Evolution und Robotik

Teilnehmer:

David Schwalb	Herder-Oberschule
Elmo Feiten	Herder-Oberschule
Johannes Dyck	Heinrich-Hertz-Oberschule
Konrad Seifert	Heinrich-Hertz-Oberschule
Elisa Pipping	Herder-Oberschule
Hans Dettmar	Herder-Oberschule
Benjamin Schneider	Herder-Oberschule

Gruppenleiter:

Manfred Hild	Humboldt-Universität

Die Evolutionsmechanismen der Natur (natürliche Auslese, Kreuzung und Mutation) lassen sich auch in der Robotik verwenden und können mobile autonome Roboter mit interessanten Verhaltensweisen hervorbringen.

Die Kursteilnehmer beschäftigten sich mit der Funktionsweise von künstlicher Evolution und der Dynamik rekurrenter, neuronaler Netze – das heißt deren Attraktoren und Bifurkationen. In einer Simulationsumgebung evolvierten sie Roboter mit verschiedenen Verhaltensweisen (z.B. Hindernisvermeidung und Tropismen) und analysierten die Netze einzelner Individuen.

Zum Schluss des Kurses übertrugen sie die Netze auf einen realen Roboter und verglichen jeweils dessen Verhaltensweise mit ihren theoretischen Überlegungen.

Evolution und Robotik

1 Einführung

Ursprünglich wurden Roboter designt, um Arbeiten zu verrichten, die Menschen nicht übernehmen können oder um menschliche Arbeitskräfte zu ersetzen. Sie sind meist hoch spezialisiert und besitzen besondere Eigenschaften, wie sehr viel Kraft oder eine große Genauigkeit. Dafür sind sie jedoch sehr unflexibel und können nur zu dem ihnen zugedachten Zweck eingesetzt werden.

Im Gegensatz dazu sind die im weiteren Verlauf betrachteten Roboter autonom, d.h. sie sind selbstständig und besitzen eine eigene Energiequelle, Sensorik, Motorik und sind intelligent. Aufgrund technischer Einschränkungen verfügen solche Maschinen jedoch nur über die weniger komplexen Aspekte der Intelligenz, sind aber dennoch innerhalb gewisser Grenzen in der Lage, einzuschätzen, Probleme zu lösen, zu reagieren, zu planen und zu urteilen. Höhere Formen der Intelligenz, wie der „gesunde Menschenverstand“, Kreativität und ästhetisches Empfinden bleiben ihnen jedoch bislang verschlossen.

Es gibt zwei grundlegende Ansätze, intelligentes Verhalten zu simulieren. Im reaktiven Modell besitzt der Roboter keine Orientierung und kennt seine Umgebung nicht. Er reagiert spontan auf den aktuellen Input und betreibt keine langfristige Planung. Deliberativ aufgebaute Roboter besitzen ein Weltmodell, planen voraus, benötigen dafür aber viel Zeit und sind im Vergleich zum reaktiven Modell träge. Beide Ansätze haben ihre Vor- und Nachteile, somit ist ein Kompromiss zwischen beiden die beste Lösung. Da es sehr schwer

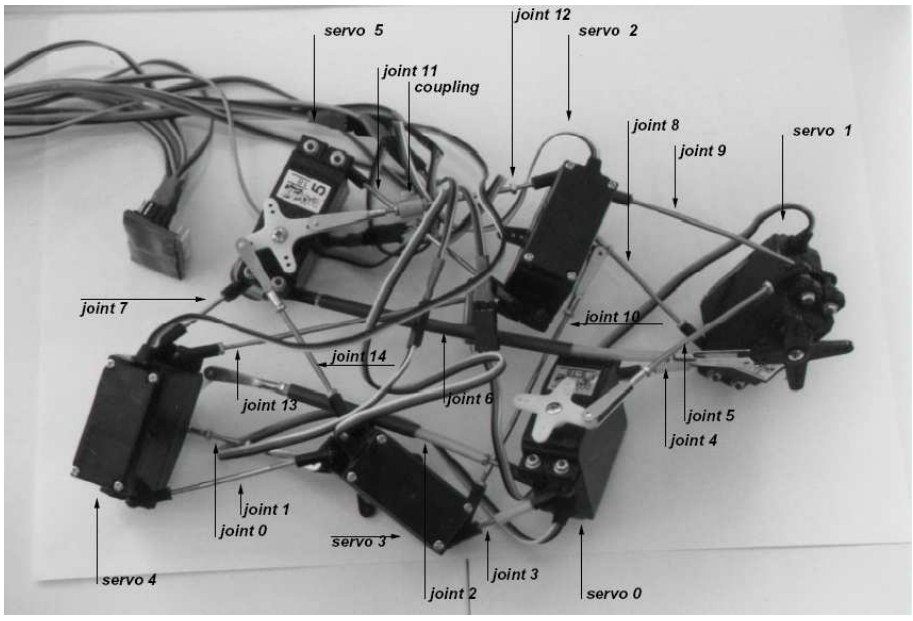


Abbildung 1.1: Hier sind sechs miteinander verbunden Servomotoren zu sehen, welche sich allein durch die Motorenbewegungen fortbewegen sollen. Durch die künstliche Evolution lernt dieses Gebilde bereits nach kurzer Zeit zu „laufen“.

ist das Verhalten solcher intelligenten, autonomen Roboter explizit zu programmieren orientiert man sich an der Natur und macht sich das Prinzip der Evolution zu Nutzen. Es gibt beispielsweise Aufgaben, die so komplex sind, dass es nahezu unmglich ist sie durch gezielte berlegungen programmiertechnisch zu lösen (siehe Abbildung 1.1).

2 Künstliche Evolution

Um solche Roboter herzustellen, bedient man sich einer Simulation des natürlichen Evolutionsprozesses, basierend auf Selektion, Mutation und Rekombination. Hierbei steht das Verhalten der Roboter für den Phänotypen in der Biologie und der Vektor, welcher das Verhalten definiert, findet seine Entsprechung in dem Genotypen. Konkret wird jedem Roboter ein sogenannter „Fitness-Wert“ zugeordnet, der seine Fähigkeit widerspiegelt, seine Aufgabe effefktiv zu erledigen. Anhand dieses Wertes wird bestimmt, welche Genotypen in die nächste Generation übernommen werden - diejenigen mit den niedrigsten Fitness-Werten fallen der Selektion zum Opfer.

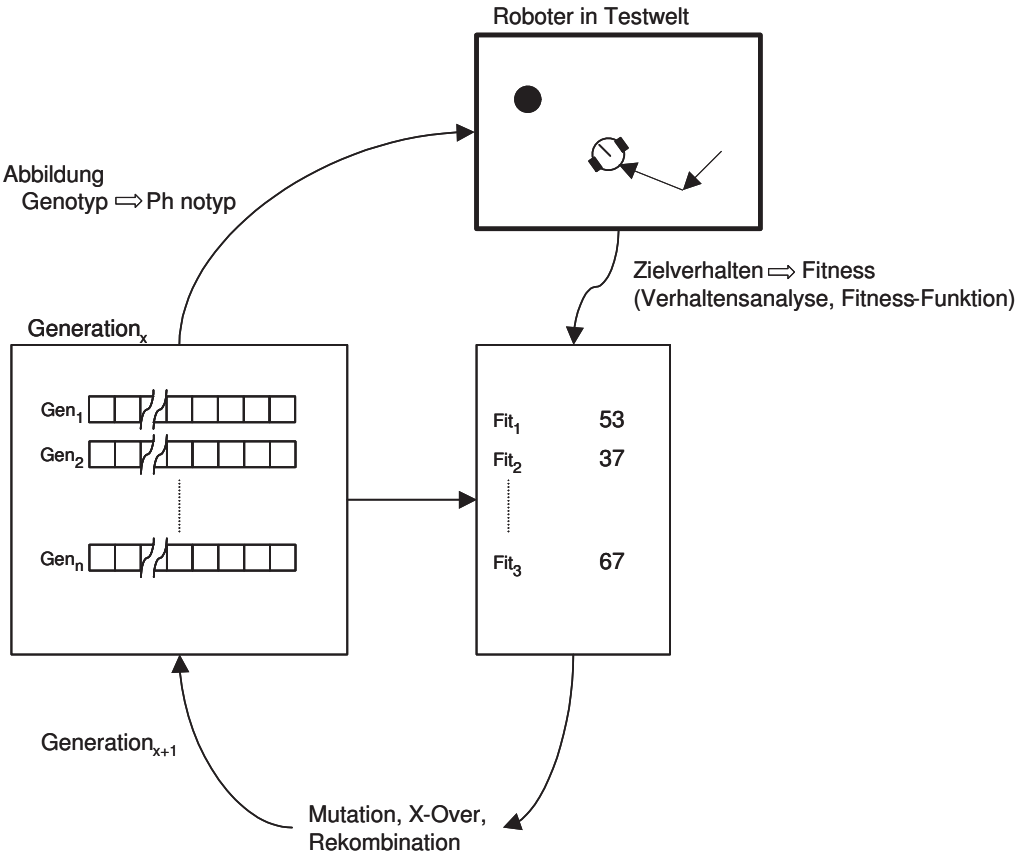


Abbildung 2.2: Künstliche Evolution

Rekombination bedeutet, dass zwei Genotypen an derselben Stelle zertrennt und die hinteren Teile vertauscht werden. Durch Mutation werden schließlich einige der Genotypen zufällig an einer Stelle verändert, bevor die neue Generation von Robotern erschaffen wird.

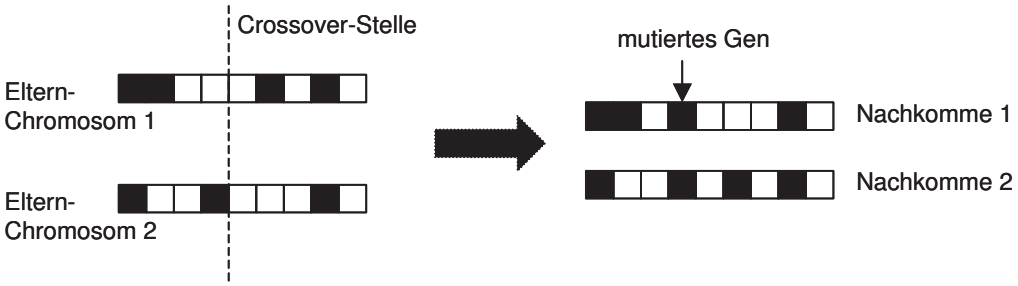


Abbildung 2.3: Rekombination und Mutation

3 Neuronale Netze

Ein neuronales Netz (wie beispielsweise auch das menschliche Gehirn) besteht aus mehreren Neuronen. Ein Neuron wiederum fungiert als signalverarbeitendes Element, das von außen oder von anderen Neuronen Signale empfängt und sie verarbeitet bzw. weiterleitet.

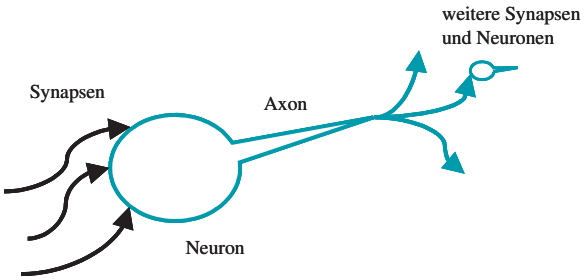


Abbildung 3.4: Schematische Darstellung eines Neurons

Die Funktionsweise ist wie folgt vorstellbar: Ein Neuron wird von einer oder mehreren Synapsen (die wiederum von anderen Neuronen kommen) umlagert. Durch diese Synapsen können elektrische Impulse an das Neuron weitergeleitet werden, die erregend oder hemmend wirken. Wird das Neuron ausreichend erregt, so „feuert“ es, das heißt, es sendet seinerseits einen elektrischen Impuls, der durch sein Axon geht und sich dann auf die Synapsen verteilt, die den Imupls wiederum an weitere Neuronen weiterleiten können.

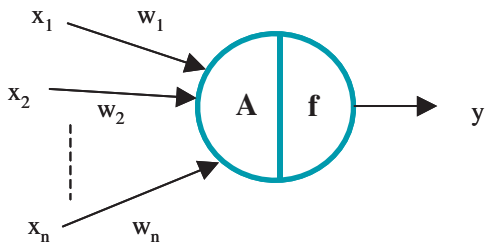


Abbildung 3.5: Mathematisches Modell eines Neurons

Um sich das Prinzip eines neuronalen Netzes für die künstliche Evolution zu Nutze zu machen, benutzt man ein sehr abstraktes Modell-Neuron. Das Bild zeigt dieses Modell-Neuron, das n x-Werte als Input bekommt $(x_1, x_2, \ldots, x_n)$, $x_n \in (-1, +1)$. Jeder dieser Werte verfügt über eine Gewichtung $(w_1, w_2, \ldots, w_n)$. Hierbei benötigt man lediglich Werte in kleineren Bereichen, so also zum Beispiel $w_i \in [-10, 10]$. Aus diesen Werten berechnet sich die sogenannte Aktivität **A** des Neurons, nämlich durch:

$$A := \sum_{i=1}^{n} x_i w_i \tag{1}$$

Die Ermittlung des Output-Wertes, der weitergegeben wird, findet durch eine geeignete Funktion f statt, also: $y := f(A)$.

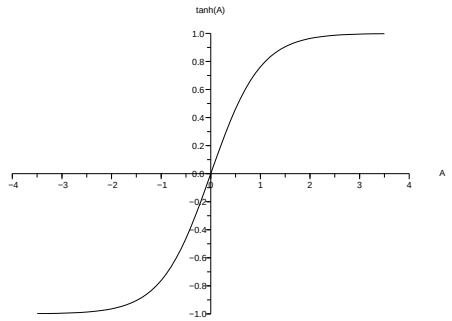


Abbildung 3.6: Graph des Tangens Hyperbolicus

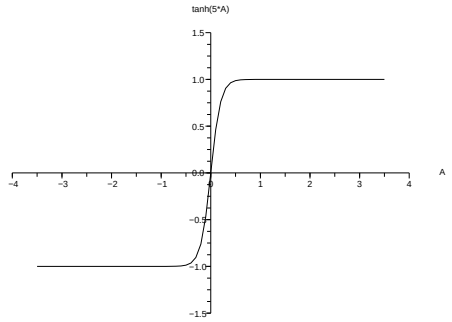


Abbildung 3.7: Graph des 5-fach in A-Achse gestauchten Tangens Hyperbolicus

Hierfür eignet sich besonders die oben abgebildete Funktion Tangens Hyperbolicus (tanh), da sie sehr vielseitig einsetzbar ist. Durch Änderung der Gewichtungen $w_1, w_2, \ldots, w_n$ kann die Funktion stark verändert werden. Wählt man ein hohes w, wird die Funktion soweit gestaucht, dass sie im Extremfall zu einer Sprungfunktion wird, die der Signum-Funktion ähnelt. Wählt man das w dagegen sehr klein, so entsteht im Prinzip eine Null-Funktion. Zusätzlich hat die tanh-Funktion in einer sehr kleinen x-Umgebung um 0 einen nahezu linearen Verlauf.

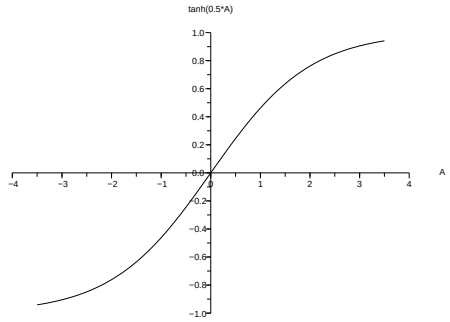


Abbildung 3.8: Graph des 5-fach in A-Achse gestreckten Tangens Hyperbolicus

Somit erhält man ein Neuronenmodell, das wie folgt aussieht:

$$t \in \mathbb{N} \qquad \text{die Zeit} \tag{2}$$

$$x_i(t) \in (-1;1) \qquad \text{Signal des Neurons } x_i \text{ zum Zeitpunkt } t \tag{3}$$

$$w_{ij} \in \mathbb{R} \qquad \text{Verbindungsgewicht von Neuron j zu Neuron i} \tag{4}$$

Es gibt Input-Neuronen (Sensorsignale), darunter optional ein Bias-Neuron (mit konstantem Wert 1) und Output-Neuronen (Motorsignale), sowie Hidden-Neuronen (alle anderen).

$$n \qquad \text{Anzahl aller Neuronen} \tag{5}$$

$$m \qquad m < n \text{ Anzahl der Input-Neuronen (inkl. Bias)} \tag{6}$$

$$\tag{7}$$

$$x_i(t+1) := \tanh\left(\sum_{j=1}^{n} w_{ij} x_j(t)\right), \qquad i = (m+1)\ldots n. \tag{8}$$

Zwar besteht das menschliche Gehirn, wie bereits erwähnt, aus mehreren hundert Millionen Neuronen, jedoch kann man schon mit nur zwei Neuronen viel anfangen.

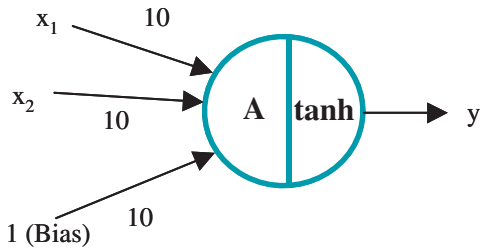


Abbildung 3.9: Neuronale Realisierung der boolschen „Oder“-Funktion

Wählt man (wie im obigen Bild) beispielsweise die Gewichtungen der zwei Input-Werte sehr hoch, so wird die tanh-Funktion derart gestaucht, dass sie im Prinzip zu einer Sprungfunktion wird. Zusätzlich existiert in diesem speziellen Anwendungsfall noch ein sogenanntes Bias-Neuron, das unabhängig von äußeren Werten stets den Wert 1 an das Neuron weiterliefert und über eine Gewichtung von 10 verfügt. Es ergibt sich:

$$A = 10 * x_1 + 10 * x_2 + 10 \quad (9)$$

$$y = f(A) = \tanh(10 * x_1 + 10 * x_2 + 10) \quad (10)$$

Stellt man sich nun vor, dass die Input-Werte nur -1 oder +1 annehmen können, ergeben sich vier Möglichkeiten:

x_1	x_2	y
-1	-1	$\tanh(-10 - 10 + 10) = \tanh(-10) = -1$
-1	+1	$\tanh(-10 + 10 + 10) = \tanh(10) = +1$
+1	-1	$\tanh(10 - 10 + 10) = \tanh(10) = +1$
+1	+1	$\tanh(10 + 10 + 10) = \tanh(30) = +1$

Tabelle 1: Wertetabelle: logisches „Oder“

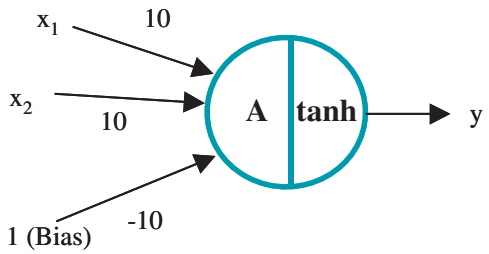


Abbildung 3.10: Neuronale Realisierung der boolschen „Und“-Funktion

Ebenso ist selbstverständlich auch eine Schaltung für das logische „Und“ möglich. Dafür wird lediglich die Gewichtung des Bias verändert und zwar zu -10. Es ergibt sich folgendes:

$$A = 10 * x1 + 10 * x2 - 10 \quad (11)$$

$$y = f(A) = \tanh(10 * x1 + 10 * x2 - 10) \quad (12)$$

Um diese einfachen logischen Schaltungen darzustellen, benötigt man also tatsächlich nur ein Neuron. Die Darstellung des Xor (exklusives Oder) ist zwar ein wenig aufwändiger, ist aber trotzdem möglich. Außerdem ist durch die Kombination zweier Neuronen noch der Aufbau von Hochpassfiltern, Tiefpassfiltern, Verstärkern, An-/Ausschaltern und weiteren Funktionen möglich.

Trotzdem stellt sich natürlich die Frage, welchen konkreten Vorteil neuronale Netze bieten. Ebenso wie die künstliche Evolution ist auch das Prinzip eines neuronalen Netzes

x_1	x_2	y
-1	-1	$\tanh(-10-10-10) = \tanh(-30) = -1$
-1	+1	$\tanh(-10+10-10) = \tanh(-10) = -1$
+1	-1	$\tanh(10-10-10) = \tanh(-10) = -1$
+1	+1	$\tanh(10+10-10) = \tanh(10) = +1$

Tabelle 2: Wertetabelle: logisches „Und“

der Natur entnommen, da es offensichtlich funktioniert. Außerdem ist ein neuronales Netz sehr einfach zu evolvieren, da nur wenige Parameter existieren, nämlich die Gewichtungen $w_1, w_2, \ldots, w_n$. Tatsächlich enthalten die Genotypen als Informationen lediglich die Gewichtungen der einzelnen Verbindungen.

4 Die Fitnessfunktion

Wenn man eine bestimmte Anzahl von Individuen künstlich evolviert, so bedient man sich einer sogenannten Fitnessfunktion, welche die Leistung der einzelnen Individuen bestimmt und somit hilft, die besten bzw. leistungsfähigsten einer Generation zu selektieren und in die nächste zu übertragen. Das Resultat einer künstlichen Evolution hängt sehr stark von der Form dieser Fitnessfunktion ab, welche sich mit Hilfe der Abbildung 4.11 beschreiben lässt. Die Kategorien lassen sich hierbei wie folgt gegeneinander abgrenzen:

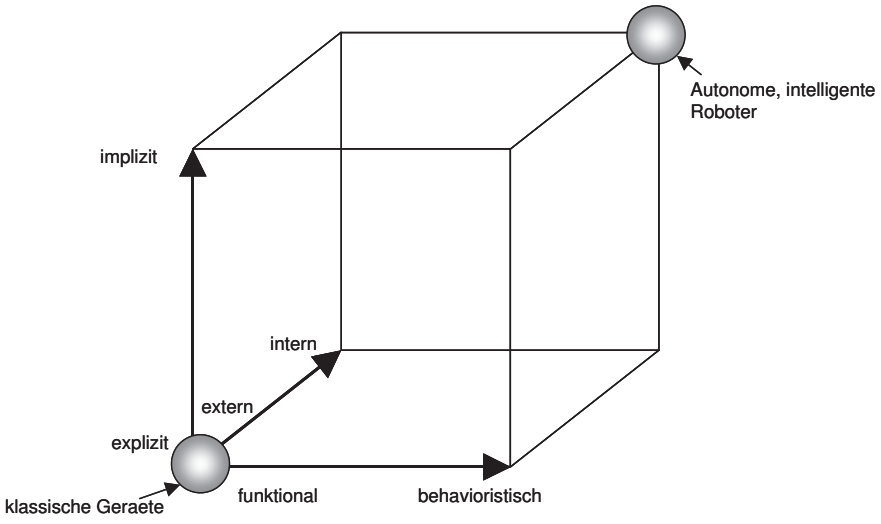


Abbildung 4.11: Implementationsmöglichkeiten der Fitnessfunktion

- Intern: die Fitnesswerte werden direkt aus den neuronalen Signalen berechnet.
- Extern: die Fitnesswerte werden mit Hilfe der Beurteilung eines externen Beobachters erstellt.

- Behavioristisch: allein das Resultat ist für den Fitnesswert ausschlaggebend (wenn die Aufgabe also lautet, von Punkt A nach Punkt B zu gelangen, so ist die Art und Weise der Fortbewegung nicht von Interesse; ausschlaggebend ist die benötigte Zeit).
- Funktional: die konkrete Funktionsweise wird bewertet; z.B. wird vorgeschrieben, mit welchen Bewegungen der Roboter von A nach B gelangen soll.
- Implizit: Anzahl von Variablen, Konstanten und Randbedingungen ist gering (in unserem Beispiel wäre nur die Tatsache, dass der Roboter geradeaus fährt von Bedeutung, der Abstand eines Sensors von der Wand z.B. nicht).
- Explizit: Anzahl von Variablen, Konstanten und Randbedingungen ist hoch.

Beispiel einer Fitnessfunktion

Eine Fitnessfunktion könnte man wie folgt realisieren: Die Länge des Zeitraums, in dem der Roboter geradeaus fährt, minus die Länge des Zeitraums, in dem er Kurven fährt (wobei die Länge der Kurve ebenfalls von Bedeutung ist), plus die Entfernung des Roboters von der vor ihm liegenden Wand. Je länger der Roboter also geradeaus fährt, desto höher ist der Fitnesswert des entsprechenden Individuums. Wenn wir den Abstand der Sensoren von der Wand mit Hilfe von neuen Variablen, die wir I_1 und I_2 nennen, angeben, so könnte die Fitnessfunktion auch folgendermaßen lauten:

$$f(x) = \frac{O_1 + O_2}{2} - \frac{|O_1 - O_2|}{2} - \frac{I_1 + I_2 + 2}{4}, \tag{13}$$

wobei

$$O_1 := \text{Output, der den linken Motor ansteuert,} \tag{14}$$
$$O_2 := \text{Output, der den rechten Motor ansteuert.} \tag{15}$$

Hierbei sind $I_1, I_2 \in [-1, +1]$, wobei -1 bedeutet, dass der Roboter einen großen Abstand von der vor ihm liegenden Wand hat und $+1$, dass er die vor ihm liegende Wand berührt. Diese Fitnessfunktion ist ein gutes Beispiel für eine explizite, interne und funktionale Funktion.

5 Praktisches Beispiel

Als praktische Aufgabe wollten wir einen Roboter evolvieren, welcher sich mit drei IR-Abstands-Sensoren in einer realen Welt zurecht findet. Mittels zweier Motoren, die von zwei Output-Neuronen angesteuert werden, kann er sich frei in seiner Welt bewegen und Hindernissen ausweichen. Das Experiment umfasste die folgenden Teilschritte:

1. In einem Seminarraum wurde die Welt aus drei Tischen und einem Teil der Wand aufgebaut.
2. Die virtuelle Welt wurde als Weltpolygonzug per x-y-Koordinaten in den PC eingegeben. Hierzu wurde der World-Simulator benutzt.

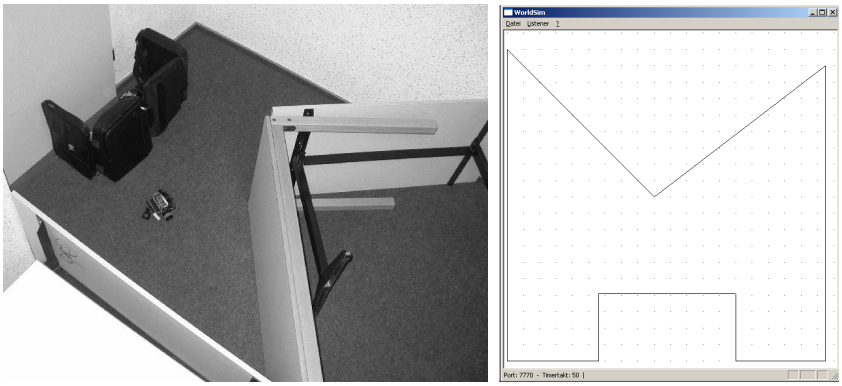


Abbildung 5.12: Aufgebaute (links) und simulierte Welt (rechts).

3. Im Robot-Simulator wurden die Maße des Roboters als x-y-Koordinaten eingegeben, wobei der Ursprung dem Drehpunkt entsprach.

4. Durch die Verknüpfung der Programme: World-Simulator, Robot-Simulator und Populations-Manager konnte die künstliche Evolution mit dem virtuellen Roboter in der virtuellen Welt durchgeführt werden.

5. Das evolvierte neuronale Netz wurde auf den realen Roboter übertragen.

6. Der Roboter wurde nun in die Testumgebung gesetzt und gestartet.

Das Resultat des Experiments: Der Roboter hat in der realen Welt exploratives Verhalten gezeigt, das heißt er fuhr in der Welt einen großen Bereich ab. Dabei stieß er nicht an Hindernisse und fand erfolgreich den Weg aus spitzen Ecken und Sackgassen.

Wie sich zeigte, konnte der Roboter Tischbeinen nicht ausweichen, da derart dünne Gegenstände nicht in der Testwelt vorhanden waren, in welcher er evolviert wurde.

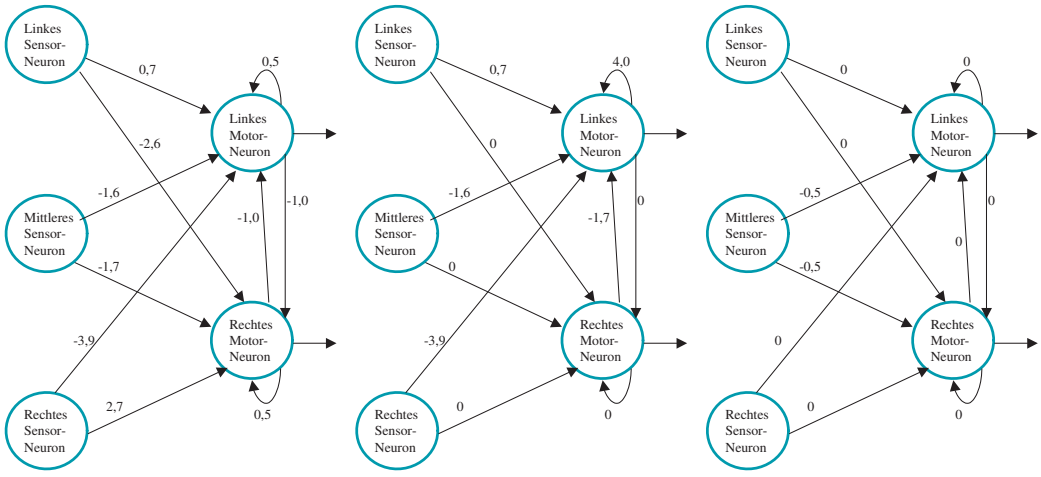


Abbildung 5.13: Modifizierte Kopien des aus der Evolution entnommenen Netzes

Um die Auswirkungen eines veränderten Genotyps auf das Verhalten des Roboters zu untersuchen, werden zum Schluss verschiedene neuronale Netze auf dem Roboter ausprobiert, die durch gezielte Modifikation aus dem ursprünglichen Netz abgeleitet wurden. In Abbildung 5.13 sind drei dieser Netze zu sehen.

Verhalten des Roboters mit dem ganz links abgebildeten Netz: Das Verhalten des Roboters veräanderte sich dahingehend, dass er bei frontalem Wandkontakt nach links anstatt nach rechts auswich. Er konnte sich aus einer spitzwinkligen Ecke nicht mehr befreien.

Verhalten des Roboters mit dem in der Mitte abgebildeten Netz: Der Roboter rotierte um sein rechtes Rad gegen den Uhrzeigersinn. Sobald der rechte Sensor dabei ein Hinderniss erfasste, wechselte der Roboter die Drehrichtung.

Verhalten des Roboters mit dem ganz rechts abgebildeten Netz: War kein Hindernis in Sicht, fuhr der Roboter geradeaus. War ein Gegenstand weniger als 10cm von ihm entfernt, dann blieb er stehen. Bei weiterer Annäherung wich er entsprechend zurück - entfernte sich das Objekt, so verfolgte er es.

Insgesamt konnte gezeigt werden, dass die künstliche Evolution neuronaler Netze geeignet ist, einen mobilen Roboter mit einer gewünschten Verhaltensweise auszustatten – sowohl in der Simulation, wie auch in der realen Welt.

Abbildung 5.14: Die Kursteilnehmer mit ihrem Roboter

Potenzsummen von ganzen Zahlen und Polynomen

Teilnehmer:

André Stenzel	Heinrich-Hertz-Oberschule
Christian Rekittke	Andreas-Oberschule
Jana Schulz	Andreas-Oberschule
Jannis Hessel	Herder-Oberschule
Konrad Steiner	Heinrich-Hertz-Oberschule
Pascal Gussmann	Heinrich-Hertz-Oberschule
Robert Altmann	Heinrich-Hertz-Oberschule

Gruppenleiter:

Olaf Teschke	Humboldt-Universität zu Berlin

Die Gruppe beschäftigte sich mit klassischen zahlentheoretischen Problemen über den ganzen Zahlen und ihrer Variante über Polynomringen. Als Motivation wurde zunächst die scheinbar einfache Gleichung $a+b=c$ betrachtet. Es stellte sich heraus, dass sich im Falle von komplexen Polynomen eine starke Aussage über die Anzahl der verschiedenen vorhandenen Nullstellen machen lässt (Satz von Mason), die weitgehende Folgerungen impliziert. So kann zum Beispiel elementar und elegant der „Satz von Fermat für Polynome“ bewiesen werden.
Auf der Suche nach vergleichbaren Resultaten in $\mathbb{Z}$ stößt man auf die *abc* - Vermutung, aus der man ebenfalls die Unlösbarkeit einer Reihe von bekannten Gleichungen folgern könnte.
Danach wurde das Problem der Darstellbarkeit von Zahlen und Polynomen als Summen von Potenzen untersucht. Mit klassischen Methoden wurden der zwei-Quadrate-Satz und der vier-Quadrate-Satz von Lagrange bewiesen und das analoge Problem für höhere Potenzen (Waring-Problem) diskutiert. Für Polynome stellt sich heraus, dass dieses Problem eine anschauliche geometrische Interpretation besitzt und auf eine Frage über Sekantenvarietäten zurückgeführt werden kann. Eine Dimensionsbestimmung liefert dann ein umfassendes Resultat.

1 $a + b = c$

Im ersten Teil beschäftigen wir uns mit Gleichungen der Form $a+b=c$ und stellen fest, dass damit starke Einschränkungen für die Zahl der verschiedenen Primfaktoren von $a \cdot b \cdot c$ verbunden sind. Für Polynome werden wir eine erstaunliche Abschätzung überraschend elementar beweisen und daraus z.B. den Satz von Fermat für Polynome folgern.

1.1 Der Satz von Mason und Satz von Fermat für Polynome

1.1.1 Grundlegende Begriffe

Es sei $f \in \mathbb{C}[x]$, d.h. f ist ein Polynom mit komplexen Koeffizienten. Nach dem Fundamentalsatz der Algebra zerfällt es in ein Produkt von Linearfaktoren

$$f(x) = c \cdot \prod_{i=1}^{m} (x - \alpha_i).$$

Dies lässt sich auch schreiben als

$$f(x) = c \cdot \prod_{i=1}^{r} (x - \alpha_i)^{m_i},$$

dabei ist m_i die Vielfachheit der Nullstelle β_i. Der Grad des Polynoms $\deg(f)$ ergibt sich dann als

$$\deg(f) = \sum_{i=1}^{r} m_i.$$

Wir schreiben weiterhin $n_0(f)$ für die Anzahl der verschiedenen Nullstellen

$$n_0(f) := r.$$

Damit gilt natürlich $n_0(f) \leq \deg(f)$. Andererseits können beide Zahlen natürlich erheblich differieren, z.B. hat $(x-1)^{1001}$ einen hohen Grad, aber $n_0 = 1$. Ziel ist es, unter bestimmten Voraussetzungen auch eine Abschätzung in die andere Richtung zu erhalten.

1.1.2 Hilfsätze

Der ggt (größter gemeinsamer Teiler) von Polynomen ist das Produkt aller gemeinsamen Primfaktoren, in unserem Fall (da ja komplexe Polynome immer eine Nullstelle haben) der gemeinsamen Linearfaktoren.

Lemma 1.1. *Sei f aus $\mathbb{C}[x]$, dann gilt* $\deg[\mathrm{ggt}(f, f')] = \deg(f) - n_0(f)$.

Beweis: Sei α_i eine Nullstelle von f mit Vielfachheit m_i. Dann ist

$$f = (x - \alpha_i)^{m_i} \cdot g, \quad x - \alpha_i \nmid g$$

$$f' = m_i \cdot (x - \alpha_i)^{m_i - 1} \cdot g + (x - \alpha_i) \cdot g'$$

$$f' = (x - \alpha_i)^{m_i - 1} \cdot [m_i \cdot g + (x - \alpha_i) \cdot g']$$

Offensichtlich gilt also:

$$(x - \alpha_i)^{m_i - 1} \nmid m_i \cdot g + (x - \alpha_i) \cdot g'$$

Jeder Linearfaktor $(x - \alpha_i)$ kommt also in der Linearfaktorzerlegung von f' genau $m_i - 1$ mal vor.

$$\Rightarrow \deg[\mathrm{ggt}(f, f')] = (m_1 - 1) + (m_2 - 1) + ... + (m_r - 1)$$

$$\Rightarrow \deg[\mathrm{ggt}(f, f')] = \sum_{i=1}^{r} (m_i - 1) = -r + \sum_{i=1}^{r} m_i = \deg(f) - n_0(f)$$

Lemma 1.2. *Seien f, g aus $\mathbb{C}[x]$, dann gilt:*

$$n_0(f \cdot g) \leq n_0(f) + n_0(g)$$

Die Gleichheit gilt genau dann, wenn der ggt *von f und g gleich* 1 *ist.*

1.1.3 Satz von Mason - Beweis nach der Variante von Noah Snyder(2000)

Satz 1.3 (Mason 1983)**.**

$$\max\{\deg(f), \deg(g), \deg(h)\} \leq n_o(f \cdot g \cdot h) - 1$$

Beweis (nach der Variante von Noah Snyder, 2000): Wir benutzen die einfache Identität

$$\begin{aligned} f \cdot g' - f' \cdot g &= f \cdot g' + f \cdot f' - f \cdot f' - f' \cdot g \\ &= f(g' + f') - f'(f + g) = f \cdot h' - f' \cdot h. \end{aligned}$$

Es sind $f \cdot g' \neq 0$ und $f' \cdot g \neq 0$, da f und g nicht konstant sind. Außerdem gilt $f \cdot g' - f' \cdot g \neq 0$, da sonst f das Produkt $f' \cdot g$ teilen würde. Da f und g teilerfremd sind, müßte $f \mid f'$ gelten, was schon wegen des kleineren Grades unmöglich ist.

Nun gilt:

$$\begin{aligned} \operatorname{ggt}(f, f') &\mid f \cdot g' - f' \cdot g \\ \operatorname{ggt}(g, g') &\mid f \cdot g' - f' \cdot g \\ \operatorname{ggt}(h, h') &\mid f \cdot g' - f' \cdot g \end{aligned}$$

Da f, g, h paarweise teilerfremd sind, gilt:

$$\begin{aligned} &\operatorname{ggt}(f, f') \cdot \operatorname{ggt}(g, g') \cdot \operatorname{ggt}(h, h') \mid f \cdot g' - f' \cdot g \\ \Rightarrow\ &\deg[\operatorname{ggt}(f, f') \cdot \operatorname{ggt}(g, g') \cdot \operatorname{ggt}(h, h')] \leq \deg(f \cdot g' - f' \cdot g) \end{aligned}$$

Wir erhalten die Ungleichung

$$\begin{aligned} &\Rightarrow \deg[\operatorname{ggt}(f, f'] \cdot \operatorname{ggt}(g, g') \cdot \operatorname{ggt}(h, h')] \leq \deg(f \cdot g') = \deg(f) + \deg(g) - 1 \\ &\Rightarrow \deg[\operatorname{ggt}(f, f')] + \deg[\operatorname{ggt}(g, g')] + \deg[\operatorname{ggt}(h, h')] \leq \deg(f) + \deg(g) - 1 \\ &\overset{\text{Lemma 1.1}}{\Rightarrow} \deg(f) - n_0(f) + \deg(g) - n_0(g) + \deg(h) - n_0(h) \leq \deg(f) + \deg(g) - 1 \\ &\Rightarrow \deg(h) \leq n_0(f) + n_0(g) + n_0(h) - 1 \end{aligned}$$

Da f, g und h immer noch teilerfremd sind, gilt:

$$\overset{\text{Lemma 1.2}}{\Rightarrow} \deg(h) \leq n_0(f \cdot g \cdot h) - 1$$

Da nun gilt $f + g = h \Leftrightarrow f + (-h) = -g \Leftrightarrow g + (-h) = f$, lässt sich analog zeigen:

$$\begin{aligned} \deg(f) &\leq n_0(f \cdot g \cdot h) - 1 \\ \deg(g) &\leq n_0(f \cdot g \cdot h) - 1 \end{aligned}$$

Da die Ungleichungen für $\deg(f), \deg(g)$ und $\deg(h)$ gelten, gilt schließlich auch:

$$\max\{\deg(f), \deg(g), \deg(h)\} \leq n_0(f \cdot g \cdot h) - 1$$

1.1.4 Die FERMATsche Gleichung für Polynome

Es seien f, g, h paarweise teilerfremde, nicht konstante Polynome in $\mathbb{C}[x]$.

Auf die Gleichung

$$f^n + g^n = h^n$$

ist der Satz von Mason also anwendbar mit

$$\deg(f^n) \leq n_0(f \cdot g \cdot h) - 1$$

$$\deg(g^n) \leq n_0(f \cdot g \cdot h) - 1$$

$$\deg(h^n) \leq n_0(f \cdot g \cdot h) - 1$$

da, wenn die Ungleichung für den maximalen Grad gilt, sie natürlich auch für alle Grade einzeln gilt.
Desweiteren lassen sich auch die rechten Seiten mit Lemma 2 als $n_0(f) + n_0(g) + n_0(h) - 1$ darstellen. Durch Summation erhalten wir

$$\begin{aligned} \deg(f^n) + \deg(g^n) + \deg(h^n) &\leq 3(n_0(f) + n_0(g) + n_0(h) - 1) \\ \Rightarrow n \cdot (\deg(f) + \deg(g) + \deg(h)) &\leq 3(\deg(f) + \deg(g) + \deg(h) - 1) \end{aligned}$$

Somit kann die FERMATsche Gleichung für Polynome für $n \geq 3$ keine Lösungen besitzen.

Für $n = 2$ gibt es übrigens die klassische Lösung $(x^2-1)^2+(2x)^2 = (x^2+1)^2$. Man erhält sie zum Beispiel, indem man die rationalen Punkte auf dem Einheitskreis durch Geraden mit rationalem Anstieg durch $(0, 1)$ parametrisiert (dies liefert zudem auch alle Pythagoräischen Tripel, d.h. alle ganzen Zahlen a, b, c mit $a^2 + b^2 = c^2$).

1.2 Die *abc*-Vermutung

Wir versuchen, für ganze Zahlen mit $a + b = c$ eine Art „Mason"-Satz zu formulieren.
Es seien $a, b, c \in \mathbb{Z}$ paarweise teilerfremd.

Als „Grad" einer ganzen Zahl wird der Betrag der Zahl gewählt. Dies kann z.B. durch die Division mit Rest motiviert werden - bei Polynomen hat der Rest einen kleinerer Grad als der Divisor, bei ganze Zahlen einen kleineren

Betrag.

Wodurch könnte n_0 ersetzt werden? Bei Polynomen wird die Anzahl der verschiedenen Nullstellen, also der Primfaktoren gezählt. Wir sind dort in der einfachen Situation, dass alle Primfaktoren denselben Grad haben. Dagegen treten im allgemeinen in der Primzerlegung

$$a = \prod_{i=1}^{n} p_i^{m_i}$$

Faktoren mit unterschiedlichem Betrag auf. Wir definieren daher

$$N_0 := \prod_{i=1}^{n} p_i,$$

also das Produkt der verschiedenen Primfaktoren.

1. Versuch:
Gilt $\max\{|a|, |b|, |c|\} \leq N_0(a \cdot b \cdot c)$ für ganze Zahlen?
Gegenbeispiel:

$$2^{10} + 1 = 1025 = 5^2 \cdot 41$$

$$N_0(a \cdot b \cdot c) = 2 \cdot 1 \cdot 5 \cdot 41 = 410 < 2^{10} + 1$$

2.Versuch:
Ausgleich durch eine Konstante $K \in \mathbb{R}$

$$\max\{|a|, |b|, |c|\} \leq K \cdot N_0(a \cdot b \cdot c)$$

Gegenbeispiel:

$$a = 3^{2^n}, b = -1$$

Es gilt ferner für alle n, dass $2^n | (3^{2^n} - 1)$ (induktiv leicht zu zeigen).

$$\begin{aligned} \Rightarrow N_0(3^{2^n} - 1) = N_0(2^n \cdot \frac{3^{2^n} - 1}{2^n}) &\leq N_0(2^n) \cdot N_0(\frac{3^{1^n} - 1}{2^n}) \leq 2 \cdot \frac{3^{2^n} - 1}{2^n} \\ \Rightarrow N_0(a \cdot b \cdot c) &\leq 1 \cdot 3 \cdot 2 \cdot \frac{3^{2^n}}{2^n} \\ \Rightarrow \max\{|a|, |b|, |c|\} = 3^{1^n} &\leq k \cdot 3 \cdot 2 \cdot \frac{3^{2^n} - 1}{2^n} \end{aligned}$$

$\Rightarrow$ Es existiert kein $k \in \mathbb{R}$, sodass dies für alle n gilt!

Die Ungleichung würde aber stimmen, wenn die rechte Seite in eine minimal höhere Potenz als 1 erhoben würde. Dies führt zu folgender

Vermutung (*abc*-Vermutung). *Seien* $a, b, c \in \mathbb{Z}$, $a + b = c$.

$$\forall\ \epsilon > 0\ \exists\ K(\varepsilon): \quad \max\{|a|, |b|, |c|\} \leq K(\varepsilon) N_0(a \cdot b \cdot c)^{1+\epsilon}$$

Bisher sind keine Gegenbeispiele bekannt, und es gibt eine Reihe von Ergebnissen, die auf die Richtigkeit der Vermutung hindeuten.

1.3 Die Anwendungen der *abc*-Vermutung

1.3.1 Die FERMATsche Gleichung

Seien $x, y, z \in \mathbb{Z}$ mit $x^n + y^n = z^n, n \in \mathbb{N}$.

Aus der *abc*-Vermutung folgt:

$$\begin{array}{rrcl}
 & \max\{|x^n|, |y^n|, |z^n|\} & \leq & K(\varepsilon) \cdot N_0(x^n \cdot y^n \cdot z^n)^{1+\varepsilon} \\
\Rightarrow & |x^n| \cdot |y^n| \cdot |z^n| & \leq & (K(\varepsilon) \cdot N_0(x \cdot y \cdot z)^{1+\varepsilon})^3 \\
\Rightarrow & (|x \cdot y \cdot z|)^n & \leq & K(\varepsilon)^3 \cdot N_0(x \cdot y \cdot z)^{3+3\varepsilon} \\
\Rightarrow & (|x \cdot y \cdot z|)^n & \leq & K(\varepsilon)^3 \cdot (|x \cdot y \cdot z|)^{3+3\varepsilon} \\
\Rightarrow & (|x \cdot y \cdot z|)^{n-3\varepsilon-3} & \leq & K(\varepsilon)^3
\end{array}$$

(1) Wenn $x \cdot y \cdot z \geq 2$, so muss offensichtlich ein n_0 existieren, sodass gilt:

$$\forall n \geq n_0 : (|x \cdot y \cdot z|)^{n-3\varepsilon-3} > K(\varepsilon)^3.$$

Für genügend große n gibt es also nur die trivialen Lösungen mit $x \cdot y \cdot z = 0$.

(2) Fixiert man $n \geq 4$, kann die Gleichung nur endlich viele Lösungen haben, da ab einer gewissen Größe von $|x \cdot y \cdot z|$ die Ungleichung nicht mehr erfüllt wird.

1.3.2 Die CATALANsche Gleichung

$\forall x, y, n \in \mathbb{N}^* : x^n - y^m = 1$ hat für $m, n \geq 2$ und $y \neq 0$ keine Lösung außer $3^2 - 2^3 = 1$.

Annahme: Die CATALANsche Gleichung ist erfüllt.

(1) $m = n = 2$ ist nicht möglich, da die Differenz zweier Quadrate $\neq 0$ immer größer als 1 ist.

(2) Sei $m > 2 \vee n > 2$
Mit der *abc*-Vermutung folgt dann:

$$\max\{x^n, y^m\} \leq k(\epsilon) \cdot N_0(x^n \cdot y^m)^{1+\varepsilon}$$

$$\begin{aligned}
\Rightarrow m \cdot \ln(y) < n \cdot \ln(x) &\leq (1+\varepsilon) \cdot \ln(N_0(x^n \cdot y^m)) + \ln(K(\varepsilon)) \\
\Rightarrow \qquad n \cdot \ln(x) &\leq (1+\varepsilon) \cdot \ln(N_0(x \cdot y)) + \ln(K(\varepsilon)) \\
&\leq (1+\varepsilon) \cdot \ln(x \cdot y) + \ln(K(\varepsilon)) \\
m \cdot \ln(y) &\leq (1+\varepsilon) \cdot \ln(N_0(x \cdot y)) + \ln(K(\varepsilon)) \\
&\leq (1+\varepsilon) \cdot \ln(x \cdot y) + \ln(K(\varepsilon))
\end{aligned}$$

$\Rightarrow m \cdot n \cdot (\ln(x) + \ln(y)) \leq (1+\varepsilon) \cdot (m+n) \cdot (\ln(x) + \ln(y)) + (m+n) \cdot \ln(K(\varepsilon))$
$m \cdot n > m + n$ gilt nach Voraussetzung (da $m = n = 2$ nicht möglich ist). Demnach ist diese Ungleichung nur für endlich viele x, y erfüllt.
Also kann auch die CATALANsche Gleichung höchstens von endlich vielen x, y erfüllt werden.

1.3.3 Die Gleichung $x^l + y^m = z^n$

Seien $x, y, z \in \mathbb{Z}$ mit $x^l + y^m = z^n, l, m, n \in \mathbb{N} \setminus \{0, 1, 2, 3\}$.

Aus der *abc*-Vermutung folgt:

$$\begin{aligned}
\max\{|x^l|, |y^m|, |z^n|\} &\leq K(\varepsilon) \cdot N_0(x^l \cdot y^m \cdot z^n)^{1+\varepsilon} \\
|x^l| \cdot |y^m| \cdot |z^n| &\leq (K(\varepsilon) \cdot N_0(x \cdot y \cdot z)^{1+\varepsilon})^3 \\
|x|^l \cdot |y|^m \cdot |z|^n &\leq K(\varepsilon)^3 \cdot N_0(x \cdot y \cdot z)^{3+3\varepsilon}
\end{aligned}$$

Fixiert man $l, m, n \geq 4$, kann die Gleichung nur endlich viele Lösungen haben, da ab einer gewissen Größe von $|x \cdot y \cdot z|$ die Ungleichung sonst nicht mehr erfüllt wird. Ebenso sieht man, dass für hinreichend große l, m, n keine Lösung mehr existieren kann.

2 Das Waring-Problem für ganze Zahlen

In diesem Abschnitt beschäftigen wir uns mit der Frage, ob man eine ganze Zahl n als Summen von Potenzen

$$n = a_1^k + \ldots a_g^k$$

darstellen kann und wieviele Summanden man dazu benötigt. Wir beginnen mit dem Fall $k = 2$, also Summen von Quadraten.

2.1 Der Zwei-Quadrate-Satz

Welche Zahlen lassen sich als Summe zweier Quadrate schreiben? Wir probieren z.B.

$$2 = 1^2 + 1^2, \quad 20 = 4^2 + 2^2, \quad 65 = 7^2 + 4^2 = 8^2 + 1^2.$$

Es fällt auf, dass die Zahlen, die sich als Summe zweier Quadrate schreiben lassen, nur bestimmte Primfaktorzerlegungen haben. Der Grund dafür ist das folgende Lemma:

Lemma 2.1. $(A^2 + B^2) \cdot (U^2 + V^2) = (AU + BV)^2 + (AV - BU)^2$

Diese Identität kann leicht durch Ausmultiplizieren berechnet werden. Man kann sie aber auch interpretieren als die Betragsgleichung $|z_1||z_2| = |z_1 z_2|$ für komplexe Zahlen $z_1 = A + Bi$, $z_2 = V + Ui$.

Korollar 2.2. *Das Problem der Darstellung als zwei Quadrate kann auf das Problem für Primfaktoren reduziert werden.*

Welche Primzahlen sind als Summe zweier Quadrate darstellbar? Z.B. sind

$$2 = 1^2 + 1^2 \ ; \ 53 = 2^2 + 7^2 \ ; \quad \text{aber } 19 = 3^2 + 3^2 + 1^2$$

Wir beobachten, dass alle Rest 1 bei der Division durch 4 lassen. Tatsächlich gilt:

Satz 2.3. *Sei p eine ungerade Primzahl. p ist die Summe von zwei Quadraten genau dann, wenn gilt:*

$$p \equiv 1(\mod 4)$$

Beweis: Wir beweisen zunächst die einfache Richtung. Wenn p die Summe von 2 Quadraten ist, dann muss eines der beiden Quadrate gerade und eines ungerade sein. Selbiges gilt somit auch für die Zahlen selber. Da gilt:

$$a = 2n+1 \Leftrightarrow a \equiv 1(\mod 4) \vee a \equiv 3(\mod 4) \Leftrightarrow p^2 \equiv 1(\mod 4)$$

$$b = 2k \Leftrightarrow bp \equiv 0(\mod 4) \vee b \equiv 2(\mod 4) \Leftrightarrow p^2 \equiv 0(\mod 4)$$

Somit gilt:

$$p \equiv a^2 + b^2(\mod 4) \Leftrightarrow p \equiv 0 + 1(\mod 4) \Leftrightarrow p \equiv 1(\mod 4)$$

Wenn p die Summe zweier Quadrate ist, ist es auch kongruent zu $1 \mod 4$. Für den Beweis der Umkehrung benötigen wir folgendes Lemma:

Lemma 2.4. *(i) (Satz von Wilson) p ist eine Primzahl genau dann, wenn*

$$(p-1)! \equiv -1(\mod p)$$

(ii) Sei p eine Primzahl, dann ist $((\frac{p-1}{2})!)^2 \equiv (-1)^{\frac{p+1}{2}}(\mod p)$

Beweis: (i) Jede Restklasse $\neq 0$ ist modulo einer Primzahl p invertierbar. Im Produkt $(p-1)!$ erscheint damit also zu jeder Restklasse ihre Inverse. Diese kürzen sich zu 1 mit Ausnahme von ± 1, die ihre eigenen Inversen sind. Daher ist das Produkt für Primzahlen -1.
Umgekehrt ist der $\text{ggt}((p-1)!, p) \geq 2$, wenn p keine Primzahl ist.

$$(ii)\ (p-1)! \equiv (-1)\cdot 1\cdot \ldots \cdot(-\frac{p-1}{2})(\frac{p-1}{2}) \equiv ((\frac{p-1}{2})!)^2 \cdot (-1)^{\frac{p-1}{2}} \overset{(i)}{\equiv} -1 (\mathrm{mod}\, p).$$

Insbesondere gibt es für Primzahlen $p \equiv 1(\mod 4)$ eine Zahl x mit $x^2 + 1 = Mp$ (dabei ist offenbar $M \leq p-1$), also haben wir ein Vielfaches von p als Summe zweier Quadrate $A^2 + B^2$ dargestellt. Wir beginnen nun

ein Reduktionsverfahren, um M zu verkleinern. Dabei setzen wir $U \equiv A(\mod M)$ und $V \equiv B(\mod M)$. Es gilt dann $U^2+V^2 = Mr$ mit $r \leq M-1$. Wir verwenden nun erneut die Quadrateidentität 2.1 und erhalten

$$(A^2+B^2)\cdot(U^2+V^2) = (AU+BV)^2+(AV-BU)^2 = M^2rp.$$

Desweiteren ist offenbar $AU+BV \equiv A^2+B^2 \equiv 0(\mod M)$ und $AV-BU \equiv AB-BA \equiv 0(\mod M)$, daher teilt M die Zahlen $AU+BV$ und $AV-BU$, und es gilt

$(\frac{AU+BV}{M})^2+(\frac{AV-BU}{M})^2 = rp$, wir haben also ein kleineres Vielfaches

von p als Summe zweier Quadrate dargestellt. Der Schluß auf $M=1$ folgt durch Iteration.

Korollar 2.5. *Jede Primzahl $p \equiv 1(\mod 4)$ läßt sich eindeutig als Summe zweier Quadrate darstellen.*

Die Existenz einer Darstellung $p = A^2+B^2$ folgt aus dem eben bewiesenem Satz, die Eindeutigkeit aus der Zerlegung $A^2+B^2 = (A+Bi)(A-Bi)$ und der Tatsache, dass in $\mathbb{Z}+i\mathbb{Z}$ die Primzerlegung eindeutig ist (letzteres kann man z.B. aus der Existenz einer Division mit Rest bzgl. des Betrags folgern).

Als Folgerung aus der Quadratidentität erhalten wir den zwei-Quadrate-Satz für beliebige natürliche Zahlen:

Korollar 2.6. *$n \in \mathbb{N}$ mit der Primzerlegung $n = \prod_{i=1}^{r} p_i^{m_i}$ ist als Summe zweier Quadrate darstellbar genau dann, wenn jedes $P_i \equiv 3(\mod 4)$ nur mit gerader Vielfachheit vorkommt.*

2.2 Satz von Lagrange

Mit Hilfe der Ideen des vorigen Abschnitts beweisen wir nun analog den

Theorem 2.7 (Satz von Lagrange)**.** *Jede natürliche Zahl ist Summe von vier Quadraten.*

Beweis: Der Beweis verläuft analog zum zwei-Quadrate-Satz. Zunächst wird eine Identität benutzt, die das Problem auf Primfaktoren zurückführt. Dann finden wir durch Restklassenüberlegung eine Quadratzerlegung eines

Vielfachen. In einem Abstiegsverfahren, das wiederum die Identität benutzt, können wir dann dieses Vielfache auf eins reduzieren.
Es gilt

$$(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = z_1^2 + z_2^2 + z_3^2 + z_4^2$$

mit

$$\begin{aligned} z_1 &= x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 \\ z_2 &= x_1y_2 - x_2y_1 + x_3y_4 - x_4y_3 \\ z_3 &= x_1y_3 - x_3y_1 + x_4y_2 - x_2y_4 \\ z_4 &= x_1y_4 - x_4y_1 + x_2y_3 - x_3y_2. \end{aligned}$$

Diese Identität kann man übrigens analog zur zwei-Quadrate-Identität als Betragsgleichung von Produkten auffinden, in disem Fall im Schiefkörper der Quaternionen.
Ferner ist $2 = 1^2 + 1^2 + 0^2 + 0^2$. Es genügt also zu zeigen, daß jede Primzahl Summe von 4 Quadraten ist.
Es sei also p eine ungerade Primzahl. Die $\frac{p+1}{2}$ Zahlen x^2 mit $0 \leq x \leq \frac{p-1}{2}$ sind paarweise inkongruent $\mod p$, ebenso die $\frac{p+1}{2}$ Zahlen $-1 - y^2$ mit $0 \leq y \leq \frac{p-1}{2}$. Da es genau p Restklassen $\mod p$ gibt, dies aber insgesamt $p + 1$ Zahlen sind, gibt es ein x, sodass $x^2 \equiv -1 - y^2 (\mod p)$ ist. Ein Vielfaches von p lässt sich also in der Form $1 + x^2 + y^2$ darstellen:

$$m \cdot p = 0^2 + 1^2 + x^2 + y^2$$

Darin ist $0 < m < p$ (wegen $1 + x^2 + y^2 < 1 + 2\left(\frac{p}{2}\right)^2 < p^2$). Es sei m_0p das kleinste Vielfache von p, welches sich in der Form

$$m_0p = x_1^2 + x_2^2 + x_3^2 + x_4^2$$

mit $0 < m_0 < p$ darstellen lässt, worin x_1, x_2, x_3, x_4 nicht alle durch p und auch nicht durch m_0 teilbar sind. Angenommen, es sei $m_0 > 1$. Aus der Minimaleigenschaft von m_0 folgt, dass dann m_0 ungerade sein muss. Wäre m_0 gerade, so hätten wir

$$x_1 + x_2 + x_3 + x_4 \equiv x_1^2 + x_2^2 + x_3^2 + x_4^2 \equiv 0 (\mod 2)$$

d.h., die x_i sind zu je zweien kongruent $\mod 2$. Es sei o.B.d.A. $x_1 \equiv x_2 (\mod 2)$ und $x_3 \equiv x_4 (\mod 2)$. Dann haben wir die Darstellung

$$\frac{m_0}{2} \cdot p = \left(\frac{x_1 + x_2}{2}\right)^2 + \left(\frac{x_1 - x_2}{2}\right)^2 + \left(\frac{x_3 + x_4}{2}\right)^2 + \left(\frac{x_3 - x_4}{2}\right)^2$$

im Widerspruch zur Minimalität von m_0. Man kann also

$$x_i = b_i \cdot m_0 + y_i$$

mit $(i = 1, 2, 3, 4)$ setzen, wobei b_i so gewählt wurde, dass $|y_i| < \frac{1}{2}m_0$ ist. Da x_1, x_2, x_3, x_4 nicht alle durch m_0 teilbar sind, ist wenigstens ein $y_i > 0$. Somit ist

$$0 < y_1^2 + y_2^2 + y_3^2 + y_4^2 < 4 \cdot \left(\frac{1}{2}m_0\right)^2 = m_0^2.$$

Aus $x_i = b_i m_0 + y_i$ folgt andererseits

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 \equiv 0(\mod m).$$

Daher ist

$$y_1^2 + y_2^2 + y_3^2 + y_4^2 = m_1 \cdot m_0$$

mit $0 < m_1 < m_2$. Es folgt demnach die Darstellung

$$m_0^2 m_1 p = z_1^2 + z_2^2 + z_3^2 + z_4^2.$$

Jedes der z_i ist darin aber wegen $x_i \equiv y_i(m_0)$ durch m_0 teilbar, also $z_i = m_0 t_i$. Es folgt

$$m_1 p = t_1^2 + t_2^2 + t_3^2 + t_4^2$$

mit $0 < m_1 < m_0 < p$, was ein Widerspruch zur Minimaleigenschaft von m_0 ist. Es muss somit $m_0 = 1$ sein, q.e.d.

2.3 Das Waring-Problem

In seinem Buch *Meditationes Algebraicae* schrieb Waring 1770: „Jede Zahl ist Summe von neun Kubikzahlen, neunzehn Biquadraten und so weiter... “ Zuvor hatte Lagrange bereits bewiesen, dass jede Zahl als Summe von vier Quadratzahlen darstellbar ist (siehe 2.7). Daher wurde nach Waring das folgende Problem benannt:

Problem 2.8 (Waring-Problem). *Gibt es zu jedem Exponenten k eine kleinste Zahl $g(k)$, so dass jede natürliche Zahl n als Summe*

$$n = a_1^k + \cdots + a_{g(k)}^k$$

von k-ten Potenzen darstellbar ist?

Hilbert bewies 1909, dass ein solches $g(k)$ für alle $k \in \mathbb{N}$ existiert. Die Bestimmung der Zahlen $g(k)$ erwies sich als deutlich schwieriger und wurde erst in den letzten Jahrzehnten abgeschlossen. Warings Vermutung, dass $g(3) = 9$ und $g(4) = 19$ ist, konnten gezeigt werden. Allgemein gibt es drei Fälle, wobei im Hauptfall das folgende Resultat gilt:

Theorem 2.9. *Sei $k > 4$ und $2^k \cdot \{(\frac{3}{2})^j\} + [(\frac{3}{2})^j] \leq 2^k$. Dann ist*

$$\Rightarrow g(k) = 2^k + [(\frac{3}{2})^k] - 2.$$

Der Beweis ist allerdings extrem schwierig.

Bei weiteren Untersuchungen stellt man fest, dass oft nur wenige Zahlen wirklich $g(k)$ Summanden benötigen. So sind im Fall $k = 3$ die Zahlen

$$23 = 2 \cdot 2^3 + 7 \cdot 1^3 \text{ und } 239 = 5^3 + 3 \cdot 3^3 + 4 \cdot 2^3 + 1^3$$

die einzigen Zahlen, die neun Kuben benötigen. Weitere fünfzehn Zahlen benötigen acht Kuben (die größte ist 8042), alle anderen höchstens sieben. Dies führt auf das bisher ungelöste

Problem 2.10 (Großes Waring-Problem). *Finde die Zahl $G(k)$, das ist die kleinste Anzahl, so dass sich fast alle natürlichen Zahlen (d.h. bis auf endlich viele) als Summe von $G(k)$ k-ten Potenzen schreiben lassen.*

Dieses Problem ist noch deutlich schwerer und weitgehend ungelöst. So ist z.B. unklar, ob $G(3) = 7$. Bisher weiß man nur, dass $G(2) = 4$ und $G(4) = 16$.

3 Das Waring-Problem für Polynome

In diesem Abschnitt betrachten wir das analoge Problem für Polynome aus $\mathbb{C}[x]$, d.h. die Frage, wann man ein Polynom als Summe von Potenzen anderer Polynome darstellen kann. Wir werden im ersten Teil schnell sehen, dass diese Frage relativ leicht beantwortbar, aber nicht besonders interessant ist. Deshalb werden wir uns auf das Problem der Darstellung als Potenzsummen von linearen Polynomen einschränken. Es erweist sich, dass dies eine schöne geometrische Interpretation und eine anschauliche Antwort besitzt.

3.1 Potenzsummen beliebiger Polynome

Wir beginnen wieder mit dem quadratischen Fall. Dieser ist besonders einfach wegen der Identität

$$P^2 = (\frac{P+1}{2})^2 + (i\frac{P-1}{2})^2,$$

d.h. jedes komplexe Polynom P (mit beliebig vielen Variablen!) ist Summe zweier Quadrate. Stimmt das auch für Kuben? Wir untersuchen dies im einfachen Fall des Polynoms x.

Satz 3.1. *x ist nicht Summe zweier Kuben.*

Beweis: Allgemein lässt sich jede Summe zweier Kuben darstellen als:

$$A^3 + B^3 = (A+B) \cdot (A+\xi B) \cdot (A+\xi^2 B),$$

wobei ξ eine dritte Einheitswurzel $\neq 1$ ist. Damit müsste x als Polynom 1. Grades darstellbar sein als Produkt dreier Faktoren. Von diesen Faktoren muss demnach einer ersten Grades und zwei konstant sein. Da dafür offensichtlich A und B konstant sein müssten (leicht nachzurechnen), führt dies automatisch zu einem Widerspruch, da dann $A^3 + B^3 \neq x$.
Übrigens könnten wir den Beweis auch mit dem Satz von Mason führen. Die Anwendung auf $x = A^3 + B^3$ liefert nämlich $n_0(A) + n_0(B) + 1 - 1 \geq \max\{\deg(A^3), \deg(B^3)\}$, also $\deg(A) + \deg(B) \geq 3\max\{\deg(A), \deg(B)\}$. Dies ist unmöglich für nichtkonstante A und B.

Damit kann nicht jedes Polynom als Summe zweier Kuben dargestellt werden. Wir können aber durch Untersuchung der Darstellung von x zeigen, dass jedes Polynom Summe dreier Kuben ist; es gilt nämlich

$$\left(\frac{x}{6}+1\right)^3 + \left(\frac{x}{6}-1\right)^3 + \left(\frac{-x}{\sqrt{108}}\right)^3 = x.$$

Durch Substitution erhält man also für jedes beliebige komplexe Polynom P (mit beliebig vielen Variablen!):

$$\left(\frac{P}{6}+1\right)^3 + \left(\frac{P}{6}-1\right)^3 + \left(\frac{-P}{\sqrt{108}}\right)^3.$$

Demnach ist jedes Polynom als Summe von drei Kuben darstellbar.

Analog kann man x auch als Summe von höheren Potenzen darstellen, man sieht allerdings bereits hier, dass diese nicht sehr ergiebig, sondern im Gegenteil sogar sehr kompliziert sind. Als abstraktes Resultat mag eine solche Darstellung befriedigen, aber das Resultat sieht sehr unschön und willkürlich aus. Eine solche Zerlegung sagt uns nicht über die Eigenschaften von P, zumal die einzelnen Summanden höheren Grad als P haben. Effizienter wäre eine Darstellung als Potenzsumme von Polynomen mit kleinerem Grad, am besten von linearen. Dies soll uns im folgenden beschäftigen. Dazu führen wir zunächst den Begriff eines homogenen Polynoms ein.

3.2 Homogenisierung

Definition 3.2. *Ein Polynom $P \in \mathbb{C}[x_0, \ldots, x_n]$ heisst homogen vom Grad d, wenn $P(\lambda x_0, \ldots, \lambda x_n) = \lambda^d P(x_0, \ldots, x_n)$ für beliebige $\lambda \in \mathbb{C}$ gilt.*

Bemerkung: Diese Eigenschaft ist äquivalent zu der Tatsache, dass in der Darstellung $P = \sum_{i_0,\ldots,i_n} x_0^{i_0} \cdot x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}$ für alle Monome gilt: $i_0 + \cdots + i_n = d$. Homogene Polynome haben den Vorteil, dass sich bei Summation der Grad nicht verändern kann (es sei denn, es ergibt sich das Nullpolynom). Wir können durch eine leichte Modifikation jedes Polynom aus $\mathbb{C}[x_1, \ldots, x_n]$ in eine homogene Form überführen, indem wir einfach in jedem Monom den Grad durch Multiplikation mit Potenzen einer zusätzlichen Variablen x_0 auffüllen. So wird zum Beispiel aus $3x_1^2 + 4x_1 + 1$ das Polynom $3x_1^2 + 4x_1x_0 + x_0^2$. Offenbar ist diese Zuordnung umkehrbar eindeutig.
Im Folgenden wollen wir daher die Darstellbarkeit von homogenen Polynomen (mit mehreren Variablen) in der Form $P = \sum_{j=1}^{k} L_j^d$, wobei L_j Linearformen sind, untersuchen.

3.3 Quadriken

Wie im Fall ganzer Zahlen betrachten wir zunächst den quadratischen Fall. Es sei also $P(x_0, \ldots, x_n) = \sum_{i,j=0}^{n} a_{ij} x_i x_j$. Man kann P auch als Produkt

$$P(x_0, \ldots, x_n) = (x_0, \ldots, x_n) \cdot A \cdot (x_0, \ldots, x_n)^T$$

darstellen, wobei $A = (a_{ij})_{i,j=1}^{n}$ die Koeffizientenmatrix ist. Der folgende Satz aus der linearen Algebra gibt uns dann die Lösung unseres Problems:

Satz 3.3. *Nach einer linearen Koordinatentransformation ist A eine Matrix, die auf der Haupdiagonalen nur Nullen und Einsen und ansonsten nur Nullen enthält. Mit anderen Worten, $P = L_1^2 + \cdots + L_{k+1}^2$ mit $k \leq n$.*

Beweis: Der symmetrische Gauß-Algorithmus (d.h. nach jeder Zeilenoperation wird dieselbe Operation in den Spalten ausgeführt) liefert in den komplexen Zahlen das gewünschte Ergebnis.
Beispiel: $P(x_0, x_1) = x_0^2 + 4x_0x_1 + 3x_1^2$ lässt sich darstellen als $(x_0 + 2x_1)^2 + (ix_1)^2$.

3.4 Der (projektive) Raum der Polynome vom Grad d

Nun betrachten wir homogene Polynome beliebigen Grades. Wir überlegen uns leicht mit kombinatorischen Argumenten:

Proposition 3.4. *Es gibt $\binom{n+d}{d}$ Monome in den Variablen $x_0, \ldots, x_n$ vom Grad d.*

Beweis: Wir wiederholen zunächst einige Basisresultate der Kombinatorik.

$$n! = 1 \cdot 2 \cdot \ldots \cdot n$$

wird benutzt, um die Anzahl aller möglichen Permutationen einer n-elementigen Menge zu berechnen. Mit dem Binomialkoeffizienten

$$\binom{n}{k} = \frac{n!}{d! \cdot (n-k)!}$$

kann man bestimmen, wie viele Möglichkeiten es gibt, aus einer n-elementigen Menge k verschiedene Elemente auszuwählen.
Die analoge Frage nach der Anzahl der Möglichkeiten, aus einer n-elementigen Menge k Elemente auszuwählen, wobei auch mehrfache Auswahl desselben Elements zugelassen ist, ergibt die Formel

$$\binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)! \cdot k!}.$$

Das Bilden der von Monomen vom Grad d aus den $n+1$ Variablen $x_0, \ldots x_n$ entspricht einer Auswahl von d Elementen aus $(n+1)$, wobei ein Element öfters ausgewählt werden kann. Dies sind also nach den vorherigen Überlegungen genau $\binom{n+1+d-1}{d}$ Stück.

Zur Abkürzung setzen wir ab jetzt $\mathbf{N} := \binom{n+d}{d}$. Wir können nun ein Polynom $\sum_{i_0+\cdots+i_n=d} a_{i_0\ldots i_n} x^{i_0\ldots i_n}$ vom Grad d auch als Vektor mit N Koordinaten $(\ldots, a_{i_0\ldots i_n}, \ldots)$ auffassen und somit den Vektorraum $\mathbb{C}^{\mathbf{N}}$ als Raum der Polynome vom Grad d interpretieren. Da allerdings für die Darstellung als Potenzsumme $L_1^d + \cdots + L_k^d$ konstante Vielfache irrelevant sind, ist es natürlicher, die Polynome als Punkte im projektiven Raum $\mathbb{P}^{N-1}$ zu betrachten (also den Raum der Geraden in $\mathbb{C}^{\mathbf{N}}$ durch den Nullpunkt). Die Koeffizienten des Polynoms ergeben dann die homogenen Koordinaten $[\cdots : a_{i_0\ldots i_n} : \ldots]$, z.B. entspricht $x_0^2 + 4x_0x_1 + 3x_1^2$ dem Punkt $[1:4:3]$ im $\mathbb{P}^2$.

3.5 Die Veronese-Abbildung

Im Fall von Polynomen höheren Grades $d \geq 3$ können wir das Waring-Problem auf eine geometrische Fragestellung zurückführen. Dazu interpretieren wir die homogenen Polynome vom Grad d als Punkte im affinen Raum $\mathbb{C}^{\mathbf{N}}$ bzw. im projektiven Raum $\mathbb{P}^{N-1}$ ($N = \binom{n+d}{d}$). Eine natürliche Frage ist dann: Welche geometrische Form hat die Teilmenge der Potenzen von Linearformen L^d?

Man berechnet mit Hilfe der multinomischen Formel

$$(a_0x_0 + \cdots + a_nx_d)^d = \sum_{i_0+\ldots i_n=d} \frac{d!}{i_0! \ldots i_n!} \prod_{j=1}^{n} (a_jx_j)^{i_j},$$

dass solche Potenzen Koordinaten der Form $[\cdots : \frac{d!}{i_0!\ldots i_n!} \prod_{j=1^n} (a_j)^{i_j} : \cdots]$ im $\mathbb{P}^{N-1}$ haben. Nach einer koordinatenweisen Stauchung mit dem Faktor $\frac{d!}{i_0!\ldots i_n!}$ entspricht dies gerade dem Bild der sogenannten Veronese-Abbildung

$$\begin{array}{rcl} \nu_d : \mathbb{P}^n & \to & \mathbb{P}^{N-1} \\ [a_0 : \cdots : a_n] & \mapsto & [a_0^d : \cdots : a_n^d], \end{array}$$

wobei die Koordinaten im Bildraum gerade alle Monome vom Grad d durchlaufen.

Beispiel: Ein einfaches Beispiel hierfür ist die rationale Normkurve (auch „verdrehte Kubik“ genannt):

$$\begin{array}{rcl} \nu_3 : \mathbb{P}^1 & \to & \mathbb{P}^3 \\ [a_0 : a_1] & \mapsto & [a_0^3 : a_0^2a_1 : a_0a_1^2 : a_1^3]. \end{array}$$

Man sieht übrigens leicht, dass diese Kurve gerade durch die Gleichungen

$$xw = yz, \quad xz = y^2, \quad yw = z^2$$

gegeben ist (wenn $[x : y : z : w]$ die Koordinaten des $\mathbb{P}^3$ bezeichnen).
Für allgemeine n, d sind die Bilder von ν_d (dies sind die sogenannten „Veronese-Varietäten ") komplizierter, aber ebenfalls als Nullstellenmengen von Polynomen beschreibbar. Die Veronese-Abbildung erweist sich außerdem als eineindeutig. Insbesondere ist das Bild $\nu(\mathbb{P}^n)$ ebenfalls n-dimensional.

3.6 Sekantenvarietäten

Der nächste Schritt geht von der natürlichen Frage aus, welche Punkte im $\mathbb{P}^{N-1}$ Polynomen der Form $L_1^d + L_2^d$ entsprechen. Man rechnet direkt aus, dass dies gerade die Punkte sind, die auf der Geraden durch die Punkte L_1^d und L_2^d liegen. Dasselbe gilt für $k+1$ Summanden:

Satz 3.5. *Ein Polynom ist als Summe $L_1^d + \ldots L_{k+1}^d$ darstellbar, genau dann wenn der zugehörige Punkt im $\mathbb{P}^{N-1}$ in dem Raum liegt, der durch die Punkte $L_1^d, \ldots, L_{k+1}^d$ auf der Veronese-Varietät aufgespannt wird.*

Geraden durch zwei Punkte der Menge $\nu_d(\mathbb{P}^n)$ bezeichnet man als Sekanten. Dementsprechend liegt die folgende Definition nahe:

Definition 3.6. *Die Sekantenvarietät $Sec_k\nu_d(\mathbb{P}^n)$ ist die Menge der Punkte aller k-dimensionalen Räume, die durch $k+1$ Punkte auf $\nu_d(\mathbb{P}^n)$ gehen, zuzüglich der Grenzwerte dieser Punkte.*

Warum nehmen wir die Grenzwerte hinzu? Dies sind Punkte, die auf Tangenten der Veronese-Varietät liegen, und bilden in diesem Sinne eine Abschließung der Sekantenräume (wir „füllen die Lücken"). Man überlegt sich leicht, dass diese Menge relativ klein ist (es gibt viel mehr Sekanten als Tangenten). Der Vorteil dieser Abschliessung ist, dass dann $Sec_k\nu_d(\mathbb{P}^n)$ ebenfalls durch polynomiale Gleichungen im $\mathbb{P}^{N-1}$ beschrieben werden kann.
Durch diese Überlegungen haben wir folgende Modifikation des Waring-Problems erreicht:

Problem 3.7. *Für welche k ist $Sec_k\nu_d(\mathbb{P}^n) = \mathbb{P}^{N-1}$?*

Dabei ist zu bemerken, dass dies dem „großen “ Waring-Problem entspricht, nämlich der Frage, mit wieviel Summanden wir „fast alle“ Polynome darstellen können (mit Ausnahme der Punkte auf den Tangenten). Wir können die Frage sogar noch etwas weiter vereinfachen. Das Problem ist nämlich äquivalent zur Gleichheit der Dimension, d.h. wir prüfen

$$\dim Sec_k\nu_d(\mathbb{P}^n) = N-1?$$

Warum reicht dies aus? Der Grund ist, dass eine polynomiale Gleichung sofort die Dimension um 1 reduziert. Wenn also die Dimension der Sekantenvarietät $N-1$ ist, heisst das nichts anderes, als dass das zugehörige polynomiale Gleichungssystem leer ist, mit anderen Worten sie ist der ganze Raum $\mathbb{P}^{N-1}$.

3.7 Dimensionszählung und der Satz von Alexander und Hirschowitz

Wir überlegen uns nun, welche Dimension die Sekantenvarietät haben kann. Man überlegt sich schnell die folgende Abschätzung:

Lemma 3.8. $\dim Sec_k\nu_d(\mathbb{P}^n) \leq (k+1)n + k$

Beweis: Wenn wir $k+1$ Punkte auf einem n-dimensionalen Raum unabhängig voneinander auswählen, entspricht dies $(k+1)n$ Parametern. Hinzu kommt die Dimension k des k-ten Sekantenraumes.

Leider kann die Dimension echt kleiner als die rechte Seite werden, zum Beispiel wenn $\nu_d(\mathbb{P}^n)$ Geraden enthält (dann kommt durch die Sekante keine Dimension hinzu) oder wenn die Menge zu „flach “ ist (z.B. schon in einem m-dimensionalem ($m < N-1$) projektiven Unterraum des $\mathbb{P}^{N-1}$ enthalten wäre). Dies ist etwa für Quadriken der Fall (Satz 3.3). Dort hatten wir gesehen, dass wir $n+1$ Summanden benötigen. Würde hingegen im Lemma für $d = 2$ Gleichheit gelten, hätten wir $\dim Sec_k\nu_2(\mathbb{P}^n) = (k+1)n + k = \binom{n+2}{2} - 1$, also $k+1 = \frac{n+2}{2} < n+1$.
Dagegen gibt es für höhere Grade das schöne Resultat, dass in fast allen Fällen Gleichheit gilt.

Theorem 3.9 (Satz von Alexander-Hirschowitz). *Sei $d \geq 3$. Dann ist*

$$\dim Sec_k\nu_d(\mathbb{P}^n) = \min\{N-1, (k+1)n + k\}$$

für alle Tripel (n, d, k) mit Ausnahme der vier Fälle

$$(n, d, k) = (4, 3, 6), (2, 4, 4), (3, 4, 8), (4, 4, 13).$$

Durch Umstellen folgern wir:

Korollar 3.10. *Fast alle homogenen Polynome vom Grad $d \geq 3$ in $\mathbb{C}[x_0, \dots x_n]$ lassen sich als Summe $L_1^d + \cdots + L_{k+1}^d$ schreiben, wenn*

$$k + 1 \geq \frac{\binom{n+d}{d}}{(n+1)}$$

ist, mit Ausnahme der Fälle $(n, d) = (4, 3), (2, 4), (3, 4), (4, 4)$.

In den Ausnahmefällen stellt sich heraus, dass wir mit einem zusätzlichen Summanden auskommen (d.h. $8, 6, 10$ bzw. 15). Das $k + 1$ kommt in den Summationsindex, da ja einer k-Sekante $k + 1$ Summanden entsprechen (etwa im Geradenfall, also bei $k = 1$, zwei Summanden).

Beispiele:
1. Der einfachste Fall ist die verdrehte Kubik im $\mathbb{P}^3$, also $n = 1$, $d = 3$. Wir erhalten $k + 1 = 4/4 = 1$, also die Geradensekanten (und -tangenten) der Kurve füllen den Raum. Daher ist fast jedes homogene Polynom 3. Grades in zwei Unbekannten als Summe zweier Potenzen darstellbar. Äquivalent dazu ist die Aussage (wenn wir wieder von homogenen zu inhomogenen Polynomen übergehen, indem wir $x_0 = 1$ setzen), dass fast jedes Polynom in einer Variablen vom Grad ≤ 3 Summe zweier Kuben von Linearformen ist. Hier sehen wir auch, dass echte Ausnahmepunkte existieren, denn z.B. x ist nicht als Summe zweier Kuben darstellbar (siehe 3.1).
2. Im Fall von $n = d = 5$ und $n = 4$, $d = 6$ erhalten wir

$$k + 1 = \frac{\binom{5+5}{5}}{5+1} = \frac{252}{6} \text{ bzw. } k + 1 = \frac{\binom{4+6}{4}}{4+1} = \frac{210}{5},$$

also in beiden Fällen DIE ANTWORT:

42

The End

Numerische Simulation von Differentialgleichungen

Teilnehmer:

Paul Lofink	Heinrich-Hertz-Oberschule
Oliver Lorenz	Heinrich-Hertz-Oberschule
Bianca Mix	Georg-Forster-Oberschule
Christian Ritschel	Georg-Forster-Oberschule
Sebastian Günther	Georg-Forster-Oberschule

Gruppenleiter:

Caren Tischendorf	Technische Universität Berlin, Mitglied im DFG-Forschungszentrum MATHEON „Mathematik für Schlüsseltechnologien“

Die Gruppe beschäftigte sich mit der numerischen Lösung von Differentialgleichungen, um einen zweistufigen Operationsverstärker simulieren zu können.
Dazu wurde zunächst an einem einfachen Schaltkreis ein Gleichungsystem zur Simulation erstellt, dass in eine Differentialgleichung umgeformt wurde. Zur Lösung der Gleichung wurde schrittweise das explizite und das implizite Eulerverfahren hergeleitet.
Danach konnte gezeigt werden, dass die ermittelten Lösungen sinnvoll sind, wenn ihre Abweichungen von den tatsächlichen Werten eingeschränkt werden können.
Anschließend wurden sowohl das explizite als auch das implizite Eulerverfahren in Matlab implementiert, um die Ergebnisse zu veranschaulichen.
Dann wurde das gleiche Verfahren für die Simulation des zweistufigen Operationsverstärker angewandt.

1 Grundlagen

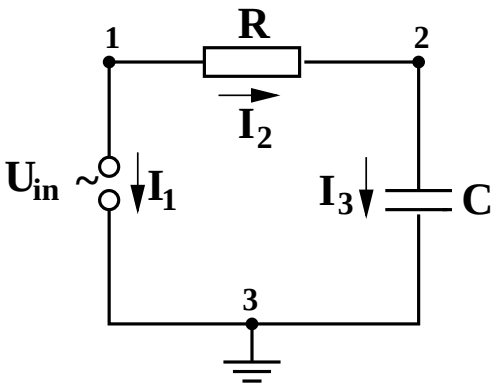


Abbildung 1: Einfache RC-Schaltung

1.1 Kirchhoff'sche Gesetze

1.1.1 Kirchhoff'sches Strom-Gesetz

Die Summe aller in einem Knoten zusammenlaufenden Ströme ist gleich Null.

$$\sum_{k=1}^{n} i_k = 0$$

Dabei haben alle vom Knoten wegführenden Ströme positives Vorzeichen und alle zum Knoten hinfließenden Ströme negatives Vorzeichen.

1.1.2 Kirchhoff'sches Spannungsgesetz

Die Summe aller zu einer Schleife gehörenden Spannungen ist gleich Null.

$$\sum_{k=1}^{n} v_k = 0$$

Dabei haben natürlicher Weise positive Spannungen positives Vorzeichen und negative Spannungen negatives Vorzeichen.

1.2 Abgeleitete Gleichungen

Aus den Kirchhoff'schen Gesetzen ergeben sich für das Schaltbild Abb.1 folgende Gleichungen:

$$\begin{aligned} I_1 + I_2 &= 0 && (1)\\ -I_2 + I_3 &= 0 && (2)\\ -U_1 + U_2 + U_3 &= 0 && (3)\\ U_1 &= U_{in} && (4)\\ R_1 &= \frac{U_2}{I_2} && (5)\\ I_3 &= C\frac{dU_3}{dt} && (6) \end{aligned}$$

Aus dem Gleichungssystem ergibt sich für die Kondensatorspannung U_3:

$$\frac{dU_3}{dt} = \frac{1}{R_1 C_1}(U_{in} - U_3)$$

Bei Gleichung (7) handelt es sich um eine Differentialgleichung. Diese ist aber nicht ohne weiteres lösbar. Die allgemeine Form einer Differentialgleichung lautet:

$$\frac{dx}{dt}(t) = \lambda x(t) + g(t) \qquad (7)$$

2 Explizites und Implizites Eulerverfahren

2.1 Explizites Eulerverfahren

Frage:
Wie kann man $x'(t) = \lambda x(t) + g(t)$ für beliebige λ und beliebige Funktionen g(t) lösen ?

Idee: Numerisch

Dafür wird das zu betrachtende Zeitintervall $[t_0, t_{end}]$ in viele kleine Zeitintervalle der Länge h zerlegt.
Für $h \to 0$ gilt $x_1 \to x(t_1)$, wobei x_1 der approximative Wert und $x(t_1)$ der tatsächliche Wert ist.

Bemerkung: (7) besitzt eine eindeutige Lösung, falls $x(0) = x_0$ ist, wobei $x_0 \in \mathbb{R}$ eine beliebige, aber feste reelle Zahl ist.

Problem: Wir suchen die eindeutige Lösung bei vorgegebenem x_0 zu den Zeit-

punkten $t_1, t_2, t_3, ..., t_{end}$!

$$\begin{aligned} x'(0) &= \lambda x(0) + g(0) \\ &= \lambda x_0 + g(0) \\ x'(t_1) &= \lambda x(t_1) + g(t_1) \\ &\approx \lambda x_1 + g(t_1) \end{aligned}$$

Wir berechnen Approximationen x_n für $x(t_n)$ mit Hilfe von

$$\frac{x_n - x_{n-1}}{h} = \lambda x_{n-1} + g(t_{n-1}) \tag{8}$$

bei x_0 startend.

$$\Leftrightarrow \qquad x_n = x_{n-1} + h(\lambda x_{n-1} + g(t_{n-1}))$$

$\frac{x_n - x_{n-1}}{h}$ ist der Differenzenquotient und daher gilt: $\frac{x_n - x_{n-1}}{h} \approx x'(t_{n-1})$

Wir betrachten folgende Differentialgleichung :

$$x'(t) = \lambda x(t) + g(t) \tag{9}$$

Satz 1: Für jedes $x_0 \epsilon \mathbb{R}$ existiert auf dem endlichen Intervall $[t_0, t_{end}]$ genau eine Lösung x von (9), das heißt,

$$x'(t) = \lambda x(t) + g(t) \qquad \forall t \epsilon [t_0, t_{end}]$$

mit der Anfangsbedingung $x(t_0) = x_0$, vorrausgesetzt,dass $g(t)$ stetig ist. Somit gilt für die Lösung, $x(t)$ ist stetig differenzierbar.

Der Satz 1 wird hier nicht bewiesen.

Satz 2: Sei g auf $[t_0, t_{end}]$ stetig und die Schrittweite h hinreichend klein. Dann gilt für die eindeutige Lösung $x(t)$ von (9) mit $x(t_0) = x_0$ und die numerische Lösung von (8):
Es existiert eine Konstante $c > 0$:

$$\max_{n=1,...,N} \mid x(t_n) - x_n \mid \leq ch$$

wobei c unabhängig von N ist. N sei die Anzahl der Teilintervalle.

2.1.1 Beweis

$$\frac{x_n - x_{n-1}}{h} = \lambda x_{n-1} + g(t_{n-1}) \quad (10)$$

$$(9) \underset{t=t_{n-1}}{\Rightarrow} \quad x'(t_{n-1} = \lambda x(t_{n-1}) + g(t_{n-1}) \quad (11)$$

1. Man bildet die Differenz aus der Verfahrensgleichung (10) und der Differentialgleichung im Punkt t_{n-1} (11).
Dann gilt für den globalen Fehler $e_n := x_n - x(t_n)$

$$e_n = (1 + \lambda h)e_{n-1} + hr_n$$

r_n ist der lokale Diskretisierungsfehler.

2. Durch rekursives Auflösen erhält man

$$e_n = (1 + \lambda h)^n e_0 + \sum_{i=0}^{n-1} (1 + \lambda h)^i h r_{n-i}$$

$$e_0 = x_0 - x(t_0) = 0 \quad \Rightarrow e_n = \sum_{i=0}^{n-1} |1 + \lambda h|^i h r_{n-i}$$

3. Dreiecksungleichung und Taylorentwicklung
Es existiert eine Konstante $c > 0$, so dass $\max\limits_n |r_n| < ch$.

$$\underset{Dreiecksungleichung}{\Longrightarrow} \mid e_n \mid \leq ch^2 \sum_{i=0}^{n-1} (1 + \lambda h)^i = ch^2 \begin{cases} n \text{ für } \mid 1 + \lambda h \mid = 0 \\ \frac{|1+\lambda h|^n - 1}{|1+\lambda h| - 1} \text{ sonst} \end{cases}$$

$$\max_N \mid e_N \mid \leq ch^2 \begin{cases} N \text{ für } \mid 1 + \lambda h \mid = 0 \\ \frac{|1+\lambda h|^N - 1}{|1+\lambda h| - 1} \text{ sonst} \end{cases}$$

4. Wir nutzen $Nh = t_n - t_0$

$$\max_n \mid e_n \mid \leq \check{c}h \text{ falls } \mid 1 + \lambda h \mid > 0 \text{ d.h. } h < \frac{-1}{\lambda} \text{ für } \lambda < 0$$

q.e.d.

2.1.2 Implizites Eulerverfahren

Wir betrachten den Punkt t_n

$$x'(t_n) = \lambda x(t_n) + g(t_n)$$
$$\approx \frac{x(t_n) - x(t_{n-1})}{h}$$

Daraus ergibt sich dann ein neues Verfahren

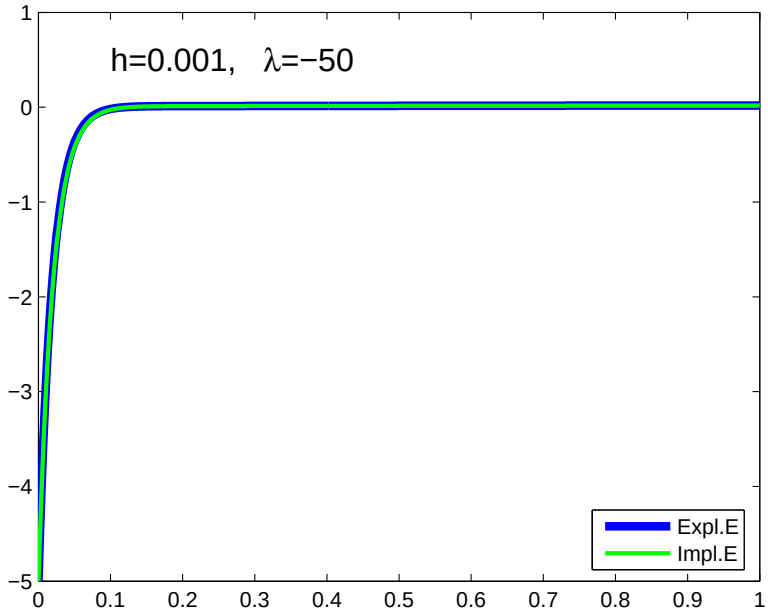
$$\frac{x_n - x_{n-1}}{h} = \lambda x_n + g(t_n)$$

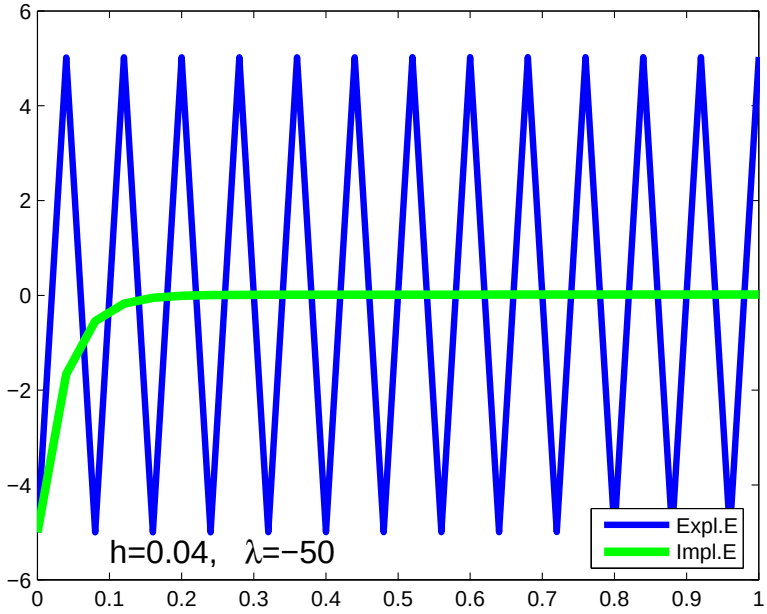
Implizites Eulerverfahren:

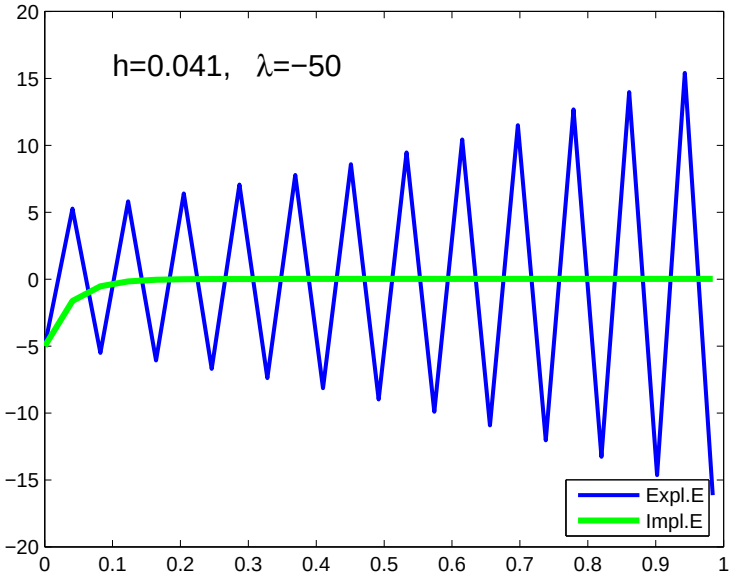
$$x_n = \frac{g(t_n)h + x_{n-1}}{1 - h\lambda}$$

2.2 Vorteile des Impliziten Eulers bezogen auf die Anwendung

Graph des expliziten und impliziten Eulers für vordefinierte h:



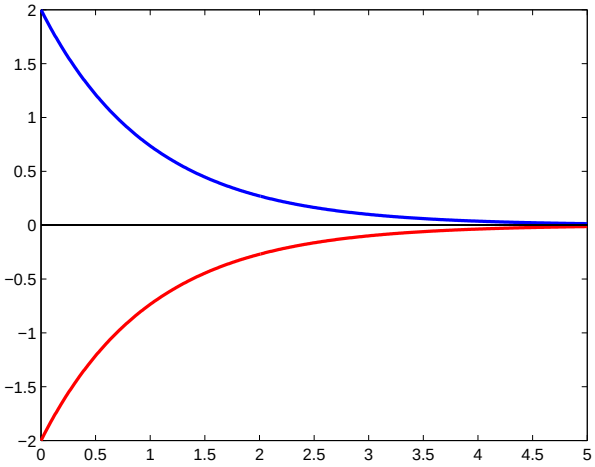




Es ist zu erkennen, dass das explizite Eulerverfahren, bei größerem h, nicht den Graph einer Exponentialfunktion liefert, der den Spannungsabfall über einem Kondensator darstellt. Es ist im zweiten Diagramm eine zwischen -5 und 5 oszillierende Funktion. Im dritten Diagramm oszilliert der Graph und divergiert außerdem. Das implizite Eulerverfahren liefert dagegen den Graph einer Exponentialfunktion. Da nur der Parameter h verändert wurde, kann man annehmen,

dass die Größe von h den Verlauf des Graphen des expliziten Eulers entscheidend beeinflusst.

$$x' = \lambda x \text{ mit } \lambda < 0$$



Für die Lösung $x(t) = x(0)e^{\lambda t}$ gilt: $|x(t)|$ ist monoton fallend.

Expliziter Euler:

$$\frac{x_n - x_{n-1}}{h} = \lambda x_{n-1} \quad \Rightarrow \quad x_n = (1 + \lambda h)x_{n-1}$$

Es gilt $|x_n| < |x_{n-1}|$ nur für $|1 + \lambda h| < 1$.

h als Schrittweite ist immer positiv, λ negativ.

$$\begin{aligned} |1 + \lambda h| < 1 \quad &\Leftrightarrow \quad 1 + \lambda h < 1 \quad \wedge \quad 1 + \lambda h > -1 \\ &\Leftrightarrow \quad \lambda h < 0 \quad \wedge \quad h < -\frac{2}{\lambda}, \quad \text{d.h. nur für } h < -\tfrac{2}{\lambda}. \end{aligned}$$

Impliziter Euler:

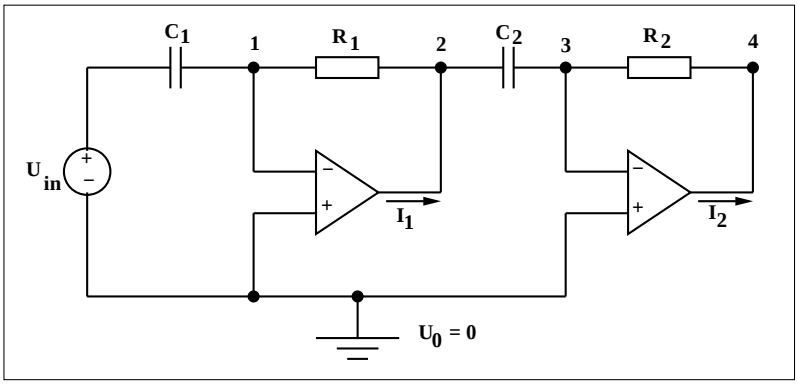
$$\frac{x_n - x_{n-1}}{h} = \lambda x_n \quad \Rightarrow \quad x_n = \frac{1}{1 - \lambda h} x_{n-1}$$

Es gilt $|x_n| < |x_{n-1}|$ für alle $h > 0$, denn

$$\frac{1}{1 - \lambda h} = \frac{1}{1 + \varepsilon} \quad < 1 \quad \text{mit } \varepsilon := -\lambda h > 0$$

Das implizite Eulerverfahren erzeugt immer abklingende Lösungen unabhängig von der Schrittweite.

3 Zweifacher Operationsverstärker



Nach dem Kirchhoff'schen Stromgesetz gilt:

$$\begin{aligned} -I_{C_1} + I_{R_1} &= 0 \\ -I_1 + I_{C_2} - I_{R_1} &= 0 \\ -I_{C_2} + I_{R_2} &= 0 \\ -I_{R_2} - I_{R_2} - I_2 &= 0 \\ I_{C_1} + I_V &= 0 \end{aligned}$$

Spannungen:

$$\begin{aligned} V_{R_1} &= U_1 - U_2 \\ V_{R_2} &= U_3 - U_4 \\ V_{C_1} &= U_5 - U_1 \\ V_{C_2} &= U_2 - U_3 \end{aligned}$$

Elementbezogene Gleichungen

$$\begin{aligned} I_{C_1} &= C_1 \frac{dV_{C_1}}{dt} = C_1(\frac{dU_5}{dt} - \frac{dU_1}{dt}) \\ I_{C_2} &= C_2 \frac{dV_{C_2}}{dt} = C_2(\frac{dU_2}{dt} - \frac{dU_3}{dt}) \\ I_{R_1} &= \frac{V_{R_1}}{R_1} = \frac{U_1 - U_2}{R_1} \\ I_{R_2} &= \frac{V_{R_2}}{R_2} = \frac{U_3 - U_4}{R_2} \\ U_2 &= -AU_1 \\ U_4 &= -AU_3 \\ U_5 &= U_{in} \end{aligned}$$

Es ergibt sich, nach mehreren Umformungen, ein Gleichungssystem mit den beiden Unbekannten U_1 und U_3:

$$\begin{aligned} -C_1(\frac{dU_{in}}{dt} - \frac{dU_1}{dt}) + \frac{1}{R_1}(U_1(A+1)) &= 0 \\ -C_2(\frac{d(AU_1)}{dt} - \frac{dU_3}{dt}) + \frac{1}{R_2}(U_3(A+1)) &= 0 \end{aligned}$$

$$\begin{aligned} &\iff && \begin{cases} \frac{dU_1}{dt} = -\frac{1+a}{R_1C_1}U_1 + \frac{dU_{in}}{dt} \\ A\frac{dU_1}{dt} + \frac{dU_3}{dt} = -\frac{1+A}{R_2C_2}U_3 \end{cases} \\ &\iff && \begin{cases} U_1' = -\frac{1+A}{R_1C_1} + U_{in}' \\ U_3' = \frac{1+A}{R_2C_2}U_3 + A\frac{1+A}{R_1C_1}U_1 - AU_{in}' \end{cases} \\ &\iff && \begin{pmatrix} U_1' \\ U_3' \end{pmatrix} = \begin{pmatrix} b_1 & b_2 \\ b_3 & b_4 \end{pmatrix} \begin{pmatrix} U_1 \\ U_3 \end{pmatrix} + \begin{pmatrix} g_1(t) \\ g_2(t) \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \overrightarrow{x}' = B\overrightarrow{x} + \overrightarrow{g}(t) \text{ mit} \quad b_1 &= -\frac{1+A}{R_1C_1} \text{ , } b_2 = 0 \text{ , } g_1(t) = U_{in}'(t) \\ b_3 &= A\frac{1+A}{R_1C_1} \text{ , } b_4 = \frac{1+A}{R_2C_2} \text{ , } g_2(t) = -AU_{in}'(t) \end{aligned}$$

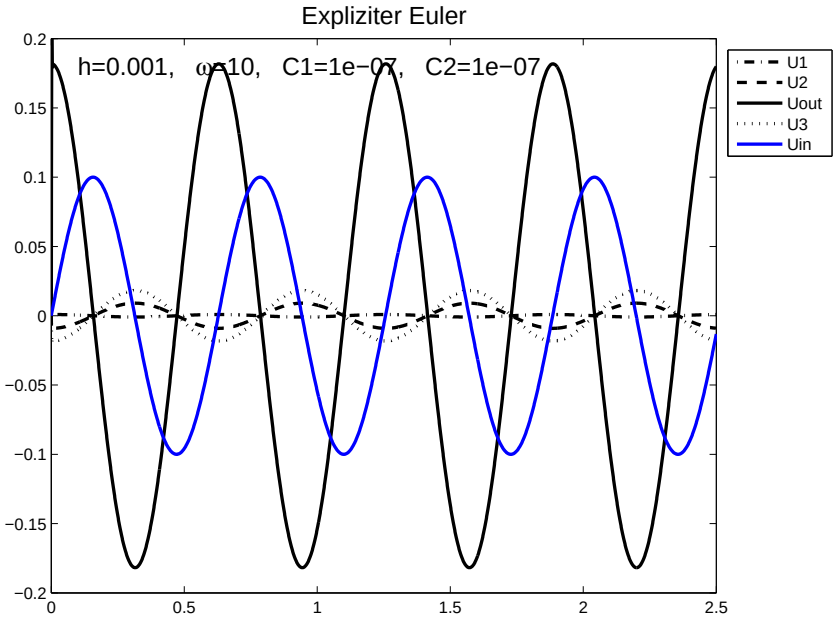
Explizites Eulerverfahren für Systeme:

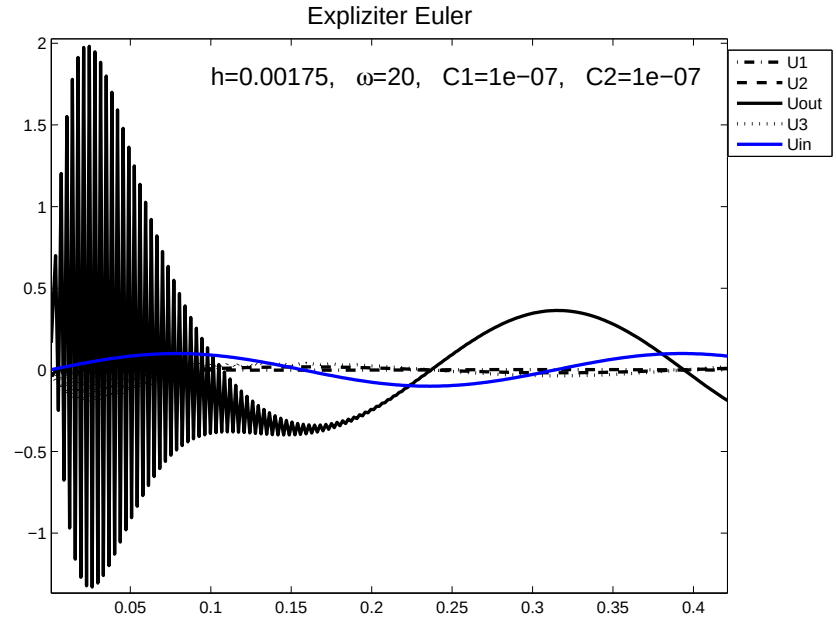
$$\begin{aligned} \frac{\overrightarrow{x_n} - \overrightarrow{x}_{n-1}}{h} &= B\overrightarrow{x}_{n-1} + \overrightarrow{g}(t_{n-1}) \\ \overrightarrow{x}_n &= \overrightarrow{x}_{n-1} + hB + h\overrightarrow{g}(t_{n-1}) \end{aligned}$$

Implizites Eulerverfahren für Systeme:

$$\begin{aligned} \frac{\overrightarrow{x}_n - \overrightarrow{x}_{n-1}}{h} &= B\overrightarrow{x}_n + \overrightarrow{g}(t_n) \\ \overrightarrow{x}_n - hB\overrightarrow{x}_n &= \overrightarrow{x}_{n-1} + h\overrightarrow{g}(t_n) \\ LGS \rightarrow \quad (I - hB)\overrightarrow{x}_n &= \overrightarrow{x}_{n-1} + h\overrightarrow{g}(t_n) \\ \overrightarrow{x}_n &= (I - hB)^{-1}(\overrightarrow{x}_{n-1} + h\overrightarrow{g}(t_n)) \end{aligned}$$

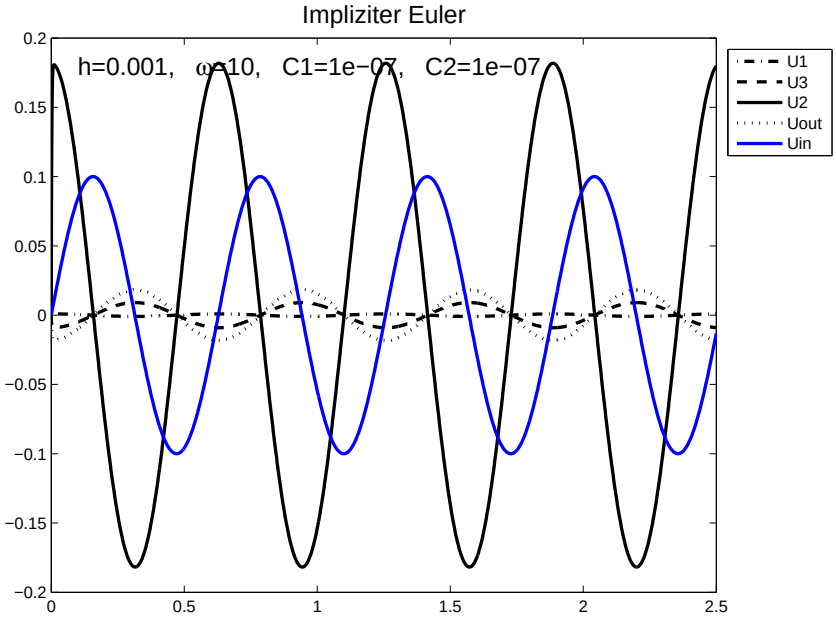
4 Ergebnisse für den zweifachen Operationsverstärker

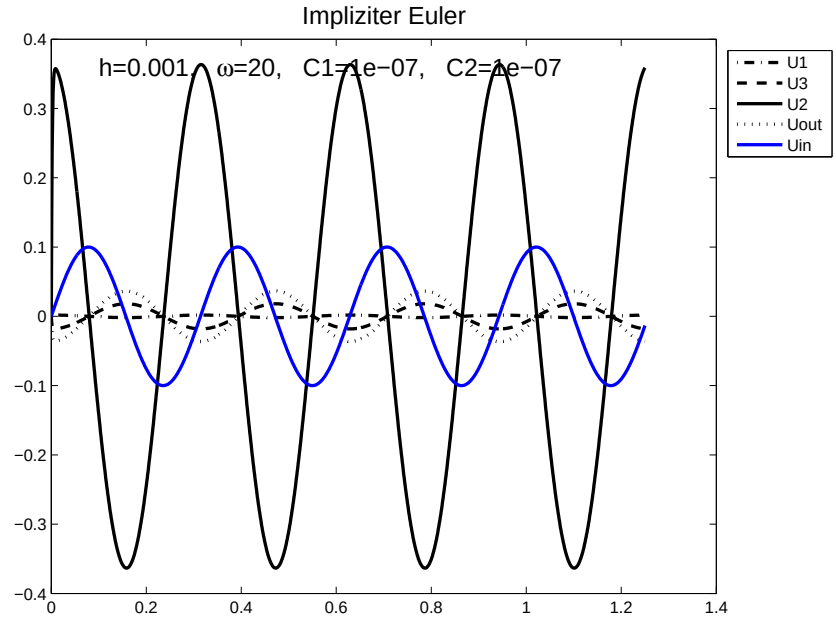




Die obigen beiden Grafiken zeigen, dass der explizite Euler für kleine Schritwweiten vernünftige Lösungen, für große Schrittweiten aber unsinnige Lösungen liefert.

Das implizite Eulerverfahren hingegen liefert auch für größere Schrittweiten vernünftige Lösungen. Insbesondere sieht man in den nächsten beiden Abbildungen, dass sich die Outputspannung $U_{out} = U_4$ verdoppelt, wenn man die Frequenz ω des Eingangssignals verdoppelt.





Die Unlösbarkeit der Gleichung fünften Grades durch Radikale

Teilnehmer:

Max Bender	Andreas-Oberschule
Marcus Gawlik	Georg-Forster-Oberschule
Anton Milge	Georg-Forster-Oberschule
Leonard Poetzsch	Georg-Forster-Oberschule
Gabor Radtke	Georg-Forster-Oberschule
Miao Zhang	Andreas-Oberschule

Gruppenleiter:

Jürg Kramer	Humboldt-Universität zu Berlin, Mitglied im DFG-Forschungszentrum MATHEON „Mathematik für Schlüsseltechnologien“

Die Gruppe beschäftigte sich mit der Frage nach der Lösbarkeit der allgemeinen Gleichung fünften Grades durch Radikale. Zunächst wurde dazu festgestellt, dass lineare, quadratische, kubische und quartische Gleichungen durch Radikale lösbar sind.
Mit Hilfe von N.-H. Abels Originalarbeit aus dem 19. Jh. erarbeitete sich die Gruppe dann das Ergebnis, dass die allgemeine Gleichung fünften Grades nicht durch Radikale lösbar ist.
Dazu musste sich die Gruppe einige Grundlagen der Gruppen- sowie der Körpertheorie erarbeiten. Speziell spielte das Verständnis der symmetrischen Gruppe S_5 von fünf Elementen eine wichtige Rolle.

Die Unlösbarkeit der allgemeinen Gleichung fünften Grades durch Radikale

1 Einleitung

Die allgemeine Gleichung eines Polynoms n-ten Grades mit den Koeffizienten $\sigma_1, \ldots, \sigma_n$ lautet

$$f(X) = X^n - \sigma_1 X^{n-1} \pm \ldots + (-1)^{n-1}\sigma_{n-1}X + (-1)^n\sigma_n.$$

Nach dem Fundamentalsatz der Algebra besitzt ein solches Polynom genau n Nullstellen $x_1, \ldots, x_n \in \mathbb{C}$. Somit lässt sich die Funktion eindeutig als Produkt ihrer Linearfaktoren darstellen

$$f(X) = (X - x_1) \cdot (X - x_2) \cdot \ldots \cdot (X - x_n).$$

Nach dem Vietaschen Wurzelsatz ergeben sich folgende Zusammenhänge zwischen den Nullstellen $x_1, \ldots, x_n$ und den Koeffizienten $\sigma_1, \ldots, \sigma_n$

$$\begin{array}{rcccc} \sigma_1 & = & \sigma(x_1, \ldots, x_n) & = & x_1 + \ldots + x_n, \\ \sigma_2 & = & \sigma(x_1, \ldots, x_n) & = & x_1 \cdot x_2 + x_1 \cdot x_3 + \ldots + x_{n-1} \cdot x_n, \\ \vdots & & & & \\ \sigma_n & = & \sigma(x_1, \ldots, x_n) & = & x_1 \cdot x_2 \cdot \ldots \cdot x_n. \end{array}$$

Aufgrund ihrer Symmetrie verändern sich die Werte dieser Funktionen beim Vertauschen (Permutieren) der Variablen nicht. Sie werden deshalb die *elementarsymmetrischen Polynome* in den Variablen $x_1, \ldots, x_n$ genannt. Von großer Bedeutung für den sich später ergebenden Widerspruch ist, dass die $x_1, \ldots, x_n$ als variabel vorausgesetzt werden. Das bedeutet, dass keine algebraischen Zusammenhänge zwischen den Variablen $x_1, \ldots, x_n$ bestehen dürfen. Als Vorbereitung auf die nachfolgende, grundlegende Definition betrachten wir das Beispiel einer quadratischen Gleichung, d.h.

$$X^2 - \sigma_1 X + \sigma_2 = 0. \tag{1}$$

Die beiden Lösungen dieser Gleichung lassen sich mithilfe der p, q-Formel berechnen zu

$$x_{1,2} = \frac{\sigma_1}{2} \pm \sqrt{\frac{\sigma_1^2}{4} - \sigma_2}.$$

Dabei sieht man, dass die auftretenden Funktionen $p = \sigma_1/2$ und $R = \sigma_1^2/4 - \sigma_2$ rationale Funktionen (in diesem Fall sogar Polynome) in den elementarsymmetrischen Polynomen σ_1, σ_2 sind.

2 Die Problemstellung

Für die Fortführung ist es zunächst wichtig, einige Bezeichnungen einzuführen. Dazu sei $\mathbb{C}[x_1, \ldots, x_n]$ die Menge der Polynome in $x_1, \ldots, x_n$ mit komplexen Koeffizienten. Ein Quotient zweier Polynome $P, Q \in \mathbb{C}[x_1, \ldots, x_n]$ heißt rationale Funktion. Entsprechend bezeichnen wir mit

$$\mathbb{C}(x_1, \ldots, x_n) := \left\{ \frac{P}{Q} \;\middle|\; P, Q \in \mathbb{C}[x_1, \ldots, x_n] \right\}$$

den Körper aller rationalen Funktionen in den n Variablen $x_1, \ldots, x_n$. Analog sei $\mathbb{C}[\sigma_1, \ldots, \sigma_n]$ die Menge der Polynome in $\sigma_1, \ldots, \sigma_n$ bzw. $\mathbb{C}(\sigma_1, \ldots, \sigma_n)$ der Körper der rationalen Funktionen in $\sigma_1, \ldots, \sigma_n$. Wir haben

$$\mathbb{C}[\sigma_1, \ldots, \sigma_n] \subseteq \mathbb{C}[x_1, \ldots, x_n]$$

und

$$\mathbb{C}(\sigma_1, \ldots, \sigma_n) \subseteq \mathbb{C}(x_1, \ldots, x_n).$$

Im Beispiel (1) wurde ein Polynom zweiten Grades mit der p, q-Formel gelöst. Dabei ergab sich die Form

$$x_1 = p + \sqrt{R} \tag{2}$$

mit $p, R \in \mathbb{C}(\sigma_1, \sigma_2)$. Wir stellen uns nun die Frage, ob i.A. die Nullstellen von f durch solche Radikale darstellbar sind. Die Form der Gleichung (2) und die Cardano-Formeln für Polynome dritten und vierten Grades motivieren dabei die folgende Definition.

Definition. Eine Nullstelle $x = x_j$ heißt durch Radikale darstellbar, falls es ein $m \in \mathbb{N}$ und rationale Funktionen $R, p, p_1, \ldots, p_{m-1} \in \mathbb{C}(\sigma_1, \ldots, \sigma_n)$ gibt, so dass

$$x = p + p_1 \sqrt[m]{R} + \ldots + p_{m-1} \left(\sqrt[m]{R} \right)^{m-1}$$

gilt; dabei gilt $\sqrt[m]{R} \notin \mathbb{C}(\sigma_1, \ldots, \sigma_n)$.

Wir können folgende Vereinfachungen erreichen:

- $p_1 = 1$.
- m eine Primzahl.

Ersetzt man nämlich R durch R/p_1^m, so erhält man $p_1 = 1$ und verändert die Voraussetzungen nicht. Auf den Fall $m =$ Primzahl sind wir geführt, indem wir eine Iteration der obigen Darstellung vornehmen.

Annahme. Um herauszufinden, ob die Nullstellen von f als Radikale darstellbar sind, treffen wir die Annahme, dass die Gleichung $f(X) = 0$ eine solche Lösung besitzt. Nach der ersten Vereinfachung ist die angenommene Lösung also von der Form

$$x_1 = p + \sqrt[m]{R} + \ldots + p_{m-1} \left(\sqrt[m]{R} \right)^{m-1}.$$

Im Weiteren wird es unser Ziel sein, einen **Widerspruch** zu dieser Annahme herzuleiten.

3 Der erste Beweisschritt

Wir schränken nun auf den Fall $n = 5$ ein; außerdem erinnern wir uns daran, dass m eine Primzahl ist. Wir nehmen also an, dass unsere Ausgangsgleichung $f(X) = 0$ eine Lösung der Form

$$x_1 = p + \sqrt[m]{R} + \ldots + p_{m-1} \left(\sqrt[m]{R} \right)^{m-1} \tag{3}$$

besitzt.
Bevor wir im ersten Beweisschritt fortfahren, haben wir noch die m-ten Einheitswurzeln ζ^j einzuführen; dabei ist

$$\zeta = e^{2\pi i/m} = \cos\left(\frac{2\pi}{m}\right) + i \cdot \sin\left(\frac{2\pi}{m}\right)$$

und $j = 0, \ldots, m-1$ ist. Wir erinnern daran, dass die m Einheitswurzeln $1, \zeta, \zeta^2, \ldots, \zeta^{m-1}$ im Einheitskreis der komplexen Ebene (mit dem Ursprung als Zentrum) ein regelmäßiges m-Gon einbeschreiben.

Nach langwierigen Überlegungen, welche aber im wesentlichen nur Methoden der Linearen Algebra und die Berechnung des größten gemeinsamen Teilers von Polynomen benötigen, stellen wir fest, dass mit der Lösung (3) gleichzeitig auch

$$\begin{aligned} x_2 &= p + \zeta \sqrt[m]{R} + \ldots + p_{m-1} \left(\zeta \sqrt[m]{R}\right)^{m-1} \\ &\vdots \\ x_m &= p + \zeta^{m-1} \sqrt[m]{R} + \ldots + p_{m-1} \left(\zeta^{m-1} \sqrt[m]{R}\right)^{m-1} \end{aligned}$$

Lösungen der Ausgangsgleichung $f(X) = 0$ sind. Diese stellen sich als paarweise verschieden heraus, d.h. wir haben somit m verschiedene Lösungen der Ausgangsgleichung. Da nun aber $f(X)$ ein Polynom vom Grad $n = 5$ ist, kann es höchstens 5 verschiedene Nullstellen haben, d.h. wir haben $m \leq 5$. Da m aber eine Primzahl ist, haben wir den Beweis auf die Fälle

$$m = 2, m = 3, m = 5$$

reduziert. Diese gilt es im Folgenden zu untersuchen. Dabei beschränken wir uns hier auf die Betrachtung des Falls $m = 5$; der Fall $m = 2$ lässt sich analog behandeln. Schließlich lässt sich zeigen, dass der Fall $m = 3$ gar nicht auftreten kann.

4 Der zweite Beweisschritt

Wir gehen aus von den m Gleichungen

$$\begin{aligned} x_1 &= p + \sqrt[m]{R} + \ldots + p_{m-1} \left(\sqrt[m]{R}\right)^{m-1}, \\ x_2 &= p + \zeta \sqrt[m]{R} + \ldots + p_{m-1} \left(\zeta \sqrt[m]{R}\right)^{m-1}, \\ &\vdots \\ x_m &= p + \zeta^{m-1} \sqrt[m]{R} + \ldots + p_{m-1} \left(\zeta^{m-1} \sqrt[m]{R}\right)^{m-1}. \end{aligned}$$

Indem wir

$$y_1 = p,\ y_2 = \sqrt[m]{R},\ y_3 = p_2 \left(\sqrt[m]{R}\right)^2, \ldots,\ y_m = p_{m-1} \left(\sqrt[m]{R}\right)^{m-1}$$

setzen, erhalten wir das lineare Gleichungssystem

$$\begin{array}{ccccccccccc}
y_1 & + & y_2 & + & y_3 & + & \dots & + & y_m & = & x_1 \\
y_1 & + & \zeta y_2 & + & \zeta^2 y_3 & + & \dots & + & \zeta^{m-1} y_m & = & x_2 \\
\vdots & & & & & & & & & & \\
y_1 & + & \zeta^{m-1} y_2 & + & \zeta^{2(m-1)} y_3 & + & \dots & + & \zeta^{(m-1)^2} y_m & = & x_m.
\end{array}$$

Nach der bekannten Theorie der linearen Gleichungssysteme sind die Lösungen $y_1, y_2, \dots, y_m$ gegeben als Quotienten von Polynomen in den Koeffizienten des Gleichungssystems, d.h. $y_1, y_2, \dots, y_m$ sind rationale Funktionen in $x_1, x_2, \dots, x_n$. Insbesondere stellen wir fest

$$\sqrt[m]{R} \in \mathbb{C}(x_1, \dots, x_n).$$

Zum Beispiel berechnet man im Fall $m = 3$ leicht

$$\sqrt[3]{R} = \frac{x_1 + \zeta^2 \cdot x_2 + \zeta \cdot x_3}{3}.$$

5 Der dritte Beweisschritt

Die symmetrische Gruppe S_n vom Index n ist definiert als die Menge aller Permutationen von n Elementen. Eine Permutation π ist dabei einfach eine Anordnung der n Elemente, wobei jedes Element in der neuen Anordnung genau einmal vorkommen soll und je zwei verschiedenen Elementen auch zwei verschiedene Bildelemente zugeordnet werden. Wir können sie also auch als injektive Abbildung der Menge der n Elemente auf sich selbst auffassen. Eine Permutation ist also insbesondere bijektiv. Wir arbeiten im Folgenden mit den natürlichen Zahlen von 1 bis n als Elementen. Diese Permutationen finden ihre Anwendung nun bei Polynomen und rationalen Funktionen in den n Variablen $x_1, \dots, x_n$. Wir lassen dabei eine Permutation auf die Indizes wirken. Sei also $\pi \in S_n$ eine Permutation. Dann definieren wir

$$g^\pi(x_1, \dots, x_n) := g(x_{\pi(1)}, \dots, x_{\pi(n)}),$$

wobei $g \in \mathbb{C}(x_1, \dots, x_n)$ ist. Nun definieren wir weiter die "Wertemenge" von g durch

$$W(g) := \{g^\pi \mid \pi \in S_n\}.$$

Eine rationale Funktion g mit $W(g) = \{g\}$, also $|W(g)| = 1$, nennen wir eine symmetrische rationale Funktion. Man überzeugt sich leicht davon, dass man Permutieren und Addieren bzw. Multiplizieren vertauschen kann, da diese beiden Prozesse völlig unabhängig voneinander sind. Induktiv erhält man dann, dass dies sogar für mehrere Summanden und Faktoren einer Summe bzw. eines Produkts gilt. Als Spezialfall, bei dem die Faktoren alle gleich sind, erhalten wir somit folgendes Vertauschungsgesetz für Potenzen:

$$(g^\pi)^r = (g^r)^\pi \quad (r \in \mathbb{N}).$$

Daraus folgt insbesondere für $r = m = 5$

$$(g^\pi)^5 = \left(g^5\right)^\pi = \left(\left(\sqrt[5]{R}\right)^5\right)^\pi = R^\pi = R.$$

Letzteres Gleichheitszeichen gilt wegen der Voraussetzung $R \in \mathbb{C}(\sigma_1, \ldots, \sigma_n)$. Somit ist wie $\sqrt[5]{R}$ auch $\left(\sqrt[5]{R}\right)^\pi$ eine Lösung der Gleichung

$$z^5 - R = 0.$$

Diese Lösungen sind aber als fünfte Wurzeln alle von der Form

$$z_j = \zeta^j \cdot \sqrt[5]{R} \quad (j \in \{0, \ldots, 4\}).$$

Daraus folgt also

$$\left(\sqrt[5]{R}\right)^\pi = \zeta^j \cdot \sqrt[5]{R}$$

mit einem $j \in \{0, \ldots, 4\}$. Mithilfe von etwas Gruppentheorie (Operationen, Bahnen und Stabilisatoren) kann man zeigen, dass auch wirklich jedes der z_j $(j = 0, \ldots, 4)$ als Bild angenommen wird, d.h. dass es zu jedem dieser z_j auch wirklich eine Permutation $\pi \in S_n$ gibt mit

$$\left(\sqrt[5]{R}\right)^\pi = \zeta^j \cdot \sqrt[5]{R}.$$

Somit folgern wir aus den obigen Überlegungen und dem letzten Argument, dass

$$W\left(\sqrt[5]{R}\right) = \left\{\sqrt[5]{R}, \zeta \cdot \sqrt[5]{R}, \ldots, \zeta^4 \cdot \sqrt[5]{R}\right\}$$

gilt und erhalten dann $\left|W\left(\sqrt[5]{R}\right)\right| = 5$.

6 Der Widerspruch

Wir erinnern zunächst nochmals daran, dass wir mit der allgemeinen Gleichung fünften Grades arbeiten, d.h. wir nehmen an, dass die Nullstellen $x_1, \ldots, x_5$ unabhängige Variablen sind, also keiner algebraischen Gleichung (mit komplexen Koeffizienten) genügen.

Ohne Beweis zitieren wir den folgenden Satz.

Satz. Ist $g \in \mathbb{C}(x_1, \ldots, x_5)$ eine rationale Funktion mit der Eigenschaft $|W(g)| = 5$, so ist g (ohne Beschränkung der Allgemeinheit) von der Form

$$g = g(x_1, \ldots, x_5) = q_0 + q_1 \cdot x_1 + q_2 \cdot x_1^2 + q_3 \cdot x_1^3 + q_4 \cdot x_1^4$$

mit symmetrischen Funktionen $q_0, q_1, \ldots, q_4 \in \mathbb{C}(\sigma_1, \ldots, \sigma_5)$.

Nach dem zweiten Beweisschritt ist $\sqrt[5]{R} \in \mathbb{C}(x_1, \ldots, x_5)$. Da nach dem dritten Beweisschritt weiter $|W(\sqrt[5]{R})| = 5$ gilt, können wir den vorhergehenden Satz anwenden. Wir finden, dass $\sqrt[5]{R}$ die Gestalt

$$\sqrt[5]{R} = q_0 + q_1 \cdot x_1 + q_2 \cdot x_1^2 + q_3 \cdot x_1^3 + q_4 \cdot x_1^4$$

hat, wobei $q_0, q_1, \ldots, q_4 \in \mathbb{C}(\sigma_1, \ldots, \sigma_5)$ sind.
Wir wenden nun die Permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$$

auf die rationale Funktion $\sqrt[5]{R}$ an und erhalten

$$\begin{aligned} \left(\sqrt[5]{R}\right)^\pi &= \left(q_0 + q_1 \cdot x_1 + q_2 \cdot x_1^2 + q_3 \cdot x_1^3 + q_4 \cdot x_1^4\right)^\pi \\ &= q_0^\pi + q_1^\pi \cdot x_2 + q_2^\pi \cdot x_2^2 + q_3^\pi \cdot x_2^3 + q_4^\pi \cdot x_2^4 \\ &= q_0 + q_1 \cdot x_2 + q_2 \cdot x_2^2 + q_3 \cdot x_2^3 + q_4 \cdot x_2^4 ; \end{aligned}$$

dabei haben wir insbesondere beachtet, dass $x_1^\pi = x_2$ ist und dass die Funktionen $q_0, q_1, \ldots, q_4$ symmetrisch sind.

Nach dem dritten Beweisschritt wissen wir andererseits, dass ein $j \in \{0, \ldots, 4\}$ existiert, so dass

$$\left(\sqrt[5]{R}\right)^{\pi} = \zeta^j \cdot \sqrt[5]{R}$$

gilt. Zusammen mit der vorhergehenden Rechnung erhalten wir somit die Relation

$$\begin{aligned} &q_0 + q_1 \cdot x_2 + q_2 \cdot x_2^2 + q_3 \cdot x_2^3 + q_4 \cdot x_2^4 = \\ &\zeta^j \cdot q_0 + \zeta^j \cdot q_1 \cdot x_1 + \zeta^j \cdot q_2 \cdot x_1^2 + \zeta^j \cdot q_3 \cdot x_1^3 + \zeta^j \cdot q_4 \cdot x_1^4 \,. \end{aligned}$$

Zusammengenommen genügen $x_1, \ldots, x_5$ somit der polynomialen Gleichung

$$\begin{aligned} &q_0 \,(1 - \zeta^j) + q_1 \,(x_2 - \zeta^j \cdot x_1) + q_2 \,(x_2^2 - \zeta^j \cdot x_1^2) + \\ &q_3 \,(x_2^3 - \zeta^j \cdot x_1^3) + q_4 \,(x_2^4 - \zeta^j \cdot x_1^4) = 0 \,. \end{aligned}$$

Dies stellt aber einen **Widerspruch** zur angenommenen algebraischen Unabhängigkeit der Variablen $x_1, \ldots, x_5$ dar.

Damit ist – zumindest im Fall $m = 5$ – gezeigt, dass die allgemeine Gleichung fünften Grades sich nicht mithilfe von Radikalen lösen lässt.