tilde{c}_\sigma) = R(c_1, \dots, c_\rho)$, i.e.,

$$C(\tilde{v}_{r-1,c_\rho - c_{\rho-1} - 1}) = R(C(v_{r-1,c_\rho - c_{\rho-1} - 1})) = R(c_1, \dots, c_\rho) = (\tilde{c}_1, \dots, \tilde{c}_\sigma),$$

as desired.
To show that this works we must consider from 3.2.3

a) the "base case; $j = 0$";

b) the addition which is made in the "inductive step; $e \to e + 1$".

a) "Base case; $j = 0$". We need for $0 \le e \le n - \rho$ some $\tilde{w}_e \in \mathbf{U}_{\ge 0}^{C\in\mathbf{C}}$ with $C(\tilde{w}_e) = R(e+1, \dots, e+\rho)$. Take $\tilde{w}_e := u_{e+1} \in \mathbf{U}_{\ge 0}^{C\in\mathbf{C}}$ from the assumption of 3.2.9 for which $C(u_{e+1}) = R(e+1, \dots, e+\rho)$ holds.
b) "Inductive step; $e \to e+1$". In 3.2.3 w_e, $v_{c_{j+1}-c_j+e} \in \mathbf{U}_{\ge 0}^{L=\rho}$ are provided and w_{e+1} is written as their linear combination with nonnegative coefficients. With the abbreviation $v := v_{c_{j+1}-c_j+e}$ we get

$$w_{e+1} := v + y \cdot w_e \qquad \text{with} \qquad y := -\frac{\pi_{c_{j+1}(v)}(v)}{\pi_{c_{j+2}(w_e)}(w_e)} > 0.$$

In 3.2.9 we must distinguish two cases at this point. Equivalently, $\tilde{w}_e$, $\tilde{v}_{c_{j+1}-c_j+e} \in \mathbf{U}_{\ge 0}^{C\in\mathbf{C}}$ are provided (again, we write $\tilde{v} := \tilde{v}_{c_{j+1}-c_j+e}$) with $C(\tilde{w}_e) = R(C(w_e))$ and $C(\tilde{v}) = R(C(v))$. And a $\tilde{w}_{e+1} \in \mathbf{U}_{\ge 0}^{C\in\mathbf{C}}$ shall be found with $C(\tilde{w}_{e+1}) = R(C(w_{e+1}))$. It holds true that

$$\begin{aligned}
C(w_e) &= (c_1, ..., c_j, c_{j+1}, c_{j+1} + e + 1, c_{j+1} + e + 2, ..., c_{j+1} + e + \rho - j - 1),\\
C(v) &= (c_1, ..., c_j, \qquad c_{j+1} + e + 1, c_{j+1} + e + 2, ..., c_{j+1} + e + \rho - j - 1, c_{j+1} + e + \rho - j),\\
C(w_{e+1}) &= (c_1, ..., c_j, c_{j+1}, \qquad\qquad c_{j+1} + e + 2, ..., c_{j+1} + e + \rho - j - 1, c_{j+1} + e + \rho - j).
\end{aligned}$$

Thus $L(w_e) = L(v) = L(w_{e+1}) = \rho$, while their restrictions have shorter length;

$$L(\tilde{w}_e) =: \rho - r_1 \le \rho, \qquad L(\tilde{v}) =: \rho - r_2 \le \rho, \qquad L(\tilde{w}_{e+1}) =: \rho - r_3 \le \rho;$$

($\tilde{w}_{e+1}$ is not yet defined, but this is what it is supposed to be). I.e., $C(\tilde{w}_e)$ arises from $C(w_e)$ by deleting the r_1 leftmost entries, etc..
Case 1; $\pi_{c_{j+1}(v)}(\tilde{v}) = 0$. Take $\tilde{w}_{e+1} := \tilde{v}$. It holds that $r_2 \ge j+1$ and $c_{1+r_2}(v) - 1 > z_{r_2}$. And since $c_k(w_{e+1}) = c_k(v)$ for $k > j+1$ follows $R(C(w_{e+1})) = R(C(v)) = C(\tilde{v}) = C(\tilde{w}_{e+1})$, as desired.
Case 2; $\pi_{c_{j+1}(v)}(\tilde{v}) \ne 0$. In this case holds that $r_2 < j+1$, i.e., $c_k(v) - 1 \le z_{k-1}$ for $k > j+1$. Thus $c_k(w_e) - 1 < c_k(w_{e+1}) - 1 = c_k(v) - 1 \le z_{k-1}$, and

$$\tilde{C} := (c_{r_1+1}, \dots, c_j, c_{j+1}, \qquad c_{j+1} + e + 2, \dots, c_{j+1} + e + \rho - j) \in \mathbf{C}_{(z_1,\dots,z_{\rho-1})}.$$

And since $c_l(w_e) = c_l(w_{e+1})$ and $c_l(v) \ge c_l(w_{e+1})$ for $l \le j+1$ follows $r_1 = r_3 \le r_2$. From these considerations follows $\pi_{c_{j+2}(w_e)}(\tilde{w}_e) \ne 0$, as well; thus

$$\tilde{w}_{e+1} := \tilde{v} + y \cdot \tilde{w}_e \qquad \text{with} \qquad y := -\frac{\pi_{c_{j+1}(v)}(\tilde{v})}{\pi_{c_{j+2}(w_e)}(\tilde{w}_e)} > 0;$$

as

$$C(\tilde{w}_{e+1}) \subset \tilde{C} \in \mathbf{C}_{(z_1,\dots,z_{\rho-1})}. \tag{3.2.14}$$

(Equality holds if none of the coordinates of $\tilde{v}$ and $\tilde{w}_e$ besides $\pi_{c_{j+1}(v)}(\tilde{v})$ resp. $\pi_{c_{j+2}(w_e)}(\tilde{w}_e)$ have been extinguished by the addition.) From (3.2.14) and property II. we get

$$C(\tilde{w}_{e+1}) = \tilde{C}.$$

Therefore $\tilde{w}_{e+1} \in \mathbf{U}_{\ge 0}^{C\in\mathbf{C}}$ with alternating signs of its nonzero coordinates and $C(\tilde{w}_{e+1}) = R(C(w_{e+1}))$, as desired. □

Theorem 3.2.7 implies the nice

Corollary 3.2.10. *Let $r \geq 1$ and $1 \leq c_1 < \cdots < c_r$, $0 \leq z_1 < \cdots < z_r$ be naturals. Then we have $c_j - 1 \leq z_j$ for all j if and only if*

$$\begin{vmatrix} p_{c_1-1}(z_1) & \cdots & p_{c_r-1}(z_1) \\ \vdots & \ddots & \vdots \\ p_{c_1-1}(z_r) & \cdots & p_{c_r-1}(z_r) \end{vmatrix} \neq 0.$$

Chapter 4

Conclusion and further work

In this work we present algebraic certificates for the fact that

$$\mathbf{V}(f(a), f'(a), \ldots, f^{(n)}(a)) \geq \mathbf{V}(f(b), \ldots, f^{(n)}(b))$$

where $f \in \mathbf{Q}[X]$ denotes a polynomial of degree n over an ordered field $\mathbf{Q}$, and $a < b \in \mathbf{Q}$. We next aim at an certificate for the fact that

$$f(0)f(t) > 0$$

stipulating that

$$\mathbf{V}(f(0), \ldots, f^{(n)}(0)) = \mathbf{V}(f(1), \ldots, f^{(n)}(1)),$$

$f(0)f(1) > 0$ and $0 < t < 1$. I.e., we would like to give an algorithm which receives as input data:
$n \in \mathbb{N}$ and two sequences of sign conditions

$$(\sigma_0, \ldots, \sigma_n),\ (\tilde{\sigma}_0, \ldots, \tilde{\sigma}_n) \in \{-1, 0, +1\}^{n+1} \tag{4.0.1}$$

with $\sigma_0\tilde{\sigma}_0 > 0$ and

$$\mathbf{V}(\sigma_0, \ldots, \sigma_n) = \mathbf{V}(\tilde{\sigma}_0, \ldots, \tilde{\sigma}_n).$$

And calculates:
For $0 \leq i \leq n$ some polynomials z, z_i, $\tilde{z}_i \in \mathbb{Z}[X]$ such that for all $0 < t < 1$

$$\sigma_0 z(t) > 0, \qquad \sigma_i z_i(t) \geq 0, \qquad \tilde{\sigma}_i \tilde{z}_i(t) \geq 0 \tag{4.0.2}$$

and for all t

$$z(t)f(t) = \sum_i z_i(t)\frac{f^{(i)}(0)}{i!} + \sum_i \tilde{z}_i(t)\frac{f^{(i)}(1)}{i!}$$

for every polynomial $f \in \mathbf{Q}[X]$ of degree n over an arbitrary ordered field $\mathbf{Q}$.
We are optimistic to find an algorithm with a complexity exponential in n and a bound for the degree of the polynomials in (4.0.2) about 2^n. A concrete algorithm with polynomial complexity would be very interesting.
In case $\sigma_i = \tilde{\sigma}_i$ for all i in (4.0.1) the certificate is related to Thom's lemma [CR].

Bibliography

[Akr] Akritas Alkiviadis G., *Reflections on a pair of theorems by Budan and Fourier.* University of Cansas **22**.

[BCR] Bochnak J., Coste M., Roy M.-F., *Real Algebraic Geometry.* Erg. Math. Grenzgeb. (3) **36**, Springer, Berlin (1998).

[Bem] Bembé D., *An algebraic certificate for Budan's theorem.* Journal of Pure and Applied Algebra **215**, 1360–1370 (2011).

[BG] Bembé D., Galligo A., *Virtual roots of real polynomials and fractional derivatives.* ISSAC'11 (2011), to appear.

[Bor] Borowczyk J., *Sur la vie et l'œuvre de François Budan (1761-1840).* Historia mathematica **18**, No. 2, 129–157 (1991).

[BP] Bini D., Pan V., *Polynomial and Matrix Computation.* Birkhäuser, Boston (1994).

[BPR] Basu S., Pollack R., Roy M.-F., *Algorithms in Real Algebraic Geometry.* Algo. Comp. in Math. (10), Springer, Berlin (2006).

[Bud] Budan de Boislaurent, *Nouvelle méthode pour la résolution des équations numériques d'un degré quelconque.* Paris (1822). Contains in the appendix a proof of Budan's theorem edited by the Académie des Sciences (1811).

[CLLR] Coste M., Lajous T., Lombardi H., Roy M.-F., *Generalized Budan-Fourier theorem and virtual roots.* Journal of Complexity **21**, 479–486 (2005).

[CLR] Coste M., Lombardi H., Roy M.-F., *Dynamical method in algebra: Effective Nullstellensätze.* Annals of Pure and Applied Logic **111**, 203–256 (2001).

[CR] Coste M., Roy M.-F., *Thom's lemma, the coding of real algebraic numbers and the topology of semi-algebraic sets.* Journal of Symbolic Computation **5**, 121–130 (1988).

[Fou1] Fourier J., *Sur lusage du théorème de Descartes dans la recherche des limites des racines.* Le Bulletin des Sciences par la Société Philomathique de Paris, Œuvres II, 291–309, Paris: Gauthier-Villars (1820, 1890) 156–165, 181–187.

[Fou2] Fourier J., *Analyse des équations déterminées.* C. L. M. H. Navier., Éd. Paris: Firmin Didot (1831).

[GG] Gathen J., Gerhard J., *Modern Computer Algebra.* Cambridge University Press, Cambridge (2003).

[GLM] González-Vega L., Lombardi H., Mahé L., *Virtual roots of real polynomials.* J. Pure Appl. Algebra **124**, 147–166 (1998).

[GLRR] González-Vega L., Lombardi H., Recio T., Roy M.-F., *Spécialisation de la suite de Sturm et sous-résultants.* I. RAIRO Informatique théoretique et Applications **24**, No 6, 561–588 (1990).

[Hol] Hollkott A., *Finite Konstruktion geordneter algebraischer Erweiterungen von geordneten Grundkörpern.* Dissertation, Hamburg (1941).

[Kri] Krivine J.-L., *Anneaux préordonnés.* Journal d'Analyse Mathématique **12**, 307–326 (1964).

[Lom1] Lombardi H., *Effective real nullstellensatz and variants.* Effective Methods in Algebraic Geometry, 263–288. Eds. Mora T., Traverso C.. Birkhäuser (1991). Progr. in Math. **94**.

[Lom2] Lombardi H., *Une borne sur les degrés pour le Théorème des zéros réel effectif.* Real Algebraic Geometry, 323–345. Proceedings, Rennes (1991). Lecture Notes in Mathematics No 1524. Eds. Coste M., Mahé L., Roy M.-F.. Springer (1992).

[LR1] Lombardi H., Roy M.-F., *Constructive elementary theorie of ordered fields.* Effective Methods in Algebraic Geometry, 249–262. Eds. Mora T., Traverso C.. Birkhäuser (1991). Progr. in Math. **94**.

[LR2] Lombardi H., Roy M.-F., *An elementary recursive bound for effective real Nullstellensatz.* In preparation.

[PD] Prestel A., Delzell C., *Positive Polynomials.* Monographs in Mathematics, Springer, Berlin (2001).

[Sche] Scheiderer C., *Positivity and sums of squares: A guide to recent results.* Preprint (2007).

[Schm] Schmüdgen K., *The K-moment problem for compact semialgebraic sets.* Mathematische Annalen **289**, 203–206 (1991).

[Schr] Schrijver A., *Theory of integer and linear programming.* John Wiley, New York (1985).

[Schw] Schweighofer M., *An algorithmic approach to Schmüdgen's Positivstellensatz.* Journal of Pure and Applied Algebra **166**, 307–319 (2002).

[Seg] Segner J. A. von, *Démonstration de la règle de Déscartes, pour connaître le nombre de racines affirmatives et négatives qui peuvent se trouver dans les équations.* Mémoires de l'Académie Royale des Sciences de Berlin **12**, 292–299 (1756).

[Ste] Stengle G., *A Nullstellensatz and a Positivstellensatz in semialgebraic geometry.* Mathematische Annalen **207**, 87–97 (1974).

[Vin1] Vincent M., *Sur la résolution des équations numériques.* Journal de mathématicques pures et appliquées **44**, 235–372 (1836).

[Vin2] Vincent M., *Addition à une précédante note relative à la résolution des équations numériques.* Journal de mathématicques pures et appliquées, 235–243.

Summary

In this work we present two algebraic certificates for Budan's theorem. Budan's theorem claims the following. Let R be an ordered field, $f \in \mathrm{R}[X]$ of degree n and $a, b \in \mathrm{R}$ with $a < b$. Then the number of sign changes in the sequence $(f(b), f'(b), \ldots, f^{(n)}(b))$ is not greater than the number of sign changes in the sequence $(f(a), f'(a), \ldots, f^{(n)}(a))$. This enables us to count real roots in a similar way to the real root counting by Sturm's theorem. (Budan's count of real roots is today known as "Budan-Fourier count" which, indeed, counts so called virtual roots which comprehend the real roots.) An algebraic certificate for Budan's theorem is a certain kind of proof which leads from the negation of the assumption to the contradictory algebraic identity $0 > 0$. The algorithm for our first certificate is based on the historical proof by Budan which uses only combinatorial arguments. It has a complexity exponential in the degree of f. The algorithm for the second certificate is based on mixed Taylor series and polynomials $\prod_{k=0}^{i-1}(X - k) \in \mathbb{R}[X]$ and shows a smaller complexity: The main calculation is solving a linear system; this is polynomial in the degree of f.

Keywords: Constructive real algebra, real closure, real root counting, Sturm's theorem, Budan-Fourier theorem, virtual roots.

MSC: 12.D.10, 12.J.15, 14.Q.20.