

"Mathematik ist wie Sauerstoff. Wenn sie da ist, bemerkt man es nicht. Wenn sie nicht da wäre, könnte man nicht ohne sie leben."

Erfolgsformeln

Anwendungen der Mathematik

Matthias Ehrhardt

Michael Günther

Wil Schilders

Erfolgsformeln

Anwendungen der Mathematik

Matthias Ehrhardt

Michael Günther

Wil Schilders

September 2021

Inhaltsverzeichnis

Interviews:

Top-Leute zur angewandten Mathematik

Interview mit Dr. Thomas Hahn	6
Interview mit Dr. Markus Hoschek	44
Interview mit Dr. Margrit Klitz	68
Interview mit Dr. Michael Meister	81
Interview mit Prof. Dr.-Ing. Stefan Kurz	102
Interview mit Prof. Dr. Anita Schöbel	135
Interview mit Prof. Dr. Helga Rübsamen-Schaeff	148
Interview mit Prof. Dr. Volker Mehrmann	170
Interview mit Prof. Dr. Andreas Schuppert	196

Vorwort

5

Epidemiologie

Mathematik zur Vorhersage der Ausbreitung einer ansteckenden Krankheit	8
Die Verlässlichkeit von COVID-19 Schnelltests	11
Robuste Inzidenzzahlen	16
Mathematische Modellierung und optimale Kontrolle von Epidemien	20
Mathematische Modellierung von Impfprogrammen	21
Mit Mathematik zur individuellen Therapie	29

Medizin

Intelligenterer Planung macht das Gesundheitswesen effizienter und sicherer	33
Bessere Bilder	36
Der Kopf isst mit	38
Die optimale Dosisfindung von Antibiotika	40

Chemie

Mathematik in der Chemie: Die Suche nach Synthesewegen für neue Medikamente	46
Quasikristalle	49

Verkehr

Best-Price Ticketingsysteme: Erst fahren, dann zahlen!	50
Projekt Ride-Hailing Wuppertal	54
Das E-Bike und sein Antrieb	57
Die überraschende Verzögerung	59
Wie man Quietschen und Schwingungen bei Bremsen unterdrückt	61
Der Krankenwagen trifft in 15 Minuten ein	64
Optimales Boarding am Flughafen	66

Umwelt

Mathematik in der Modellierung zur Systemanalyse	70
Klimawandel: Die optimale CO ₂ -Steuer	72

Energie

Layout großer Photovoltaikkraftwerke	74
Optimierte Auslegung von Gasheizgeräten	77
Statistik bestimmt die Wahl der Windturbine	79

Gesellschaft

Zuwanderung und Integration in sich verändernden Gesellschaften	84
Mathematische Modellierung von Radikalisierungsprozessen	89
Faire Schulplatzvergabe: von Boston in die ganze Welt	93
Ausfallwahrscheinlichkeiten für Verschwörungstheorien	96

Kommunikation

Schnelle Satelliten-Kommunikation mit π	99
Verkehrsknoten von digitalen Datenpaketen	100

Sicherheit

Die Mathematik des Drängelns	104
Mit mathematischer Optimierung zu sicheren und klimafreundlichen Flügen	106
Weniger Flugzeugabstürze aufgrund besserer Wahrscheinlichkeitsmodelle	110
Digitale Schutzmechanismen knacken, um sie sicherer zu machen	112
Muster in PIN-Codes	114
Das Benfordsche Gesetz – Wie Mathematik Zahlenfälscher entlarvt	115

Künstliche Intelligenz

Wie funktioniert Autonomes Fahren ?	117
Mathematische Architekturen für Neuronale Netze	122
Wie soziale Netzwerke bei der Behebung von Softwarefehlern helfen	127
Die Mathematik hinter der Filmauswahl	130
Sprachen durch Zählen von Wörtern bändigen	132

Produktion

Der Digitale Zwilling – Den Betrieb von Anlagen neu denken	136
--	-----

Finanzmathematik

Mit Mathematik die Finanzmärkte zähmen	139
Modellierung Negativer Zinsen – Nur Bares ist Wahres?	141
Beipackzettel für Finanzinstrumente: PRIIPs!	145

Hightech

Berechnung von Kamerabildern	150
Der Heilige Gral für Software	152

Kriminologie

Verbrechen durch Mathematik verstehen	154
Das Haus des Täters: Was ist Geographic Profiling?	158
Wie gross sind die Chancen, dass die DNA-Spur des Verdächtigen gefunden wird?	160

Musik

Wie erkennt Shazam ein Lied so schnell?	164
Wie fair ist der Gesangswettbewerb?	166

Kunst

Origami-Mathematik im Museum für Moderne Kunst	172
Liebeslyrik: Petrarca und Laura – Chaos in Liebesaffären	174
Der Zahn der Zeit	178

Sport

Wie man beim Beachvolleyball öfter gewinnt	181
Daten- und modellgestützte Talentförderung von Nachwuchsfußballern	186
Der optimale Freiwurf im Basketball	191
Das Geheimnis hinter einem erfolgreichen Endspurt im Radsport	194

Games

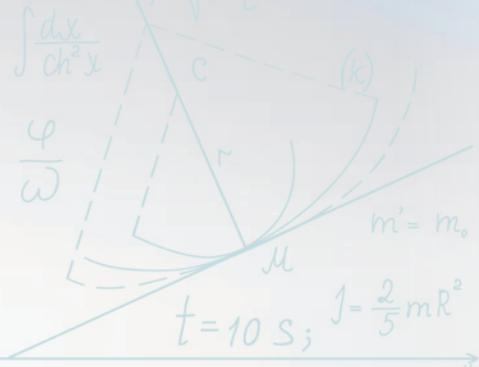
Cyberpunk 2077: Die Mathematik hinter der Computergrafik	198
Sudoku – Spaß, aber auch ernsthafte Mathematik	201
Deal or no Deal?	202

Danksagungen

204

$$= \pi \rho \left[\int_0^R R^3 dz - 2 \int_0^R R^2 x^2 dz \right] + \int_0^R x^3 a \left[\frac{dx}{\cos x} \right] \left[\frac{1+n^2 x^2}{e^2 - m^2} + \frac{1+(n+1)^2 x^2}{e^2 - m^2} \right]$$

$$M = 0, \quad \pi \rho \left[R^5 - \frac{2}{3} R^5 + \frac{1}{5} R^5 \right] + \frac{1}{15} = \rho R^5 \frac{1}{2} m A^2 \frac{\pi^2}{16} = 250 \text{ J}$$



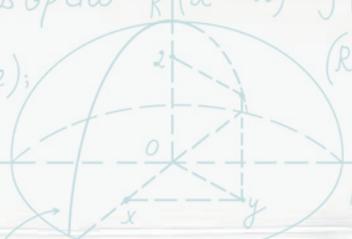
$$\int \frac{dx}{\sqrt{x^2 + 2x + 3}}$$

$$\frac{1}{\sqrt{3}} \int r' \cos \theta \rho d\omega$$

$$\frac{x}{(x^2+a)^2} = \frac{2(n+1)x^2}{(x^2+a)^{n+1}} = \frac{x}{(x^2+a)^n}$$

$$R^2 - z^2 = \frac{2}{5} \rho \pi R^5$$

$$S = ? \quad 2 \int t i + j \frac{2}{3} t$$



$$F_z = m' \frac{dz}{dt} = m' \frac{dz}{dt}$$

$$\frac{d\theta}{dt} F_z = \frac{1}{h} \sum m \vartheta_z^2 D$$

$$S = x_2 A - RT \ln \frac{V_2}{V_1} F_m = \frac{\mu_0}{4\pi} \frac{g_1 \vartheta_1 g_2 \vartheta_2}{r^2} \frac{1}{m'}$$

$$+20\% \quad +15\%$$

$$2 \frac{d\vartheta}{dt} \frac{g_1 \vartheta_1 g_2 \vartheta_2}{r^2} = \frac{g_1 \vartheta_1 g_2 \vartheta_2}{r^2} \frac{1}{m'}$$

$$\frac{1}{3} \alpha R^3 h \frac{1}{3} \pi r^2 \quad \vartheta_0 = 2i + k$$



$$-10\% \quad -10\%$$

$$\frac{1}{3} \rho \int r^3 \frac{1}{2} (3 \cos^2 \theta - 1/2) \rho dt \vartheta' \quad Q_{1,2} = Q \vartheta_{1,2} \quad A_{1,2} = \vartheta RT_2 \ln \frac{V_2}{V_1} - \vartheta RT_1 \ln \frac{V_1}{V_2}$$

$$F = \frac{\mu_0}{4\pi} \frac{g_1 \vartheta_1 g_2 \vartheta_2}{r^2} \quad F = \frac{\mu_0}{4\pi} \frac{g_1 \vartheta_1 g_2 \vartheta_2}{r^2} \quad F = \frac{\mu_0}{4\pi} \frac{g_1 \vartheta_1 g_2 \vartheta_2}{r^2}$$

Vorwort

Dieses Buch soll ein wenig von der Mathematik zeigen, der wir jeden Tag begegnen. Es stellt sich heraus, dass Mathematik im Sport, in der Medizin und in der Wettervorhersage, Verkehr, vorhanden ist und sogar zur Lösung von Mordfällen beitragen kann. Das Buch zeigt auch, dass die Mathematik in der Wirtschaft wichtig ist, wo sie oft unsichtbare Beiträge zu sichtbaren Erfolgen leistet. Beispielsweise trägt die Mathematik zu Innovationen in allen Spitzensektoren bei, sie ist eine sog. Schlüsseltechnologie. Darüber hinaus sprechen neun prominente Vertreter über die Rolle der Mathematik in ihrer Arbeit und in ihrem Privatleben.

*„Wir wenden täglich Mathematik an.
Um das Wetter vorherzusagen, Zeit zu messen,
Geldgeschäfte abzuwickeln. Mathematik ist
mehr als nur Formeln und Gleichungen. Sie ist
Vernunft. Mit ihr kann unser Verstand die
größten Rätsel lösen, die wir kennen.“*

(OFF-TEXT CHARLIE EPPES IN TV-SERIE NUMB3RS)

In diesem Buch wollen wir demonstrieren, dass Mathematik viel mehr ist als pures Rechnen, einfaches Einsetzen von Werten in Formeln, Lösen von Knobelaufgaben oder etwa Kurvendiskussion. Mathematik ist in vielen Bereichen unseres Lebens enthalten: sie modelliert und simuliert den Verlauf der Corona-Pandemie, sie beschreibt Auswirkungen

und Effektivität von Impfstrategien, sie steuert Roboter, sie optimiert Produktionsabläufe und Verkehrsströme und vieles mehr.

Wenn Sie mehr über dieses Buch wissen möchten, zögern Sie bitte nicht, uns zu kontaktieren. Auf der Website www.erfolgsformeln.uni-wuppertal.de finden Sie einen Überblick über die Mathematiker, die zu diesem Buch beigetragen haben, sowie weitere Literatur zu den einzelnen Kapiteln dieses Buchs. Wir haben den einzelnen Kapiteln die drei Schwierigkeitsgrade leicht '*',mittel '**', schwer '***' zugeordnet.

Die Idee zu diesem Buch ist entstanden aus einer Broschüre 'Succesformules' von niederländischen Mathematikern unter der Leitung von Wil Schilders, der im Wintersemester 2020/21 Mittelsten-Scheid Gastprofessor an der Bergischen Universität war. An dieser Stelle möchten wir uns ganz herzlich für die Unterstützung durch die Bergische Universität Wuppertal bedanken.

Wir wünschen Ihnen viel Spaß beim Lesen der Kapitel und Interviews in diesem Buch.

Wuppertal, September 2021

Matthias Ehrhardt
Michael Günther
Wil Schilders

Dr. Thomas Hahn

Eine kurze Zusammenfassung Ihrer Karriere

Dr. Thomas Hahn ist seit 2011 Chief Expert Software bei der Siemens AG. Nach dem Studium der Informatik an der Friedrich-Alexander-Universität Erlangen trat er 1986 in das Unternehmen ein. Thomas Hahn hatte verschiedene Aufgaben im Bereich der Fertigung wie z.B. als Leiter der Softwareentwicklung für industrielle Automatisierungssysteme oder im Bereich der Mobilität wie z.B. als Entwicklungsleiter für Verkehrsleitsysteme. Zusätzlich zu seiner aktuellen Position als Chief Expert Software war er von 2011 bis August 2013 Leiter des Bereichs Business Analytics and Monitoring.

Über seine Tätigkeit bei Siemens hinaus ist Thomas Hahn Mitglied bzw. Vorstandsmitglied in verschiedenen Gremien, unter anderem im Openlab CERN und im Lenkungsausschuss Plattform Industrie 4.0. Er ist Vizepräsident der OPC Foundation, stellvertretender Vorsitzender von GAIA-X AISBL, Vorsitzender des Labs Network Industrie 4.0, Vorstandsmitglied des bayerischen „KI-Rat“ und Präsident der Big Data Value Association.



Was hat Mathematik für Ihre Karriere bedeutet?

Mathematik als Grundlagenwissenschaft war mir insofern wichtig, als sie hilft, strukturiert und logisch zu denken und zu argumentieren. Und das sind wichtige Eigenschaften, die man in der Wirtschaft beherrschen muss, auch wenn es überhaupt nicht um mathematische Zusammenhänge geht, wie beispielsweise Geschäftsmodelle, Entwicklungsprozesse oder Software-Architekturen. Man muss andere von seinen eigenen Ideen überzeugen oder Ideen Anderer kritisch reflektieren, und eine systematische Vorgehensweise, wie man sie in der Mathematik gelernt hat, hilft hier einfach enorm.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihrer Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Bei Siemens spielt die Mathematik unterschiedliche Rollen: Zum ersten gibt es einige Mathematiker bei Siemens, die in Ihren Gebieten absolute Experten sind und in der weltweiten Community auch entsprechend anerkannt sind. Sie agieren hier meistens als Treiber für dieses Gebiet aus einer gewissen Anwendungsperspektive heraus.

Außerdem gibt es viele Ingenieure bei Siemens, die Aufgaben durchführen, die ein fundiertes mathematisches Verständnis erfordern, beispielsweise wenn es um die Entwicklung spezifischer Simulationsmodelle oder Simulationsrechnungen geht. Hier ist dann neben dem methodischen Grundverständnis ein Verständnis sowohl der eingesetzten Werkzeuge als auch des Anwendungskontextes notwendig.

Und schließlich gibt es Mathematiker, die mit dem Einstieg in die Firma ihre Beschäftigung mit der Mathematik als solche beendet haben, aber das methodische Wissen, das sie sich im Studium angeeignet haben, jetzt in der Praxis einsetzen.

Sind, in Ihrer Hinsicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Ich möchte diese Frage gerne nutzen, um grundsätzlich auf die Brücke zwischen Universität und Wirtschaft einzugehen. Das, was ich sage, trifft nicht nur auf die Mathematik, sondern auch die Naturwissenschaften und Ingenieurwissenschaften zu. In der Wirtschaft ist es äußerst wichtig, dass man wirtschaftliche Zusammenhänge versteht. Das betrifft nicht nur betriebswirtschaftliche oder volkswirtschaftliche Zusammenhänge, sondern insbesondere auch

die Auswirkungen von technischen Entscheidungen auf das Geschäft. Man kann lange und im Detail über technische Lösungen diskutieren, aber letztlich wird man als Firma mit einer technischen Lösung nur dann Erfolg haben, wenn man einem Kunden den Nutzen erklären kann, den ihm die technische Lösung stiftet, und was das Alleinstellungsmerkmal der technischen Lösung gegenüber Wettbewerbern ist. Diese Kenntnis wird aus meiner Sicht an den Universitäten zu wenig vermittelt. Ich sehe hier aber auch die Wirtschaft in der Pflicht, sich geeignet in den Lehrbetrieb an den Universitäten einzubringen, um den Studenten solche Zusammenhänge – am besten dann auch illustriert an konkreten Beispielen – zu vermitteln.

Wenn man sich das Verhältnis zwischen PraktikantInnen bzw. WerkstudentInnen bei uns im Hause und der Anzahl der Studierenden insgesamt anschaut, wäre es insbesondere toll, wenn noch mehr Mathematiker bereits im Studium die Chance nutzen würden, praxisnahe Erfahrungen zu sammeln.

Was raten Sie jungen Leuten hinsichtlich Mathematik?

Mein erster Hinweis an junge Leute ist immer, dass sie unterscheiden müssen zwischen „Rechnen“, das man in der Grundschule und den ersten Klassen der weiterführenden Schulen lernt, und „Mathematik“, mit der man eigentlich erst in den letzten Schulklassen, aber spätestens im Studium konfrontiert wird. Für einige ist das erste Semester Mathematik an der Universität einfach ein Kulturschock.

Mein zweiter Hinweis ist, dass man unterscheiden muss zwischen der Mathematik als Theorie und der Mathematik als Anwendung in verschiedenen Domänen. In der Wirtschaft ist überwiegend die Anwendung der Mathematik von Bedeutung. Anwendung bedeutet hierbei oft die Realisierung von mathematischen Algorithmen in Software, insofern kann ich jedem nur empfehlen, frühzeitig programmieren zu lernen und gerade in diesem Bereich möglichst viel Erfahrung zu sammeln.

Und mein letzter Hinweis ist, dass man in der Mathematik auch nicht immer alles verstehen muss. Ich habe selbst als studentische Hilfskraft ein Mathematikbuch Korrektur gelesen, und einen Beweis darin habe ich bis heute nicht verstanden, obwohl ich damals die Argumentationskette mehrfach mit dem Professor durchgegangen bin. Nun, ich habe auch ohne Verständnis dieses Beweises meinen Weg gefunden.



Die Ausbreitung der aktuellen COVID-19 Pandemie lässt sich mit einem fast 100 Jahre alten Modell beschreiben, das auf Differentialgleichungen beruht.

Mathematik zur Vorhersage der Ausbreitung einer ansteckenden Krankheit

Frau Dr. Maria Vittoria Barbarossa, Gruppenleiterin am Frankfurt Institute for Advanced Studies, erinnert sich an die Anfänge der COVID-19 Pandemie:

Als ich diesen Beitrag schreibe ist genau ein Jahr vergangen seitdem ich als Mathematikerin meine Arbeit zur Bekämpfung der Pandemie aufgenommen habe. SARS-CoV-2 hatte angefangen sich rasch durch Europa zu verbreiten und in Deutschland gab es um die 400 festgestellte Infektionen, dazu waren eine Handvoll Menschen verstorben. Ich erinnere mich an das Telefonat mit meinem Institutsdirektor, der sagte in etwa 'wir stehen vor einer Europäischen Krise, wir müssen diesen Virus loswerden, Du kannst die Ausbreitung von Epidemien modellieren – tu irgendetwas!'. Es war Anfang März 2020 als wir mit Kollegen verteilt zwischen Heidelberg, München, Jülich und Frankfurt – alle in Homeoffice – anfangen, die damals noch wenig organisierten Informationen und Daten zum neuen Coronavirus einzusammeln. Wir haben in sehr kurzer Zeit mathematische Modelle aufgestellt, um anhand der ersten Daten Szenarien zur Virusaus-

breitung darzustellen und Prognosen für die Auswirkung möglicher Kontrollmaßnahmen (damals hatten wir noch keinen Impfstoff zur Verfügung!) zu machen. Im Lauf des Jahres wurden die Modelle immer weiter entwickelt und verfeinert, um neue Erkenntnisse und Daten zu integrieren, sowie um neue Fragen zu beantworten – zum Beispiel zu Virusvarianten oder zu der optimalen Verteilung der Impfstoffe.“

Mathematische Modellierung mit dem SIR-Modell

„Vielleicht gehe ich besser erstmals einen Schritt zurück und erkläre überhaupt, was es heißt ein mathematisches Modell für die Ausbreitung einer ansteckenden Krankheit aufzustellen. Lasst uns an eine Krankheit denken, z.B. die klassische Grippe, verursacht durch einen saisonalen Grippevirus, der von Person zu Person übertragen wird. Wenn wir ein mathematisches Modell bilden, sollten wir uns auch überlegen, welche Fragen wir durch das Modell beantworten wollen. Zum Beispiel, anhand von Daten zum Verlauf einer Grippe-

welle in den vergangenen Jahren möchten wir vorhersagen, wie die Grippewelle in kommenden Herbst verlaufen wird. Zu jedem Zeitpunkt während der Grippewelle möchten wir z.B. die Anzahl der infizierten Personen kennen. Wir wollen auch wissen, wie viele aus der Krankheit genesen sind, weil diese Menschen nach der Infektion volle Immunität gegen die Krankheit entwickelt haben. Der Einfachheit halber unterscheiden wir erstmal nicht zwischen Personen, die genesen sind, und Personen, die an der Krankheit sterben – auch diese können sich nicht mehr mit der Krankheit infizieren und können auch keine weitere Person anstecken. Wir betrachten dann eine geographische Region und nehmen an, dass die Bewegung in diese Region hinein und aus dieser heraus vernachlässigt werden kann. Wir vernachlässigen insbesondere Geburten und Todesfälle für die Zeit, die wir beobachten wollen. Dann sollten wir die wesentlichen Modellkomponenten (Zustände und Prozesse) definieren, die wir unbedingt brauchen. Die (konstante) Gesamtbevölkerung N in der betrachteten Region wird dann in drei verschiedenen Gruppen unterteilt:

- Die anfälligen Individuen (S), d.h. diejenigen, die keinen Schutz gegen die Krankheit haben und angesteckt werden können,
- Die infizierten Individuen (I), die krank und ansteckend sind, und
- Die genesenen Individuen (R), die wieder gesund (oder verstorben) sind und weder anstecken noch angesteckt werden können.

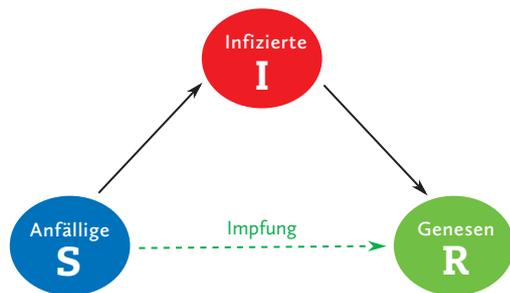


Abbildung 10: Schematische Darstellung des SIR-Modells mit Übergängen zwischen der anfälligen (S), infizierten (I) und genesenen (R) Gruppe der Bevölkerung.

Die Änderung der anfälligen Population ist zu jedem Zeitpunkt t durch die Ableitung $S'(t)$ gegeben. Nehmen wir an, jedes infizierte Individuum hat eine feste Anzahl κ an Kontakten pro Tag, und dass die Wahrscheinlichkeit bei einem Kontakt die Infektion zu übertragen ρ sei. Wenn wir eine homogene Vermischung der Bevölkerung annehmen, dann ist für jede anfällige Person die Wahrscheinlichkeit sich

an jedem Tag zu infizieren gegeben durch $\beta I(t)/N$, wobei $\beta = \kappa\rho$. Die neu infizierten Individuen dürfen sich nicht mehr in der anfälligen Gruppe S aufhalten und werden stattdessen dazu beitragen, dass die infizierte Population I zunimmt. Damit haben wir, dass

$$S'(t) = -\frac{\beta S(t)I(t)}{N}.$$

Die infizierten Personen $I(t)$ genesen mit einer Rate γ : das bedeutet, die durchschnittliche Infektionsdauer ist $1/\gamma$ Tage, also

$$I'(t) = \frac{\beta S(t)I(t)}{N} - \gamma I(t).$$

Die Änderungsrate der genesenen Gruppe $R(t)$ ist entsprechend

$$R'(t) = \gamma I(t).$$

Diese drei Differentialgleichungen (also Gleichungen welche die Änderungsraten der beobachteten Zustände S , I und R beschreiben) bilden das bekannte *SIR-Modell*, das vor fast 100 Jahren zum ersten Mal vorgeschlagen wurde. Mithilfe numerischer Methoden kann man das Modell am Computer simulieren und damit z.B. die Anzahl der infizierten Personen in der Zeit abbilden. Dabei sehen wir auch den typischen Verlauf einer saisonalen Grippewelle, die mit wenigen infizierten Personen anfängt. Die Anzahl der Infizierten (und kurz drauf die der Genesenen) nimmt in der Zeit zu, dabei nimmt die Anzahl der anfälligen Personen ab, bis zu dem Zeitpunkt, wo ein Peak in der Zahl der Infektionen erreicht wird, und die Epidemie langsam ausläuft (siehe Abbildung 11).“

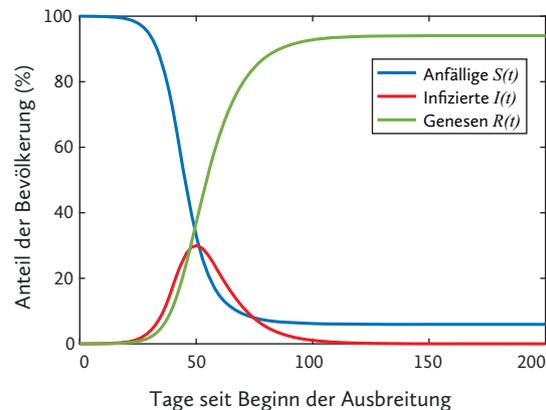


Abbildung 11: Typischer Verlauf einer Infektionswelle nach dem SIR-Modell.

Die Basisreproduktionszahl

„Außerdem lässt sich aus diesem Modell die *Basisreproduktionszahl*

$$R_0 = \frac{\beta}{\gamma}$$

bestimmen. Diese Zahl sagt uns, wie viele Personen im Durchschnitt ein infizierter Mensch anstecken würde, wenn dieser in einer völlig anfälligen Population eintrifft. Die Basisreproduktionszahl variiert von Krankheit zu Krankheit (sogar von Modell zu Modell), ist z.B. zwischen 2 und 3 für saisonale Grippe und kann so hoch wie 15-18 bei Masern sein. Diese Zahl ist besonders wichtig, denn durch $1 - 1/R_0$ lässt sich derjenige Anteil der Bevölkerung abschätzen, der für die sogenannte *Herdenimmunität* nötig ist. Wenn wir die Gruppe der infizierten Personen in der Zeit für unterschiedliche Werte von R_0 abbilden, sehen wir, dass eine Reduktion von R_0 zu Kurven führt, die einen niedrigen und verschobenen Peak haben (siehe Abbildung 12)“.

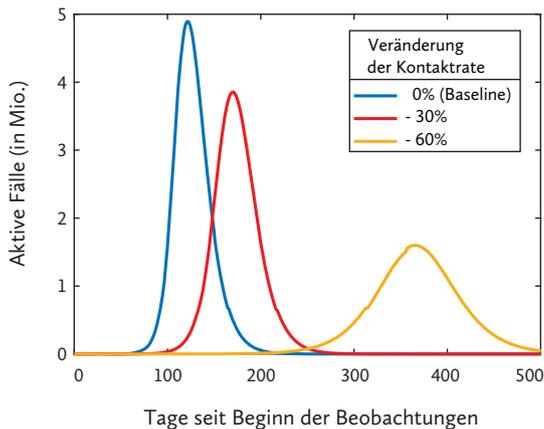


Abbildung 12: Auswirkung der Reduktion der Kontaktrate (oder äquivalent Reduktion von R_0) auf die Anzahl der aktiven Fälle. Eine niedrige Kontaktrate führt zu einem niedrigen und verschobenen Peak.

Erweiterungen des SIR-Modells

Zurück zur COVID-19 Modellierung: Frau Barbarossa erklärt: „Im wesentlichen kann man auch ein relativ einfaches Modell wie das SIR verwenden, um den Verlauf der Pandemie zu beschreiben. In der Praxis möchte man allerdings mehr Informationen aus dem Modell gewinnen, so dass man eher mit einer erweiterten Version des SIR-Modells arbeitet, wo man auch die Möglichkeit hat z.B. nicht-erkannte Infektionen (weil mild, asymptomatisch, oder einfach nie getestet worden), hospitalisierte Fälle, Personen die einen Aufenthalt auf der Intensivstation bedürfen und auch solche die an der Krankheit sterben, zu verfolgen. Schauen wir wieder auf Abbildung 12 und denken zurück an die ergriffenen Kontrollmaßnahmen, die immer wieder zur Kontaktreduktion im Lauf der COVID-19 Pandemie führten, dann verstehen wir was man damit erreichen wollte: die Anzahl der Infektionen möglichst lange Zeit auf niedrigen Werte belassen, damit möglichst alle kranken Personen medizinisch gut versorgt werden konnten. Auch die Wirkung eines Impfstoffs kann man in einem solchen Modell integrieren (vgl. Abbildung 10). Mathematisch gesehen verschiebt man durch die Impfung Menschen von der anfälligen Gruppe S zu einer Gruppe der geimpften Personen, die sich im besten Fall so gut wie die genesene Gruppe R verhält, also nicht zur Verbreitung der Krankheit beiträgt.“

Frau Dr. Maria Vittoria Barbarossa erinnert sich an ihre Zeit als Studentin: „Als ich im Herbst 2003 als Erstsemesterin in den Hörsäle der Fakultät für Mathematik saß, hatte ich keinen Plan was ich im Studium gelernt hätte und vor allem was ich nach dem Bachelorabschluss machen würde. Im 5. Semester durfte ich als Erasmus Student an der TU München studieren, wo ich zum ersten Mal von Anwendungen der Mathematik in der Biologie und Lebenswissenschaften hörte. Das hat mich fasziniert, ich habe das Thema nicht mehr lassen wollen.“

Schnell konnte ich begreifen, dass die Anwendungen in sehr unterschiedlichen Richtungen gehen können. Um eine Vorstellung zu geben, in den Jahren seit meiner Masterarbeit durfte ich mit Mikrobiologen, Biochemikern, Immunologen, Epidemiologen und Intensivmediziner zusammenarbeiten. Die Herausforderung dabei ist nicht so sehr die Mathematik. Viel mehr finde ich es spannend, mich in andere Fächer einzuarbeiten und dabei zu lernen, eine gemeinsame Sprache zu definieren, um mich mit Kollegen aus anderen Fachrichtungen auszutauschen.“

Die Verlässlichkeit von COVID-19 Schnelltests

Ein mathematischer Satz aus dem 18. Jahrhundert hilft uns die Aussagekraft von COVID-19 Schnelltests zu untersuchen.



COVID-19 Schnelltests, auch „Lateral-Flow-Tests“ (LFTs) genannt, werden eingesetzt, um Mitarbeiter vor Arbeitsbeginn zu testen, Besuche in Schulen, Geschäften, Restaurants, Theatern zu ermöglichen oder 'körpernahe Dienstleistungen', wie etwa Friseure, in Anspruch nehmen zu können.

Aber wie sicher ist so ein Test? Wie stark kann man sich darauf verlassen? Nehmen wir dazu ein Beispiel vom Robert Koch-Institut¹. In einer Niedrig-Inzidenz Situation (5 von 10.000 Getesteten sind tatsächlich infiziert) machen alle 10.000 Menschen einen Test und 200 fallen positiv aus, obwohl die Person nicht infiziert ist. Man spricht dann von einer 'falsch-positiv' Rate von 200/10.000, also 2 %. Wenn also der Schnelltest, nur zweimal in 100 Fällen ein falsches Positiv liefert, wie groß ist dann die Chance, dass man bei einem positiven Test tatsächlich infiziert ist? Ist 98 % richtig? Die Frage war etwas gemein, weil dem Getesteten etwas Hintergrundinformation fehlt, ohne die, wie wir sehen werden, diese Frage nicht beantwortet werden kann.

Bei der Beantwortung hilft uns ein mathematischer Satz aus dem 18. Jahrhundert. Der Satz von Bayes (auch: Formel von Bayes), benannt nach dem englischen Mathematiker Thomas Bayes (1701-1761), dient der Berechnung *bedingter Wahrscheinlichkeiten*. Die Formel von Bayes lautet in mathematischer Notation

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Hierbei ist $P(A|B)$ die (bedingte) Wahrscheinlichkeit des Ereignisses A unter der Bedingung, dass B eingetreten ist.

In unserem Fall (siehe RKI-Quelle, linke Seite), werden 10.000 Menschen getestet, die nach dem Zufallsprinzip ausgewählt wurden. Es sind 5 von 10.000 Personen mit COVID-19 infiziert. Man spricht hier von der „Vorwahrscheinlichkeit“: die Prävalenz oder die Hintergrundrate in der Bevölkerung, sie ist hier 0,05 %. Der Schnelltest identifiziert 4 von ihnen korrekt. Es gibt 9.995 nicht infizierte Menschen und der Schnelltest identifiziert 9.795 von ihnen korrekt.

Eigenschaften des Schnelltest: Wenn man infiziert ist, wird er dies in $4/5=80\%$ der Fälle richtig sagen; wenn man nicht infiziert ist, wird er dies in $9.795/9.995$, also ca. 98 % der Fälle richtig sagen. Aber das bedeutet auch, dass der Schnelltest, obwohl er in 80 % der Fälle die richtige

Antwort gibt, 200 Menschen gesagt hat, dass sie infiziert sind, obwohl sie nicht infiziert sind.

Zurück zur Mathematik und dem Satz von Bayes. Als Ereignisse definieren wir A : infiziert und B : positiver Test und erhalten die Wahrscheinlichkeiten

$$P(A) = 5/10.000, \quad P(B) = 204/10.000, \quad P(B|A) = 4/5.$$

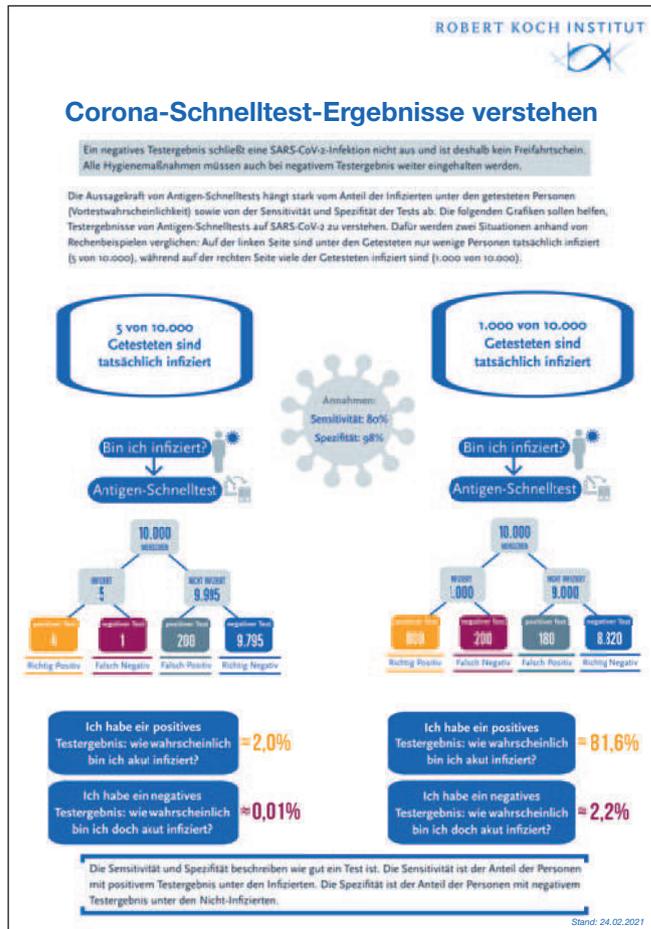
und wir erhalten

$$P(A|B) = \frac{\frac{4}{5} \cdot \frac{5}{10.000}}{\frac{204}{10.000}} = \frac{4 \cdot 5}{5 \cdot 204} = \frac{2}{102} \approx 2\%.$$

Wenn man also ein positives Ergebnis erhält, liegt in diesem Fall die Chance, tatsächlich infiziert zu sein, nur bei ca. 2 %. Ohne die Vorwahrscheinlichkeit zu kennen, weiß man nicht, wie wahrscheinlich es ist, dass ein Ergebnis falsch oder richtig ist. Als Übung könnte man das Ergebnis von der RKI-Quelle, rechte Seite, bestimmen; bei dieser Hoch-Inzidenzlage ist jede 10. Person infiziert und es ergeben sich ganz andere Zahlen!

Diese Zahlen gelten natürlich nur, wenn man die Bevölkerung wirklich nach dem Zufallsprinzip getestet wird. Benutzen die Menschen die Schnelltests, weil sie einen Verdacht haben (z.B. weil sie Symptome aufweisen), dann wäre auch die Vorwahrscheinlichkeit entsprechend höher.

¹ Robert-Koch Institut, Infografik: Corona-Schnelltest-Ergebnisse verstehen, 25. Feb. 2021. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Infografik_Antigentest_PDF.pdf, Seiten 47–49



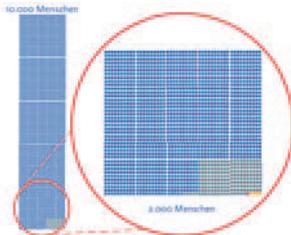
Infografik Corona-Schnelltest, Seite 1 (Quelle: Robert Koch-Institut)



Corona-Schnelltest-Ergebnisse verstehen

10.000 Testergebnisse

5 von 10.000 Getesteten sind tatsächlich infiziert



Testergebnisse einordnen

Positive Tests (204), davon

● infiziert: 4 (Richtig Positiv)
● nicht infiziert: 200 (Falsch Positiv)
Wahrscheinlichkeit, dass ein positiver Test richtig liegt (4/204) = 1,96%

Negative Tests (9.796), davon

● nicht infiziert: 9.795 (Richtig Negativ)
● infiziert: 1 (Falsch Negativ)
Wahrscheinlichkeit, dass ein negativer Test richtig liegt (9.795/9.796) = 99,99%



10.000 Testergebnisse

1.000 von 10.000 Getesteten sind tatsächlich infiziert



Testergebnisse einordnen

Positive Tests (980), davon

● infiziert: 800 (Richtig Positiv)
● nicht infiziert: 180 (Falsch Positiv)
Wahrscheinlichkeit, dass ein positive Test richtig liegt (800/980) = 81,6%

Negative Tests (9.020), davon

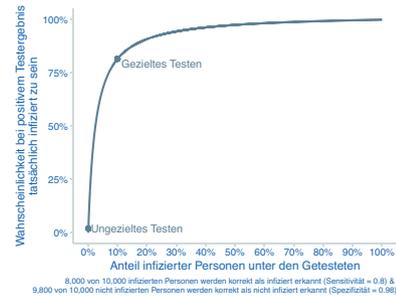
● nicht infiziert: 8.800 (Richtig Negativ)
● infiziert: 200 (Falsch Negativ)
Wahrscheinlichkeit, dass ein negativer Test richtig liegt (8.800/9.020) = 97,6%



Stand: 24.02.2021



Corona-Schnelltest-Ergebnisse verstehen



Effektive Testansätze stehen im Zentrum der Bekämpfung von SARS-CoV-2. Für den Virusnachweis werden eine Vielzahl von Antigen-Schnelltests angeboten. Diese basieren auf dem Nachweis von viralem Protein in Abstrichen aus den Atemwegen. Antigen-Schnelltests können schneller durchgeführt werden als PCR-Tests. Im Vergleich zur PCR erkennen diese Schnelltests jedoch sowohl infizierte Personen schlechter (niedrigere Sensitivität) als auch nicht-infizierte Personen schlechter (niedrigere Spezifität). Die Aussagekraft von Antigen-Schnelltests hängt stark vom Anteil der Infizierten unter den getesteten Personen (Vor-testwahrscheinlichkeit) sowie von der Sensitivität und Spezifität der Tests ab. Die Rechenbeispiele oben illustrieren den Zusammenhang zwischen dem Anteil der infizierten unter den Getesteten, der Sensitivität und Spezifität der Tests und den resultierenden positiven und negativen Vorhersagewerten. Die angenommenen Werte für die Sensitivität und Spezifität der Tests sind großzügig angelegt. Der positive Vorhersagewert beziffert die Wahrscheinlichkeit, dass eine Person infiziert ist, wenn sie positiv getestet wurde. Der negative Vorhersagewert beziffert die Wahrscheinlichkeit, dass eine Person nicht infiziert ist, wenn sie negativ getestet wurde. Wenn unter den Getesteten nur wenige Personen tatsächlich infiziert sind, dann sind positive Testresultate unzuverlässig. Wenn unter den Getesteten allerdings sehr viele Personen infiziert sind, dann sind positive Testresultate zuverlässig, aber die negativen Testresultate dafür weniger. Die Aussagekraft der Tests hängt vom Testansatz und der Verbreitung des Virus ab.

Referenzen

- McDonnell, M., & Jacobs, P. (2012). Meta-analysis of the effect of natural frequency on Bayesian reasoning. *Psychological Bulletin*, 142(1), 127-139. <https://doi.org/10.1037/a0028014>
- Siefried, J., Bötcher, S., Ott, D.T., Michel, J., Weische, A., Jemp, N.A., Wierle, L.H., Auda, E.M., Jung-Sendzik, T., Dürmalk, E., Diercke, M., Haas, W., Abu-Sin, M., Schumann, T., Harmsch, D., & Müller, M. (2021). Was ist bei Antigenrass zur Eigenreinscheidung (Selbsttest) zum Nachweis von SARS-CoV-2 zu beachten? *Epidemiologisches Bulletin*, 0221(1)9. <http://dx.doi.org/10.29344/EB022109>

Stand: 24.02.2021

Robuste Inzidenzzahlen

Mit Mathematik kann man die Inzidenz stabilisieren, allerdings muss man auf Transparenz und Verständlichkeit achten, um eine breite Akzeptanz in der Bevölkerung zu bekommen.

Bei der Steuerung der (nicht-pharmazeutischen) Maßnahmen zur Bekämpfung der COVID-19 Pandemie wird standardmäßig als Orientierung die *7-Tage Inzidenz* der bestätigten Neuinfektionen (Summe der letzten 7 Tage, pro 100.000 Einwohner, auf Grundlage von Meldedaten im Zeitraum ab dem Berichtstag)²

$$\text{7-Tage Inzidenz} = \frac{\text{Anzahl der bestätigten Neuinfektionen der letzten 7 Tage}}{\frac{\text{betrachtete Einwohnerzahl}}{100.000}}$$

verwendet. Es werden also alle gemeldeten Neuinfektionen der jeweils zurückliegenden 7 Tage addiert und die Summe wird durch die betrachtete Einwohnerzahl (z.B. Stadt, Landkreis, ...) geteilt und mit 100.000 multipliziert.

Um mit diesem Inzidenzwert die Pandemie zu steuern, könnte man zum Beispiel es als Ziel setzen, diese Inzidenz jede Woche um 20 % zu senken³. Vielmehr wurden allerdings absolute Schwellwerte (z.B. 7-Tage-Inzidenz = 50, 100, usw.) verwendet, um Maßnahmen hinzuzufügen oder zu lockern (siehe z.B. die bundeseinheitliche „Notbremse“- Regelungen⁴).

Während die einen dies als ein geeignetes Frühwarnsystem betrachten, geben andere die mangelnde Robustheit der Inzidenz-Zahlen zu bedenken. So würde bei vermehrter Testung diese Inzidenz ansteigen; allerdings dann in der Regel durch die Nachverfolgung von mehr Fällen auch wieder schneller fallen. Ebenso geht die Inzidenz als Index implizit von einer Gleichverteilung der Infektionen in der getesteten Gruppe aus, was nicht immer der Fall sein muss, da sich vermehrt Leute mit einem Verdacht selber testen. Was kann man tun?

Zunächst kann man die Berechnung der Inzidenz selbst oder vergleichbarer Indizes durch eine *Ausgleichsrechnung* stabilisieren, anstatt einfach nur die Summe der letzten 7 Tage zu nehmen, um die sog. 'Wochen-Saisonalität' (Wochenend-Effekt) auszugleichen. Neben der Stabilisierung ist ein weiterer Vorteil, dass man den Wert für 'heute' erhält und nicht die Summe (also eine Art Mittelwert) der letzten 7 Tage. Zudem könnte man auch den gewählten Indexwert für einige Tage im voraus abschätzen.

Ausgleichsrechnung

Wir wollen im folgenden das Verfahren der Ausgleichsrechnung an einem konkreten Beispiel durchrechnen. Für das Beispiel wechseln wir von der Inzidenz zur sog. *Verdopplungszeit*, die aufgrund des Trends in den Daten geschätzt werden soll. Ende April 2020 war der Trend in den Daten schon klar abfallend (effektive Reproduktionszahl < 1 seit Ende März 2020) und die Verdopplungszeit wurde deutlich länger verglichen mit den Schätzungen in den ersten Wochen der Pandemie (wo wir im exponentiell wachsenden Trend waren).

Im Zuge der Corona-Pandemie meldete die Johns-Hopkins-Universität (JHU) im Zeitraum 21.-26. April 2020 für Deutschland folgende Daten:

Datum	18.04	19.04	20.04	21.04	22.04	23.04	24.04
I	143.342	145.184	147.065	148.291	150.648	153.129	154.999

wobei I die Gesamtzahl der bestätigten Infektionen bezeichnet.

Als einfache Approximation in der Anfangszeit einer Pandemie (oder dem Beginn einer Infektionswelle) kann ein exponentielles Wachstum

$$I(t) = I(t_0) e^{b(t-t_0)}, \quad t > t_0$$

mit Parameter b angenommen werden. Nun soll die *Verdopplungszeit* T_2 am 24. April 2020 in diesem Modell als Maß für die Inzidenz vorhergesagt werden. Es gilt

$$\frac{I(T_2)}{I(t_0)} = 2 \quad \text{d.h.} \quad e^{b(T_2-t_0)} = 2.$$

Somit erhalten wir die Verdopplungszeit $T_2 = t_0 + (\ln 2)/b$.

Wir betrachten nun die logarithmierten Werte der Infektionszahlen

$$L(t) := \ln I(t) = \ln I_0 + b(t-t_0)$$

und berechnen durch (lineare) Ausgleichsrechnung auf einer sinnvollen Zeitskala Polynome, z.B. $p_1(t) = a_0 + a_1(t-t_0)$. Wir bestimmen b und damit T_2 am 24. April durch Ableiten der Approximationen für $L(t)$ und vergleichen das Ergebnis mit dem der JHU (53,0 Tage)! Nummeriert man die Tage aufsteigend durch $t_0 = 0, \dots, t_6 = 6$, erhalten wir für p_1 das Problem $\|A_1 x_1 - v\|_2 \rightarrow \min$, wobei

$$A_1 = \begin{pmatrix} 1 & \dots & 1 \\ 0 & \dots & 6 \end{pmatrix}^\top, \quad x_1 = (a_0, a_1)^\top, \quad v = (\ln I(t_0), \dots, \ln I(t_6))^\top$$

Ferner ist

$$A_1^\top A_1 = \begin{pmatrix} 7 & 21 \\ 21 & 91 \end{pmatrix}, \quad A_1^\top v \approx (83.3772, 250.4968)$$

und wir erhalten als Lösung der *Normalgleichungen* $x_1 \approx 11.8719, 0.0130^\top$. Ableiten liefert $b \approx p'_1(t_6) = a_1 \approx 0.0130$, womit $T_2 = (\ln 2)/b \approx 53.1438$, also der Wert der JHU.

Verwendet man dagegen ein quadratisches Polynom $p_2(t) = \hat{a}_0 + \hat{a}_1(t-t_0) + \hat{a}_2(t-t_0)^2$, erhält man nach Ableiten $b \approx p'_2(t_6) = 2\hat{a}_2(t_6-t_0) + \hat{a}_1 = 12\hat{a}_2 + \hat{a}_1 \approx 0.0146$, womit $T_2 = (\ln 2)/b \approx 47.48$. Diese (bessere?) Vorhersage weicht stärker von der der JHU ab. Somit wird die JHU vermutlich auch einen linearen Fit mit dem Polynom p_1 verwendet haben.

Robuste Indizes mit Verzögerung

Man ist bestrebt, den Inzidenzwert oder einen entsprechenden Index so zu modifizieren, dass er für den Verlauf der Pandemie und damit für die zu ergreifenden Maßnahmen aussagekräftiger ist. Ein klassisches COVID-19 Modell wurde bereits im ersten Kapitel vorgestellt. Nun gibt es zwar robuste, verlässliche Werte, wie etwa

- Einweisungen in ein Krankenhaus (sog. Hospitalisierung)
- Anzahl COVID-19 Patienten auf Intensivstationen (intensivpflichtige Patienten)
- Anzahl verfügbarer Intensivbetten (engl. intensive care units, ICUs), inklusive der zugehörigen Fachpflegekräfte
- Anzahl der (an/mit COVID-19) verstorbenen Personen

Allerdings haben diese Werte eine große Verzögerung (z.B. 14 Tage) zu den möglichen Maßnahmen (in der Mathematik: 'Steuerungsvariablen') und sind daher als alleiniger Index nur bedingt geeignet.

Eine Lösung wäre ein Index als (gewichteter) Mittel der verschiedenen Einzelindizes, vergleichbar mit einem Portfolio von Aktien. Auch auf dem Aktienmarkt möchte man durch diese Streuung das Risiko falsch zu liegen reduzieren. Ein Nachteil dabei ist, dass dieses Vorgehen (mit den sozio-ökonomischen Konsequenzen) für die Bevölkerung weniger transparent ist und die genaue Wahl des Portfolios zu einem gewissen Grad willkürlich ist.

Eine weitere Möglichkeit, die Inzidenz stabiler gegenüber der Anzahl der Testungen zumachen, wäre die *Dunkelziffer der Neuinfektionen* abzuschätzen und zu berücksichtigen. Dazu erweitert man das SIR-Modell aus dem ersten Kapitel um weitere Kompartimente zur Berücksichtigung von robusten Daten (s.o.) und baut einen Dunkelzifferfaktor ω ein. Nun fittet man das erweiterte Modell an den Daten (Hospitalisierungsraten, Sterbezahlen) und findet auf diesem Weg den Faktor ω der am besten zu den Daten passt und verwendet ihn als Schätzung für die Dunkelziffer⁵.

Der EPG-Index

Eine weitere Alternative ist der *EPG-Index* (engl. Effective Potential Growth), den wir hier kurz aus einem Artikel zusammenfassen⁶. Wir beginnen dazu mit dem *Infizierten-Verstorbenen-Anteil* (engl. infection fatality ratio, IFR), der den Anteil der Todesfälle unter allen Infizierten angibt:

$$IFR = \frac{\text{Anzahl der an COVID-19 Verstorbenen in einem Zeitraum}}{\text{Gesamtzahl der COVID-19 Infektionen im gleichen Zeitraum}}$$

Der IFR schätzt also das Sterberisiko bei allen Infizierten (diagnostiziert, asymptomatisch, nicht getestet) ab.

Zunächst nehmen wir einen Schätzwert $IFR = 1\%$ (in Europa 0,3–3%) und unter Berücksichtigung der gemeldeten (relativ sicheren) Zahlen der Todesfälle schätzen wir den Zeitraum (*mittlere Zeit bis zum Tod* TtD (engl. time to death) auf $TtD=18$ Tage (Menschen sterben zwischen 15 und 22 Tagen nach dem Auftreten der ersten Symptome)⁷. Die geschätzte Anzahl der infizierten Personen mit COVID-19 E_t zum Zeitpunkt t , lautet:

$$E_t = \frac{d_{t+TtD}}{IFR},$$

wobei d_{t+TtD} die Anzahl der gemeldeten Todesfälle zum Zeitpunkt $t+TtD$ ist. Damit können wir die Anzahl der Fälle vor $TtD=18$ Tagen schätzen und den Wert mit der Anzahl der vor 18 Tagen entdeckten Fälle vergleichen, wodurch wir einen diagnostischen Prozentsatz erhalten.

Allerdings ist dies eine unrealistische untere Grenze, da niemand am ersten Tag der Symptome einen PCR-Test durchführt oder zum Arzt geht. Ebenso können externe Faktoren, wie die Verfügbarkeit von PCR-Tests, Sättigung des Gesundheitssystems, usw. die Diagnose verzögern. Diese *Verzögerung bis zur Diagnose* (DD, engl. delay to detection, 4-14 Tage) entspricht der Zeit zwischen dem Auftreten der ersten Symptome und der Meldung durch das Gesundheitssystem. Diese Zeit ist stark länderspezifisch.

Die Autoren Català, Pino, Marchena, et al. schlagen folgende allgemeine Prozedur für die Berechnung des *Prozentsatzes der Diagnose* (%D) vor:

- Mithilfe der obigen Formel kann mit TtD und IFR die Anzahl der Fälle geschätzt werden (alternativ mittels eines gefitteten Modells, siehe voriger Abschnitt)
- Berechnung der zeitlichen Korrelation (Zusammenhang) zwischen gemeldeten Fällen C_t und gemeldeten Todesfällen d_t . Das Maximum der Korrelation bestimmt die Zeit zwischen (Bestätigungs) Diagnose und Tod (engl. Diagnosis to Death, DtD, 4-14 Tage). Die Verzögerung in der Diagnose DD ergibt sich als $DD=TtD - DtD$.
- Auswertung des prozentualen Anteils der Diagnose %D basierend auf geschätzten (1.) und gemeldeten (2.) Fällen.

Sobald die Diagnoserate %D bekannt ist, ist es einfach, eine reale Inzidenz zu ermitteln, die nicht mehr von (länderspezifischen) Zeit-

verzögerungen bis zur Diagnose abhängt. Das Niveau der Diagnose und der realen Inzidenz ist nützlich für politische Entscheidungsträger. Da die Diagnoserate %D stark von dem länderspezifischen Wert für DD abhängt, definieren wir auch die *Delay-to-Detection-Diagnoserate* (DD-DR) als die Diagnoserate %D, die unter Verwendung einer unterschiedlichen Zeitverzögerung zwischen dem Auftreten von Symptomen und Erkennbarkeit für jedes Land berechnet wird.

Mit Hilfe der DD-DR kann man berechnen, wie viele unentdeckte Personen in den letzten 14 Tagen zur Zahl der Infizierten $I_{t,14}$ hinzugekommen sind. Somit kann man die 14-Tage Befallsrate besser abschätzen. Die Befallsrate A (engl. attack rate) ist der Anteil einer Population, die in einem gewissen Zeitraum an der Infektion erkrankt. Die ECDC (Europäisches Zentrum für die Prävention und die Kontrolle von Krankheiten) benutzt in diesem Zusammenhang den Zeitraum der letzten 14 Tage⁸. Das wollen wir auch so tun. Die 14-Tage-Befallsrate zum Zeitpunkt t ist definiert als

$$A_{t,14} = \frac{\text{exponierte Empfängliche einer Gruppe, die in den letzten 14 Tagen erkrankten}}{\text{exponierte Empfängliche in dieser Gruppe insgesamt, in den letzten 14 Tagen}}.$$

Anstatt der (*effektiven*) Reproduktionszahl R_t , die man von erweiterten SIR-Modellen abschätzt, verwenden wir als empirische Alternative ρ_t , die Anzahl der heute entdeckten neuen Fälle N_t geteilt durch die Anzahl der neuen Fälle, die vor fünf Tagen entdeckt wurden: N_t/N_{t-5} . Die starken Schwankungen dieser Größen erfordern jedoch, dass wir gemittelte Werte verwenden:

$$\rho_t = \sum_{i=-n_d}^{n_d} \frac{N_{t+i}}{N_{t+\tau+i}}.$$

Konkret verwenden wir wegen der ausgeprägten Wochen-Saisonalität Werte über sieben Tage ($n_d = 3$) und die Verzögerung $\tau = 5$ Tage, welches ungefähr die Zeit ist, in der infizierte Personen Symptome entwickeln, falls sie welche entwickeln, d.h. wir nehmen

$$\rho_{t,7} = \frac{N_{t-3} + N_{t-2} + N_{t-1} + N_t + N_{t+1} + N_{t+2} + N_{t+3}}{N_{t-8} + N_{t-7} + N_{t-6} + N_{t-5} + N_{t-4} + N_{t-3} + N_{t-2}}.$$

Diese Rate ist kleiner als 1, wenn die Zahl der neuen Fälle abnimmt und größer als 1, wenn die Anzahl der Fälle zunimmt. Durch Mittelungsprozesse (oder besser Ausgleichsrechnung, s.o.) kann dieser Wert $\rho_{t,7}$ weiter stabilisiert werden und wird mit $\bar{\rho}_{t,7}$ bezeichnet.

Die Autoren Català, Pino, Marchena, et al. schlagen den folgenden täglichen EPG-Index vor:

$$\text{EPG}_t = \bar{\rho}_{t,7} A_{t,14}.$$

Der EPG-Index ist einfach die Multiplikation der (gemittelten) Wachstumsrate der Krankheit $\bar{\rho}_{t,7}$ mit der Schätzung der Befallsrate $A_{t,14}$, beide ausgewertet zum richtigen Zeitpunkt t in der jüngsten Vergangenheit.

Der ungünstigste Fall ist einer, bei dem sowohl $A_{t,14}$, als auch $\bar{\rho}_{t,7}$ groß sind. Das bedeutet, dass man eine große Population mit der Krankheit und viel Ausbreitung früher. Die beste Situation ist ein niedriger Wert der Geschwindigkeit und eine geringe Anzahl von aktiven Fällen. Eine große Anzahl von $A_{t,14}$, mit niedrigem $\bar{\rho}_{t,7}$ oder ein großes $\bar{\rho}_{t,7}$ mit niedrigem $A_{t,14}$ sind potenziell gefährliche Situationen.

² Robert Koch-Institut: COVID-19-Dashboard, <https://corona.rki.de/>

³ www.wa.de/politik/corona-notbremse-inzidenz-michael-meyer-hermann-marius-lanz-zdf-hamburg-90483279.html

⁴ www.bundesregierung.de/breg-de/suche/bundesweite-notbremse-1888982

⁵ A. Kergaßner, C. Burkhardt, D. Lippold, et al., Memory-based meso-scale modeling of COVID-19, *Comput. Mech.* 66 (2020), 1069–1079. <https://doi.org/10.1007/s00466-020-01883-5>.

⁶ Martí Català, David Pino, Miquel Marchena, et al., Robust estimation of diagnostic rate and real incidence of COVID-19 for European policymakers, *PLOS*, January 7, 2021 <https://doi.org/10.1371/journal.pone.0243701>

⁷ F. Zhou, T. Yu, R. Du, G. Fan, Y. Liu, et al., Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study, *Lancet* 395 (2020), 1054–1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)

⁸ European Centre for Disease Prevention and Control, Rapid Risk Assessment: Coronavirus disease 2019 (COVID-19) in the EU/EEA and the UK – tenth update. 11th June, 2020. verfügbar: <https://www.ecdc.europa.eu/en/publications-data/rapid-risk-assessment-coronavirus-disease-2019-covid-19-pandemic-tenth-update>.

Mathematische Modellierung und optimale Kontrolle von Epidemien

Mathematische Modelle sind unerlässlich, um den Verlauf von Epidemien und die Wirkung von Interventionsmaßnahmen vorherzusagen. Die Theorie der optimalen Steuerung erlaubt es, effiziente Interventionsstrategien direkt zu berechnen.

Dr. Markus Kantner und Dr. Thomas Koprucki vom Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin beschreiben die Anwendung der Theorie der optimalen Steuerung auf die COVID-19 Pandemie:

Die Weltgesundheitsorganisation schätzt, dass es etwa 200 Zoonosen gibt, d.h. Infektionskrankheiten, die durch Anpassung von Krankheitserregern aus dem Tierreich auf den Menschen überspringen. Manche dieser neuen Erreger sind harmlos, andere hingegen können katastrophale Epidemien verursachen. Eine wichtige Größe, um die Übertragungsfähigkeit einer neuen Infektionskrankheit zu bestimmen, ist die *Basisreproduktionszahl* R_0 . Sie gibt an, wie viele Personen ein Infizierter im Durchschnitt ansteckt, wenn in der Gesamtbevölkerung keinerlei Immunität vorherrscht. Der Zahlenwert der Basisreproduktionszahl ist keine Naturkonstante, sondern hängt sowohl von den biomedizinischen Eigenschaften des Erregers als auch vom Grad der Vernetzung der Gesellschaft und deren gesundheitlichem Grundzustand ab.

Ist die Basisreproduktionszahl kleiner als 1, sinkt die Zahl der Infizierten in der Bevölkerung mit der Zeit und das Infektionsgeschehen

kommt zum Erliegen. Ist sie größer als 1, kommt es zu einem exponentiellen Anwachsen der Infektionszahlen („Schneeballeffekt“), siehe Abbildung 16. Im Fall der COVID-19 Pandemie beispielsweise, reichen die Schätzungen der Basisreproduktionszahl von etwa 2.4 bis 3.5, wodurch es praktisch weltweit zu einem rasanten exponentiellen Anwachsen der Infektionszahlen kam.

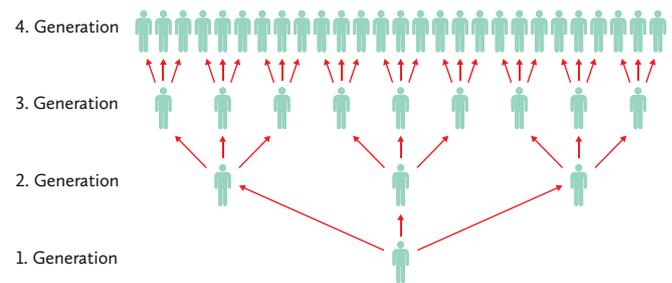


Abbildung 16: Exponentielles Wachstum einer unkontrollierten Epidemie mit Basisreproduktionszahl $R_0 = 3$.

Nicht-pharmazeutische Interventionen

Bei neu auftretenden Infektionskrankheiten existieren in aller Regel zu Beginn einer Epidemie weder wirksame medizinische Therapien noch Impfstoffe. Um das exponentielle Wachstum abzuschwächen, müssen in einer solchen Situation nicht-pharmazeutische Gegenmaßnahmen ergriffen werden. Die Bandbreite solcher nicht-pharmazeutischen Maßnahmen ist riesig und reicht von intensiver Handhygiene und dem Tragen von medizinischen Schutzmasken bis hin zu „Lockdown-Maßnahmen“ wie dem Verhängen von Ausgangssperren, Schulschließungen und dem Verbot von Großveranstaltungen. Zusätzlich können Tests und Quarantäne von Infizierten und deren Kontaktpersonen die *effektive Reproduktionszahl* weiter senken, bis diese schließlich unter 1 fällt und die Infektionsketten abbrechen.

Jedoch kommt es nach der Aufhebung dieser Interventionsmaßnahmen jederzeit wieder zu einem exponentiellen Anwachsen des Infektionsgeschehens, solange in der Gesamtbevölkerung keine „Herdenimmunität“ vorherrscht (siehe Kapitel COVID-19 Modellierung, Seite 8). Wenn es nicht gelingt den Erreger auszurotten, verschaffen nicht-pharmazeutische Maßnahmen somit insbesondere Zeit um Medikamente und Impfstoffe zu entwickeln.

Auf der anderen Seite sind diese Maßnahmen für die Gesellschaft unter Umständen mit hohen Belastungen und Kosten verbunden. Dies betrifft z.B. den wirtschaftlichen Schaden durch die Schließung von Geschäften als auch psychologische Auswirkungen langandauernder Kontaktbeschränkungen. Nicht-pharmazeutische Maßnahmen müssen also koordiniert eingesetzt werden, um eine wirksame Eindämmung der Infektionsdynamik zu erreichen und zugleich die sozio-ökonomischen Schäden so gering wie möglich zu halten. Wie aber sollen solche Maßnahmen sinnvoll koordiniert und ihre Auswirkungen prognostiziert werden?

Mathematische Modellierung: Die Epidemie mit Mathematik verstehen

Eine Antwort auf diese Frage können mathematische Modelle liefern. Die Dynamik von Infektionskrankheiten kann auf verschiedene Art und Weise modelliert werden. Die Skala reicht von einfachen Modellen (*Differentialgleichungen*), die den Verlauf einer Epidemie für größere Bevölkerungsgruppen im Mittel wiedergeben, bis hin zu sehr komplexen und detaillierten Modellen, die das Infektionsgeschehen auf der Ebene individueller Akteure beschreiben (*agentenbasierte, stochastische Modelle*).

Letztere benötigen eine Vielzahl von Parametern und Daten (z.B. Kontaktnetzwerke, Altersstruktur, Mobilitätsdaten, Gesundheitsdaten etc.), um genaue Vorhersagen treffen zu können, wohingegen einfache Modelle bereits mit relativ wenigen, aber gut angepassten Parametern das Infektionsgeschehen adäquat wiedergeben können. Mathematische Modelle erlauben es die Auswirkungen von Schulschließungen und ähnlichen Interventionen zur Kontaktreduktion in Computersimulationen zu analysieren und zu bewerten.

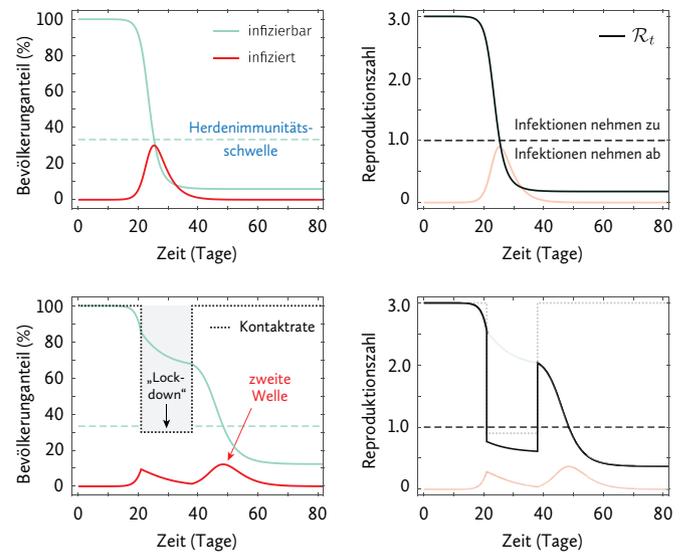


Abbildung 17: Oben: Verlauf einer unkontrollierten Epidemie und zugehörige effektive Reproduktionszahl. Nachdem der Anteil der Infizierbaren unter die Herdenimmunitätsschwelle (gestrichelte Linie) gefallen ist, beginnen die Infektionszahlen zu sinken. Unten: Simulation eines kurzen Lockdowns durch vorübergehende Kontaktreduktion. Nach Beendigung der Maßnahme kommt es zu einer zweiten Infektionswelle.

Wir wollen das Grundprinzip nicht-pharmazeutischer Interventionsmaßnahmen am Beispiel des einfachen SIR-Modells illustrieren, das wir bereits im vorhergehenden Kapitel kennengelernt haben (siehe Kapitel COVID-19 Modellierung, Seite 8). In diesem Modell wird die Gesamtbevölkerung in drei Gruppen eingeteilt: *S* (*susceptible*) für die anfänglich Gesunden und Infizierbaren, *I* (*infectious*) für die Infizierten und Infektiösen und *R* (*removed*) für die Genesenen und Verstorbenen. In der Simulation wird für eine bestimmte Zeitspanne die Kontaktrate β reduziert, wobei in diesem einfachen Modell jedoch offen bleibt, durch welche konkreten Maßnahmen dies erreicht wird.

Abbildung 17 zeigt die Auswirkungen von Maßnahmen zur Kontaktreduktion. Es zeigt sich, dass die Infektionen nach der Aufhebung der Maßnahmen wieder anwachsen („zweite Welle“), solange die Herdenimmunitätsschwelle noch nicht erreicht wurde (d.h. die Zahl der infizierbaren Personen ist noch zu groß, gestrichelte Linie). Solche „An-Aus-Interventionen“ lassen sich auch beliebig oft hintereinanderschalten, siehe Abbildung 18, wodurch aufflammende Infektionswellen immer wieder unterbrochen werden.

Theorie der optimalen Steuerung: Die Epidemie mit Mathematik in den Griff bekommen

Das obige Beispiel zeigt, dass im Fall von einfachen An-Aus Strategien die Epidemie immer wieder aufflammt und weitere Interventionen notwendig werden. Wiederholte Shutdowns führen jedoch höchstwahrscheinlich zu geringer Akzeptanz in Bevölkerung und Wirtschaft. Wie aber könnte eine Interventionsstrategie aussehen, die diese Probleme vermeidet?

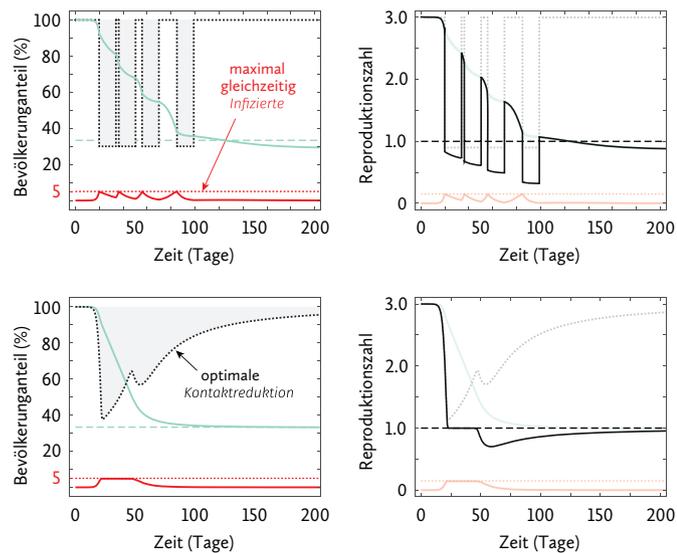


Abbildung 18: Oben: An-Aus-Kontrolle einer Epidemie mit der Nebenbedingung, dass der Anteil der Infizierten an der Gesamtbevölkerung niemals mehr als 5 % betragen soll. Unten: Die optimal gesteuerte Kontaktreduktion stellt sicher, dass die Infektionszahlen niemals die kritische Obergrenze überschreiten, während die Maßnahmen graduell gelockert werden können.

Die *Theorie der optimalen Steuerung* bietet einen systematischen Zugang zur Berechnung des idealen zeitlichen Verlaufs der mittleren Kontaktreduktion, wobei sowohl die sozio-ökonomische Kosten als auch weitere Nebenbedingungen berücksichtigt werden können. Solche Nebenbedingungen sind z.B. die begrenzte Zahl der verfügbaren intensivmedizinischen Behandlungsplätze, die in keinem Fall überschritten werden dürfen. Im Mittelpunkt der Methode steht dabei, neben dem Modell für die Infektionsausbreitung, ein sogenanntes *Zielfunktional*, das sowohl die Ziele der Intervention als auch deren (abstrakte) Kosten widerspiegelt. Mit Methoden aus der *Variationsrechnung* lassen sich somit Bedingungen (in Form von mathematischen Gleichungen) ableiten, aus denen dann der optimale zeitliche Verlauf der Maßnahmen zur Kontaktreduktion berechnet werden kann.

Abbildung 18 zeigt den Verlauf einer *optimal gesteuerten* Epidemie im SIR-Modell, bei der die nicht-pharmazeutischen Maßnahmen so koordiniert sind, dass eine bestimmte Obergrenze von gleichzeitig Infizierten (höchstens 5 % der Bevölkerung) niemals überschritten wird. Dabei können die Maßnahmen mit der Zeit graduell gelockert werden, weil die stetig zunehmende Immunität in der Bevölkerung kontinuierlich mehr Kontakte erlaubt ohne gleich wieder in ein exponentielles Wachstum zurückzufallen. Die Intervention endet mit dem Erreichen der Herdenimmunitätsschwelle, sodass nachfolgende Infektionswellen ausgeschlossen sind. Die sozio-ökonomischen Gesamtkosten der optimal gesteuerten Intervention sind deutlich geringer als im Fall wiederholter Lockdowns (An-Aus-Strategie).

In ähnlicher Weise können solche mathematischen Techniken auf weitaus komplexere epidemiologische Modelle übertragen werden, die als Grundlage für die politische Entscheidungsfindung dienen können. Weiterhin können Impfkampagnen und deren optimale Koordination mit Hilfe von mathematischen Methoden geplant werden (siehe nachfolgendes Kapitel über Impfstrategien, Seite 21). Mathematische Modellierung von Epidemien ist somit unerlässlich um Prognosen über die zukünftige Infektionsdynamik zu erhalten und frühzeitig geeignete Gegenmaßnahmen einleiten zu können.

Übrigens, die gleichen mathematischen Techniken der optimalen Steuerung werden auch im Kapitel für die Bestimmung einer optimalen CO₂-Steuer benutzt.

Mathematische Modellierung von Impfprogrammen



Impfungen sind der beste Weg, um Virusepidemien zu verhindern. Aber eine falsche Strategie bei der Impfung der Bevölkerung kann katastrophale Folgen haben.

Prof. Dr. med. Helmut Brunner von der Universität Düsseldorf und dem Lehrstuhl Gesundheitsökonomie der Bergischen Universität Wuppertal und Prof. Dr. Matthias Ehrhardt von der Bergischen Universität Wuppertal beschreiben in diesem Kapitel ihre Erfahrungen mit der mathematischen Modellierung von Impfungen.

Impfprogramme gegen Pocken, Diphtherie, Kinderlähmung, Gehirnhautentzündung, Masern, Mumps, Röteln, Keuchhusten und Tetanus haben sich als sehr wirksam erwiesen. Trotzdem ist die Ausrottung der Zielkrankheit bisher nur bei Pocken gelungen. Die Impfung von Erwachsenen erweist sich oft als weniger erfolgreich, da der Wechsel auf andere Serotypen des Erregers (HPV-Impfung) oder Varianten (Mutanten) des Erregers (Antigen-Drift oder Shift bei Influenza) den Impfschutz schwächen können.

Immunologische Grundlagen

Zur Abwehr von Infektionskrankheiten stehen die angeborene (unspezifische) und die erworbene (spezifische) Immunität zur Verfügung. Das spezifische Immunsystem mit seiner Fähigkeit zum „immunologischen Gedächtnis“ bildet die Voraussetzung für eine

Immunsierung durch Impfung. Hauptakteure der spezifischen Immunität sind B-Lymphozyten, die sich zu Plasmazellen differenzieren und Antikörper produzieren können, sowie T-Lymphozyten, die die zelluläre Abwehr repräsentieren. Beim ersten Kontakt der B-Zellen mit Antigenen von Infektionserregern werden vorwiegend Antikörper der Klasse IgM, beim zweiten Kontakt Antikörper der IgG-Klasse gebildet. Diese sind oft lebenslang nachweisbar und gewährleisten über B-Gedächtniszellen (memory cells) einen dauerhaften Schutz. Da viele Antigene die B-Zellen nicht direkt aktivieren können, ist eine Kooperation von Makrophagen und B-Lymphozyten mit den T-Lymphozyten (CD4, CD8-Zellen) notwendig.

Aktive Immunisierung

Bei der aktiven Immunisierung werden spezifische Antigene von Infektionserregern intramuskulär oder subkutan verabreicht, um die Produktion von Antikörpern zu induzieren. Meist wird mehrmals geimpft (Boosterung), um die Antikörpertiter anzuheben und einen lange-anhaltenden Schutz zu gewährleisten. Lebendimpfstoffe (abgeschwächte, noch vermehrungsfähige Infektionserreger) und Totimpfstoffe (Komponentenimpfstoffe und Ganzkeim-Impfstoffe) werden verwendet. Neue innovative Impfstoffe auf gentechnischer Basis (mRNA) werden erst seit wenigen Monaten eingesetzt und müssen in ihrer Praxistauglichkeit bei der Bekämpfung der SARS-Cov-2-Pandemie unter Routinebedingungen geprüft werden.

Durch Impfprogramme sollen Krankheits- und Todesfälle an Infektionskrankheiten bei Kindern und Jugendlichen vermieden werden. Der erhebliche nationale und internationale Erfolg der Prävention durch aktive Immunisierung ist vielfach belegt. Abbildung 19 zeigt am Beispiel einer bakteriellen Krankheit, der eitrigen Meningitis (Gehirnhautentzündung) durch *Haemophilus influenzae* Typ b, und einer Viruskrankheit Poliomyelitis (Kinderlähmung) die Abnahme der Erkrankungshäufigkeit nach der Einführung von Impfprogrammen.⁹

Andererseits haben Impfskeptiker behauptet, dass die allgemeine Sterberate, Mortalität, durch Impfprogramme nur wenig gesenkt werden könne, da bereits andere Maßnahmen, wie bessere Behandlungsmethoden, Antiinfektiva, Hygiene, Ernährung und weitere sozioökonomische Faktoren den Rückgang der Mortalität an Infektionskrankheiten erklären. Die Sterblichkeit an den genannten Erkrankungen sei daher bereits vor Einführung der Impfungen zurückgegangen. Die Hinweise dafür, dass *Aktive Impfungen* zu den wichtigsten und wirksamsten Maßnahmen der Prävention vor Todesfällen

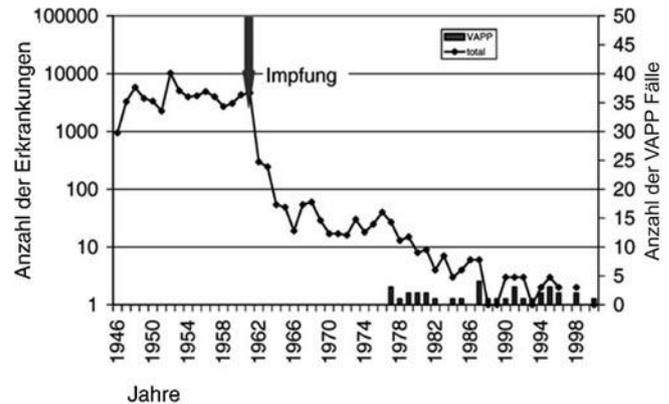
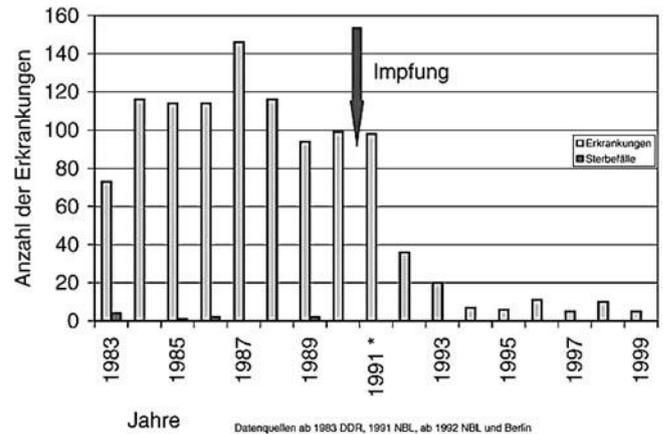


Abbildung 19: Oben: *Haemophilus influenzae* Meningitis in der ehemaligen DDR und den neuen Bundesländern 1983-1999 ab 1991 Bundesrepublik Deutschland, unten: Erkrankungen an Poliomyelitis, VAPP: Vaccine Associated Paralytic Poliomyelitis (Eigene Darstellung nach Heiningner 2004)

bei Kindern und Jugendlichen gehören, sind erst vor einigen Jahren unter anderem durch Untersuchungen aus den Niederlanden belegt worden¹⁰. Dabei wurden mathematische Modelle unter Anderem zur Ermittlung der verlorenen Lebensjahre eingesetzt. Die Sterblichkeit von Kindern und jungen Erwachsenen bis zum Ende des 20. Lebensjahres ist in der Zeit von 1903 bis 1992 exponentiell gesunken (mit einer Halbwertszeit von 19 Jahren) und zwar auch in der Zeit vor der Einführung der ersten Impfung. Die Autoren kamen aber zu dem Ergebnis, dass Impfungen gegen Kinderkrankheiten im 20. Jahrhundert in den Niederlanden 6.000 bis 12.000 Kindern das Leben gerettet haben. Dabei wurden z.B. Regressionsmodelle und Zeitreihen folgender Form eingesetzt:¹¹

Seien Y_t die beobachtete Anzahl gemeldeter Fälle im Monat t , $t = \{1, \dots, n\}$ und n die Gesamtzahl an Monaten vor der Impfung. Y_t folge einem Poisson-Prozess

$$Y_t \sim \text{Poisson}(\mu_t).$$

Das Regressionsmodell kann beschrieben werden als

$$\log(\mu_t) = \log(p_t) + \beta_0 + \beta_1 t + \sum_{j=1}^k \left[a_j \sin\left(\frac{2\pi t}{12\tau_j}\right) + \gamma_j \sin\left(\frac{2\pi t}{12\tau_j}\right) \right] + x_t,$$

wobei ρ_t die Allgemeinbevölkerung der 0-bis 20-Jährigen zum Zeitpunkt t als „Offset“ angegeben wird (die Bevölkerung mit dem höchsten Infektionsrisiko), β_0 ist das Intercept, β_1 ist der Koeffizient des säkularen Trends (transformiert, um eine prozentuale Änderung anzugeben: $\tilde{\beta}_1 = 1 - \exp(\beta_1) \times 100$). Saisonalität und mehrjährige Zyklen werden als die Summe von k „Oberschwingungen“ mit Frequenzen von τ Jahren eingegeben. τ ist ein Satz von ganzen Zahlen basierend auf den dominanten Frequenzen in der Zeit vor der Impfung und enthält mindestens einen saisonalen Term, d.h. $\tau = 1$. Der Term x_t ist die logarithmierte Inzidenzrate der Infektion (Kapitel "Robuste Inzidenzzahlen") und kann nicht beobachtet werden. Er beschreibt den latenten Autokorrelationsprozess definiert als

$$x_{t > 1} \sim \text{Normal}(\rho x_{t-1}, \sigma^2), \quad x_1 \sim \text{Normal}(0, \sigma^2(1 - \rho^2)^{-1}).$$

Das Modell wurde in einem „Bayes-Framework“ definiert, in dem die Prioritäten für die marginale Varianz und ρ definiert sind als

$$\sigma^2(1 - \rho^2) \sim \text{Gamma}(1, 10^{-5}), \quad \log\left(\frac{1 + \rho}{1 - \rho}\right) \sim \text{Normal}(0, 0.15).$$

Quantifizierung der Wirksamkeit von Impfstoffen

Wie bei allen Medikamenten ist auch für Impfstoffe die Bewertung der Wirksamkeit und der Verträglichkeit in randomisierten, kontrollierten, klinischen Studien, das bedeutet unter „optimalen Bedingungen“, erforderlich. In vielen Ländern ist zusätzlich eine Kosten-Nutzen-Bewertung für die Zulassung notwendig. Auch für die Bundesrepublik Deutschland ist dies in nicht zu ferner Zukunft zu erwarten. Die Wirksamkeit von Impfstoffen in klinischen Studien (engl. vaccine efficacy) VE, ist nicht zu verwechseln mit der Impfstoffeffektivität (engl. vaccine effectiveness) VE_{eff} , die Wirksamkeit des Impfstoffes unter klinischen Alltagsbedingungen. Die VE wird als Reduktion des relativen Risikos, RRR, bei Geimpften im Vergleich zu Nichtgeimpften berechnet, an der Zielkrankheit zu erkranken. Die RRR wird also aus dem Relativen Risiko, RR, einer Ansteckung bei geimpften und ungeimpften Personen abgeleitet. Wir vergleichen den Anteil der ungeimpften Personen mit dem Anteil der Geimpften unter den Erkrankten:

$$VE = 1 - \frac{AR_v}{AR_{nv}} = \frac{AR_{nv} - AR_v}{AR_{nv}},$$

$$RR = \frac{AR_v}{AR_{nv}}, \quad VE = 1 - RR = RRR.$$

Es bedeuten dabei

AR_v : Absolutes Risiko, attack rate, vakzinierter Personen, mit einem Krankheitserreger infiziert zu werden, gegen den der Impfstoff schützen soll.

AR_{nv} : Absolutes Risiko nicht geimpfter Personen mit einem Erreger infiziert zu werden.

RR: Relatives Risiko

RRR: Relative Risiko-Reduktion

Ein Impfstoff muss sich unter Routinebedingungen, also unabhängig von den „optimierten“ Bedingungen eines klinischen Versuchs bewähren. Impfungen mit den Vakzinen von AstraZeneca und Biontech/Pfizer haben die Zahl der Hospitalisierungen und Todesfälle in Großbritannien deutlich gesenkt¹². England hatte als erstes Land am 8. Dezember 2020 mit einer Impfkampagne begonnen. Diese Studie vom April 2021 zeigt, dass es bereits 10 bis 13 Tage nach der 1. Impfstoffdosis zu einem Rückgang der Mortalität und der Krankenhausbehandlungen wegen COVID-19 um 70 % (95%-Konfidenzintervall 59 bis 78 %) kam. Die Schutzwirkung stieg 14 Tage nach der 2. Dosis auf 89 % (85 bis 93 %): $VE_{\text{eff}} \approx 70 - 89$ %.

Die Reproduktionszahl R

Von besonderer Bedeutung für die Beurteilung der Dynamik einer Epidemie ist die Reproduktionszahl R . Sie gibt die durchschnittliche Anzahl der Sekundärfälle wieder, die ein Infizierter in einer empfänglichen Population zu einem bestimmten Zeitpunkt einer Epidemie erzeugen kann. Anders gesagt, beschreibt R , wie viele Menschen ein Infizierter im Mittel in einem bestimmten Zeitraum ansteckt. Zu Beginn einer Epidemie sprechen wir von R_0 , später von effektiver Reproduktionszahl, R_{eff} . Mit Hilfe der Reproduktionszahl R sind Schätzungen der Geschwindigkeit der Ausbreitung einer übertragbaren Krankheit sowie des Anteils der Bevölkerung, der nach überstandener Krankheit, durch sog. „stille Feiung“ (Immunantwort des Körpers bei einer symptomlosen Infektion) oder durch Impfung immun sein muss, um die weitere Ausbreitung der Epidemie zu verlangsamen oder zu stoppen.

R kann nicht allein als Maß für die Notwendigkeit oder Wirksamkeit von Maßnahmen herangezogen werden. Wichtig sind auch die absolute Zahl der Neuinfektionen, die kumulative Inzidenz, IC, oder die Inzidenzdichte, ID, im Bezugszeitraum sowie die Infektionsdosis und die daraus resultierende Schwere der Krankheitsverläufe. Man versucht, die absolute Zahl der Neuinfektionen so gering zu halten, dass eine effektive Nachverfolgung der Kontaktpersonen von Infizierten möglich ist und die Intensivstationen in Krankenhäusern nicht überlastet werden. R_0 ist für einen Erreger keine weltweit gültige Konstante, da sie von Umweltbedingungen und ethnischen Einflüssen abhängig ist. Mathematisch hängt R_0 auch noch vom zugrundegelegten Modell ab (siehe Kapitel zur COVID-19 Modellierung).

Herdenimmunität

Impfprävention gewährleistet Individualschutz, d.h. den Schutz des Individuums und Kollektivschutz gegen Infektionskrankheiten, die von Mensch zu Mensch übertragen werden, wenn die Durchimpfungsrate in der Bevölkerung hoch genug ist. Diese sog. Herdenimmunität beruht darauf, dass auch ungeimpfte Personen mit einer geringeren Wahrscheinlichkeit dem Infektionserreger ausgesetzt sind. Der Anteil der Personen, die in einer Bevölkerung geimpft sein müssen, um Herdenimmunität zu erzielen, ist für verschiedene Infektionskrankheiten unterschiedlich. Für die Diphtherie werden 80 % angenommen, für Mumps 90 % und für Masern 92 bis 95 %. In einfachen Modellen wird eine „effektive“ Herdenimmunität, HI, errechnet, wenn ein bestimmter Anteil der Bevölkerung (HI-Schwellenwert) über eine

schützende körpereigene Abwehrleistung verfügt. Wir nehmen an, dass nach dem Erreichen des Schwellenwertes Infektionsketten unterbrochen werden und Personen, die nicht geimpft werden können oder wollen, indirekt geschützt sind. Für die Herdenimmunität HI gilt

$$HI = \frac{1}{VE} \left(1 - \frac{1}{R_0} \right).$$

Beispiel COVID-19 und der Impfstoff BNT162b2 mRNA¹³

$$AR_v = \frac{8}{18.198} = 0,00044$$

$$AR_{nv} = \frac{162}{18.325} = 0,00884$$

$$RR = \frac{0,00044}{0,00884} \approx 0,05$$

$$VE = 1 - 0,05 = 0,95.$$

Der Impfstoff wurde also in einem klinischen Versuch an mehr als 36.000 Personen geprüft und erzielte eine Wirksamkeit von ca. 95%.

R_0 wurde zu Beginn der SARS-Cov-19-Pandemie auf ca. 3,0 geschätzt. Daraus lässt sich der Schwellenwert für die Herdenimmunität wie folgt berechnen:

$$HI = \frac{1}{0,95} \left(1 - \frac{1}{3} \right) \approx 0,7.$$

Dieses Ergebnis bedeutet, dass ca. 70 % der Bevölkerung, die noch keine COVID-19-Infektion überstanden haben, geimpft werden müssten, um eine ausreichende Herdenimmunität zu erzeugen. COVID-19-Genesene, die eine schützende Immunität entwickelt haben, können davon abgezogen werden.

Unspezifische Schutzeffekte von Impfungen

Nach bisherigem Verständnis schützt ein Impfstoff nur spezifisch vor einer bestimmten Infektion. Inzwischen legen mehrere Beobachtungen nahe, dass Lebendimpfstoffe, z.B. der orale Impfstoff



gegen Poliomyelitis, die BCG-Impfung gegen Tuberkulose und die Masern-Impfung die Geimpften nicht nur spezifisch gegen die Zielkrankheiten immunisieren, sondern auch Schutzwirkungen gegen andere Krankheiten zur Folge haben und sogar Todesfälle verhindern können. Diese breiten vorteilhaften unspezifischen Wirkungen von Impfstoffen auf das Immunsystem könnten bei einer Pandemie nützlich sein und möglicherweise das Risiko für einen schwerwiegenden Verlauf von COVID-19 verringern. Zur mathematischen Modellierung der unspezifischen Wirkung der BCG-Impfung haben wir Vorschläge publiziert¹⁴.

Ökonomische Aspekte der Prävention durch Impfungen

Seit Jahrzehnten steigen in den Industrieländern die Kosten der Gesundheitssysteme stärker als das Volkseinkommen (BIP). Daher stößt die Einführung weiterer, relativ teurer Impfungen trotz des beachtlichen medizinischen Nutzens und der Kosteneffektivität, die u.a. für die Impfung gegen *Haemophilus influenzae b* gezeigt werden konnte, auf ethische und ökonomische Probleme, insbesondere dann, wenn die impfpräventablen Erkrankungen zwar schwer verlaufen, aber selten vorkommen. Auch diese Fragen können in mathematische Prognosemodelle integriert werden. Pocken, Poliomyelitis und Masern sind Beispiele weltweit auftretender Infektionskrankheiten, die durch Impfungen ausgerottet oder in ihrer Inzidenz sehr stark

reduziert worden sind. Obwohl es gegen diese Krankheiten keine gut wirksamen Medikamente gibt, haben sie ihren Schrecken in der Bevölkerung verloren, da sie bei ausreichend geimpften Personen praktisch nicht auftreten. Da die Menschheit eine hohe Populationsdichte erreicht hat, muss auch in Zukunft mit Pandemien durch Infektionskrankheiten mit neuen Erregern gerechnet werden.

⁹ U. Heininger, Risiken von Infektionskrankheiten und der Nutzen von Impfungen, Bundesgesundheitsbl. Gesundheitsforsch. – Gesundheitsschutz 47 (2004), 1129–1135.

¹⁰ M. van Wijhe, S.A. McDonald, H.E. de Melker, et al., Effect of vaccination programmes on mortality burden among children and young adults in the Netherlands during the 20th century: a historical analysis, The Lancet Infectious Diseases 16 (2016), 592–598.

¹¹ M. van Wijhe, A.D. Tulen, H. Korthals Altes, et al., Quantifying the impact of mass vaccination programmes on notified cases in the Netherlands, Epidemiology and Infection 146 (2018), 716–722.

¹² J. Lopez Bernal, N. Andrews, C. Gower, et al., Effectiveness of the Pfizer-BioNTech and Oxford-AstraZeneca vaccines on covid-19 related symptoms, hospital admissions, and mortality in older adults in England: test negative case-control study, BMJ 373 (2021), n1088

¹³ F.P. Polack, S.J. Thomas, N. Kitchin, et al., Safety and efficacy of the BNT162b2 mRNA Covid-19 Vaccine, The New England Journal of Medicine 383, (2020), 2603–2615.

¹⁴ S. Treibert, H. Brunner, M. Ehrhardt, Compartment models for vaccine effectiveness and non-specific effects for Tuberculosis, Mathematical Biosciences and Engineering 16 (2019), 4250–4298.

Mathematik ermöglicht Kulturveranstaltungen

Mithilfe einer einfachen mathematischen Formel kann man bestimmen, wie viele Menschen bei Kulturveranstaltungen in Innenräumen dabei sein dürfen.

Während der Corona-Pandemie mussten im Rahmen der sozialen Distanzierung viele Kulturbetriebe wie Theater, Oper, Kino usw. schließen. Im Rahmen der Lockerungsmaßnahmen stellt sich nun die Frage, wieviele Personen sich unter welchen Umständen (Maske, Belüftung, ...) wieder in bestimmten Räumen treffen können und so der Kulturbetrieb wieder aufgenommen werden kann.

Eine einfache Formel erlaubt eine Berechnung von Personenzahlen, die bei der Durchführung von Veranstaltungen in geschlossenen Räumen als Orientierung dienen kann¹⁹. Diese Zahl kann man auch als Orientierung für andere geschlossene Räume, wie etwa Schulen, Universitäten, Gaststätten, usw. nehmen.

Prof. Dr. Martin Kriegel von der TU Berlin entwickelte eine mathematische Formel für die Virus-Variante Alpha (B117) unter der Annahme, dass die (effektive) Reproduktionszahl $R < 1$ ist²⁰. Die maximal mögliche Personenanzahl P_{\max} berechnet sich nach

$$P_{\max} = \frac{\dot{V}_{\text{zu}}}{f_L \cdot f_M \cdot 105 \frac{\text{m}^3}{\text{h}} \cdot t_A}$$

Die Parameter auf der rechten Seite erklären sich dabei wie folgt:

\dot{V}_{zu} virenfreier Zuluftvolumenstrom (in der Regel Außenluft) in m^3/h (Wert aus der Bedienungsanleitung der Belüftungsanlage)

f_L Belüftungsart:

- $f_L = 1$ für Belüftung nach Mischluftprinzip (Frischluft kommt oben in den Raum)
- $f_L = 0,7$ Belüftung nach Quilluftprinzip (Frischluft kommt unten in den Raum)

f_M Faktor für Maskentragen:

- $f_M = 1$ niemand trägt eine Maske
- $f_M = 0,5$ alle Zuschauer tragen medizinischen Mund-Nasen-Schutz
- $f_M = 0,2$ alle Zuschauer tragen FFP2-Maske

t_A Aufenthaltsdauer, also die Dauer der Veranstaltung, in Stunden h

Der Faktor $105 \text{ m}^3/\text{h}^2$ dient nur der Skalierung und damit sich ein einheitenloses Ergebnis für die Personenzahl ergibt.

¹⁹ Hygienerahmenkonzept Juni für Kultureinrichtungen im Land Berlin, Stand 7.6.2021, https://www.berlin.de/sen/kulteu/aktuelles/corona/20210607_hrk_juni_finalv3.pdf

²⁰ Es gibt auch eine ausführliche Web-App der TU Berlin <https://hri-pira.github.io/>.

Mit Mathematik zur individuellen Therapie

Mathematik & Pharmazie? Das ist eine faszinierende Kombination mit spannenden Problemstellungen von großer Relevanz!

Durch digitale mobile Geräte wird es auch im Gesundheitsbereich immer einfacher, Daten zu erheben. Aber Daten alleine reichen nicht aus, um Einsichten und neues Wissen zu generieren. Man muss sie interpretieren, verstehen und aus großen Datenmengen das Wesentliche herausfiltern. Und manchmal hat man auch nur sehr wenige Daten und muss eine Entscheidung treffen, wie zum Beispiel bei der Individualisierung von Therapien in der Onkologie. Dabei spielt die Mathematik eine wichtige Rolle. Sie hilft, aus den Daten vorangegangener Studien und Therapien zu lernen, um die Behandlung für zukünftige Patienten zu verbessern. Ein Bericht aus der (wissenschaftlichen) Praxis von Corinna Maier (AbbVie Deutschland GmbH & Co. KG) und Prof. Dr. Wilhelm Huisinga (Mathematische Modellierung & Systembiologie; Institut für Mathematik der Universität Potsdam).

Bevor ein neues Medikament auf den Markt kommt, durchläuft es einen langen Forschungs- und Entwicklungsprozess. Und am Ende prüfen Zulassungsbehörden, ob es den hohen Anforderungen an die Sicherheit und Wirksamkeit genügt. Einen sehr guten Überblick dazu findet man z.B. online im Science Media Center Germany¹⁵. Aber woher weiß man eigentlich, dass ein zugelassenes Medikament bei allen (zukünftigen) Patienten/innen sicher und wirksam ist?

Während der Entwicklungsphase ist es ein wichtiges Ziel, dieser Fragestellung nachzugehen. Dazu werden in klinischen Studien unterschiedliche Daten zu dem Wirkstoff erhoben, beispielsweise wieviel Wirkstoff befindet sich 1, 2, 12, 24 Stunden nach der Verabreichung im Blut. Welche Nebenwirkungen traten auf? Tritt der erwartete Effekt ein? Gibt es dabei Unterschiede zwischen Frauen und

Männern? Das sogenannte therapeutische Fenster gibt den Bereich an, in dem der Wirkstoff sicher und wirksam ist (siehe Abbildung 27). Aber Menschen sind verschieden, und damit ist auch das therapeutische Fenster individuell unterschiedlich. Eine der vielen Herausforderungen in der Entwicklung neuer Wirkstoff-Therapien ist es, dass man das individuelle therapeutische Fenster nicht kennt. Man weiß aber, welche Wirkung und Nebenwirkungen in klinischen Studien bei Probanden/innen und Patienten/innen aufgetreten sind. Aus diesen Informationen versucht man sich ein Bild zu machen. Es gibt Medikamente, bei denen die „Schnittmenge“ aller individuellen therapeutischer Fenster hinreichend groß ist (Abbildung 27-C). Hier ist es vergleichsweise einfach, eine einzelne therapeutische Dosis für alle Patienten/innen zu finden. Andere Medikamente haben enge individuelle therapeutische Fenster (Abbildung 27-D). Dann wird es notwendig, die Dosis für einzelne Gruppen anzupassen oder für jede/n einzelne/n Patientin/en eine individualisierte Dosis zu finden.

Die Abbildung 27-D deutet noch auf einen anderen wichtigen Aspekt hin, um den es schwerpunktmäßig in diesem Beitrag gehen soll: Auch wenn es gelungen ist, eine Dosis zu finden, die für alle im therapeutischen Fenster liegt, also sicher und wirksam ist, so sieht man deutlich, dass es für die meisten Patienten/innen nicht unbedingt die optimale Dosis sein muss (hellgrüner Punkt in Abbildung 27-D). Wie kann man aber das Ziel der Arzneimitteltherapie, sichere und wirksame Dosierungen für die optimalen individuelle Behandlung zu finden, realisieren? Hier nimmt gerade in der Anwendung ein Ansatz „Fahrt auf“, der im Englischen *model-informed precision dosing* heißt und mit Modellunterstützter Dosis-Individualisierung übersetzt werden könnte.

Ein Beispiel aus der Onkologie

Paclitaxel ist ein Wirkstoff, der zur Behandlung einer Lungenkrebsart eingesetzt wird, indem er die Zellteilung von Krebszellen hemmt. Als unerwünschte Nebenwirkung hemmt er auch die Zellteilung gesunder Körperzellen. Dies kann zu schweren Nebenwirkungen führen und dazu, dass die Dosis reduziert werden muss. Eine dosislimitierende Nebenwirkung ist die Neutropenie, eine drastische Reduzierung der Neutrophilen (Zellen des Immunsystems) mit potentiell lebensbedrohlichen Konsequenzen. Um diese (Neben)Effekte in der Therapieplanung zu berücksichtigen, nutzt man Modelle, d.h. vereinfachte Abbilder der Wirklichkeit, welche die für die Fragestellung wesentlichen Aspekte nach aktuellem Wissensstand abbilden.

Abbildung 28 zeigt eine schematische Darstellung eines Modells für eine Nebenwirkung von Paclitaxel auf das Immunsystem. Der linke Teil der Graphik beschreibt (sehr vereinfacht) die Prozesse, denen Paclitaxel im Körper des Patienten unterliegt: die Aufnahme per Infusion, die Verteilung im Körper sowie die Elimination. Der rechte Teil der Graphik beschreibt (ebenfalls sehr vereinfacht) die Reifungsprozesse von Blutzellen im Knochenmark. Die Nebenwirkung von Paclitaxel, der hemmende Effekt auf die Teilung der Stammzellen und Vorläuferzellen, ist durch rote Pfeile dargestellt. Da es eine gewisse Zeit (die sog. Maturationszeit) dauert, bis sich aus den Vorläuferzellen im Knochenmark reife Neutrophilzellen entwickeln, die im Blut zirkulieren, manifestiert sich der Nebeneffekt erst mit Verzögerung, d.h. Tage nach der Gabe des Wirkstoffs, im Blut.

Der Brückenschlag zur Mathematik

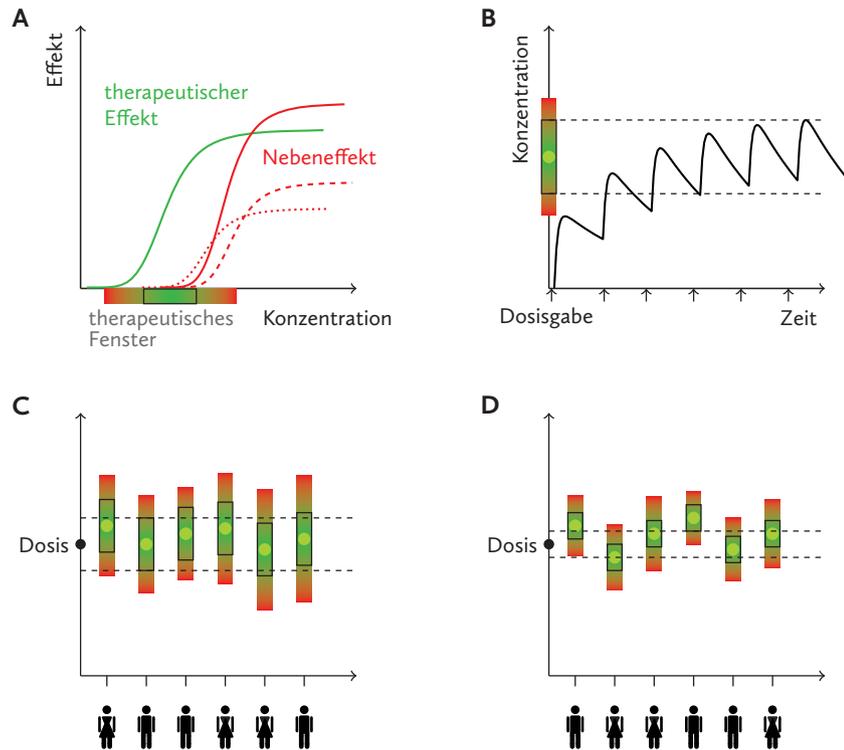
Mathematisch kann man solche Darstellungen in Funktionsgleichungen übersetzen. Diese hängen von dem Zustand des Patienten ab, z.B. der Konzentration der Stammzellen und Vorläuferzellen im Knochenmark sowie der Neutrophilzellen im Blut. Und von gewissen Parametern, nennen wir sie ϑ (ein kleines griech. Theta), wie z.B. der Maturationszeit oder wie effektiv der Wirkstoff das Zellwachstum hemmt. Nehmen wir für einen Augenblick an, wir hätten all diese Informationen, dann könnten wir das Modell nutzen, um basierend auf dem aktuellen Patientenzustand x_0 zum Zeitpunkt t_0 (Zeitpunkt der Dosierung) und den individuellen Parametern ϑ den Zustand $x(t_0 + h)$ des Patienten nach h Tagen voraussichtlich sinken wird:

$$x(t_0 + h) = f(h, x_0; \vartheta, D) + \epsilon_k.$$

Hierbei ist f die Modellvorhersage-Funktion und D die Wirkstoff-Dosis, die verabreicht wurde. Wichtig ist auch der Term ϵ_k ; er drückt Ungenauigkeiten unseres Modells, z.B. durch Modellfehler oder Messungenauigkeiten etc. aus. Durch ϵ_k können wir also keine genauen Vorhersagen machen, sondern müssen mit einer gewissen unsicheren Vorhersage leben. So, wie man es auch erwarten würde bei der Komplexität des Problems. Der vorhergesagte Patientenzustand $x(t_0 + h)$ beinhaltet alle Aspekte des Modells und kann genutzt werden, um weitere Informationen zu berechnen, z.B. wie tief die Neutrophilkonzentration N_{12} nach $h = 12$ Tagen im Blut gesunken ist:

$$N_{12} = g(x(t_0 + 12)); \quad t_0 \hat{=} \text{Zeitpunkt der letzten Dosierung}$$

Abbildung 27: Das therapeutische Fenster.
 (A) Auftreten des therapeutischen Effekts sowie der Nebeneffekte in Abhängigkeit der Wirkstoffkonzentration.
 (B) Wirkstoffkonzentration als Funktion der Zeit bei einer Mehrfachdosierung; in dem Beispiel bleiben die Konzentrationen ab der vierten Dosis im therapeutischen Fensters.
 Individuelle therapeutische Fenster mit großer (C) und kleiner (D) Schnittmenge in der Patienten/innen-Population.

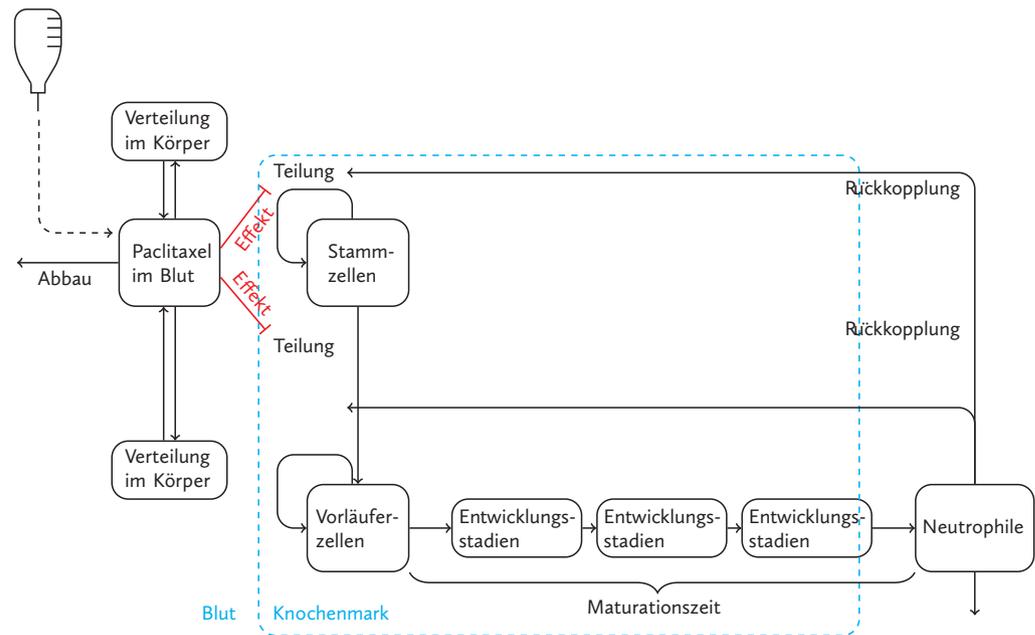


Hierbei nutzt man eine weitere Funktion g zur Informationsextraktion. Damit können wir also bis auf eine gewisse Restunsicherheit vorhersagen, wie tief die Konzentration der Neutrophilen im Blut fällt, wenn man die Wirkstoff-Dosis D verabreicht. Natürlich wird man gleichzeitig auch den Effekt auf das Tumorwachstum modellieren und berücksichtigen. Dieser Aspekt ist im obigen Modell zwecks Vereinfachung nicht dargestellt. Die Schwierigkeit besteht nun darin, die richtige Dosis zu finden, um einerseits die Krebszellen an der Vermehrung zu hindern und andererseits dafür zu sorgen, dass die Neutrophilen nicht zu sehr reduziert werden, damit der Patient noch ausreichend vor Infektionen geschützt ist.

Die bisherige Geschichte hat nur einen kleinen Haken: In der klinischen Realität kennen wir den Patientenzustand x_0 und seine individuellen Parameter ϑ nicht. Auch hier kann die Mathematik, genauer die Statistik, wieder weiterhelfen. Wir wissen zwar noch nicht, wie das Medikament bei unserem Patienten wirkt, aber wir wissen, wie es

bei vielen anderen Patienten/innen gewirkt hat. Wir können diese Information aus vorherigen klinischen Studien nutzen, um damit eine Wahrscheinlichkeitsverteilung $p(x_0, \vartheta)$ auf den Zuständen und Parameterwerten zu schätzen. Diese gibt an, wie häufig ein gewisser Zustand bzw. Parameterwert bisher beobachtet wurde. Zu Therapiebeginn t_0 nehmen wir an, dass die einzelnen Zustände und Parameterwerte genauso wahrscheinlich den Zustand unseres Patienten beschreiben, wie $p(x_0, \vartheta)$ es angibt. Man kann auch sagen: $p(x_0, \vartheta)$ drückt unser Wissen/Unwissen über den anfänglichen Patientenzustand x_0 und die individuellen Parameter ϑ aus. Dies erlaubt zu Therapiebeginn bereits Modellvorhersagen, welchen Verlauf der Neutrophilkonzentration wir aufgrund bestimmten Patienten-Charakteristika (z.B. Alter, Gewicht) erwarten. Im Laufe der Therapie erhobene individuelle Patienten-Daten, z.B. die Neutrophilkonzentration durch Blutproben, werden dann genutzt, um den Zustand unseres Patienten immer besser und genauer beurteilen zu können.

Abbildung 28:
Pharmakologisches Modell für die Aufnahme, Verteilung & Elimination (Teil links von den roten Pfeilen) sowie den Effekt des Wirkstoffes (Teil rechts von den roten Pfeilen). So ein Modell kann man in mathematische Gleichungen übersetzen: jeder Kasten entspricht einer Variablen, jeder Pfeil einem Term in den zugrundeliegenden mathematischen Gleichungen. Die graphische Darstellung erlaubt Pharmazeuten/innen und Mathematiker/innen, auf einer gemeinsamen Grundlage über das Modell zu sprechen.



Modell-unterstützte Dosis-Individualisierung

Wie kann nun die Ärztin eine informierte Dosisentscheidung auf Basis dieser zwei unterschiedlichen Quellen (a-priori Modellvorhersage, individuelle Patienten-Daten) treffen? Hier kommt nun wieder die Mathematik ins Spiel: In der Bayes'schen Statistik nutzt man dazu folgende Formel:

$$\underbrace{p(\vartheta, x_0|y)}_{\text{Wissensupdate}} = \frac{\overbrace{p(y|x_0, \vartheta)}^{\text{indiv. Daten}} \cdot \overbrace{p(x_0, \vartheta)}^{\text{Vorwissen}}}{\underbrace{p(y)}_{\text{Normalisierung}}}$$

um das allgemeine Vorwissen von vielen vorangegangenen Patienten mit individuellen Patientendaten y zu verknüpfen und das Wissen über unseren Patienten zu aktualisieren. Den Teilbereich der Mathematik, der sich mit der Verknüpfung von Modellen und Daten beschäftigt, nennt man Datenassimilation. Entwickelt wurden Datenassimilationstechniken in der Meteorologie (Wettervorhersage) sowie der Geologie. Ähnlich kann zur individuellen Therapieplanung die

Modellvorhersage mit jeder gemessenen Neutrophilkonzentration des Patienten verbessert werden und verschiedenen Szenarien für mögliche zukünftige Dosisgaben generiert werden. In Abbildung 29-A sieht man die gemessenen Neutrophilkonzentrationen (Kreuze). Der grüne Bereich zeigt die Simulation (inkl. verbleibender Unsicherheit) basierend auf dem aktualisierten Wissen. Die kleinen schwarzen Pfeile geben an, wann der Wirkstoff dosiert wurde: der rechte Teil ohne Kreuze ist also eine Vorhersage in die Zukunft (basierend auf der bisherigen Dosis). Die roten Bereichen geben den Grad der Nebenwirkung an; je dunkler das Rot, desto schwerer der Grad. Der Grad der Nebenwirkung bezieht sich dabei auf die niedrigste Neutrophilkonzentration nach der Wirkstoffgabe. Wie bereits erwähnt, sind alle Vorhersagen mit Unsicherheit behaftet. In Abbildung 29-C ist u.a. die Wahrscheinlichkeit dargestellt, mit der man Nebenwirkungen vom Grad 0-4 erwartet (also z.B. mit 60 % Grad 3). Man kann auch vorhersagen, wann der Patient sich wieder von den Nebenwirkungen erholt hat, d.h. wann die Neutrophilkonzentration wieder nahe am Ausgangslevel ist. Das ist für die Therapieplanung in einem Krankenhaus von großem Interesse, denn die nächste Dosis kann erst verabreicht werden, wenn die Level wieder hoch genug sind.

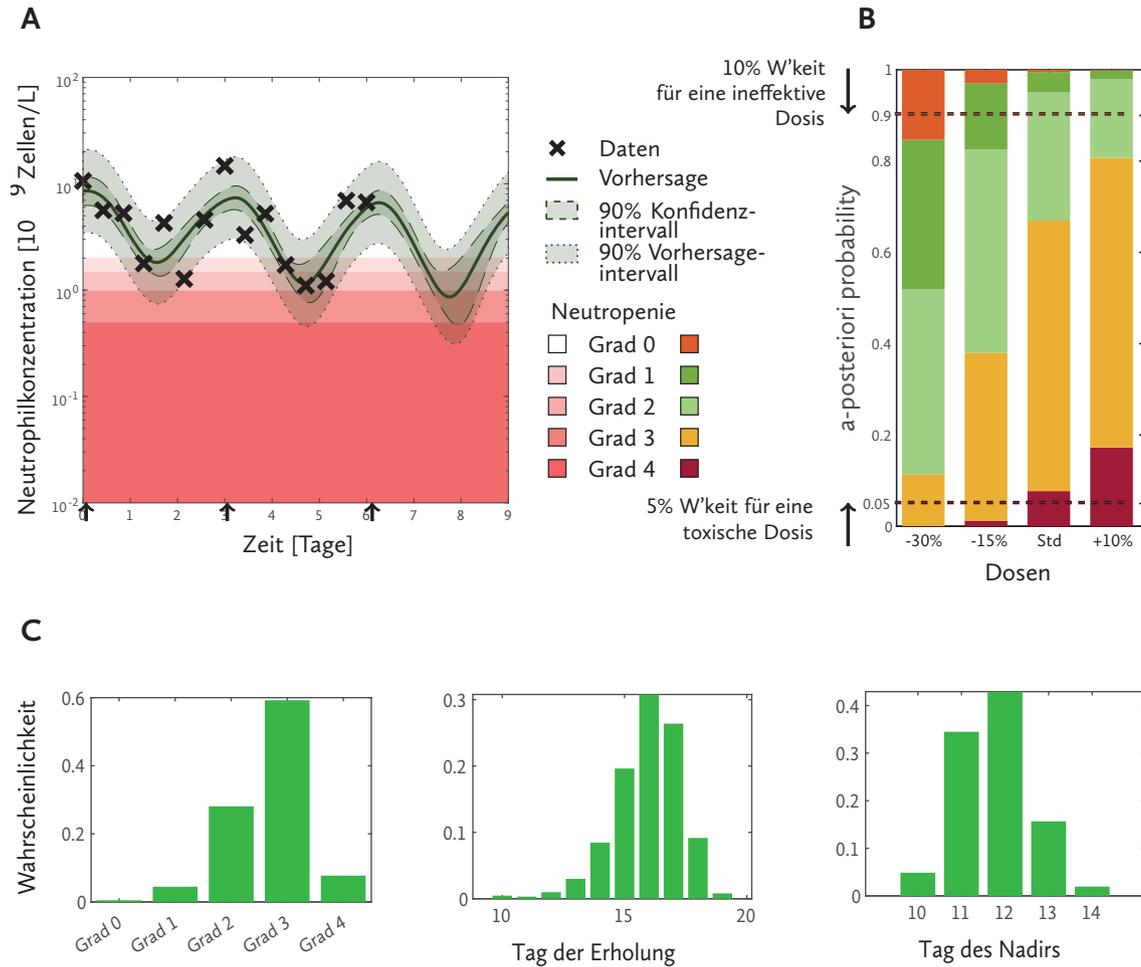


Abbildung 29: Unterstützung bei der Entscheidungsfindung für die optimale individuelle Dosis.

- (A) Individuelle Daten (Kreuze) und Modellvorhersage als Funktion der Zeit.
 (B) Vorhersage der Neutropeniegade für die bisherige Dosis und Alternativen (2x reduzierte Dosis, 1x erhöhte Dosis).
 (C) Mit dem Modell vorhergesagte Wahrscheinlichkeit der Neutropeniegade, des Tages der Erholung sowie des Tages mit der niedrigsten Neutrophilkonzentration (siehe auch Gl. mit N12 weiter vorne im Text).¹⁶

Die Ärztin muss bei der Therapieplanung viele unterschiedliche Aspekte berücksichtigen (Allgemeinzustand des Patienten, Veränderung der Tumorgröße, andere Nebenwirkungen, etc). Bezüglich der Neutropenie würde sie die Dosis so wählen, dass weder Grad 4 auftritt (da lebensbedrohlich), noch Grad 0 (da dann der Wirkstoff zu gering dosiert wäre, um das Tumorwachstum hinreichend zu hemmen). In Abbildung 29-B wurde das Modell genutzt, um für vier verschiedene Dosen die Wahrscheinlichkeit für das Auftreten von Grad 0 und Grad 4 (hell- und dunkelrot) sowie Grad 1-3 (grün und gelb) vorherzusagen. Die beiden horizontalen, gestrichelten schwarzen Linien sollen einen Zielbereich angeben: Grad 0 mit weniger als 10 % und Grad 4 mit weniger als 5 % Wahrscheinlichkeit. Unter den Vorhersagen der vier gezeigten Dosen würde man nach diesen Kriterien also die zweite von links auswählen, d.h. die bisherige Dosis um 15 % reduzieren. Derartige Vorhersagen stellen eine evidenzbasierte Entscheidungsunterstützung für die Ärztin dar, da sie Informationen aus den unterschiedlichen Quellen (klinische Studien, Patientendaten, Literatur) integrieren.

Und es gibt noch viele weitere, spannende & wichtige Probleme zu lösen: Wie kann man allgemeine Dosierungsempfehlungen aus der Gesamtheit der individuellen Dosierungen ableiten? Wie kann man Wissen aus der Therapieindividualisierung vorangegangener Patienten/innen nutzen, um die Therapie in Zukunft noch besser & schneller individualisieren zu können, ohne dabei die personenbezogenen Patientendaten teilen zu müssen? An derartigen Forschungsfragen wird z.B. im Rahmen des Promotionsprogramms PharMetriX: Pharmacometrics & Computational Disease Modelling Berlin/Potsdam¹⁷ sowie des DFG-Sonderforschungsbereiches zum Thema Datenassimilation¹⁸ an der Universität Potsdam geforscht. In der pharmazeutischen Industrie beschäftigen sich spezielle Pharmakometrie-Abteilungen mit der Analyse der Daten aus klinischen Studien, um die Entscheidungsfindung im Entwicklungsprozess eines Medikaments zu unterstützen.

Wer an der Mathematik mit Anwendung Pharmazie Interesse hat, sollte ein grundständiges Mathematik-Studium wählen, z.B. mit Nebenfach in Richtung Biologie, oder falls möglich Pharmazie. Wichtig ist eine solide Grundausbildung; spezialisieren kann man sich später. Am wichtigsten aber ist die Motivation und Begeisterung für die Mathematik und ein echtes Interesse an der Anwendung.

¹⁵ <https://www.sciencemediacenter.de/alle-angebote/factsheet/details/news/anzneimittel-von-der-entwicklung-bis-zur-zulassung/>

¹⁶ Abbildung basiert auf Artikel von C. Maier et al., Bayesian Data Assimilation to Support Informed Decision Making in Individualized Chemotherapy (CPT:PSP), 2020.

¹⁷ <http://www.phametrx.de>

¹⁸ <http://www.sfb1294.de>



Intelligentere Planung macht das Gesundheitswesen effizienter und sicherer

Der Gesundheitssektor sieht sich mit substantiellen Herausforderungen wie hohem Arbeitsdruck, langen Wartezeiten und hohen Kosten konfrontiert. Jedoch kann eine bessere Planung auf Grundlage von Wahrscheinlichkeiten diese Probleme lösen.

In den nächsten zwei Jahrzehnten werden die Gesundheitskosten schätzungsweise von zehn auf zwanzig Prozent unseres Brutto-sozialprodukts steigen; eine enorme Steigerung der Ausgaben. Besonders teuer ist der OP-Bereich in Krankenhäusern, der mehr als 40 % aller Krankenhauskosten verursacht. Ein Grund dafür ist die Komplexität der Operationen, bei denen die richtige Ressource in der richtigen Menge zur richtigen Zeit für den richtigen Patienten vorhanden sein muss. Der Personalmangel führt oft zu Situationen, in denen Operationssäle frei bleiben, was das Krankenhaus noch mehr belastet: Die angesammelten Verzögerungen am Tag erzwingen Überstunden. Das hat nicht nur Kosten für das Krankenhaus zur Folge, sondern erzeugt auch Stresssituationen für Personal und Patienten und wirkt sich negativ auf die Qualität der Versorgung aus.

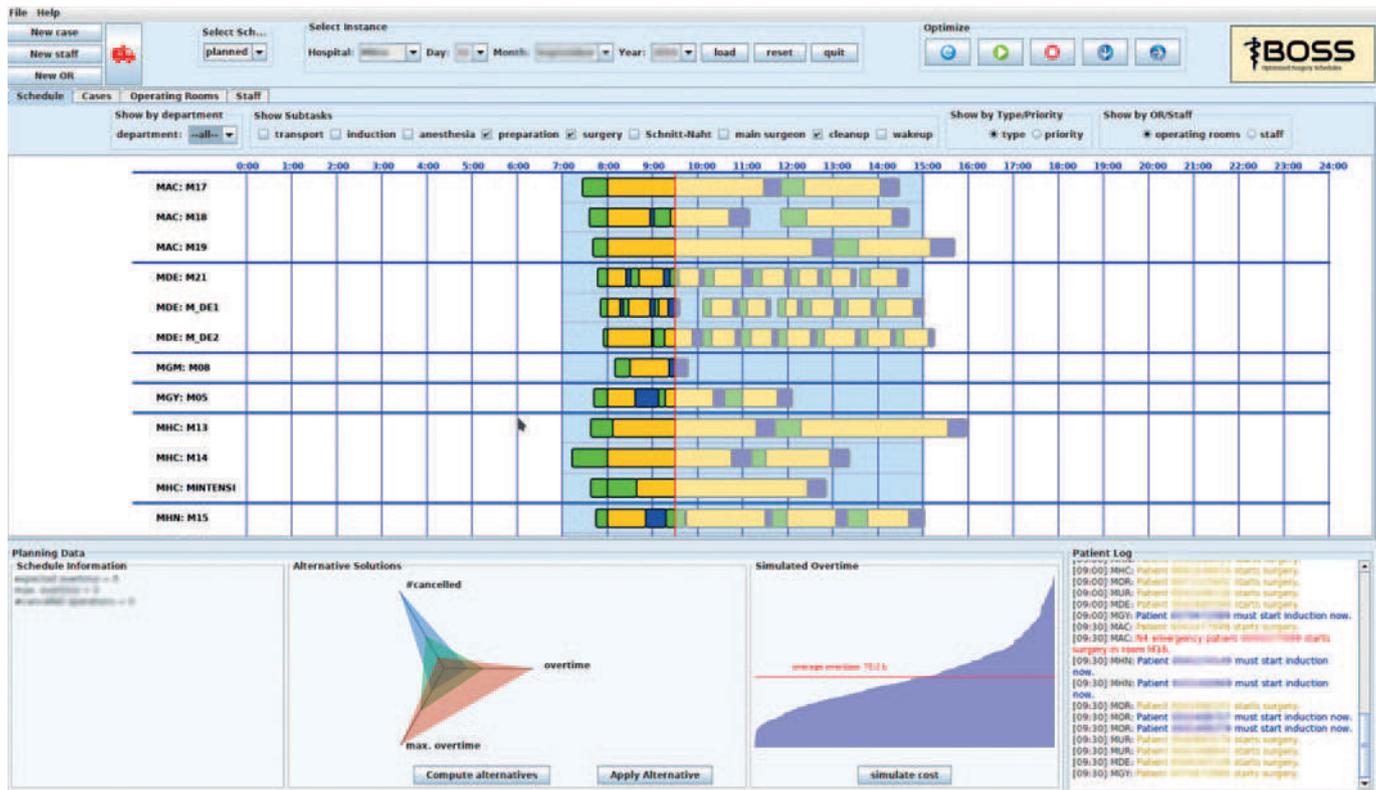


Abbildung 20: Prototyp der entwickelten Planungssoftware.

OP-timierung

Dennoch sind Prof. Ralf Borndörfer (FU Berlin & ZIB, Zuse-Institut Berlin), Dr. Guillaume Sagnol (TU Berlin) und Alexander Tesch (ZIB) überzeugt, dass die Mathematik helfen kann, dieses Problem zu lösen, indem die personellen und materiellen Ressourcen besser genutzt werden. In Zusammenarbeit mit der Charité-Universitätsklinikum Berlin führten sie ein Projekt durch, in dem ein Algorithmus zur Erstellung von optimierten OP-Plänen anhand von individuellen Zielkriterien und Nebenbedingungen entwickelt wurde; siehe Abbildung 20. Ziel ist eine effizientere Zuordnung, Sequenzierung und Steuerung von OP-Prozessen bezüglich der gegebenen Personal- und Sachressourcen. Doch was kann ein Algorithmus besser als ein erfahrener OP-Planner? Betrachten wir ein Beispiel: Angenommen, zehn Operationen müssen für den morgigen Tag auf drei OP-Säle

verteilt und sequenziert werden. Für dieses kleine Szenario gibt es bereits rund 250 Mrd. verschiedene OP-Pläne. Bei 55 Operationen und 15 OP-Sälen sind es ca. $1,96 \cdot 10^{87}$ verschiedene Pläne, was ungefähr der Anzahl an Atomen im Universum entspricht. Und dabei ist die Unsicherheit in den OP-Dauern noch gar nicht mit eingerechnet.

Demnach gibt es weitaus mehr mögliche OP-Pläne, als der Mensch bewerten kann. Ein Algorithmus ist da wesentlich schneller, auch wenn das Testen aller möglichen OP-Pläne selbst auf einem Supercomputer Jahre dauern würde. Und genau hier kommt die Mathematik ins Spiel: Mit cleveren Methoden kann die Suche deutlich eingeschränkt werden, sodass am Ende eventuell nur ein paar Tausend OP-Pläne evaluiert werden müssen. Auf einem heutigen Standard-PC geht das innerhalb von Sekunden oder Minuten, je nach

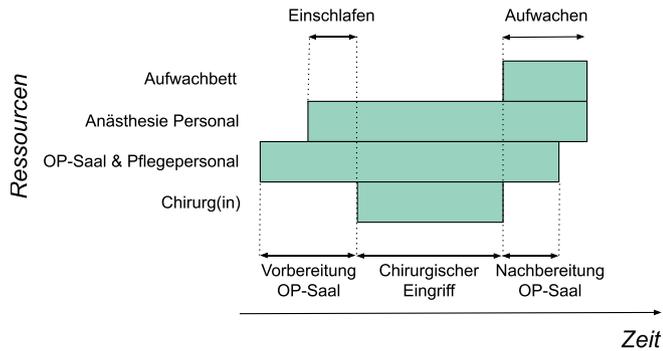


Abbildung 21: Tetris-Darstellung einer OP.

Größe des Problems. Der reine Zeitaufwand für die Erstellung eines OP-Plans kann damit deutlich reduziert werden.

Schematisch können wir uns diese Planungsprobleme als ein Tetris-Spiel vorstellen (siehe Abbildung 21), bei dem die Form der Spielsteine den Bedarf an verschiedenen Ressourcen im Laufe der Zeit darstellt. Das Ziel ist es, die Teile so gut wie möglich zusammenzufügen, und zwar nach Kriterien, die ausgewählt und angepasst werden können (Anzahl der Überstunden, Leerlaufzeit des Personals, Wartezeit der Patienten, Stabilität des Zeitplans,...). Nun stellt sich die Frage, wie ein „optimaler“ OP-Plan aussieht. Jede beteiligte Person hat wahrscheinlich eine andere Vorstellung des idealen Zeitplans: Der Planer möchte die OP-Ressourcen bestmöglich nutzen, das Personal möchte pünktlich Feierabend machen, und der Patient erwartet eine bestmögliche Behandlungsqualität. In der Praxis wird es jedoch keinen OP-Plan geben, der die Wünsche jedes Einzelnen maximal befriedigt, da die Zielkriterien oft im Konflikt zueinander stehen. Algorithmen erlauben es, OP-Pläne bezüglich einer gewünschten Gewichtung der Zielkriterien zu optimieren. Dadurch können OP-Pläne generiert werden, die bestmögliche Kompromisse zwischen mehreren Zielkriterien darstellen. Entspricht der berechnete OP-Plan nicht dem gewünschten Ergebnis, kann die Gewichtung angepasst werden, bis man eine zufriedenstellende Lösung erhält.

Zufall und Reaktivität

Unsichere Ereignisse haben einen starken Einfluss auf die operative OP-Planung. Selten werden OP-Pläne genauso durchgeführt wie am Vortag geplant. Einer der Hauptgründe ist die oft hohe Varianz in den OP-Dauern, die eine exakte Planung erschwert. Die Mathe-

matik hilft dabei, die OP-Dauern vorherzusagen. Dazu bedarf es vor allem eins: Daten, je mehr desto besser. Bei ausreichend großem Datensatz kann eine Wahrscheinlichkeitsverteilung für jede OP-Dauer berechnet werden. Eine Wahrscheinlichkeitsverteilung ist dabei viel aussagekräftiger als ein fixer Durchschnittswert, da Unsicherheit zu verschiedenen Graden auftritt. Das heißt, ein OP-Plan, der bezüglich fixer OP-Dauern auf dem Blatt gut aussieht, könnte in der Realität nicht umsetzbar sein. Durch häufige Simulationen des OP-Tages können Algorithmen die vorhandene Unsicherheit im System besser quantifizieren. Dadurch können wiederum bessere Planungsentscheidungen getroffen werden, ähnlich dem Risikomanagement einer guten Partie Poker.

Ein wichtiger Punkt beim Umgang mit Unsicherheiten ist die Suche nach reaktiven Strategien, die sich an die beobachteten Informationen anpassen. Eine optimale Strategie könnte z.B. wie folgt aussehen: „Wenn die OP in Raum A zuerst beendet wird, führe OP X in Raum A aus; andernfalls führe OP X in Raum B aus.“ Eine solche Anpassungsfähigkeit ist jedoch in einem Krankenhaus nicht immer möglich, z.B. um eine Chirurg-Patienten-Bindung zu gewährleisten. Dieser Gewinn durch Anpassungen kann auch theoretisch untersucht werden. So hat das Team von Dr. Sagnol bewiesen, dass bei einem Kriterium mit fixen Kosten pro OP-Saal und variablen Kosten für Überstunden eine einfache nicht-adaptive Strategie existiert, die höchstens $100 e^{-1} \% \approx 36 \%$ verliert, im Vergleich zu einer optimalen adaptiven Strategie. Wenn andererseits das Ziel ist, alle Operationen so früh wie möglich zu beenden, dann kann jede Strategie, die eine Umplanung mit einer gewissen Verzögerung berücksichtigt, beliebig schlecht sein. Jedoch wächst der Verlust der Optimalität sehr langsam (und zwar doppelt logarithmisch) mit der Anzahl der Operationssäle.

Bessere Bilder

Neue Berechnungsmethoden machen medizinische Scans schneller und genauer. Sie benötigen weniger Fotos, um ein gutes Bild zu produzieren. Auf diese Weise helfen diese mathematischen Methoden bei der Erforschung von Knochenerkrankungen.

Die Bevölkerung altert und Knochenkrankheiten treten immer häufiger auf. Zum Beispiel leidet jede dritte Frau über sechzig Jahre an Osteoporose, einem Zustand, bei dem der Knochen langsam zerfällt. Forscher versuchen genau herauszufinden, was bei dieser Krankheit mit dem Knochen passiert. Sie wollen immer und immer wieder den gleichen Knochen scannen, aber bei all diesen Scans wird die Strahlungsmenge auf den Körper unakzeptabel hoch. Der Mathematiker Joost Batenburg, Forscher am Zentrum für Mathematik und Informatik und Professor in Leiden und Antwerpen, entwickelt deshalb Methoden, die mit weniger Strahlung arbeiten: „In kleinen Studien ist es nun gelungen, die Strahlendosis um den Faktor drei zu reduzieren.“

Arbeiten mit Vorkenntnissen

Niederländische Krankenhäuser führen jedes Jahr mehr als eine Million CT-Scans durch. Ein CT-Scanner macht Röntgenaufnahmen aus verschiedenen Winkeln, allesamt lose Körperteile. Danach besteht der Trick darin, aus diesen zweidimensionalen Fotografien das vollständige dreidimensionale Bild zu rekonstruieren.

Der Körper besteht aus allen möglichen Geweben, jedes mit seiner eigenen Dichte, die im Scan alle eine unterschiedliche Graufarbe erhalten. Doch bei der diskreten Tomographie, dem Fachgebiet von Joost Batenburg, geht man davon aus, dass es nur eine begrenzte Anzahl von unterschiedlichen Dichten gibt: „Das funktioniert zum Beispiel gut mit Knochen. Es hat überall etwa die gleiche Dichte

und ist in kleinem Maßstab sehr porös. Sie sehen auf den Scans ein Objekt mit vielen Löchern und brauchen nur zwei Farben: es gibt Knochen oder keinen Knochen.“

Durch die Nutzung dieses Vorwissens über die Struktur der Knochen nimmt die Gesamtzahl der möglichen Lösungen drastisch ab: Statt allerlei Grautöne verwenden Sie jetzt nur noch Schwarz und Weiß. Das macht es viel einfacher, aus allen Möglichkeiten die richtige Lösung zu finden. Auf diese Weise ist es möglich, aus weniger Fotos das richtige dreidimensionale Bild zu berechnen. Der Patient liegt kürzer im Scanner und erhält weniger schädliche Strahlung.

Obwohl es sehr nützlich ist, mit wenigen Messdaten arbeiten zu können, scheint das Problem mit nur zwei Dichten mathematisch schwieriger zu lösen zu sein. Batenburg: „Beim Standardproblem kann die Lösung alle möglichen Graustufen haben. Es ist ein kontinuierliches Problem, und Sie können differenzieren, integrieren, was immer Sie wollen. Sobald man sich mit diskreten Problemen beschäftigt, verliert man all diese schönen Analysen.“ Dies erhöht die Berechnungszeit erheblich. Um eine Vorstellung vom Ausmaß dieser Probleme zu vermitteln: Ein Scan besteht schnell aus 2000 x 2000 x 2000 Elementen. Daraus ergibt sich ein System von acht Milliarden Vergleichen mit acht Milliarden Unbekannten. Um damit wirklich schneller rechnen zu können, ist ein tiefes mathematisches Verständnis erforderlich.

Die goldene Mitte

Andere medizinische Bilder haben genau die gleichen Probleme. In der Wissenschaft sind in den letzten Jahren leistungsfähige Algorithmen entwickelt worden, die wenig Daten verarbeiten können, aber viel Rechenzeit benötigen. Das macht sie in der Praxis unbrauchbar. Mit diesen Methoden kann ein Patient schneller aus dem Scanner herauskommen, müsste dann aber tagelang auf die Ergebnisse warten, weil der Computer noch rechnet.

Die altmodischen Methoden sind schnell, aber sie brauchen viel mehr Bilder, um ihre Arbeit zu tun. Joost Batenburg arbeitet mit Doktoranden an einer goldenen Mitte: Er passt die alten Berechnungsmethoden so an, dass sie ihre Geschwindigkeit mit der neuen Technik kombinieren, die mit weniger Fotos gute Ergebnisse erzielt. „Der Standardalgorithmus aus den siebziger Jahren ist eine *gefilterte Rückprojektion* und enthält, wie der Name schon sagt, einen Filter. Wir passen nur diesen Filter an und lassen den Rest des Algorithmus



Abbildung 22: Dieser Abschnitt des Oberschenkelknochens (einer Maus) ist das Ergebnis einer diskreten tomographischen Rekonstruktion auf der Grundlage von Daten aus dem Scan.

unverändert. Das macht es einfach, den Code zu ändern, der sich bereits in allen möglichen Geräten befindet.“ Mit dieser Lösung kann eine ganze Reihe von Methoden, die zunächst nur wissenschaftlich sinnvoll waren, in der Praxis angewendet werden. Auf diese Weise kann die Berechnung bis zu hundertmal schneller durchgeführt werden.

Die neuen Methoden werden in der Praxis noch nicht eingesetzt. Wenn Ärzte von ihnen hören, fragen sie, ob alles bereits klinisch getestet wurde. Ohne diese Tests ist es ihnen nicht erlaubt, die Technik anzuwenden. Die medizinische Welt ist konservativ, und Veränderungen gehen nur langsam vonstatten. Es dauert viele Jahre, bis Innovationen in der Praxis zum Einsatz kommen. Batenburg: „Sie müssen sich viel Zeit lassen, aber ich rechne damit, dass die Krankenhäuser unsere Methoden innerhalb von zehn Jahren anwenden werden.“

Der Kopf isst mit

Der Appetit auf Nahrung sinkt bereits Wochen vor dem Beginn einer Diät, steigt aber auch Wochen vor dem Ende der Diät wieder an. Dies ist das Ergebnis des ersten mathematischen Modells des Ess-Sinns.

Als der Mensch noch als Jäger und Sammler lebte, war es klug sich satt zu essen, wenn es gerade einen Überfluss an Nahrung gab. Knappheit war damals die Regel, Überfluss die Ausnahme. Aber jetzt, da der westliche Mensch ständig einen Überfluss an Lebensmitteln hat, ist es sehr unklug, diesen Überfluss jedes Mal zu essen, egal wie schmackhaft unser Appetit ist. Für viele Menschen ist das schwierig, wenn man bedenkt, dass in Deutschland rund ein Viertel der Erwachsenen (23 % der Männer und 24 % der Frauen) stark übergewichtig (adipös) ist²¹.

Die Liste der Diäten zur Bekämpfung all dieses Übergewichts geht in die Dutzende: von Atkins über Montignac und Sonja Bakker bis hin zur salzlosen Diät oder Intervallfasten. Die Menschen, die eine Diät machen, führen oft wie ein Buchhalter Buch darüber, wie viele Kalorien sie täglich zu sich nehmen und häufig auch, wie viel sie durch tägliche Bewegung und Sport verbrennen.

Jojo-Kurven

Es fiel dem Mathematiker Johan Grasman von der Universität Wageningen auf, dass Menschen zwar Buch führen, aber eigentlich keine

Ahnung haben, wie sich ihr Appetit auf Essen mit der Zeit entwickelt. „Ob wir essen wollen oder nicht, ist zum Teil ein physischer Prozess, zum Teil aber auch ein mentaler Prozess, eine bewusste Entscheidung“, sagt Grasman. „Ich wollte eine Formel entwickeln, die den Wunsch zu essen beschreibt. Wenn wir wissen, wie sich der Wunsch zu essen verändert, dann können wir vielleicht besser darauf reagieren, wie eine Person sich am besten aufstellen kann.“

Menschen, die an Gewicht zunehmen, nehmen mehr Kalorien auf, als sie verbrennen. Die Differenz wird im Fettgewebe gespeichert. Die Fettzellen bilden das Hormon Leptin. Wenn der Körper genügend Fett produziert hat, gibt Leptin das Signal, weniger zu essen. Leptin ist ein wichtiges, aber nicht das einzige Hormon, das zur Esslust beiträgt.

Der entscheidende neue Faktor in dem Modell ist der „Drang“. Unter diesem Begriff fasst Grasman alle Reize (außer Leptin) zusammen, die dazu führen, dass jemand isst oder in manchen Fällen nicht isst. Das können physiologische Reize sein, die den Körper zum Essen veranlassen (z.B. Insulin), oder psychologische Reize wie der Wunsch, Gewicht zu verlieren.

Grasman modellierte die sog. 'Linien' als ein System von *Differentialgleichungen*, die den Wunsch zu essen („Drang“), die Kalorienaufnahme und den Kalorienverbrauch als Funktion der Zeit beschreiben. Eine (gewöhnliche) Differentialgleichung enthält unbekannte Funktionen und deren Ableitungen, meist bezüglich der Zeit.

Die Variablen in seinem Modell sind die Menge der im Fettgewebe gespeicherten Kalorien und der Pegel des Plasma-Leptins. Das Modell hat als Input einen Antrieb, der die Nahrungsaufnahme steuert. Dieser Antrieb besteht aus einer Sammlung von physiologischen und psychologischen Anreizen, zu essen oder mit dem Essen aufzuhören.

Um einige der Parameter in seinem Modell abzuschätzen, verwendete er Daten aus der wissenschaftlichen Literatur von zwei Frauen, die über einen Zeitraum von zehn Jahren fünfmal versuchten ihr Gewicht zu reduzieren. Das Ergebnis waren die bekannten und ungesunden Jo-Jo-Kurven: Gewichtsverlust – Zunahme – Verlust – Zunahme des Körpergewichts über diesen Zehnjahreszeitraum. In der Folge konnte Grasman mit seiner Formel den Wunsch zu essen über den gleichen Zeitraum von zehn Jahren berechnen. Nach einer Phase des Abnehmens fallen Diätwillige oft in alte Essgewohnheiten zurück. Mit dem Modell von Grasman könnte man besser verstehen, wie der Essensdrang eines Menschen funktioniert, und evtl. daraus neue Strategien ableiten.

Abnehmen beginnt im Kopf

Die Schlussfolgerungen erwiesen sich als überraschend. Grasman: „Zunächst einmal scheint der Appetit auf Nahrung Wochen vor Beginn der Diät nachzulassen. Anscheinend hatten diese Frauen schon seit geraumer Zeit darüber nachgedacht. Zweitens stellt sich heraus, dass der Appetit schon Wochen und manchmal sogar Monate vor dem Ende der Diät wieder rasch zunimmt. Anscheinend ist der Verstand wieder schwach geworden. Diese beiden Schlussfolgerungen zeigen, wie wichtig der Verstand ist. Denken Sie nur an das 'Trostessen': Wenn es Ihnen eines Tages nicht schmeckt, können Sie eine Süßigkeit essen, damit Sie sich besser fühlen.“

Laut Grasman sind die psychologischen Prozesse, die an den Linien beteiligt sind, viel größer als die Rolle der physischen Prozesse: „Das Hormon Leptin reduziert die Lust zu essen nur um wenige Prozent. Die Ernährung reduziert die Esslust um etwa 10 %. Aber der Rest, mehr als 80 % des Verlangens zu essen, hat mit dem zu tun, was der Verstand will.“

Traditionell ließen die Physiologen den Verstand aus, wenn sie die Linien betrachteten, und die Psychologen ließen die Physiologie aus. Grasman ist der erste, der Körper und Geist gleichzeitig modelliert. Er hat bereits Ideen, wie man den Esssinn weiter aufbrechen kann, so dass er besser zeigen kann, welchen Beitrag die verschiedenen Hormone leisten, wie Kohlenhydrate und Fette getrennt voneinander zur Gewichtszunahme beitragen und vielleicht sogar psychologische Prozesse entwirren kann.

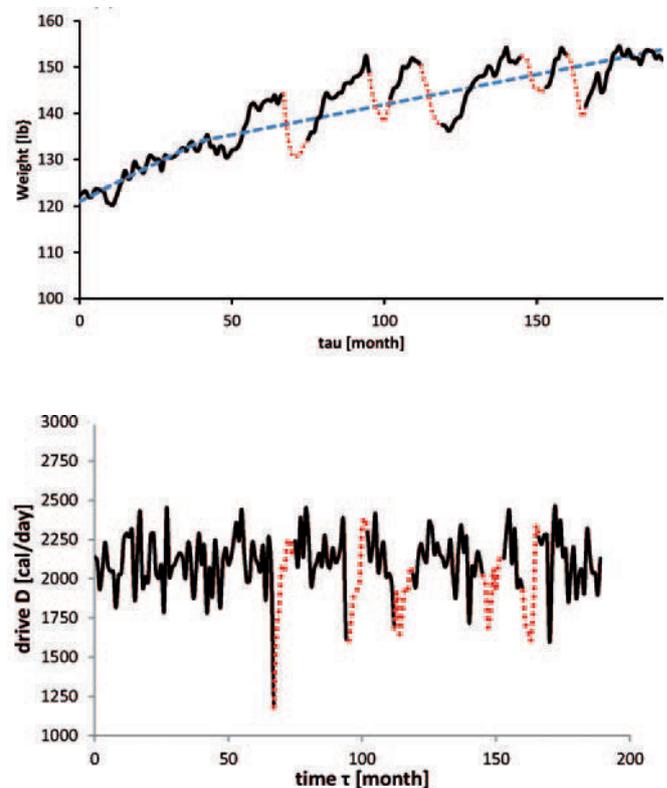


Abbildung 23: Zwei Diagramme zu Gewicht und 'Esslust'. Das erste zeigt das Gewicht als Funktion der Zeit. Das zweite zeigt den Wunsch zu essen als Funktion der Zeit. Die roten Punkte sind die Daten bei den bewussten Versuchen, Gewicht zu verlieren.

²¹ https://www.rki.de/DE/Content/Gesundheitsmonitoring/Themen/UEbergewicht_Adipositas/UEbergewicht_Adipositas_node.html



Die richtige Dosierung von Antibiotika spielt bei der Behandlung von infektiösen bakteriellen Krankheiten eine wichtige Rolle, da sowohl eine zu geringe, als auch zu hohe Dosierung das Ergebnis der Behandlung negativ beeinflussen kann.

Die optimale Dosisfindung von Antibiotika

Niels Neveling hat in seiner Bachelorarbeit an der Bergischen Universität Wuppertal untersucht, wie man mathematisch die optimale Dosisfindung von Antibiotika bestimmen kann²². Wie oft sollte man welche Dosis eines Wirkstoffs verabreichen?

Eine zu niedrige Dosis kann den Erreger nicht vollständig eliminieren und somit unter Umständen zur Selektion von resistenten Bakterien führen. Andererseits liefert eine zu hohe Dosis unter Umständen kein verbessertes Ergebnis oder schädigt sogar den Makro-Organismus (z.B. Menschen), da manche Nebenwirkungen dosisabhängig sind. Daher ist es wichtig, eine optimale Dosisfindung durch ein mathematisches Modell zu unterstützen und die Effektivität der Dosierung zu optimieren.

Zunächst beschreibt die Pharmakokinetik (PK) die zeitabhängige Verteilung des Wirkstoffs, also der antimikrobiellen Substanz, im Organismus. Daraus lassen sich dann wichtige Indizes ableiten, nämlich die maximale Konzentration (C_{max}), die Fläche unter der Serumkonzentration-Zeit-Kurve (engl. area under the curve, AUC), also die Belastung des zugeführten Wirkstoffs in einem bestimmten Zeitintervall und die Halbwertszeit $T_{1/2}$. Die ersten beiden Indizes hängen von der täglichen Dosis und dem Dosis-Intervall ab, wobei Letzteres vom Wirkstoff und dem Organismus, in das der Wirkstoff appliziert wird abhängt.

Die Pharmakodynamik (PD) beschreibt den Effekt des Wirkstoffs, z.B. die Abtötungsrate des Erregers, verglichen mit dessen Konzen-

tration im Körper. Die minimale Hemmkonzentration (MIC) gibt an, ab welcher Konzentration die Substanz (sog. Bakteriostatikum) jegliches Wachstum der bakteriellen Population hemmt. Der MIC-Wert beschreibt zwar die Potenz eines Wirkstoffs, liefert jedoch keine Informationen über dessen zeitabhängige Aktivität. Die Zeitabhängigkeit der Wirkung eines Antibiotikums lässt sich mit mehrfacher Probenentnahmen ermitteln und die so gewonnenen Keimzahlreduktionskurven werden als „Abtötungskurven“ (engl. „time-kill curves“) bezeichnet.

Die Dosierungsintensität

Diese sog. PK/PD-Parameter C_{max} , AUC, MIC stellen den antibakteriellen Effekt mit der Zeit der Exposition in Verbindung und quantifizieren die Aktivität des Antibiotikums. Die C_{max}/MIC -Rate gibt das Verhältnis der höchsten Konzentration zur minimalen Hemmkonzentration wieder. $T > MIC$ ist der prozentuale Anteil an einem Dosierungsintervall, in dem die Konzentration des Wirkstoffs über der MIC ist. Schließlich gibt der AUC/MIC (bzw. $24\text{-}AUC/MIC$) das Verhältnis der Menge des Wirkstoffs, welchem das Bakterium über 24 Stunden ausgesetzt ist, und der MIC wieder. Mithilfe von Differentialgleichungen kann man den Zusammenhang zwischen der Konzentration des Antibiotikums und dem eintretenden Effekt modellieren und mit einer Integration über ein Dosierungsintervall bekommt man die Dosierungsintensität D . Diese Dosierungsintensität D ist ein Indikator für die Potenz eines Antibiotikums gegen einen Erreger und erlaubt Rückschlüsse auf die Abtötungseigenschaften des Antibiotikums.

Aus der Interaktion des Wirkstoffs mit dem Organismus über sog. Rezeptoren lässt sich folgende Differentialgleichung herleiten. Diese setzt die Konzentration des Wirkstoffs und die Konzentration der Wirkstoff-Rezeptor-Verbindung (nur in dieser Verbindung ist eine Wirkung auf den Erreger möglich) in Zusammenhang. Assoziation beschreibt die Verbindung von γ Wirkstoff Molekülen mit einem Rezeptor und Dissoziation die entsprechende Auflösung der Verbindung.

$$\dot{v}(t) = \overbrace{k_{+1} \cdot c^\gamma(t) \cdot [r_0 - v(t)]}^{\text{Assoziation}} - \overbrace{k_{-1} \cdot v(t)}^{\text{Dissoziation}} \quad \text{mit} \quad v(0) = 0.$$

Unter der Annahme, dass obige Differentialgleichung in kürzester Zeit in ein Gleichgewicht übergeht, kann die Konzentration des Wirkstoff-Rezeptor Komplexes in Abhängigkeit von der Wirkstoff-Konzentration bestimmt werden. Die Dosierungsintensität D ergibt sich dann als Mittelwert der Konzentration des Wirkstoff-Rezeptor Komplexes in einem bestimmten Zeitintervall τ .

$$D = \frac{1}{\tau} \int_0^\tau \frac{k_k \cdot c(t)^\gamma}{c_{50}^\gamma + c(t)^\gamma} dt.$$

Die drei Parameter (K_k , C_{50} , γ) beschreiben das pharmakodynamische Profil eines Wirkstoffs und können aus experimentellen Daten ermittelt werden.

Zur Berechnung dieser Dosierungsintensität D ist es nötig, den Abbau des Antibiotikums im Körper mithilfe eines Multi-Kompartiment Modells (ähnlich wie das SIR-Modell aus dem ersten Kapitel), zu modellieren. Die Kompartimente stellen hierbei die Teile des Körpers außerhalb des Blutkreislaufs dar, die der Wirkstoff erreicht und dort unter bestimmten Annahmen abgebaut wird. Mittels Ausgleichsrechnung (siehe Kapitel über robuste Inzidenzzahlen) werden die Funktionen anhand gegebener Messdaten (gewonnen durch Bluttests) parametrisiert. In Abbildung 24 ist für die drei Dosierungsintervalle $\tau \in \{24, 12, 6\}$ der Konzentrationsverlauf exemplarisch dargestellt. Abbildung 25 zeigt das analoge Ergebnis für ein 2-Kompartimentmodell. In Abbildung 26 sind die Kurven der Dosierungsintensitäten zu den drei Profilen dargestellt, wobei wieder entweder das Dosierungsintervall oder die tägliche Dosis festgehalten wird.

Im Rahmen der Bachelorarbeit wurde der Effekt des Antibiotikums Ciprofloxacin auf Populationen von Pseudomonaden (verursacht unter anderem eine Lungenentzündung) unterschiedlicher Empfindlichkeit im Reagenzglas untersucht. Dazu wurden zu unterschiedlichen Anfangskonzentration die Dosierungsintensitäten bestimmt. Eine Dosierung des zehnfachen MIC-Wertes der jeweiligen Population stellte sich hierbei als effizient heraus, unabhängig von der Empfindlichkeit der Erregerpopulation.

Die individuell optimale Dosis

Der für eine therapeutische Wirkung maßgebliche PK/PD Parameter ist sowohl das Verhältnis AUC/MIC, als auch C_{\max}/MIC . Diese Parameter können naturgemäß aber nur nach Applikation der Substanz als Tablette oder Injektion gemessen werden, lassen sich aber durchaus in einem im Reagenzglas Versuch nachstellen/simulieren. In den dieser Bachelorarbeit zugrundeliegenden Versuchen wurden die Indikatorstämme, die unterschiedliche MIC-Werte aufwiesen, konstanten Präparatkonzentrationen gegenüber exponiert. Als optimal für die Erreichung eines guten Therapieergebnisses, d.h. klinischen Heilung des Patienten und Vermeidung einer Resistenzentwicklung, gilt ein C_{\max}/MIC Verhältnis von 10. Diese theoretisch notwendige Dosissteigerung muss dann mit dem Nebenwirkungsprofil der Substanz abgeglichen werden und muss mit den zugelassenen Dosen übereinstimmen.

Berechnungen wie diese können gegebenenfalls auf eine notwendige Dosissteigerung hinweisen und somit dazu beitragen, die Therapie zu individualisieren, indem die für eine Keimabtötung optimale Konzentration ermittelt wird. Ebenso wird die Therapie ökonomisiert, da die Dosis individuell angepasst werden kann und zu hohe Dosen vermieden werden können. Weiterhin können unwirksame Dosen identifiziert werden, weil mit den zugelassenen Dosis schemata ein für die effektive Therapie notwendige C_{\max}/MIC -Rate nicht erreichbar ist. Schließlich kann man so auch einem Therapieversagen vorbeugen und eine Resistenzentwicklung des bakteriellen Erregers aufgrund zu niedriger Exposition vermindern.

²² Niels Neveling, Mathematische Modellierung von Pharmakokinetik und Pharmakodynamik antibiotischer Substanzen, Bachelorarbeit, Bergische Universität Wuppertal, 2015.

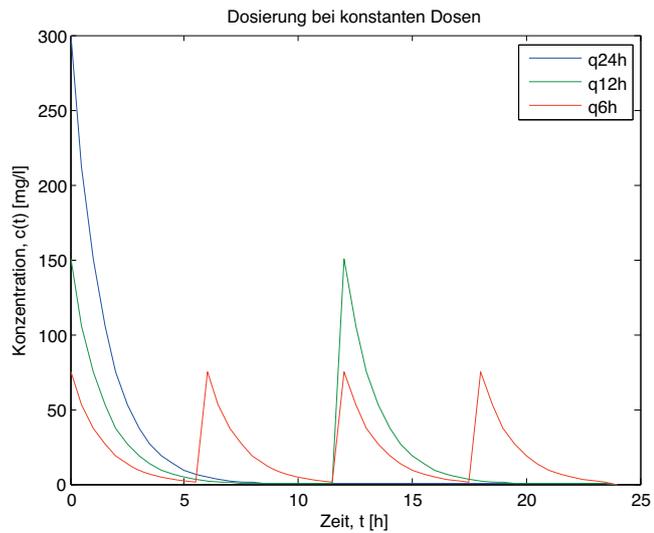


Abbildung 24: Pharmakokinetik im 1-Kompartiment Modell

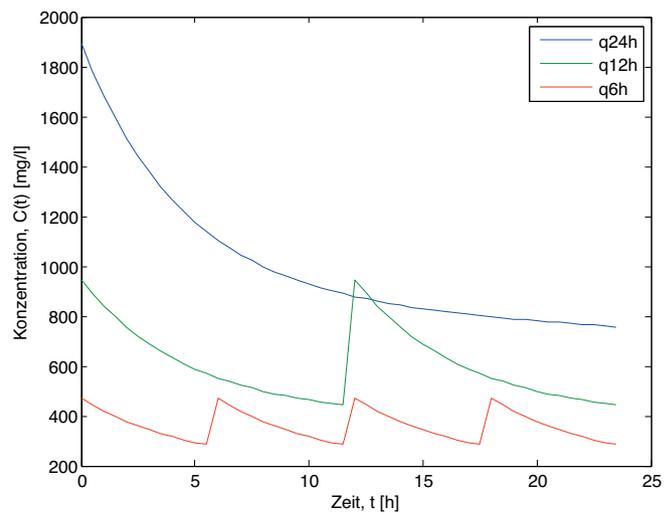


Abbildung 25: Pharmakokinetik im 2-Kompartiment Modell

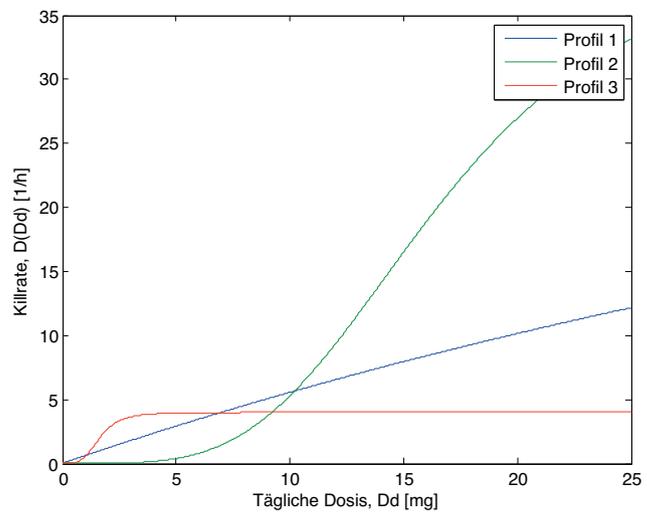
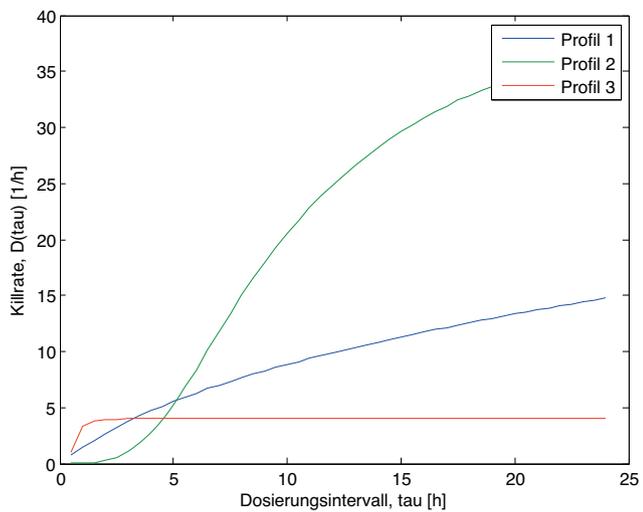


Abbildung 26: Links: 10mg tägliche Dosis (D_d), Rechts: 5h Dosierungsintervall (τ)

Dr. Markus Hoschek

Eine kurze Zusammenfassung Ihrer Karriere

Ich habe Mathematik und Elektrotechnik an der TU Darmstadt und dem ENSIMAG in Grenoble, Frankreich, studiert und danach bis zur Promotion 1999 als wissenschaftlicher Mitarbeiter der TU Darmstadt und der Siemens AG München gearbeitet. Anschließend war ich als Berater der Firma IBM international in der Beratung von Banken bei der Entwicklung von Risikomodellen und der Akquisition von Unternehmen tätig.

Seit 2008 bin ich Vorstandsmitglied der HEAG Holding AG – Beteiligungsmanagement der Wissenschaftsstadt Darmstadt (HEAG). Die HEAG versteht sich als Multidienstleister und ist zum einen Holding und Führungsgesellschaft des HEAG-Konzerns mit Sitz in Darmstadt, einer Bilanzsumme von rund 2.917 Millionen Euro und etwa 3.000 Beschäftigten im Jahr 2019. Zudem ist sie das Beteiligungsmanagement der Wissenschaftsstadt Darmstadt. Die HEAG unterstützt dabei die wirtschaftlichen Beteiligungen Darmstadts und verankert vor allem bei den mehr als 100 Mehrheitsbeteiligungen mit rund 8.000 Beschäftigten die städtischen Ziele. Somit ist die HEAG der zentrale und maßgebliche Berater der Wissenschaftsstadt Darmstadt in allen wirtschaftlichen Angelegenheiten der Stadtwirtschaft.

Im Zusammenhang mit meiner Vorstandstätigkeit bei der HEAG bin ich Mitglied in zahlreichen Aufsichtsgremien und Beiräten sowie in Verbänden der Kommunalwirtschaft. Außerdem bin ich seit 2010 Vorstandsvorsitzender der Bürgerstiftung Darmstadt.

Was hat Mathematik für Ihre Karriere bedeutet?

Mathematik war schon immer meine Leidenschaft. Sie ist für mich Verständnis der Welt, der Unternehmen und Sprache: Mathematik beschreibt viel eindeutiger als Worte. Dies gilt auch umgekehrt – was ich mathematisch erfasst und formuliert habe, lässt sich sehr prägnant in Worte fassen. Mathematik schult natürlich analytische Fähigkeiten und logisches Denken, aber auch die Fähigkeit, die entscheidenden Entwicklungen in komplexen Systemen zu erkennen und zu fokussieren. Das hat mir unabhängig von der jeweiligen Branche an



allen meinen Stationen geholfen, auch wenn ich keine klassische (Ingenieurs-) wissenschaftliche Karriere begonnen habe.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihre Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Ehrlicherweise muss man sagen, dass die mathematischen Fragen bei uns selten über ein ordentliches Schulniveau hinausgehen. Wir entwickeln keine mathematischen Werkzeuge, nur in wenigen Bereichen, wie dem Stromeinkauf oder der Netzsteuerung, werden komplexere Simulationen genutzt. Dennoch spielt Mathematik eine wichtige Rolle für die HEAG, sowohl in unserer Funktion als Konzernmutter des HEAG-Konzerns als auch als Teilnehmungsmanagement der Wissenschaftsstadt Darmstadt. Berichte zur wirtschaftlichen Lage, Investitionen, Planzahlen und Prognosen, Personalangelegenheiten, Einkauf von Waren und Dienstleistungen, Spenden, Sponsoring und die Prozesse der Digitalisierung der Arbeitswelt – zum Verständnis und zur Umsetzung all dieser Beispiele benötigen wir Mathematik. Es ist unabdingbar, diese Themen komplett zu durchdringen und auch in einfachen Worten erläutern zu können – und dabei hilft die Präzision und Fokussierung der Mathematik. All das kann man von der HEAG auch auf die gesamte Branche übertragen.

Eine neuere Entwicklung ist erfreulicherweise, dass wir vermehrt Mathematiker und Mathematikerinnen in Geschäftsleitungsfunktionen haben. Wo vor zehn Jahren noch eher JuristInnen und BWLer zu finden waren, kommen heute Mathematikerinnen und Mathematiker in die Verantwortung.

Wie sieht ein typischer Arbeitstag aus?

Mein typischer Arbeitstag ist sehr vielfältig. Termine mit den Eigentümern der HEAG – das sind größtenteils politisch Verantwortliche – und den Geschäftsleitungen unserer Teilnehmungsunternehmen gehören ebenso dazu wie die Arbeit in Aufsichtsgremien und interne Abstimmungsrunden mit Abteilungen oder Fachbereichen. Dazu kommt dann noch das Beantworten von E-Mails, das Lesen von Berichten usw. Die allermeisten der Gespräche finden auf Grund der aktuellen Situation zurzeit digital statt – auch das wäre ohne Mathematik unmöglich.

Sind, aus Ihrer Sicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Auf der einen Seite beobachte ich die positive Entwicklung, dass sich vermehrt Mathematikerinnen und Mathematiker um Positionen in

Geschäftsleitungen bewerben und sich auch in diesen Funktionen wiederfinden.

Auf der anderen Seite fällt es Mathematikern häufig noch schwer, sich so auszudrücken, dass nicht nur Fachpublikum mitgenommen wird. Mit der Bürgerstiftung Darmstadt organisieren wir beispielsweise seit vielen Jahren eine Kindervorlesung an der TU Darmstadt für Kinder im Grundschulalter. Während die Expertinnen und Experten anderer Naturwissenschaften die Kinder mit einem Feuerwerk beeindruckender Simulationen und Experimente gefesselt haben, hat ein wirklich guter Professor der Mathematik die ganze Vorlesung mit Tafel und Kreide bestritten.

Darüber hinaus hat die Mathematik meiner Meinung nach teilweise auch ein Image-Problem. Ich habe mehrfach erlebt, dass mathematische Kongresse von Politikerinnen oder Politikern eröffnet wurden, die sich dann darin „sonnten“, Mathe in der Schule nie verstanden zu haben. Solche Situationen sind unwürdig für die Politik und sollten auch genau so benannt werden.

Auch könnten Mathematikerinnen und Mathematiker vermehrt Stellung zu aktuellen, gesellschaftspolitisch wichtigen Themen beziehen. So vermisse ich in der aktuellen COVID-Diskussion eloquente MathematikerInnen, die sich zur Lage und den sich daraus ergebenden Entwicklungsvarianten äußern.

Was raten Sie jungen Leuten hinsichtlich Mathematik?

Mathematik ist die Grundlage für alle Ingenieurs-, natur-, sozialwissenschaftlichen und mathematischen Studiengänge sowie für zahlreiche Ausbildungsberufe. Ein Verständnis für Zahlen und mathematische Zusammenhänge zu entwickeln hilft aber in nahezu allen Lebensbereichen weiter. Das beginnt im Alltag natürlich mit dem Dreisatz, mit den Mengenangaben beim Kochen oder bei Rabatten im Supermarkt. Mathematisches Verständnis hilft, die Prozesse der allgegenwärtigen Digitalisierung besser zu verstehen. Man denke nur an die Suchalgorithmen bei Google oder an QR-Codes. Auch Verschlüsselungsverfahren – ob beim Handy, Computer oder mobilen Bezahlen sind ohne Mathematik undenkbar. Es ist gut, diese grundsätzlich verstehen und einordnen zu können. Ich kann nur jedem und jeder empfehlen, sich mit Mathematik zu beschäftigen und sich auch mal durchzubeißen, wenn es etwas schwieriger wird. Je besser man Mathematik beherrscht, umso besser kann man sich die Welt erklären.

Mathematik in der Chemie: Die Suche nach Synthesewegen für neue Medikamente

Nicht nur Zuschauer der Serie *Breaking Bad* sondern auch viele Wissenschaftler halten chemische Synthese für eine Art Kunstform. Gute Chemische Synthesewege zu finden, ist eine große Herausforderung bei der Suche nach neuen Wirkstoffen in Medizin oder Pflanzenschutz. Dass Mathematik und künstliche Intelligenz hier weiterhelfen können, überrascht vielleicht.

Dr. Georg Mogk, Principal Expert Applied Mathematics, Bayer AG, Leverkusen und Priv. Doz. Dr. Thomas Mrziglod, Head of Applied Mathematics, Bayer AG, Leverkusen geben ein Beispiel zur Rolle der Mathematik in der Chemie.

Mathematik haftet der Ruf an, sehr abstrakt und vielleicht auch etwas weltfremd zu sein. Das dem nicht so ist, wollen wir im Folgenden am Beispiel der Chemie zeigen. Hier kann Mathematik dabei helfen, denkbare Moleküle (z.B. als Wirkstoffe von Medikamenten) in die Realität zu holen. Das Bild des Chemikers, der in seinem Labor herumwerkelt und dabei sensationelle neue Wirkstoffe findet, gehört immer mehr der Vergangenheit an. Tatsächlich werden potenzielle Wirkstoffe am Computer auf eine bestimmte Wirkung hin designt. Hierbei werden Algorithmen verwendet und dabei oft hunderttausende Substanzen generiert, die eine bestimmte Wirkung erzielen könnten.



Abbildung 30: Synthese Roboter bei Bayer

Für die vielversprechendsten Kandidaten muss dies dann natürlich noch im Labor überprüft werden. Dazu muss jeder Wirkstoffkandidat in ausreichenden Mengen synthetisiert werden.

Die Vorhersage, ob sich ein Molekül tatsächlich synthetisieren lässt, ist allerdings sehr schwierig. Experten schätzen, dass auf ein synthetisierbares Molekül mindestens eine Milliarde Moleküle kommen, die zwar existieren könnten, aber mit dem heutigen Stand der Chemie nicht synthetisiert werden können. Dies führt uns zu zwei Fragen: *kann man vorhersagen, ob ein am Rechner designedes Molekül synthetisierbar ist?* Und wenn ja, *wie kann es synthetisiert werden?* Das Konzept, mit dem man diese Fragen beantworten kann, heißt Retro-Synthese und wurde ab den späten 1960er Jahren von dem amerikanischen Chemiker Elias James Corey entwickelt.

Bei der Retro-Synthese geht man gedanklich jede Bindung eines Moleküls durch, scheidet sie durch und fragt sich, durch welches Reaktionsschema die Bindung entstehen könnte. Weiter fragt man, wie die Ausgangsprodukte (auch Edukte genannt) aussehen müssten, damit bei einer Vorwärts-Synthese genau die gesuchte Bindung entsteht. Rekursiv wendet man nun das gleiche Verfahren auf die gerade ermittelten Edukte an, bis man schließlich bei kaufbaren Molekülen oder in einer Sackgasse landet. Das Ergebnis ist dann der sogenannte Synthesebaum, der alle denkbaren Synthesewege enthält. Der Chemiker muss dann *nur* aus dem Synthesebaum die beste Syntheseroute auswählen. Für das Konzept der Retro-Synthese erhielt Corey 1990 den Nobelpreis für Chemie.

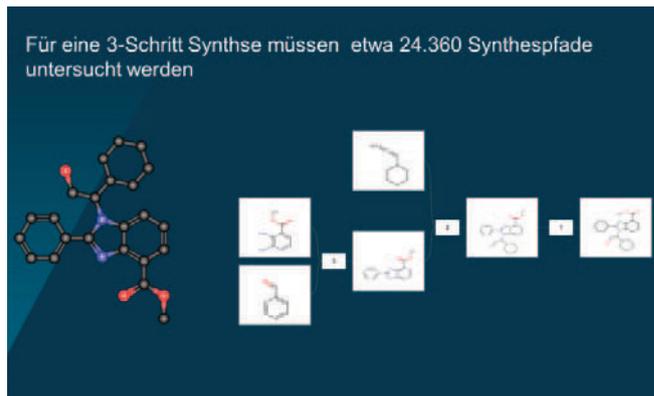


Abbildung 31: Drei-stufige Syntheseroute

So einfach das Konzept sich anhört, so hat es doch seine Tücken, die sich in zwei Kernfragen zusammenfassen lassen:

- Wo und wie kann man überhaupt ein Molekül schneiden?
- Die Synthesebäume können sehr groß werden. Gibt es einen Syntheseweg, der ausschließlich bei kaufbaren Bausteinen landet?

Hier kommt nun Mathematik ins Spiel. Da für beide Fragen extrem viele Möglichkeiten bestehen, bieten sich Computer an, um diese anzugehen. Regelbasierte Ansätze für die erste Frage stoßen auf Grund des explorierenden chemischen Wissens schnell an ihre Grenzen. Ein modernerer Ansatz ist es, den Computer die Schneideregeln selbst lernen zu lassen. Ein Update des Systems bei Weiterentwicklung der Chemie ist dann sehr einfach: Man lässt den Computer auf den neuen Daten lernen. Innerhalb der Bayer AG verfolgen wir diesen Ansatz. Dafür haben wir 19 Mio. veröffentlichte Reaktionen der letzten 250 Jahre und unsere eigenen Laborjournale genutzt, um ein Künstliches Neuronales Netz zu trainieren.

Wie der Computer Reaktionen lernt

Bevor der Computer lernen kann, sind aber einige Hindernisse zu überwinden: Die chemische Nomenklatur für Moleküle und Reaktionen ist für Menschen sehr gut zugänglich, sie muss aber erst in eine für Computer geeignete Darstellung von Molekülen überführt werden. Auch bei den zu lernenden Reaktionen gibt es Herausforderungen: In Publikationen fokussieren sich Forscher auf das Wesentliche. „Unwichtige“ Nebenprodukte, wie Wasser oder CO₂ werden oft nicht dokumentiert. Oder in den Reaktionen spielen doch plötz-

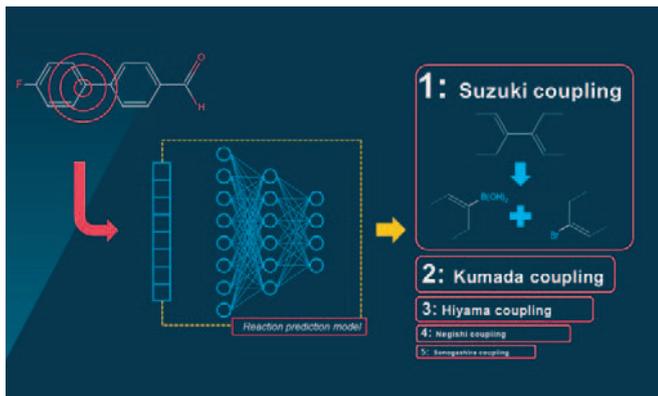


Abbildung 32: Ein neuronales Netz zur Vorhersage geeigneter Reaktionsmechanismen

lich Atome aus dem Lösungsmittel eine Rolle, die in den Reaktionsgleichungen aber nicht auftauchen. Kurz: in vielen publizierten Reaktionen stimmen die Massenbilanzen nicht. Dies bereitet Rechnern beim richtigen Verstehen von Reaktionen große Probleme und man muss mit intelligenten mathematischen Algorithmen aushelfen (Stichwort *Atom-Mapping*).

Eine weitere Herausforderung liegt in dem Publikationsverhalten, da üblicherweise nur funktionierende Reaktionen veröffentlicht werden. Wenn wir Menschen lernen, lernen wir aber auch aus dem, was nicht funktioniert. Das gilt auch für Computer: „Der Mensch lernt aus eigenen Fehlern, Computer aus den Fehlern anderer“. Dabei geht es vor allem darum, *sinnvolle* nicht funktionierende Reaktionen zu finden. Das sind solche, bei denen einem Experten nicht sofort klar ist, dass die Reaktion nicht funktioniert. Solche Reaktionen finden wir beispielsweise in den Laborjournalen von Bayer. Bayer ist über 150 Jahre alt und die chemische Forschung bei Bayer hat in den letzten Jahren etwa 12 Mio. elektronische Laborjournaleninträge hervorgebracht. Diese müssen allerdings ausgelesen und interpretiert werden. Für die automatische Interpretation der Laborjournaleninträge verwenden wir ebenfalls ein neuronales Netz. Das Neuronale Netz liest alle 12 Mio. Laborbucheinträge und bewertet jeden einzelnen, ob die dort beschriebene Reaktion funktioniert hat oder nicht.

Nach all diesen Vorbereitungen haben wir einen maschinen-lesbaren Datensatz aus funktionierenden und nicht-funktionierenden Reak-

tionen. Mit diesen Daten trainieren wir ein Neuronales Netz für eine sogenannte Ein-Schritt-Retrosynthese.

Damit hätten wir eine moderne Antwort auf die erste der beiden obigen Fragen, *wo schneide ich?* Kommen wir nun zu der zweiten Frage: *Wie finden wir einen Syntheseweg, der bei bekannten oder kaufbaren Ausgangsstoffen landet?* Bei der Beantwortung dieser Frage haben wir mit der sogenannten *kombinatorischen Explosion* zu kämpfen. Auch wenn wir uns hier mit der Synthese von sogenannten *small molecules* beschäftigen, können diese bis zu 100 Verbindungen zwischen Schweratomen haben. Nehmen wir beispielsweise ein Molekül mit 100 Verbindungen, welches wir in 10 Schritten synthetisieren wollen. Dann müssten wir $5 \cdot 10^{21}$ mögliche Syntheserouten untersuchen. Zum Vergleich das Universum existiert erst seit ca. $7,8 \cdot 10^{18}$ Sekunden. Aber reichen 10 Stufen?

Effiziente Algorithmen

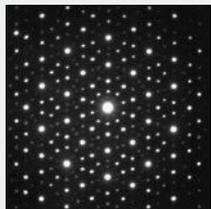
Um dieser enormen Vielfalt Herr zu werden, wenden wir Algorithmen an, die für Strategie-Spiele, wie Schach oder GO, entwickelt wurden. Wie bei Strategiespielen ist es im Allgemeinen nicht möglich, den gesamten Baum zu erkunden. Daher werden effiziente Algorithmen benötigt, um diese Art von Bäumen zu erkunden und möglichst schnell gute Lösungen zu liefern.

Eines der schönen Dinge der Mathematik ist, dass sie einen immer wieder überrascht, wo man Ergebnisse aus anderen Teildisziplinen der Mathematik anwenden kann. Hier verwenden wir Ergebnisse der Spieltheorie für die chemische Synthese.

An dem Beispiel der Retro-Synthese sieht man gut, wie die Arbeit eines modernen Mathematikers aussieht: Sie ist interdisziplinär und teamorientiert. Nur durch die Zusammenarbeit verschiedener Disziplinen können Projekte wie dieses erfolgreich umgesetzt werden. In dem Projektteam finden sich neben Mathematikern und Chemikern auch Chem-Informatiker und Computer Linguisten. Die Probleme müssen in Mathematik übersetzt werden, werden dann mit Mathematik gelöst, und die Lösung muss zurück in die Fachsprache der Fragesteller übertragen werden.

Quasikristalle

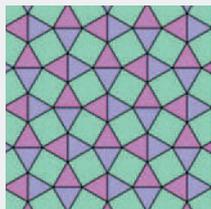
1. Ein überraschendes Muster



Als Dan Shechtman 1982 dieses Muster unter seinem Mikroskop sah, traute er seinen Augen nicht. Zu diesem Zeitpunkt untersuchte er die atomare Struktur eines Kristalls, genauer gesagt einer Aluminium-Mangan-Mischung. Alle bekannten Kristalle hatten eine regelmäßige Struktur, in der sich das Muster immer und immer wieder wiederholte. Doch nun sah Schechtman etwas ganz anderes: jedes Mal zehn Punkte in immer größer werdenden Kreisen. Eine Kristallstruktur, die in der Theorie überhaupt nicht existieren könnte. Denn alle Experten waren sich einig, dass sich die Atome in Kristallen in einem sich wiederholenden Gitter befinden, in dem die Abstände zwischen den Atomen immer gleich sind.

Als Schechtman seine Entdeckung veröffentlichte und ihr den Namen Quasikristall gab, sagte der zweifache Nobelpreisträger Linus Pauling, dass keine Quasikristalle existierten, nur Quasi-Wissenschaftler.

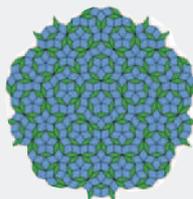
2. Flache Füllungen



Mathematiker haben jahrhundertlang die Muster untersucht, in denen atomare Kristalle normalerweise vorkommen. Auch wenn sie selbst dies Mosaikearbeit nannten: Mit welchen Formen kann man ebene Flächen vollständig ausfüllen, ohne Löcher oder Risse zu hinterlassen?

Das funktioniert zum Beispiel mit allen Quadraten oder mit regelmäßigen Sechsecken. Auch Kombinationen verschiedener Formen sind möglich: zum Beispiel mit Dreiecken und Quadraten wie diesem. Jedenfalls gibt es alles ordentliche Muster, die sich immer und immer wieder wiederholen. Genau wie Chemiker ihre Kristallstrukturen mochten.

3. Ein nicht-periodisches Muster



Lange Zeit fragten sich Mathematiker, ob es auch möglich sei, eine endliche Menge von Kacheln herzustellen, mit denen man einen unendlich großen Bereich so abdecken könne, dass sich das Muster nie wiederholt. Mitte der 1970er Jahre kam Roger Penrose zu einer erstaunlich eleganten Antwort: Er arbeitet mit nur zwei verschiedenen rautenförmigen Kacheln. Auf diese Weise können Sie ein unendliches Muster erzeugen, das sich nie wiederholt.

4. Die große Verbindung



Ein Gitter aus Atomen an den Eckpunkten einer Penrose-Kontrolle liefert unter einem Mikroskop genau das Muster, das Dan Schechtman gesehen hat. Es dauerte bis 1984, bis alle Teile des Puzzles an ihren Platz waren und die Menschen sahen, dass Schechtman Recht hatte: Es gibt Quasi-Kristalle. Schließlich änderte die International Union of Crystallography sogar die offizielle Definition eines Kristalls: ein sich wiederholendes Muster war nicht länger notwendig. 2011 wurde Dan Scherman für seine Entdeckung der Quasikristalle mit dem Nobelpreis für Chemie ausgezeichnet. Stähle mit diesen unregelmäßigen Kristallen haben jetzt auch verschiedene Anwendungen. Sie sind zum Beispiel ziemlich verschleißfest und werden in Rasierklingen und chirurgischen Instrumenten verwendet.

Das Reisen mit dem öffentlichen Personenverkehr wird mit modernen Ticketing-Systemen leicht gemacht: Einchecken per App, um eine Fahrtberechtigung zu erhalten, und dann einfach losfahren. Der garantiert günstigste Preis für die Gesamtheit aller Reisen des Tages, auch über Tarifgrenzen und Verkehrsmittel hinweg, wird im Nachhinein berechnet.



Best-Price Ticketingsysteme:

Erst fahren, dann zahlen!

Frau Dr. Petra Bauer arbeitet bei Siemens Technology, der Zentralabteilung für Forschung und Entwicklung der Siemens AG, in einem Team von Experten für Mathematische Optimierung. Sie berichtet hier über Siemens Next Generation Ticketing, eine moderne Ticketing-Lösung von Siemens Mobility, für die sie mathematische Algorithmen entwickelt hat.

„Welche Fahrkarte benötige ich, um von Zürich Stadelhofen nach Winterthur zu fahren? Gibt es für Hin- und Rückfahrt eine Fahrkarte, die günstiger ist als zwei Einzelfahrten? Möglicherweise fahre ich abends noch zu einer Kollegin nach Oerlikon. Lohnt sich in diesem Fall eine Tageskarte? Wie wahrscheinlich ist es, dass ich sie tatsächlich besuche?“

Wer kennt nicht diese oder ähnliche Überlegungen vor Antritt einer Fahrt mit öffentlichen Verkehrsmitteln, insbesondere in Städten oder Regionen, deren Tarifsystem uns nicht vertraut ist. Mit modernen Ticketingsystemen ist die Nutzung öffentlicher Verkehrssysteme ganz einfach: für jede Fahrt an der Starthaltestelle per Ticketing-App einchecken, einsteigen und, evtl. nach Umstiegen, an der Zielhaltestelle aussteigen. Mathematische Verfahren übernehmen die Identifikation der durchgeführten Fahrten und die Berechnung des günstigsten Gesamtpreises.

Check-In/Be-Out-Systeme

Die oben beschriebene Variante eines Ticketingsystems ist ein so genanntes Check-In/Be-Out-System (CiBo), bei dem der Nutzer das System über den Fahrtantritt informiert (Check-In), das System jedoch das Ende der Fahrt erkennt (Be-Out).

Gesammelte Daten

Um die Fahrten identifizieren und den günstigsten Preis berechnen zu können, werden während einer Fahrt, beginnend mit dem Check-In, Daten durch die Ticketing-App gesammelt, an das Hintergrundsystem sicher übermittelt und anschließend ausgewertet.

Die wichtigsten der gesammelten Daten sind GPS-Koordinaten, die Auskunft darüber geben, zu welcher Zeit der Fahrgast an welchem Ort war. Sind die GPS-Koordinaten genau, alle Verkehrsmittel pünktlich und haben diese ausreichend zeitlichen und räumlichen Abstand, so geben die GPS-Daten schon recht gut Aufschluss.

Was aber, wenn U-Bahnen Teil des Verkehrssystems sind und dort der GPS-Empfang eingeschränkt ist, oder wenn wir entscheiden müssen, ob ein sich langsam fortbewegender Fahrgast einen Teil der Strecke läuft oder in einem Fahrzeug sitzt, das im Stau steht?

Bei der Beantwortung dieser und ähnlicher Fragen helfen zwei weitere Datenquellen: im Untergrund liefern an den Stationen installierte Bluetooth-Low-Energy (BLE) Beacons Auskunft, kleine Funksender, die Informationen aussenden, die von in der Nähe befindlichen Smartphones empfangen werden können. Scheut man die Kosten nicht, so können BLE Beacons auch in Fahrzeugen installiert werden, um z.B. besser entscheiden zu können, ob ein Fahrgast innerhalb oder außerhalb eines Fahrzeuges war oder ob er das Fahrzeug gewechselt hat. Zudem liefern Sensoren des Smartphones Daten, die Rückschlüsse auf Bewegungsmuster zulassen. Man erhält so Wahrscheinlichkeiten für einzelne Fortbewegungsarten („Activities“), wie beispielsweise Laufen oder Fahrradfahren.



Abbildung 33: Einfach per Ticketing-App einchecken und losfahren.

Fahrtenrekonstruktion

GPS-Koordinaten, Sichtungen von BLE-Beacons und Activity-Daten werden als „Felddaten“ bezeichnet. Unter Heranziehung der Infrastrukturdaten (Lage der Haltestellen) und der Fahrpläne berechnen Verfahren der kombinatorischen Optimierung und der Stochastik die gemäß Datenlage wahrscheinlichste Abfolge von Fahrplanfahrten (z.B. S8, Zürich Hbf bis Oerlikon, planmäßige Abfahrt 9:00, Id 000011:18830:101) und sonstigen Fortbewegungsarten (wie z.B. Laufen beim Umsteigevorgang).

Die besondere Herausforderung liegt in der Verarbeitung sehr großer Datenmengen und im Umgang mit Unsicherheiten. Nebel, ein Nutzer, der etwas zu spät eincheckt, kurzfristig verlegte Haltestellen,

z.B. wegen eines Unfalls, sind Fehlerquellen, auf die die Verfahren robust reagieren müssen. Der Fahrgast kennt die eine, richtige Lösung und diese muss mit sehr hoher Zuverlässigkeit gefunden werden.

Bestimmung des günstigsten Preises

Der Reiz dieser modernen Ticketingsysteme liegt in Ihrer Einfachheit für den Nutzer („einfach losfahren“) und im Versprechen, im Nachhinein den günstigsten Preis verrechnet zu bekommen.

Eine Menge von Fahrten kann mit unterschiedlichen Kombinationen von Tickets abgedeckt werden (Einzelfahrten, Tageskarten, länger gültige Zonentickets, Kurzstreckentickets, Sonderangebote auf bestimmten Teilstrecken etc.). Die Zahl der Möglichkeiten wird schnell sehr groß, doch mit Verfahren der kombinatorischen Optimierung kann, dank mathematischer Theorie, schnell eine beweisbar optimale Lösung, d.h. die preislich günstigste Kombination von Tickets, gefunden werden.

Reale Implementierungen von BiBo, CiCo und BiBo-Systemen

Neben den CiBo-Systemen gibt es auch CiCo-Systeme (Check-In/Check-Out) und sogar BiBo-Systeme (Be-In/Be-Out) sind möglich. Pauschal kann man sagen: je weniger der Fahrgast tun muss, desto mehr Equipment (z.B. Beacons) ist erforderlich, um die erforderliche Zuverlässigkeit des Systems zu erreichen.

Im April 2018 sind zeitgleich zwei von Siemens entwickelte Systeme in Betrieb gegangen: das weltweit erste BiBo-System auf den Strecken der Südostbahn Schweiz und ein schweizweites CiCo-System (inkl. Skilifte!). Die als „Proof of Concept“ dienenden Implementierungen haben im Echtbetrieb Ihre Eignung und Zuverlässigkeit bewiesen. Die Weiterentwicklung des CiCo-Systems durch die Siemens-Gesellschaften HaCon und eos.uptrade hat in der CiBo-Variante im Oktober 2020 in Osnabrück (VOS) den Betrieb aufgenommen und steht jedem Fahrgast zur Verfügung.

Es ist zu erwarten, dass in den kommenden Jahren viele weitere solcher Systeme, entwickelt durch unterschiedliche Unternehmen, den Betrieb aufnehmen.

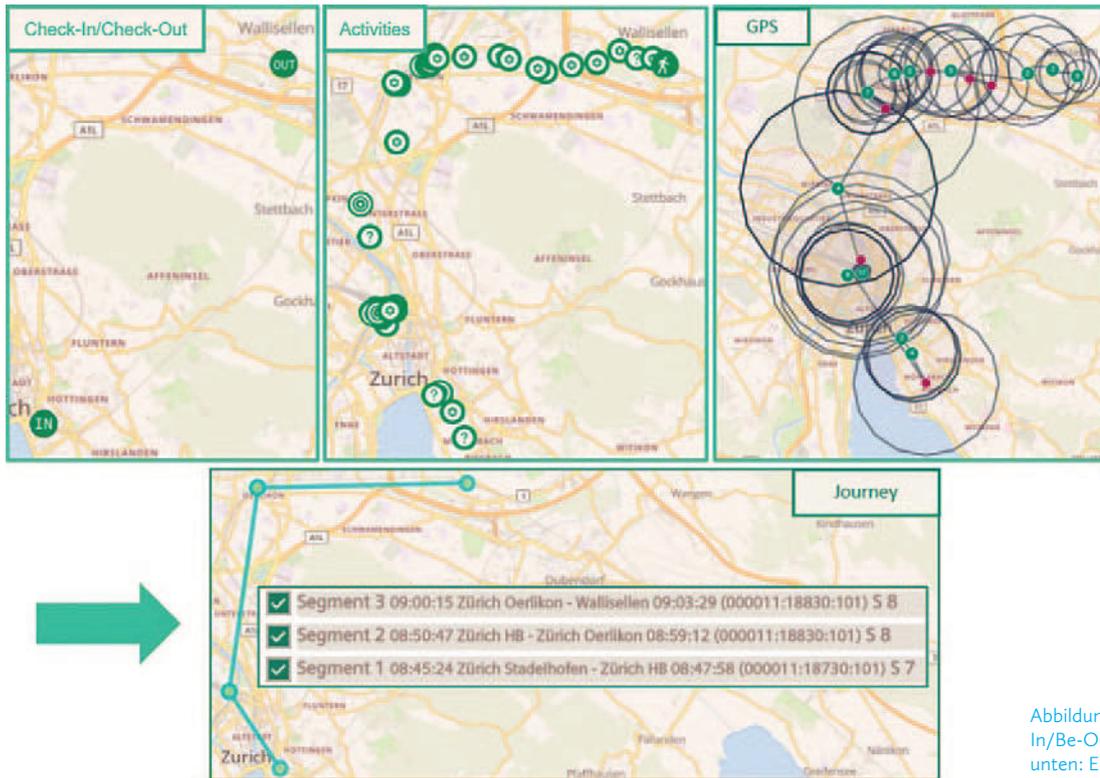


Abbildung 34: Oben Felddaten (Check-In/Be-Out, Activities, GPS-Koordinaten); unten: Ergebnis der Fahrrekonstruktion

Exkurs: Die kombinatorische Optimierung

Sowohl die Fahrrekonstruktion als auch die Bestimmung des besten Preises werden durch Verfahren der kombinatorischen Optimierung berechnet. Diese Verfahren ermöglichen es, dass oftmals unter Trillionen möglicher Lösungen für eine kombinatorische Fragestellung die optimale Lösung gefunden wird.

Ein berühmtes Beispiel hierfür ist das Traveling Salesman Problem. Ein Handlungsreisender muss die kürzeste Rundreise durch eine gegebene Menge von Städten finden. Schon bei 65 Städten hat er mehr Möglichkeiten ($64!/2$) als es geschätzt Atome im Universum gibt. Alle Möglichkeiten zu enumerieren ist also ausgeschlossen.

Die Komplexitätstheorie teilt die kombinatorischen Optimierungsprobleme in unterschiedliche Komplexitätsklassen ein. Für die Klasse der polynomiell lösbaren Probleme weiß man z.B., dass es schnelle

Algorithmen gibt, die diese Probleme optimal lösen. Für die so genannten NP-schweren Probleme (zu ihnen gehört das Traveling Salesman Problem) nimmt man an, dass es keine Verfahren gibt, die diese Probleme generell schnell lösen. Jedoch zeigt sich, dass man auch für diese Komplexitätsklasse mit Hilfe mathematischer Verfahren in der Praxis oft optimale oder gute Lösungen (unter Angabe einer Gütegarantie) für konkrete Problemstellungen berechnen kann.

Die kombinatorische Optimierung hat viele reale Anwendungen in Logistik und Verkehr, Produktion, Kommunikation, Energieversorgung und anderen Bereichen.

Projekt Ride-Hailing Wuppertal



Im Rahmen des Verbundprojekts bergisch.smart.mobility forschen wir in Zusammenarbeit mit den Wuppertaler Stadtwerken an intelligenter Routenplanung für Ride-Hailing Services in Wuppertal. Wir nutzen Methoden der multikriteriellen, diskreten Optimierung, um kosteneffiziente und kundenfreundliche Touren zu bestimmen.

PD Dr. Michael Stiglmayr und Daniela Gaul M.Sc. beschreiben ihre Erfahrungen mit dem Ride-Hailing Projekt.

Ride-Hailing-Dienste werden als ein wichtiger Baustein im ökologischen Stadtverkehr der Zukunft betrachtet und stellen eine flexible Ergänzung zum klassischen öffentlichen Personennahverkehr (ÖPNV) dar. Ride-Hailing-Dienste grenzen sich dabei durch ihre adaptive Routenplanung vom linienbasierten ÖPNV ab. Vom klassischen Taxiverkehr unterscheiden sich Ride-Hailing-Dienste dadurch, dass nicht individuelle Touren gefahren werden, sondern Transportanfragen gegebenenfalls gebündelt werden und sich mehrere Fahrgäste auf Teilstrecken ein Fahrzeug teilen. In diesem Fall spricht man von *Pooling*. Auf Grund ihrer spezifischen Anforderungen erfordert die

Routenplanung für Ride-Hailing-Dienste angepasste Modelle und Algorithmen. Insbesondere die Zeitdynamik stellt dabei eine große Herausforderung an Optimierungsalgorithmen dar: Fahrthanforderungen gehen kontinuierlich ein (Online-Problem) und müssen ad-hoc beantwortet werden. Mögliche Entscheidungen sind die (optimierte) Integration der Fahrt in eine bestehende Tour (die dadurch auch dynamisch verändert werden kann), die Eröffnung einer neuen Tour, aber auch die Ablehnung der Anfrage. Vorrangiges Ziel der Routenplanung ist die zeiteffiziente Erfüllung möglichst vieler Fahrthanfragen. Unvermeidbare Wartezeiten für einzelne Fahrgäste müssen dabei möglichst präzise vorhergesagt werden, um eine entsprechende Akzeptanz zu gewährleisten. Dabei müssen wirtschaftliche Aspekte von Anfang an mitgedacht werden, da der berechnete Fahrpreis nur von Start- und Zielpunkt abhängt, während die Transportkosten von der gefahrenen Route abhängen, insbesondere also davon wie gut sich diese Anfrage in eine bestehende Tour integrieren lässt. Der Trade-Off zwischen individuellen (Nutzer-) Interessen und ökonomischen sowie ökologischen Kriterien wird mit Hilfe multikriterieller Ansätze untersucht.

Event-basierter Graph

Anstatt die Routenplanung direkt auf dem Straßennetzwerk zu bestimmen nutzen wir eine Reformulierung des Problems in einem event-basierten Graph (siehe Abbildung 35). Dadurch kann die Reihenfolge und Verteilung der Passagiere auf die Fahrzeuge von der konkreten Wegeplanung zwischen zwei Event-Knoten entkoppelt werden. Knoten im event-basierten Graph entsprechen dabei Änderungen in der Passagierbelegung eines Fahrzeugs. Da diese fest mit einem Einsteige- bzw. Aussteigeort und einem entsprechenden Zeitfenster verknüpft sind, können die paarweisen Fahrzeiten zwischen diesen vorberechnet werden.

Das Ride-Hailing-Problem lässt sich wie folgt formulieren: Insgesamt haben wir n Fahrthanfragen und m Fahrzeuge. Sei $R := \{1, \dots, n\}$ die Menge der Fahrthanfragen. Jede Anfrage i aus R enthält einen Einsteigeort i^+ und einen Aussteigeort i^- . Zu jeder Anfrage gehören außerdem die Anzahl der zu transportierenden Personen, die maximale Fahrzeit, sowie ein Einsteige- und ein Aussteigezeitfenster.

Im folgenden Beispiel gehen wir zur Vereinfachung von 3 Sitzplätzen in einem Fahrzeug aus. Der event-basierte Graph besteht aus Knoten und Kanten. Einen Knoten schreiben wir als 3-Tupel $v = (v_1, v_2, v_3)$ wobei jeder Eintrag einen Sitzplatz des Fahrzeuges beschreibt. Ist

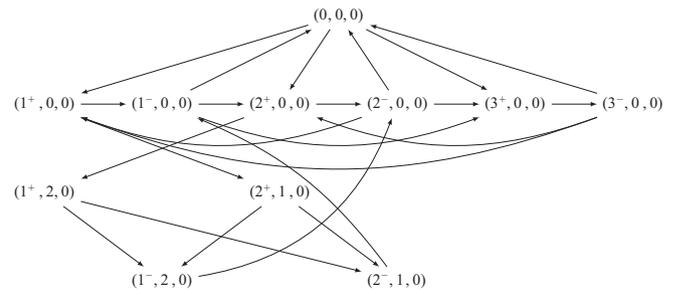


Abbildung 35: Event-basierter Graph für Fahrzeuge mit drei Sitzplätzen.

$v_j = 0$, so ist der Sitzplatz i leer. Der Knoten $v = (1^+, 2, 0)$ beschreibt ein Fahrzeug mit den Kunden 1 und 2. Die erste Koordinate des Knotens gibt dabei zusätzlich an welcher Ort zuletzt besucht wurde (in diesem Beispiel ist Kunde 1 eingestiegen). Ein Knoten wird nur zum Graphen hinzugefügt, falls die Summe der angefragten Sitzplätze die Fahrzeugkapazität nicht übersteigt. In Abbildung 35 ist der event-basierte Graph für ein Szenario mit drei Kunden abgebildet, wobei Kunden 1 und 2 beide zwei Sitzplätze anfragen.

Gemischt-ganzzahlige Optimierung

Mithilfe des event-basierten Graphen kann das Ride-Hailing-Problem als gemischt-ganzzahliges lineares Optimierungsproblem (MILP) beschrieben werden. Das bedeutet, wir beschreiben die Menge der zulässigen Lösungen für das Ride-Hailing-Problem durch lineare Ungleichungen mit ganzzahligen und kontinuierlichen reellen Variablen. Die ganzzahligen Variablen x_α modellieren dabei die diskreten Entscheidungen, in welcher Reihenfolge welches Fahrzeug welche Fahrgäste abholt bzw. aussteigen lässt. Die kontinuierlichen Variablen B_ν enthalten die Zeitpunkte, an denen ein Fahrzeug an den Knoten (Einstieg/Ausstieg) ankommt. Diese müssen innerhalb entsprechenden Zeitfenstern liegen. Neben den Nebenbedingungen, die sicherstellen, dass die Fahrgäste in den angegebenen Zeitfenster einsteigen und aussteigen, sind eine Reihe weiterer Nebenbedingungen nötig, die unter Anderem die Aufteilung der Fahrgäste auf die Fahrzeuge und die Auftragsreihenfolge der Fahrzeuge modellieren. Die Zielfunktion bewertet dabei die Qualität einer zulässigen Lösung. Im vereinfachten Modell in Abbildung 36 wird dazu die Gesamtfahrstrecke aller Fahrzeuge verwendet. Eine optimale Lösung für dieses Modell wäre also eine Tourenplanung, die alle Nebenbedingungen erfüllt und dabei die geringste Gesamtfahrstrecke aufweist.

$$\begin{aligned}
& \min \sum_{a \in A} c_a x_a \\
& \text{s. t. } \sum_{a \in \delta^{\text{in}}(v)} x_a - \sum_{a \in \delta^{\text{out}}(v)} x_a = 0 \quad \forall v \in V, \\
& \quad \sum_{a \in \delta^{\text{in}}(v)} x_a = 1 \quad \forall i \in R, \\
& \quad \sum_{a \in \delta^{\text{out}}(0)} x_a \leq |K|, \\
& \quad B_w \geq B_v + s_{v_1} + t_{(v,w)} - \tilde{M}_{v,w} (1 - x(v,w)) \quad \forall (v,w) \in A, \\
& \quad e_0 \leq B_0 \leq \ell_0, \\
& \quad e_{i^+} + (\ell_{i^-} - e_{i^-}) \left(1 - \sum_{a \in \delta^{\text{in}}(v)} x_a \right) \leq B_v \leq \ell_{i^+} \quad \forall i \in R, v \in V_{i^+}, \\
& \quad e_{i^-} \leq B_v \leq e_{i^+} + L_i + s_{i^+} + (\ell_{i^-} - e_{i^-}) \sum_{a \in \delta^{\text{in}}(v)} x_a \quad \forall i \in R, v \in V_{i^-}, \\
& \quad B_w - B_v - s_{i^+} \leq L_i \quad \forall i \in R, v \in V_{i^+}, w \in V_{i^-}, \\
& \quad x_a \in \{0, 1\} \quad \forall a \in A, \\
& \quad B_v \geq 0 \quad \forall v \in V.
\end{aligned}$$

Abbildung 36: Vereinfachtes Ride-Hailing Optimierungsproblem.

Möchte man sich zur Bewertung einer Lösung nicht auf rein ökonomische Aspekte beschränken, können neben der Minimierung der Fahrtkosten, welche die Zielfunktion in Abbildung 36 beschreibt, weitere Kriterien in Betracht gezogen werden. Dies können kundenorientierte Kriterien sein, wie zum Beispiel die Minimierung der zusätzlichen Fahrtzeit im Vergleich zu einer direkten Fahrt, oder die Wartezeit am Einsteigeort. Desweiteren können Situationen auftreten in denen einige Anfragen abgelehnt werden müssen, wenn das Bedienen dieser Anfragen für die restlichen Kunden große zusätzliche Umwege bedeutet. Ein weiteres Kriterium ist somit die Erfüllung möglichst vieler Anfragen. Bei der Betrachtung mehrerer Kriterien (Zielfunktionen) spricht man von multikriterieller Optimierung. Bei der Lösung eines multikriteriellen Optimierungsproblems möchte man sogenannte Pareto-optimale Lösungen finden. Das sind Lösungen, bei denen man sich in keinem Kriterium verbessern kann, ohne sich in einem anderen Kriterium zu verschlechtern.

Für die exakte Lösung der auftretenden gemischt-ganzzahligen Optimierungsprobleme verwenden wir Branch-and-Bound Algorithmen. Diese nutzen die rekursive Aufteilung des Entscheidungsraums (Menge der zulässigen Lösungen) hinsichtlich der ganzzahligen

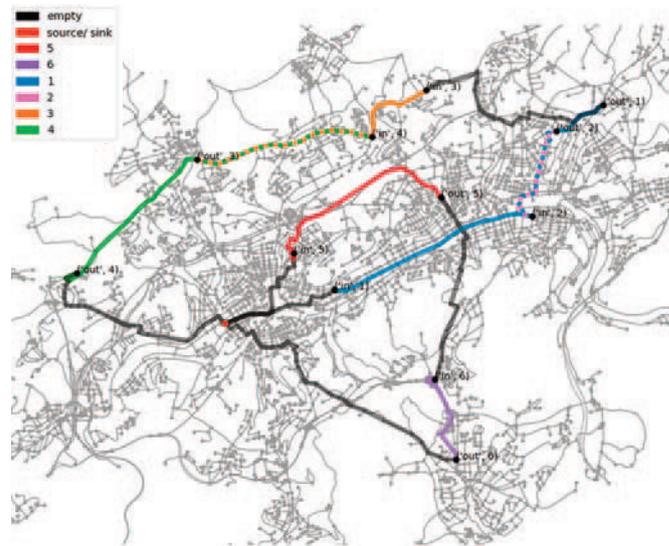


Abbildung 37: Routenplanung für 6 Kunden und 2 Fahrzeuge (auf den gestrichelten Abschnitten teilen sich mehrere Kunden ein Fahrzeug).

Variablen, um mit Hilfe von Schranken suboptimale Entscheidungen frühzeitig zu erkennen und zu verwerfen. Für ein kleines Beispiel mit sechs Kunden und zwei Fahrzeugen ist die optimale Tourenplanung in Abbildung 37 dargestellt.

Rolling-Horizon-Ansatz

Für große Instanzen kann die exakte Lösung jedoch zu zeitaufwendig sein, weswegen wir aktuell ergänzend an heuristischen Verfahren arbeiten, die schnell eine gute, aber nicht notwendigerweise optimale, Lösung generieren können. Dies ist insbesondere in der dynamischen Betrachtung des Problems relevant, wenn Fahranfragen sukzessive eintreffen, beantwortet und in die Tourenplanung integriert werden müssen. Dazu kann man die oben diskutierten statischen Modelle nutzen und diese in einen sogenannten *Rolling-Horizon-Ansatz* integrieren. Darüber hinaus liefern Lösungen des statischen Modells in Verbindung mit simulierten Fahrtanfragedaten aber auch wichtige Information für die unternehmerische Planung von Ride-Hailing-Diensten. So können etwa Fragen zu potentiellen Gebietserweiterungen oder Vergleichsstudien zur Ersetzung von Buslinien durch Ride-Hailing Angebote untersucht werden.

Das E-Bike und sein Antrieb

In einem E-Bike steckt jede Menge Mathematik aus vielen Teilbereichen.

Dr. Markus Hinterkausen, Dr. Uwe Vollmer und Dr. Uwe Iben vom Zentralbereich Forschung und Vorausbildung der Robert Bosch GmbH fassen Ihre Erfahrungen zur E-Bike Entwicklung zusammen:

„Leistungsstark, leise summend und innovativ im Design – eBikes sind echte Trendprodukte des 21. Jahrhunderts. Jedoch es gibt sie schon viel länger als wir glauben wollen. Ein erstes Patent für ein Elektrofahrrad wurde bereits im Jahr 1895 angemeldet. Die E-Mobilität wurde schon um 1835 als die Zukunfts-Mobilität in den Städten angesehen. Glatte Straßen und kurze Wege waren die Argumente. Schon 1900 wurden weitere Elektrofahrräder zum Patent angemeldet, gebaut und vertrieben. Zeitgleich zu Fahrrädern mit Dampfmaschinen, die sich jedoch wegen des hohen Gewichtes nicht durchsetzen. Aber auch das E-Fahrrad blieb eine Randerscheinung bis 2010. Das hohe Gewicht der Batterien, die geringe Reichweite und der ruckartige

Antrieb waren die Gründe. Doch mit der Lithium-Ionen-Batterie und neuen kleinen Sensoren begann ein neues Zeitalter der E-Mobilität. 2009 stieg Bosch eBike Systems als Start-up innerhalb des Bosch Konzerns in die E-Mobilität ein. Mit dem ersten eBike-Antriebssystem, das 2011 auf den Markt kam und entspanntes Cruisen in der Stadt genauso möglich macht wie sportliches Biken in der Natur, begann die Erfolgsstory. Ein großer Meilenstein war die Markteinführung des ABS für E-Bikes im Jahre 2017.“

„Eine Erfolgsstory hat viele Gründe. Einer davon ist die zielgerichtete Entwicklung der Komponenten Elektromotor, Getriebe und die Sensorik. Sie bilden das Herzstück des Antriebs, siehe Abbildung 38. Ein weiterer Grund ist die Bedienerfreundlichkeit des Systems. Hier wurde z.B. ein spezieller Bordcomputer entwickelt, der mit dem Smartphone verbunden werden kann.“

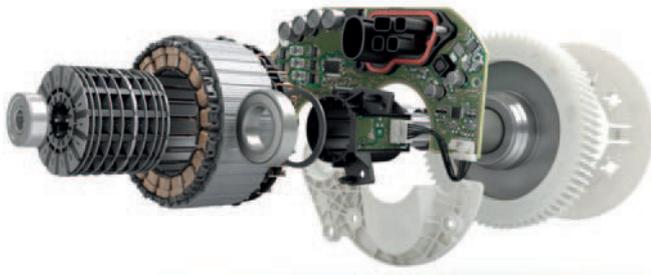


Abbildung 38: Elektromotor, Sensorik und Getriebe (Quelle: Bosch eBike Systems).

Viele mathematische Methoden

„Bei der Entwicklung des e-Bike-Antriebes und seiner Komponenten werden eine Vielzahl mathematischer Methoden angewandt. Analytische Methoden werden zur Erstauslegung des Antriebes und der E-Maschine genutzt. Mit diesen Ansätzen erhält man sehr schnell verschiedene Designs mit ihren Eigenschaften. Mathematische Verfahren, die hier genutzt werden, sind Vektor- und Matrizenrechnung aus der linearen Algebra und zumindest an einer Stelle die konforme Abbildung. Hierbei handelt es sich um winkeltreue Abbildungen, die u.a. in der Physik, Elektrotechnik oder Strömungsmechanik genutzt werden. Hier fließen erste Kundenanforderungen ein, wie z.B. Drehmomentenverlauf etc. Methoden der Stochastik helfen bei der Toleranzanalyse von Bauteilen, um Fertigungsverfahren und deren Toleranzen gleich zu Beginn in die Auslegung des Antriebsstrangs mit einzubeziehen.“

„In einem weiteren Schritt werden Finite Elemente Methoden (FEM) und Boundary Elemente Methoden (BEM) angewandt, um die Antriebsauslegung zu verfeinern und zu analysieren. Hier werden die entstehenden komplexen Gleichungssysteme numerisch gelöst. Dabei kommt üblicherweise der Ansatz der gewichteten Residuen mit Galerkin Testfunktionen zum Einsatz. Der Magnetkreis wird in einem ersten Schritt mit diesen numerischen Methoden modelliert und simuliert, um das Betriebsverhalten in verschiedenen Arbeitspunkten zu betrachten. Aus diesen Berechnungen werden Verhaltensmodelle im Zeitbereich oder im Frequenzbereich abgeleitet. Dazu dienen die Laplace- und Fourier-Transformationen. Modelle im Frequenzbereich sind vor allem für die Bewertung des akustischen

Verhaltens hilfreich, denn eine wichtige Kundenanforderung ist ein niedriges angenehmes Geräusch.“

Mathematische Optimierung

„Die Verhaltensmodelle im Zeitbereich werden z.B. zur Auslegung der Regelung des Antriebs genutzt und um Fahrzyklen zu simulieren. Im Zeitbereich können nichtlineare Effekte z.B. beim Materialverhalten berücksichtigt werden. Gradientenbasierte Methoden zur Topologie-Optimierung werden im Verlauf der Antriebsauslegung genutzt, um z.B. thermische, mechanische, elektrodynamische oder akustische Eigenschaften zu verbessern. Sie erfordern in der Regel den Einsatz von leistungsstarken Rechnern. Multizieloptimierung durch Suchen nach Pareto-optimalen Auslegungen in mehreren physikalischen Domänen wie strukturelle Festigkeit, Elektrodynamik, aber auch Herstell- und Materialkosten dienen der Findung der besten Auslegungen für den Antrieb. Die Regelungen werden im Frequenzbereich mit Hilfe der System-Beschreibung durch Laplace-Transformationen und im Zeitbereich durch Differentialgleichungen entworfen.“

„Hierzu dienen Hilfsmittel wie z.B. das Bode-Diagramm oder die Wurzelortskurve zur Bewertung der Güte und Stabilität des Regelkreises. Die Methode der Transferpfadanalyse beschreibt den Weg des Körperschalls von der Quelle bis zum Gehäuse, von dem es als Luftschall an unser Ohr dringt. Moderne Methoden aus dem Bereich der künstlichen Intelligenz (KI) wie z.B. Deep Learning unterstützen die Bewertung des Schalls hinsichtlich psychoakustischer Wahrnehmung.“

Der digitale Zwilling

„Alle diese mathematischen Methoden finden Anwendung bei der Auslegung des elektrischen Antriebes für das E-Bike. Schrittweise werden die Methoden und Verfahren verbessert und erweitert. Im Fokus stehen neben Genauigkeit, die Vorhersagbarkeit und die Rechengeschwindigkeit der Algorithmen. Dabei versucht man, die Erkenntnisse von früheren Designs in neue mit einzubringen. Stück für Stück werden die Methoden und Algorithmen zu einem digitalen Zwilling zusammengefasst, der in Zukunft den Antriebsstrang komplett digital abbilden wird, was dem Ingenieur völlig neue Möglichkeiten zur Bewertung von Funktionen und Kosten eröffnet.“

Die überraschende Verzögerung

Eine zusätzliche Straße kann zu einer längeren Reisezeit führen. Dieses kontraintuitive Ergebnis wird als Braess-Paradox bezeichnet und tritt in allen möglichen Formen auf.

Im Projekt *Networks* erforscht ein Team von Wissenschaftlern die Verkehrsnetze der Zukunft. Hierbei spielt das *Braessche Paradoxon*, aus den stochastischen Netzwerke eine grosse Rolle. Stochastischen Netzwerke sind Netzwerke, in denen der Zufall eine Rolle spielt. Bei diesen Netzwerken kann der Verkehrsfluss, genau wie beim Strassenverkehr, schlechter werden, wenn man die Kapazität des Netzwerks erhöht. Das scheint ein Paradoxon zu sein, aber sobald man die Mathematik versteht, sieht man, dass es wirklich keinen Widerspruch gibt.

Die alte Situation

Ein kleines Beispiel macht deutlich was geschieht. Zehn Fahrzeuge wollen von Quelle **Q** bis Ziel **Z** fahren. Jeder Fahrer kann zwischen zwei Routen wählen: die untere Route über **A**, oder die obere Route

über **B**. Die Fahrzeit für jede Route hängt von der Anzahl der Fahrzeuge V ab, die passieren, die genauen Zeiten sind der Zeichnung zu entnehmen. Beispielsweise kostet die Fahrt von Quelle nach **A** das Zweifache der Anzahl der Fahrzeuge, die diese Route wählen (in Minuten). Wenn alle 10 Autos diese Route wählen, dann beträgt die Fahrzeit auf diesem Stück zwanzig Minuten für alle. Wenn nur ein Auto fährt, wird es **A** in zwei Minuten erreichen.

In einer Gleichgewichtssituation sind die Fahrzeiten über die untere und obere Strecke genau gleich lang und kein Fahrer hat ein Interesse daran, die Straße zu wechseln. In diesem Fall nehmen fünf Autos die obere Route und fünf die untere Route. Diese Lösung liegt auch aufgrund der Symmetrie dieses Problems auf der Hand. Man kann leicht ausrechnen, dass in diesem Fall jeder eine Fahrzeit von 30 Minuten hat.

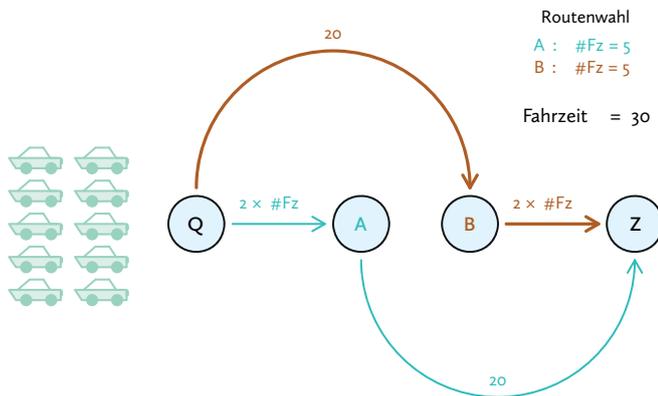


Abbildung 39: Alte Fahrsituation

“Verbesserungen”

Dann beschließt die Regierung, eine neue, schnelle Straße von **A** nach **B** zu bauen (belanglos). Bei Quelle **Q** können die zehn Fahrer nun aus drei Routen wählen: die beiden alten plus die neue Straße, die über **A** und **B** verläuft. Wenn nun eines der 5 Autos auf der unteren Strecke die neue Straße nimmt, sinkt seine Fahrzeit auf 22 Minuten. Dadurch verlängert sich die Fahrzeit für die 5 Autos auf der oberen Strecke etwas: Sie beträgt 32 Minuten.

Es ist daher klüger, wenn einer von ihnen ebenfalls den neuen Weg über **A** und **B** nimmt. Danach gibt es eine weitere Gleichgewichtssituation, in der die drei Routen jeweils gleich lange dauern. Die Fahrzeit beträgt nun 40 Minuten für alle, das sind zehn Minuten länger als vor dem Hinzufügen der zusätzlichen Straße. Zu diesem Zeitpunkt kann jedoch kein Fahrer seine Fahrzeit reduzieren, indem er individuell eine andere Route nimmt.

Dies ist ein wundersames Phänomen. Es wäre besser, wenn alle die neue Straße ignorieren würden, aber dann wird es immer einen Fahrer geben, der meint, er sei schneller unterwegs, wenn er sie nimmt. Eine Zeit lang scheint es ihm besser zu gehen, aber am Ende ist jeder das Opfer. Dieser Effekt ist unvermeidlich. Es wäre viel effizienter, wenn man zentral entscheiden könnten, welches Auto wohin fährt.

Das Braess-Paradoxon ist ein regelmäßiges Phänomen. In Boston erhielten Pendler eine längere Fahrzeit, als ein zusätzlicher Tunnel unter dem Namen Big Dig gebaut wurden. Auch das Gegenteil ist der Fall: Als New York die Verkehrsader der 42nd Street vorübergehend

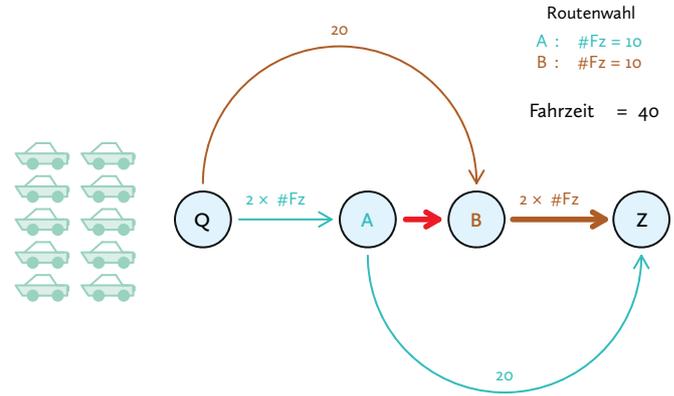


Abbildung 40: Fahrsituation mit der zusätzlichen Straße



Abbildung 41: Stau am Wuppertaler Robert-Daum-Platz. Im Hintergrund die Bergische Universität Wuppertal. Foto: Tomas Riehle (arturimages)

schloss, schien der Verkehr durch die Stadt reibungsloser zu fließen. Das Phänomen wurde auch in Stuttgart oder Seoul beobachtet und dokumentiert.

Zukünftige Pläne

Forscher wollen ein solches Verhalten in Verkehrsnetzen in den kommenden Jahren besser abbilden und auch prüfen, wie man das System verbessern kann, indem man einen kleinen Prozentsatz der Autos verfolgt. Darüber hinaus werden sie an Netzwerken für Kommunikation, Energie und Logistik arbeiten. Das Schöne dabei ist, dass alle Netzwerke im Großen und Ganzen die gleiche Struktur haben. Somit sind die anzugehenden Probleme nicht nur einem Anwendungsbereich vorbehalten.

Modellierung, Simulation und Optimierung zur Geräuschreduzierung bei Scheibenbremsen

Modellbasierte Auslegung und Verbesserung von modernen Fahrzeugen. Eine Erfolgsgeschichte der Kooperation von Mathematik und Ingenieurwissenschaften.

Moderne industrielle Produkte wie Automobile, Züge oder Flugzeuge sind hochkomplexe Systeme die auf einer Kombination von verschiedenen Gebieten der Ingenieurwissenschaften beruhen, wie z.B. Strukturmechanik, Fluidodynamik, Elektrotechnik, Hydraulik, Verfahrenstechnik. In der Auslegung und Optimierung solcher Systeme spielt die Verwendung von numerischen Simulation auf der Basis mathematischer Modelle und der Optimierung des Systemverhaltens und des Fertigungsprozesses heute eine zentrale Rolle. Wenn man derartig komplexe Systeme durch mathematische Modelle beschreiben will, ist man gezwungen bezüglich der Modellierungsgüte, zwischen Genauigkeit und Rechenzeit bei der Simulation, und der erzielbaren Optimalität Kompromisse zu machen. Die Mathematik liefert dabei die Werkzeuge dieses auf sinnvolle und rationale Weise zu tun, in dem Modelle analysiert und mathematische Algorithmen entwickelt werden, Modellierungs- und Approximations-Fehler quantifiziert werden, sowie Komplexität und Rechenzeit abgeschätzt werden.

Dieses Vorgehen bei der Behandlung eines realen Problems wird am Beispiel der Konstruktion und Verbesserung der Geräuschemission von Scheibenbremsen von Prof. Dr. Volker Mehrmann (Institut für Mathematik, TU Berlin) beschrieben. Für eine mathematisch und mechanisch detaillierte Beschreibung, siehe diesen Artikel²³.

Das Quietschen von Scheibenbremsen bei Kraftfahrzeugen (aber auch bei Bremsen von Fahrrädern oder Zügen) ist ein unangenehmes Geräusch. Eine wesentliche Ursache für das Quietschen ist eine selbst-

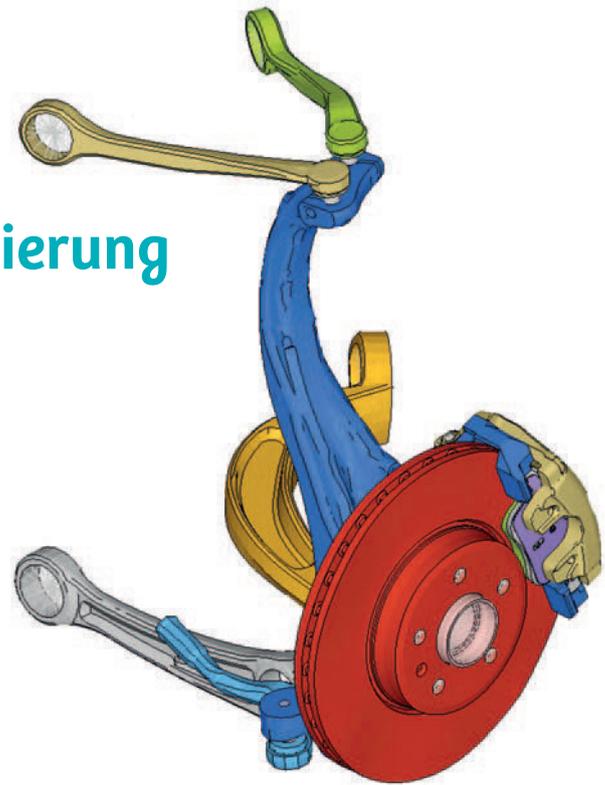


Abbildung 42: Bremsenmodell.

erregte Schwingung der Bremsscheibe, welche durch die Reibungskräfte an der Schnittstelle zwischen Bremsscheibe und Bremssattel erzeugt wird. In Abbildung 42 ist ein Modell einer gesamten Scheibenbremse inklusive der Aufhängung dargestellt.

Mathematische Modellierung von Bremsen

In der heutigen industriellen Praxis wird so eine Bremse durch Laborexperimente und durch numerische Simulation auf der Basis eines Finite-Elemente Modells analysiert und verbessert. Dazu wird die Struktur in sehr kleine Quader oder Tetraeder (Finite Elemente) zerlegt und die Auslenkung aus einem Ruhezustand, bzw. die Bewegungsgeschwindigkeiten werden durch stückweise Polynome auf diesen Finiten Elementen approximiert, siehe Abbildung 43 für so eine Zerlegung in finite Elemente. Es ergibt sich ein sehr großes System von Differentialgleichungen. Dabei werden für jedes Finite Element sechs oder mehr Freiheitsgrade (z.B. drei für die Position und drei für die Geschwindigkeit) verwendet. Für eine reale Bremse können das Mil-

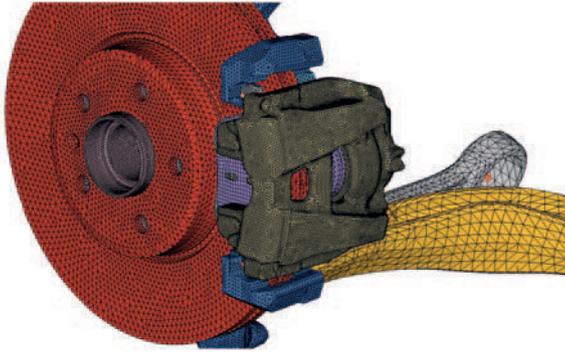


Abbildung 43: Beispiel eines Finite Elemente Modells einer Scheibenbremse.

tionen von Freiheitsgraden sein, je nachdem wie genau die Beschreibung sein soll. Ein großes Problem ist, dass sich die Scheibe natürlich sehr schnell dreht und damit die Koordinaten in einem mitbewegten Koordinatensystem betrachtet werden müssen. Neben der Modellierung der Struktur durch Finite Elemente ist aber insbesondere die Modellierung des Reibungsverhaltens ein sehr schwieriges Problem, da Reibung auf molekularer Ebene stattfindet. Daher verwendet man für die Modellierung der Reibung Ersatzmodelle, die durch Anpassung von Parametern an Messungen aus Experimenten gewonnen werden.

Insgesamt ergibt sich für die zur Zeit industriell verwendete Modellierung eine lineare Differentialgleichung zweiter Ordnung der Form

$$M_\omega \ddot{u} + D_\omega \dot{u} + K_\omega u = f,$$

wobei M_ω , D_ω , and K_ω sehr große Matrizen (Größe \approx Anzahl der Freiheitsgrade) die den Zusammenhang zwischen Beschleunigungs-, Geschwindigkeits-, und Positionskoordinaten beschreiben. Diese Koeffizientenmatrizen hängen von verschiedenen Parametern ab, z.B. Materialparametern der Bremse, Abrieb der Bremsscheibe, interner Dämpfung etc. Hier ist nur die Abhängigkeit von der Rotationsgeschwindigkeit ω der Bremsscheibe angedeutet, die in einem bestimmten Parameterbereich variiert wird. Im Gegensatz zu klassischen Modellen für Schwingungen von Balken oder Platten, müssen hier durch die Drehung der Scheibe zirkulatorische und gyroskopische Kräfte berücksichtigt werden, so dass die Matrizen D_ω und K_ω nicht-symmetrisch sind, während M_ω eine symmetrische Massenmatrix ist. Für selbsterregte Schwingungen wird zudem die äußere Kraft f vernachlässigt.

In dieser Modellbeschreibung sind natürlich eine große Anzahl an Vereinfachungen gemacht worden und ein Vergleich mit Laborexperimenten zur Evaluierung bzw. Validierung des Modells ist notwendig, bei dem auch die Güte des Modells (mehr oder weniger gut) quantifiziert wird.

Um die Dynamik des Systems zu verstehen, bzw. Schwingungsmoden der Differentialgleichung zu analysieren und zu verändern, macht man zuerst eine sogenannte Eigenwertanalyse, d.h. man berechnet auf der Basis einer Fourieranalyse, Eigenfrequenzen $\lambda(\omega)$ und Eigenvektoren $x(\omega)$, so dass für gegebenes ω die quadratische Eigenwertgleichung

$$\left(\lambda(\omega)^2 M_\omega + \lambda(\omega) D_\omega + K_\omega \right) x(\omega) = 0$$

gilt. Diese Eigenwertanalyse gilt lokal in der Nähe eines Gleichgewichtszustands, und das System ist instabil (und wird mehr oder weniger hörbar) quietschen, wenn es für bestimmte ω Eigenwerte mit positivem Realteil gibt.

Es gibt nun für dieses System $2n$ Eigenwerte (wobei n die Systemgröße ist). Es ist selbst mit heutigen Computerclustern nicht realistisch und auch nicht sinnvoll alle diese Eigenwerte und zugehörigen Eigenmoden für viele Parameterwerte ω zu berechnen um festzustellen bei welchem ω das Quietschen einsetzt und mit welcher Frequenz es quietscht, um dann das Modell auf der Basis von Variationen der Geometrie, Massenverteilung, oder von Materialparametern zu optimieren.

Reduzierte Modelle für Effiziente Berechnungen

Für die Berechnung der Eigenwerte werden moderne Algorithmen der Numerischen Linearen Algebra (sog. Krylovraumverfahren) verwendet, bei denen man Bereiche vorgeben kann in denen man Eigenwerte sucht. Das heißt man berechnet nur die Eigenwerte in einem bestimmten Bereich der komplexen Zahlenebene. Diese Methoden sind bzgl. ihrer Fehleranfälligkeit sehr gut untersucht und man kann die Fehler in den Approximationsmethoden abschätzen und auch an die Modellierungsfehler anpassen, so dass man eine Abschätzung für den Gesamtfehler von Modellierungs- und Eigenwertberechnungsfehlern durchführen kann.

Die Analyse der Eigenfrequenzen und Schwingungsmoden ist damit möglich, aber extrem rechenintensiv, insbesondere wenn dann noch über die Parameter optimiert werden soll. Aus diesem Grund wird

zuerst einmal ein sehr kleines parameterabhängiges reduziertes Modell erzeugt, welches für Parameteroptimierung verwendet werden kann. Es gibt eine Vielzahl von Modellreduktionsmethoden. Dazu wird das große Problem auf ein kleines Problem projiziert, indem die Eigenvektoren $x(\omega)$ in einem Unterraum gesucht werden, der durch die orthonormalen Spalten einer $n \times d$ Matrix Q aufgespannt wird, wobei d in einer Größenordnung von maximal einigen Hundert ist. Die klassische Methode (modal truncation) im Industrieinsatz verwendet die Eigenvektoren des Matrixpaares $\lambda M + K_1$ wobei K_1 ein symmetrischer Teil von K ist zu den größten Eigenwerten (diese sind alle real und negativ). In einem modernen Ansatz, in dem die Fehler wesentlich geringer sind, wird eine Matrix Q so berechnet, dass in einem großen Parameterbereich die relevanten Eigenwerte des reduzierten Modells (d.h. die mit positivem Realteil), diejenigen des vollen Modells gut approximieren, und dies mit einer Fehlerabschätzung. Dazu werden die entsprechenden Eigenwerte und Eigenvektoren des vollen Modells für einige zufällig gewählte ω -Werte im Untersuchungsbereich berechnet und anschließend wird mit Hilfe einer Singulärwertzerlegung (engl. Proper Orthogonal Decomposition (POD)) aus diesen ein kleiner Raum erzeugt, der im gesamten Parameterbereich verwendet werden kann.

Natürlich wird bei der Modellreduktion ein weiterer Modellreduktionsfehler gemacht, der bei einer guten Reduktionsmethode in der gleichen Größenordnung der anderen bisher gemachten Fehler liegen sollte.

Eine Illustration verschiedener Fehlermaße in Abhängigkeit von der Dimension des reduzierten Modells für zwei industrielle Finite Elemente Modelle M_1 und M_2 und das klassische sowie das neue Verfahren ist in Abbildung 44 dargestellt.

Man sieht sehr schön, dass die POD Methode, mit einem sehr kleinen d sehr kleine Fehler in den Eigenwerten err_λ und Eigenvektoren err_x produziert. Zu diesen Fehlern gibt es zudem eine Fehlerabschätzung, die es erlaubt die Größe d und die Anzahl der gewählten Auswertungspunkte adaptiv zu steuern. Im Vergleich dazu stagniert der Fehler bei der klassischen Methode auch bei Vergrößerung des Raums.

Neben der Fehlerabschätzung ist dann auch jeder Auswertungsschritt in der Parameteroptimierung sehr viel kostengünstiger (im Schnitt 30 mal schneller als mit dem klassisch reduzierten Modell), wenn man das so reduzierte Eigenwertproblem

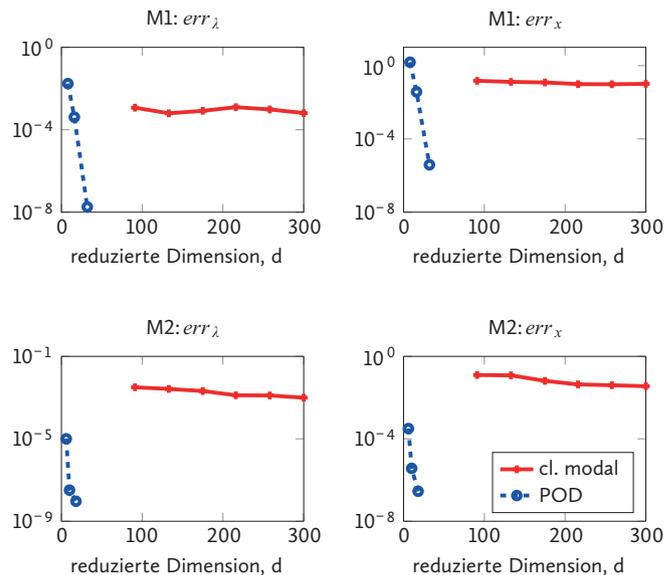


Abbildung 44: Abhängigkeit der Fehlergrößen von der reduzierten Dimension d für zwei industrielle Modelle M_1 und M_2 , sowie klassische und moderne Modellreduktionsmethoden.

$$(\lambda(\omega)^2 Q^T M_\omega Q + \lambda(\omega) Q^T D_\omega Q + Q^T K_\omega Q) x_d(\omega) = 0$$

löst. Damit können jetzt effizient die Scheibenbremsen verbessert werden und die Methode hat ebenfalls dazu geführt, dass in realen Beispielen Instabilitäten (Eigenwerte mit positivem Realteil) entdeckt wurden, die der klassische Modellreduktionsansatz nicht gefunden hatte.

Die neue Methode wurde als Open Source Python Skript implementiert und ist allgemein frei verfügbar.

An diesem Beispiel aber auch einer Vielzahl von anderen Anwendungen wird deutlich, welche Beiträge die moderne Mathematik leistet um in allen Wissenschaftsbereichen die Methoden und das Verständnis zu verbessern.

²³ N. Gräßner, V. Mehrmann, S. Quraishi, C. Schröder, U. von Wagner, Numerical methods for parametric model reduction in the simulation of disc brake squeal, Zeitschrift für Angewandte Mathematik und Mechanik Vol. 96 (2016), 1388–1405.



Der Krankenwagen trifft in 15 Minuten ein

Wie viele Krankenwagen werden benötigt, um sicherzustellen, dass niemand länger als 15 Minuten auf die Notfallversorgung warten muss? Ein mathematisches Modell von 1909 bietet eine Lösung für dieses Problem der Personaleinsatzplanung.

Bei dringenden und lebensbedrohlichen Situationen, wie Verkehrsunfällen und Herzinfarkten, muss innerhalb von fünfzehn Minuten nach der Meldung 112 ein Krankenwagen vor Ort sein. Für nicht lebensbedrohliche Situationen, wie z.B. ein gebrochenes Bein, gibt es eine maximale Annäherungszeit von einer halben Stunde.

Aber das ist alles leichter gesagt als getan. Wie viele Krankenwagen benötigt eine Stadt oder Region, um diese Anforderungen zu erfüllen? Wenn die Hälfte der Krankenwagen den ganzen Tag nichts zu tun hat, sind es eindeutig zu viele. Wenn aber die Hälfte der Krankenwagen standardmäßig zu spät kommt, gibt es eindeutig zu wenig Krankenwagen. Wo ist also das Optimum? Und wo können diese Krankenwagen am besten sein? Sicherlich nicht alle auf einem Haufen, aber vielleicht nicht alle so weit voneinander entfernt wie möglich.

Zeitlose Mathematik

Um die optimale Anzahl von Krankenwagen zu bestimmen, braucht man zwei wichtige Parameter. Dies ist die Anzahl der pro Minute eingegangenen Anrufe und die Dauer, die ein Krankenwagen zu einer bestimmten Tageszeit besetzt ist. Ambulanzdienste sammeln kontinuierlich Daten zu diesem Thema. Mathematiker können diese Daten nutzen, um die Ambulanzplanung zu optimieren.

Am einfachsten lässt sich die optimale Anzahl der Krankenwagen mit Hilfe einer Formel ermitteln, die der dänische Mathematiker Agner Erlang in zwei Artikeln 1909 und 1917 aufgestellt hat: der *Erlang-C-Formel*. Erlang entwarf mittels Wahrscheinlichkeitstheorie die Formel für die damals aufkommenden Festnetz-Telefondienste. In seiner ersten Arbeit 1909²⁴ wies er nach, dass zufällig verteilte Telefonanrufe dem *Poissonschen Verteilungsgesetz* folgen. Abhängig von der Anzahl der zu erwartenden Telefongespräche pro Minute und der durchschnittlichen Gesprächsdauer bestimmte er die Anzahl der benötigten Telefonleitungen so, dass die Wahrscheinlichkeit, dass jemand nicht anrufen kann, unter einem bestimmten Qualitätswert liegt. Seine zweite Arbeit²⁵ enthielt Formeln für Verlust und Wartezeit („Erlang-B“ und „Erlang-C“), die heute in der Theorie des Telefonverkehrs bekannt sind.

Das Schöne ist, dass dieselbe Formel von 1909/1917 heute in vielen verschiedenen Bereichen angewendet werden kann: von der Bestimmung der Anzahl der in einem Mobilfunknetz benötigten Masten über die Bestimmung der Netzkapazität für Video-Streaming bis hin zur Bestimmung der optimalen Anzahl von Krankenwagen oder der Anzahl von Mitarbeitern in einem Call-Center, um ein maximal-akzeptable Wartezeit der Kunden (ein sog. Service-Level) zu garantieren. Dies zeigt auf wunderbare Weise die Zeitlosigkeit der Mathematik.

Schattenplanung

Forscher arbeiten eng mit den Menschen vor Ort zusammen: mit Rettungsdiensten, Feuerwehren und Polizeikräften. Sie haben eine Software gebaut, mit der sie ein Jahr lang einen Schattenplan führen konnten. In dieser Schattenplanung sahen sie ihre eigene manuelle Planung, unsere mathematisch berechnete Planung und den realen Einsatz von Krankenwagen. Nach einem Jahr kamen sie zu dem Schluss, dass ihre Planung genauer war als die der Rettungsdienste. Das bedeutet, dass sie mit der gleichen Anzahl von Krankenwagen

und Personen mehr Qualität bieten oder die gleiche Qualität zu geringeren Kosten liefern können.

Die Wissenschaftler machen ihre mathematischen Modelle immer realistischer und ausgefeilter. Ein neuer Zweig dabei ist das dynamische Ambulanzmanagement. In der Regel kehren Krankenwagen nach einem dringenden Auftrag an einen festen Standort zurück. Im Allgemeinen ist es für das Dienstleistungsgewerbe besser, Krankenwagen nicht an einen festen Standort zurückzuschicken, sondern zu ermitteln, welcher der zu diesem Zeitpunkt günstigste vorübergehende Warteort ist. Dies hängt davon ab, wo sich die anderen Krankenwagen zu diesem Zeitpunkt befinden und von der zu erwartenden Anzahl von Katastrophen an einem bestimmten Ort.

Was mathematisch gesehen optimal ist, muss in der Praxis jedoch nicht immer machbar sein. Mathematisch gesehen müsste man alle Krankenwagen nach jedem Notruf ein wenig zur Einsatzzentrale umziehen, aber das ist natürlich auch unerwünscht. Durch Rücksprache mit Menschen in der täglichen Praxis werden die Modelle so genau wie möglich auf deren Bedürfnisse, wie etwa Pausenzeiten, oder die Heterogenität der Mitarbeiter abgestimmt.

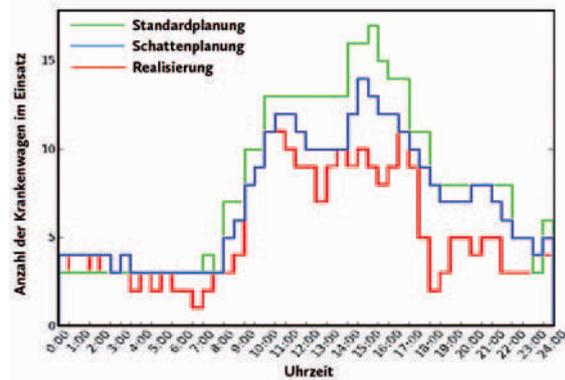


Abbildung 45: Die mathematisch berechnete Schattenplanung funktioniert besser als die in der Praxis verwendete Planung.

²⁴ A.K. Erlang, The Theory of Probabilities and Telephone Conversations, Nyt Tidsskrift for Matematik B, vol 20, 1909.

²⁵ A.K. Erlang, Solution of some Problems in the Theory of Probabilities of Significance in Automatic Telephone Exchanges, Elektrotekniker, vol 13, 1917.

Optimales Boarding am Flughafen

Mit Markov-Chain Monte-Carlo Methoden kann man das Boarding eines Flugzeuges simulieren und somit auch optimieren und damit Zeit sparen.

Wer kennt das Problem nicht: Beim Einsteigen ins Flugzeug entsteht schnell ein Stau und das sog. Boarding dauert (zu) lange. Dies nervt nicht nur die Passagiere – die Boarding-Zeit ist auch ein Wirtschaftsfaktor für die Flugunternehmen. Daher versuchen die Flugunternehmen, mit verschiedensten Boarding-Strategien die Einsteigezeit zu minimieren. Häufig sollen die Passagiere, die hinten sitzen, zuerst einsteigen. Dies ist die sog. back-to-front-Strategie. Ebenso erscheint es vernünftig, erst die Menschen mit Fensterplatz dann mit Mittelplatz und zuletzt die mit einem Gangplatz zu 'boarden', die sog. outside-in-Strategie. Man könnte auch die Passagiere in verschiedene Klassen einteilen (z.B. nach der Buchungsklasse, FrequentFlyer oder Alter usw. (sog. multiple classes-Strategie) oder auch gar keine Strategie verwenden (sog. random boarding). Natürlich kommt auch eine Kombination dieser Strategien in Frage.

Probleme beim Boarding

Es ist die übliche Leidensgeschichte für viele Reisende, die darauf warten, in ein Flugzeug einzusteigen. Die Fluggesellschaft muss dabei eine Liste von Passagieren durchgehen, die Priorität haben: Menschen mit Behinderungen (die z.B. auf einen Rollstuhl angewiesen sind), Familien mit Kindern, Passagiere der First- und Business-Class, Inhaber von Vielflieger-Karten und Passagiere, die für Prioritätsplätze bezahlt haben. Wenn die Reisenden der Economy-Class aufgerufen werden, stehen sie in der Kabine in einer Schlange und müssen warten, bis die vor ihnen Stehenden ihr Gepäck verstaut haben und sich hinsetzen. Dann sind die Gepäckfächer häufig schon voll.

Um mehr Umsatz zu generieren, haben die Fluggesellschaften neue Gebühren für Priority Boarding und Aufgabegepäck eingeführt. Letzteres hat das Problem nur noch verschlimmert, weil die Passagiere jetzt mehr Rollkoffer mit an Bord nehmen und damit Gänge beim Boarding blockieren. Laut einer Untersuchung von Boeing hat sich die Boarding-Zeit seit den 70er Jahren deutlich erhöht²⁶ und daher sind die Fluggesellschaften daran interessiert, die Boardingzeit zu reduzieren, möglichst ohne Umsatzeinbußen und Unannehmlichkeiten für die Passagiere.

Strategien für schnelleres Boarding

Die erste Idee bestand darin, für das Handgepäck eine etwas höhere Gebühr als für das Aufgabegepäck zu verlangen, und resultierte in einer kürzeren Boardingzeit, allerdings mit Unzufriedenheit bei den Kunden, die diese Preispolitik nicht nachvollziehen konnten. Daher wurden weitere Boardingstrategien entwickelt, um das Gate pünktlich zu verlassen. So kann man zuerst die Passagiere in den hinteren Reihen einsteigen lassen, oder man lässt diejenigen mit Fensterplätzen den Vorrang.

Eine weitere ausgeklügelte Strategie ist als einfallen lassen, darunter eine nicht mehr verwendete, die als „umgekehrte Pyramide“ (engl. Reverse Pyramid (RP)) bekannt ist. Das Reverse Pyramid Boarding-Verfahren wurde von Van den Briel et al.²⁷ entwickelt. Diese Methode trennt Passagiere in Boarding-Gruppen in Abhängigkeit von der Position ihrer Flugzeugsitze. Die Boarding-Gruppen werden nach einem „Diagonal Load“-Schema gebildet. Das bedeutet, dass in einer Gruppe die den meisten Passagiere Sitze im hinteren Teil des Flugzeugs, einige im Mittelteil und wenige vorne sitzen, wobei die Passagiere im mittleren/vorderen Bereich näher an den Fenstern sitzen. Innerhalb jeder

Boarding-Gruppe (3 bis 6 Gruppen) betreten die Passagiere das Flugzeug zufällig. Diese Boarding-Methode wurde zwischenzeitlich nicht mehr angewandt (weil Business-Passagiere am Ende der Pyramide keinen Platz mehr für ihr Gepäck fanden), erlebt jedoch in Zeiten der Corona-Pandemie und den damit verbundenen Distanzregeln ein Revival, da man mit dieser einfachen RP-Strategie die Passagiere gleichmäßiger in der Kabine verteilen kann und es den Passagieren ermöglicht, schneller ihre Plätze zu finden²⁸. Diese RP-Methode hilft den Passagieren auch, ihr Gepäck näher an ihren Sitzen zu verstauen, so weniger Gepäckstücke in letzter Minute aufgegeben werden müssen – eine häufige Ursache für verspätete Flüge.

Mathematik simuliert das Boarding

Durch sogenannte „Markov-Chain-Monte-Carlo Optimierungsverfahren“, ein mathematischer Algorithmus für zufällige Ereignisse, kann man den Boardingprozess simulieren und auch optimieren. So spielt der genaue Sitzplan des Flugzeugs eine Rolle, und ob es sich um ein Schmalrumpfflugzeug wie die Boeing 737 mit nur einem Gang oder um ein größeres Flugzeug für Transatlantikflüge handelt, wo die Passagiere zwei mögliche Wege haben. Besser als das Standard-Boarding ist in der Simulation das Random-Boarding, bei dem die Passagiere rein zufällig einsteigen. Ein weiterer Ansatz ist die Unterteilung in Boarding-Gruppen, die sich nach dem Zeitpunkt des Eincheckens richten. Damit wird die Verspätung durch Sitzplatzzuweisungen reduziert. Allerdings haben die meisten Passagiere bereits online einen Sitzplatz reserviert, bevor sie zum Flughafen kommen.

²⁶ https://www.boeing.com/commercial/aeromagazine/aero_01/textonly/t01txt.html

²⁷ M.H.L. van den Briel, J.R. Villalobos G.L. Hogg GL, T. Lindemann A.V. Mulé, America West Airlines Develops Efficient Boarding Strategies, *Interfaces* 35 (2005), 191–201.

²⁸ R.J. Milne, et al., Adapting the reverse pyramid airplane boarding method for social distancing in times of COVID-19, *PLoS ONE* 15(11) (2020), e0242131.

Dr. Margrit Klitz



Eine kurze Zusammenfassung Ihrer Karriere

Nach der Schule wusste ich nicht recht, ob ich Medizin oder Physik studieren sollte. Bei der Medizin reizte mich der Gedanke nah an Menschen zu sein, ihnen helfen zu können und überall auf der Welt arbeiten zu können. Eine Verbindung ergab sich dabei über die medizinische Physik, die mich sehr faszinierte. Ich entschied mich für Physik an der Universität Bonn, musste aber bald feststellen, dass ich die Mathematik schnell spannender fand. Daher wechselte ich nach einem Semester und studierte Mathematik mit dem Nebenfach Anglistik. Eine recht ungewöhnliche Kombination, aber Sprachen fielen mir immer eher leicht, und das Studium des Englischen konnte wohl nicht schaden.

Spezialisiert habe ich mich in der Mathematik nach dem Vordiplom auf die Strömungssimulation, was dann schon wieder ein großer Schritt in Richtung Physik war: Wie verhalten sich Tropfen auf Oberflächen? Wie verhält sich die Wasserströmung in einem Wehr? Wie werden Textilien umströmt, um Faserverbundstoffe herzustellen? Und Informatik: Wie kann ich all das im Computer berechnen und daraus überzeugende Bilder und wissenschaftliche Videos erstellen?

Das Feld der Strömungssimulation ist groß und vieles noch unerforscht, sodass ich nach meiner Diplomarbeit beschloss einen Doktor zu machen. Dies hat dann länger gedauert als erwartet, da während dieser Zeit meine drei Töchter zur Welt gekommen sind. Es ist nicht immer leicht, einen perfekten Zeitpunkt für die Familienplanung zu finden, aber ich würde es im Nachhinein genauso wieder machen.

Nach dem Abschluss der Doktorarbeit blieb ich noch für einige Monate als Post-Doc an der Uni, bevor ich meine Stelle als wissenschaftliche Mitarbeiterin beim Deutschen Luft- und Raumfahrtzentrum (DLR) antrat. Als stellvertretende Abteilungsleiterin und Gruppenleiterin bin ich nun auch wieder nah an den Menschen: Ich arbeite nicht nur für mich allein, sondern auch mit dem Ziel, dass sich unsere Mitarbeiterinnen und Mitarbeiter wertgeschätzt fühlen und wiederum ihre wissenschaftlichen Ziele mit möglichst wenig Störungen von außen erreichen können.

An die Uni Bonn bin ich dann auch nochmal zurückgekehrt: Ich habe ein Semester lang ein Programmierpraktikum zum Thema „Numerische Strömungssimulation“ gehalten und für die frischen Mathestudierenden eine Einführung die Programmiersprache „C“ gegeben.

Was hat Mathematik für Ihre Karriere bedeutet?

Als Mathematikerin lernt man Probleme jeglicher Art analytisch und mit viel Ausdauer zu lösen. Das hat mir sehr viel Vertrauen in meine Fähigkeiten gegeben. Da ich nicht den klassischen Weg einer Banker oder Versicherungskarriere einschlagen wollte, aber auch nicht rein wissenschaftlich arbeiten wollte, war mir schnell klar, dass ich außerhalb der Uni einiges würde dazulernen müssen, aber davor habe ich keine Angst. Im Gegenteil glaube ich, dass eine Arbeit mir so lange Spaß macht, bis sich keine Gelegenheit mehr ergibt etwas Neues zu lernen. Mathematikerin zu sein bedeutet für mich, fest zuzubeißen und erst loszulassen, wenn das Problem gelöst ist.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihrer Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Die Abteilung „High-Performance-Computing“ des Instituts für Softwaretechnologie, in der ich arbeite, ist auf Softwareprobleme spezialisiert, die man auf einem Computer alleine nicht mehr rechnen kann. Stattdessen werden viele Rechner zusammen eingesetzt, die mit viel Leistung und Kommunikation untereinander diese großen Probleme lösen. Das kannte ich bereits aus der Strömungssimulation an der Universität. Die vielfältigen Projekte, die sich beim DLR im Bereich von Luft- und Raumfahrt, aber auch Energie, Verkehr und Sicherheit ergeben, sorgen für abwechslungsreiche wissenschaftliche Projekte.

All diesen Projekten liegen meist physikalische Problemstellungen zu Grunde (Warum fliegt ein Flugzeug? Wie kann ich schnellere Züge bauen? Was passiert beim 3D- Druck?). Diese lassen sich durch ein

physikalisches Modell beschreiben. Die Sprache, in der diese Beschreibung erfolgt, ist die Mathematik, und die Lösung des Problems auf einem Computer ist die Numerik oder Numerische Simulation, welche ebenfalls ein Teilbereich der Mathematik ist. Wenn man die Sprache der Mathematik spricht, ist man also in der Lage solche physikalischen Probleme zu beschreiben und im besten Fall auch zu lösen. Das Spannende im DLR ist insbesondere, dass wir Wissenschaftlerinnen und Wissenschaftler aus vielfältigen Bereichen haben wie der Ingenieurwissenschaft, der Psychologie, der Physik und der Informatik. Alle bringen ein Expertenwissen mit, welches wir gezielt nutzen können, um die mathematischen Modelle zu verbessern und zu einer besseren Lösung zu führen.

Sind, in Ihrer Hinsicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Nicht jeder Mathematiker möchte angewandt arbeiten. Manche lösen abstrakte Probleme, erbringen komplexe Beweise und beschäftigen sich mit Problemstellungen, die nicht an einer Anwendung orientiert sind. Das hilft uns immer mehr Zusammenhänge in der Mathematik zu verstehen.

Viele Mathematiker arbeiten in Versicherungen und Banken und setzen hier gezielt ihr Wissen ein, um die Unternehmen zu stärken und zu unterstützen. In der Industrie ersetzt die numerische Simulation kostspielige physikalische Experimente.

Die große gesellschaftliche Herausforderung der jetzigen Coronapandemie gehen wir mit Hilfe von mathematischen Modellen an. Wir möchten voraussagen, wie sich die Pandemie für die einzelnen Regionen in Deutschland entwickelt, wie die Mobilität zwischen den Regionen auf die Pandemie wirkt und welche nicht-pharmazeutischen Maßnahmen einen großen Effekt haben.

Was raten Sie jungen Leuten hinsichtlich Mathematik?

Wenn ihr Spaß am Knobeln habt, Durchhaltevermögen mitbringt und mathematische Problemstellungen Euch Freude machen, dann solltet ihr Mathematik studieren. Später kann man sich weiter spezialisieren, z.B. mehr in eine physikalische oder informatische Richtung orientieren. Bei der Berufswahl werdet ihr überall dort gebraucht, wo ihr bei Problemen mit Eurer analytischen Herangehensweise unterstützen könnt.

Mathematik in der Modellierung zur Systemanalyse

Wie sollte sich das Energiesystem entwickeln? Welche Zusammenhänge gibt es zwischen unserem Konsum und Emissionen? Mathematische Modelle können in der Systemanalyse helfen, solche Fragen zu beantworten.

Die Modellierung hilft uns immer dann, wenn Zusammenhänge zu komplex sind, um sie intuitiv zu verstehen. In diesem Zusammenhang spricht man von der Systemanalyse. Zum Beispiel ist Wasserstoff ein wichtiger Baustein der Energieversorgung der Zukunft. Aber nur wenn ausreichend grüner Strom vorhanden ist kann auch grüner Wasserstoff erzeugt werden. Was ist dann eine sinnvolle Strategie, sollte man so bald wie möglich Elektrolyseure für die Wasserstoffherzeugung bauen? Welche Wechselwirkungen ergeben sich dann, wird dann der Strom in Deutschland teurer? Oder werden vielleicht sogar mehr Emissionen verursacht, wenn dadurch mehr Strom aus Kohle erzeugt oder aus dem Ausland bezogen wird? Zur Beantwortung solcher Fragen können wir Modelle nutzen, die auf mathematischen Methoden beruhen, und damit die Zusammenhänge darstellen und untersuchen.

Hier zeigen wir zwei Beispiele, wie unterschiedliche mathematische Modelle in der Systemanalyse genutzt werden: Optimierende Energiesystemmodelle zur Bestimmung kostenoptimaler Entwicklungen und Input-Output-Modellierung zur Abbildung gesamtwirtschaftlicher Auswirkungen von Konsum und Produktion.

Optimierende Energiesystemmodelle

Ziel solcher Modelle ist die Beantwortung von Fragen wie beispielsweise „Wie sollte die kostengünstigste Deckung des Strombedarfs in den nächsten Jahrzehnten gestaltet werden?“, „Wie könnte sich die Energieversorgungsstruktur verändern, wenn man annimmt, dass die Preise für Emissionszertifikate stärker steigen?“ oder „Wie werden Stromspeicher optimal genutzt?“. Dafür bauen Forscher wie Christine Krüger und José Acosta Fernandez vom Wuppertal Institut für Klima, Umwelt, Energie ein Modell auf, in dem die Bestandteile des Energiesystems durch Gleichungen dargestellt werden.

Die Stromerzeugung aus Windenergie beispielsweise ist eine Zeitreihe, die sich aus der installierten Leistung der Windenergieanlagen und dem Verlauf der Windgeschwindigkeit ergibt. Ein Heizkraftwerk ist eine Einheit, die einen Brennstoff wie Kohle umsetzt und in Strom, Wärme und Emissionen umwandelt, eine Stromleitung wird beschrieben durch die Leistung, die sie transportieren kann, und die Verluste, die dabei entstehen. Auf diese Art werden alle Zusammenhänge des Systems in Gleichungen gefasst. Auch Randbedingungen werden als Gleichungen formuliert: Der Strombedarf muss zu jeder Stunde

immer gedeckt werden, die Emissionen müssen im vorgegebenen Rahmen sinken. Ein wichtiges Element eines solchen Modells sind die Kosten: Brennstoffpreise, Preise für Emissionszertifikate, Kosten für Wartung und so weiter.

Mit diesem System aus Gleichungen wird dann ein Optimierungsproblem formuliert: Minimiere die Kosten des Gesamtsystems, und zwar so, dass die Randbedingungen eingehalten werden. Das Ergebnis ist dann das unter den getroffenen Annahmen günstigste System. Und mit diesen Annahmen kann man dann spielen: Wie hoch muss der Emissionspreis sein, damit die Stromerzeugung in 2040 klimaneutral wäre? Wie verändert sich der Ausbau und Einsatz von Kraftwerken, wenn man eine andere Entwicklung der Brennstoffpreise annimmt? Oder welche Auswirkungen hat es, wenn man Mindest- oder Maximalmengen für Wasserstoff vorgibt? Aus solchen Untersuchungen können dann zum Beispiel Energieversorger ihre Strategien ableiten oder Empfehlungen für die Politik formuliert werden.

Erweiterte Input-Output-Modelle

Modelle werden auch eingesetzt, um die verschiedenen Auswirkungen der menschlichen Produktions- und Konsumaktivitäten zu beziffern, ob erwünschte und unerwünschte. Mit Hilfe dieser Modelle werden sowohl die Ursachen als Möglichkeiten zur Veränderung des Ausmaßes dieser Auswirkungen untersucht. Außerdem werden damit geeignete Indikatoren definiert, um die Veränderungen im Lauf der Zeit zu verfolgen und dann auf dieser Basis entsprechende politische Strategien entwerfen zu können.

Für Treibhausgase ist ein solcher Indikator beispielsweise der „Carbon Footprint (CF)“. Um diesen zu berechnen, wird die „Erweiterte Input-Output-Analyse (E-IOA)“ angewendet. Aus der Perspektive des Konsums drückt der Carbon Footprint den gesamten Ausstoß von Treibhausgasen aus, der entlang der globalen Liefer- und Herstellungskette durch die Verwendung (Konsum) einer Produktgruppe verursacht wird. Es werden also alle Treibhausgase einbezogen, die unmittelbar, aber auch mittelbar entlang der gesamten Herstellungskette anfallen. Der Zusammenhang dieser Produktionsaktivitäten kann alle Wirtschaftssektoren überspannen. Diese Interdependenzen, also die wechselseitigen Zusammenhänge, der verschiedenen Wirtschaftssektoren sind das Herzstück der E-IOA. Ein Beispiel dafür: Pkw-Nutzung → Herstellung von Autos → Herstellung von verschiedenen Maschinen → Produktion von Stahl für die Autos und die Maschinen → Produktion von Kunststoffen → Abbau



Abbildung 46: Die Wechselwirkungen und Zusammenhänge im Energiesystem sind komplex und nicht immer intuitiv zu verstehen – hier können Modelle helfen.

von Mineralien, Kohle, Erdöl, Benzin → Herstellung von Lkw für den Transport der Mineralien ...

Die Grundlage der E-IOA bilden „Erweiterte Input-Output-Tabellen (E-IOT)“. Eine E-IOT ist im Prinzip eine Matrix, in deren Zeilen und Spalten die Zusammenhänge von Leistungen und Produkten in einer Volkswirtschaft detailliert abgebildet und mit Zahlen versehen werden – wie hoch ist beispielsweise die Erhöhung des Energieverbrauchs bei einer Erhöhung der Stahlproduktion um 5%? Dazu kommen sozio-ökonomische und umweltbezogene direkte Effekte (z.B. Treibhausgase). Die komplexen Zusammenhänge führen zu sehr großen Matrizen, für deren Umgang geeignete mathematische Methoden gebraucht werden.

Dies sind zwei Beispiele für die große Welt der mathematischen Modelle in der Energiesystemanalyse. Für alle Modelle gilt: Ein Modell ist nicht die Realität, es ist immer ein vereinfachtes, abstraktes Abbild. Die Kunst liegt darin, ein Modell so zu gestalten, dass es trotz dieser Einschränkung eine Antwort geben kann, die der jeweiligen Frage gerecht wird. Deswegen gibt es nicht nur eine Vielzahl an Modellen für die verschiedensten Zwecke, sondern auch meist mehrere Modelle für denselben Gegenstand.



Eine aktuelle politische Maßnahme zur Bekämpfung des Klimawandels ist die Einführung einer CO₂-Steuer. Für die Umsetzung einer CO₂-Steuer spielt die Mathematik eine zentrale Rolle.

Klimawandel: Die optimale CO₂-Steuer

Leon Hoffe und Benjamin Leonhardt sind Master-Studenten im Studiengang Wirtschaftsmathematik an der Bergischen Universität Wuppertal. Im Rahmen Ihrer Bachelorarbeiten befassten Sie sich mit analytischen und numerischen Verfahren zur Bestimmung der optimalen Bepreisung einer CO₂-Steuer.

Verantwortlich für das Klima auf unserer Erde ist der Treibhauseffekt. Ausgelöst wird der Treibhauseffekt durch die Treibhausgase. Bei ihnen handelt es sich unter anderem um Kohlenstoffdioxid (CO₂), Methan (CH₄) und Lachgas (N₂O). Seit der industriellen Revolution nimmt die Konzentration der Treibhausgase in der Luft stetig zu. Dementsprechend steigt auch die globale Mitteltemperatur seit der Industrialisierung immer weiter an. Das Ausmaß des Klimawandels ist mittlerweile so groß, dass er in diversen Bereichen der Weltbevölkerung wahrgenommen wird. Erkennlich wird dies durch politische Übereinkommen wie das Pariser Klimaabkommen, welches 2015 verabschiedet wurde. Zudem wird das Ausmaß in Nachrichten und sozialen Medien oder durch weltweite Fridays-for-Future-Demonstrationen immer präsenter. Die überwiegend jugendlichen Teilnehmer der Demonstrationen setzen sich für schnelle, umfassende und effiziente Klimaschutz-Maßnahmen ein. Nur so kann das Ziel des Pariser-Abkommens, den globalen Temperaturanstieg bis zum Jahr 2100 unter 2°C zu halten, erfüllt werden. Es hat extreme Konsequenzen, die die Lebensverhältnisse der Menschen und Tiere beeinträchtigen, wenn

die Temperatur über diesen Grenzwert hinaus ansteigt. Zum Beispiel treten Hitzewellen, Stürme, Hochwasser und andere Extremwetterereignisse in erhöhter Frequenz und Intensität auf.

Eine Steuer zum Kampf gegen den Klimawandel

Um das 2°C-Ziel zu erreichen, ist bereits in verschiedenen Ländern eine CO₂-Steuer als Klimaschutz-Maßnahme eingeführt worden. Doch wie genau hält eine solche Steuer den Klimawandel auf? Für die CO₂-Steuer ist ein Preis festgelegt, der für jede Tonne CO₂-Emissionen von Unternehmen und Privatpersonen entrichtet werden muss. Die Kosten, die so durch die Steuer entstehen, motivieren die Emittenten von Treibhausgasen dazu die ausgestoßene Menge zu verringern. Beispielsweise macht sich die Steuer durch höhere Preise für Heiz- und Kraftfahrstoffe bemerkbar. Aus diesem Grund steigt der Verbraucher auf umweltschonende Alternativen wie Wärmepumpen, Elektromobilität und die Nutzung von erneuerbaren Energien um. Je höher der Preis, desto mehr Alternativen erweisen sich als lukrativ. Je mehr umweltschonende Alternativen genutzt werden, desto größer ist die Menge, um die die jährlichen CO₂-Emissionen reduziert werden.

Aufgrund von Messungen und Datenerhebungen ist bekannt, wie viel CO₂ noch emittiert werden darf, bis der maximale Temperaturanstieg von 2°C erreicht ist. Wenn mehr CO₂-Emissionen entstehen, muss die Weltbevölkerung mit den angesprochenen Konsequenzen des Kli-

mawandels rechnen. Um diese zu verhindern, wäre es naheliegend den Preis der CO₂-Steuer exorbitant hoch festzulegen. Das würde den Umstieg auf umweltschonende Alternativen in allen Anwendungsgebieten erzwingen und eine ausreichende Verringerung der CO₂-Emissionen gewährleisten. Dadurch wäre dem Klima geholfen, aber zu welchen Kosten? Kein Staat, kein Unternehmen und keine Privatperson wird für die Rettung des Planeten in Kauf nehmen, sein finanzielles Wohlergehen aufzugeben. Aus diesem Grund steht in der CO₂-Steuer-Thematik nicht nur die Erfüllung des Pariser Abkommens im Vordergrund, sondern auch die Vermeidung von unnötigen Kosten. Diese zusätzliche Bedingung stellt die Schwierigkeit bei der Festlegung des optimalen Preises für eine CO₂-Steuer dar. Eine weitere Schwierigkeit besteht darin, dass nicht nur ein konstanter Preis für die Steuer gesucht wird, sondern ein „Preispfad“, da der Preis der CO₂-Steuer von Jahr zu Jahr voneinander abweichen kann.

Daraus ergibt sich folgende Problemstellung: Es wird der Preispfad einer CO₂-Steuer gesucht, der hoch genug angesetzt ist, um ein Klimaziel, wie das angesprochene 2°C-Ziel des Pariser Abkommens, zu erfüllen. Weiterhin soll der Preispfad so gering wie möglich sein, um unnötige Kosten bei Unternehmen und Endverbrauchern zu vermeiden. Der Preispfad einer CO₂-Steuer, der dieses Problem löst, wird daher als „optimaler Preispfad der CO₂-Steuer“ bezeichnet.

Der Zusammenhang zwischen der Mathematik und dem Aufhalten des Klimawandels erscheint nicht offensichtlich. Mit Hilfe mathematischer Methoden ist es aber möglich, den gesuchten Preispfad für eine CO₂-Steuer zu bestimmen. Bei dieser Art von Problemstellung sprechen Mathematiker von einem Optimierungsproblem. Dazu zählen unter anderem Maximierungs- und Minimierungsprobleme, wie sie aus der Schule bekannt sind. In diesem Fall handelt es sich um Letzteres, denn es geht darum, die Kosten, die durch die CO₂-Steuer anfallen, zu minimieren. Die Bedingung, dass außerdem ein Klimaziel erfüllt werden soll, stellt eine zusätzliche Restriktion des Minimierungsproblems dar. Durch den Preis der CO₂-Steuer können die Emissionen so gesteuert werden, dass die Restriktion eingehalten wird. Deswegen handelt es sich bei der Bestimmung des optimalen Preispfades für eine CO₂-Steuer, um genauer zu sein, um ein Problem der *optimalen Steuerung*.

Die optimale Steuerung der CO₂-Steuer

Es existieren mehrere mathematische Verfahren, um die gegebene Problemstellung zu lösen. Diese lassen sich in analytische und nume-

rische Verfahren aufteilen. Die analytische Lösung liefert eine fundamentale Grundidee für die numerische Lösung und beschreibt die theoretische Darstellung eines Preispfades für die CO₂-Steuer. Für die konkrete, realitätsnahe Berechnung eines Preispfades sind große Datenmengen notwendig. Dazu gehören die Emissionsdaten der vergangenen Jahre, um Prognosen für die Zukunft aufzustellen. Außerdem sind die Daten relevant, die den Einfluss der umweltfreundlichen Technologien auf die jährlichen CO₂-Emissionen darstellen. Aufgrund dieser umfassenden Mengen an Daten werden in der Praxis die Algorithmen der numerischen Verfahren häufig auf dem Computer programmiert. Das hat den Vorteil, dass betrachtete Szenarien flexibel sind und die Daten zu den CO₂-Emissionen und über die umweltfreundlichen Technologien fortlaufend anpassbar sind. Zudem können die Programme jederzeit durch neue relevante Faktoren erweitert werden. Beispielsweise ist es möglich, Lernfaktoren wie das Learning-over-Time einzubauen. Dadurch wird berücksichtigt, dass die Technologien, die die CO₂-Emissionen mindern, aufgrund exogener Faktoren, wie dem allgemeinen technischen Fortschritt, günstiger werden.

Der optimale Preispfad einer CO₂-Steuer kann auf diese Weise für beliebige Szenarien berechnet werden. Dadurch ist es theoretisch möglich, den Klimawandel aufzuhalten. In der Praxis würde das eine Zusammenarbeit aller Länder der Welt voraussetzen, um eine globale CO₂-Steuer einführen zu können. Das ist aber aufgrund geographischer, politischer und gesellschaftlicher Aspekte nicht möglich. Außerdem sind die Datenmengen, die für ein reales und weltweites Szenario erforderlich sind, zu komplex und unvollständig. Aus diesen Gründen kommt die CO₂-Steuer bisher nur in einzelnen Ländern zum Einsatz. Auch auf nationaler Ebene basieren die Berechnungen eines optimalen Preispfades auf Hochrechnungen und vielschichtigen Datenmengen. Trotzdem liefert der optimale Preispfad wichtige Anhaltspunkte und dient politischen Entscheidungsträgern als Orientierungshilfe bei der Festlegung der jährlichen Preise einer CO₂-Steuer.

Der Forschungsbereich rund um den Klimawandel ist zukunftsweisend. Die Erfolge und Erkenntnisse auf diesem Gebiet entscheiden über unser Leben und das der zukünftigen Generationen auf diesem Planeten. Die CO₂-Steuer ist eine erfolgversprechende Maßnahme, um positiven Einfluss auf den Klimawandel zu nehmen. Es bleibt spannend, welche Länder sich ebenfalls dazu entschließen, solch eine Steuer einzuführen. Aufgrund der Aktualität der CO₂-Steuer-Thematik können die Auswirkungen und Resultate dieser Maßnahme in den nächsten Jahren verfolgt werden.



Das Layout großer Photovoltaik-Kraftwerke birgt eine enorme Zahl von Freiheitsgraden bei der Auswahl, Platzierung und Verschaltung der erforderlichen Komponenten. Durch Mathematische Optimierung wird der Planungsaufwand drastisch reduziert und die Qualität der geplanten Anlage erhöht.

Layout großer Photovoltaik-Kraftwerke

Frau Dr. Petra Bauer arbeitet bei Siemens Technology, der Zentralabteilung für Forschung und Entwicklung der Siemens AG, in einem Team von Experten für Mathematische Optimierung. Das hier vorgestellte Optimierungstool zur Planung großer Photovoltaik-Kraftwerke entstand in einer Kooperation mit Siemens Energy.

Die Auslegung großer Photovoltaik-Kraftwerke ist eine komplexe Aufgabe, die erfahrene Planerinnen und Planer über Wochen hinweg beschäftigen kann. Sie beinhaltet die Platzierung und Verkabelung der erforderlichen Komponenten unter Beachtung von physikalischen Gesetzen, äußeren Rahmenbedingungen und gegebenen Zielkriterien.

Ein bei Siemens entwickeltes Softwaretool konnte den Planungsaufwand wesentlich reduzieren: nach Eingabe aller Daten berechnen Verfahren der Kombinatorischen Optimierung ein Layout in wenigen Minuten. Damit wird es möglich, Planungsvarianten zu vergleichen oder Planungen zu ändern, ohne jedes Mal Tage oder Wochen zu investieren.

Komponenten eines Photovoltaik-Kraftwerkes

Die wesentlichen Komponenten einer Photovoltaikanlage sind die Photovoltaikmodule, die unter Einstrahlung von Sonnenlicht Gleichstrom produzieren, die Tische, auf denen die Photovoltaikmodule montiert werden, Wechselrichter, die den Gleichstrom in Wechselstrom umwandeln, Kabel, die die Module untereinander verschalten sowie an den jeweils zugeordneten Wechselrichter anschließen und Anschlusskästen, in denen mehrere Kabel in jeweils einem dickeren Kabel gebündelt werden.

Die Planungsaufgabe

Die Planungsaufgabe ist, unter Einhaltung physikalischer und weiterer Restriktionen auf einem gegebenen Gelände eine Photovoltaikanlage zu errichten, die die erwartete Leistung maximiert und dies so kostengünstig wie möglich.

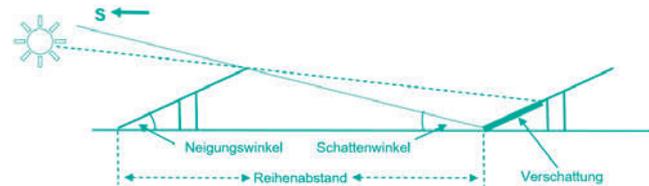
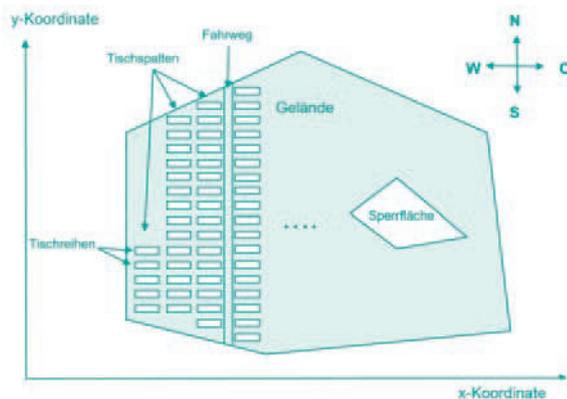
Zur Verdeutlichung der Komplexität des Problems beschreiben wir im Folgenden einige Aspekte der Aufgabe etwas detaillierter.



Das Gelände ist durch einen Polygonzug gegeben, der den Umriss beschreibt, sowie ggf. durch weitere Polygonzüge für Sperrflächen, die nicht bebaut werden dürfen. Für die Annäherung an die Geländetopographie benötigen wir außerdem Höhenangaben an ausreichend vielen Orten als Basis für eine Triangulation, also eine Approximation des Geländes durch ebene Dreiecke.

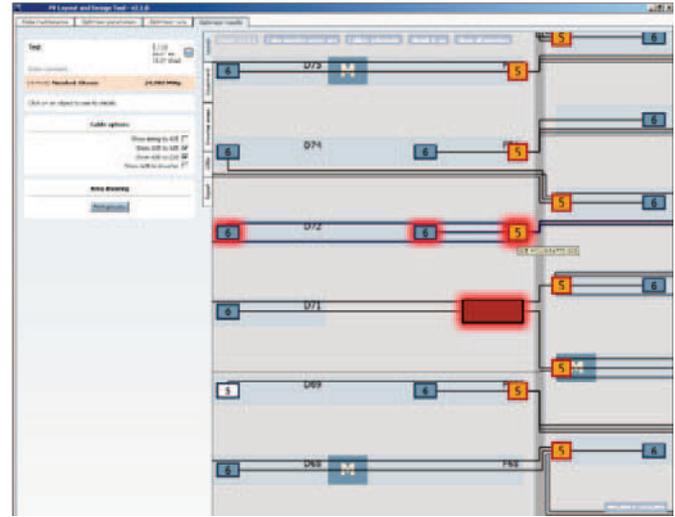
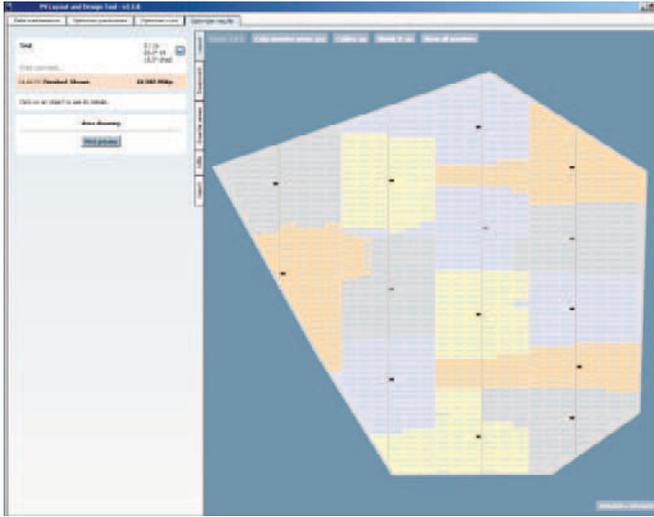
Die Solarmodule werden, nach Süden geneigt, in von Ost nach West verlaufenden Reihen aufgestellt. Fällt das Gelände nach Norden ab, muss der Abstand zur nächsten Tischreihe (von Süden aus gesehen) wegen Verschattungseffekten größer sein, als wenn das Gelände nach Norden ansteigt.

Man möchte nun möglichst viele Tische platzieren unter der Nebenbedingung, dass Fahrwege von Süd nach Nord nach parametrierbaren Vorgaben existieren und an einem ebenfalls einstellbaren Zeitpunkt im Jahr (z.B. am längsten Tag des Jahres um 12 Uhr mittags) sich die Module nicht gegenseitig verschatten. Die Reihen müssen zwischen zwei Fahrwegen auf gleicher „Höhe“ in Süd-Nord-Richtung liegen, um für Wartungsarbeiten gut zugänglich zu sein.



Zudem müssen Wechselrichter ausgewählt und platziert werden, die zusammen ausreichend Kapazität haben, um den produzierten Gleichstrom in Wechselstrom umzuwandeln. Jeder Tisch muss außerdem einem Wechselrichter zugeordnet und mit diesem in einer hierarchischen Kabelstruktur verbunden werden.

Bei der Verkabelung ist zu beachten, dass sie stets rechtwinklig verläuft und dass alle Kabel, die in Süd-Nord-Richtung verlaufen unterirdisch verlegt werden müssen, was zusätzliche Kosten verursacht. Kabeldurchmesser müssen in Abwägung zwischen Anschaffungskosten und Verlusten durch den Kabelwiderstand bestmöglich gewählt werden.



Das Optimierungstool

Diese Gesamtaufgabe ist zu komplex, um sie durch ein Optimierungsverfahren zu lösen, das alle Entscheidungen gleichzeitig fällt. Hier eine Liste der wichtigsten zu treffenden Entscheidungen:

- Lage der Fahrwege
- Verortung der Tische in Reihen und Spalten
- Auswahl der Wechselrichter
- Verortung der Wechselrichter und Zuordnung von Tischen zu Wechselrichtern
- Kabelführung unter Berücksichtigung der Grabungskosten
- Bestimmung der Kabeldurchmesser

Wir sind auf eine Dekomposition des Gesamtproblems angewiesen und lösen einzelne Aufgaben, obwohl sie nicht voneinander unabhängig sind, nacheinander. Um die damit verbundenen Abstriche in der Lösungsqualität so gering wie möglich zu halten, berücksichtigen wir in den einzelnen Teilschritten, soweit möglich, näherungsweise die Aspekte der nachfolgenden Schritte. So geht z.B. bei Bestimmung der exakten Lage der Fahrwege bereits eine Approximation an die Verkabelungskosten in die Zielfunktion ein.

Bei der Lösung der Teilprobleme kommen die Ganzzahlige Lineare Programmierung, die Dynamische Programmierung und Heuristiken zum Einsatz. Bei der Ganzzahligen Linearen Programmierung werden die Restriktionen und die Zielfunktion durch lineare Gleichungen

Abbildung 50: 25 MWp Kraftwerk mit Tischen (grau), Wechselrichtern (schwarz), Wechselrichtergebieten (bunt) und Wegen. Rechts eine Detailsicht.

und Ungleichungen in ganzzahligen Variablen modelliert und mit entsprechenden Solvern gelöst, bei der Dynamischen Programmierung wird die Lösung eines Problems auf die Lösungen zu Teilproblemen zurückgeführt, die gespeichert und dann geeignet kombiniert werden. Wenn Probleme zu komplex sind, um sie exakt ausreichend schnell zu lösen, wendet man Heuristiken an, die durch „geschickte“ Vorgehensweise praktikable Lösungen errechnen.

Das von uns entwickelte Optimierungstool erlaubt es dem Planer auf der Bedienoberfläche die Geländedaten zu importieren, die zur Verfügung stehenden Komponenten (Modultypen, Tische, Wechselrichtertypen, Kabeltypen, etc.) anzugeben und Nebenbedingungen zu formulieren.

Per Knopfdruck wird in wenigen Minuten ein Layout durch eine Hierarchie mathematischer Optimierungsverfahren berechnet, ausgewertet und an der Oberfläche dargestellt. Durch Hineinzoomen in das Layout können Details angesehen und vom Planer nach Wunsch verändert werden. Auch ein anschließender Reoptimierungsschritt, der die Wünsche des Planers berücksichtigt, kann ausgeführt werden.

Optimierte Auslegung von Gasheizgeräten

Oder: Wie komme ich schnell und komfortabel zur Wunschtemperatur? Mathematische Optimierung verbessert Komfort und Effizienz.

Matthias Stursberg, R&D Manager bei der Vaillant GmbH in Remscheid berichtet zusammen mit Tobias Suszka und Prof. Dr. Kathrin Klamroth von der Bergischen Universität Wuppertal über ihr gemeinsames Projekt.

Wie in fast allen technologischen Bereichen wird auch die Entwicklung von Gasheizgeräten immer komplexer und aufwändiger. Numerische Simulation und mathematische Optimierung haben hier zahlreiche Anwendungen, denn schon kleine Effizienzsteigerungen können zu ganz erheblichen CO₂-Einsparungen führen. Dabei soll aber auch der Komfort von Warmwassergeräten nicht zu kurz kommen, denn schließlich wollen wir nicht lange warten, bis das warme Wasser aus der Dusche kommt.

Mathematische Regelung der Wassertemperatur

Der Schlüssel dafür liegt in der Regelung der Warmwasserbereitung, also z.B. von Durchlauferhitzern für die Warmwasserbereitung für Duschen und Waschbecken. Idealerweise liefert das Heizgerät, unabhängig von der durchlaufenden Wassermenge, eine konstante Auslauftemperatur von z.B. 55°C. Diese wird dann mit dem Mischer der Dusche oder des Wasserhahns auf die Wunschtemperatur gemischt. Dabei sind einerseits zeitlich langsame Änderungen der Auslauftemperatur zu vermeiden, da sonst beim Duschen oder Händewaschen die Temperatur permanent mit dem Mischer nachgeregelt werden muss. Andererseits verursacht ein sich verändernder Wasserfluss sogenannte Temperaturüberschwinger, die den Komfort beeinträchtigen: Das Wasser wird kurz viel zu heiß – so dass man wegspringen muss – und gleich danach ist es wieder zu kalt. Das Ziel einer Optimierung ist es jetzt, mit Hilfe einer Analyse des Trade-Off zwischen den beiden Kriterien „zu schnelle“ und „zu langsame“ Regelung der Auslauftemperatur einen geeigneten Kompromiss zu bestimmen – gerade so, dass man als Nutzer weder permanent nachregeln muss, noch mit starken Temperatursprüngen zu kämpfen hat.

Hierfür werden *evolutionäre multikriterielle Optimierungsverfahren* (EMO) eingesetzt. Diese ahmen den aus der Biologie bekannten Evolutionsprozess nach. Es wird eine Population von Lösungen erzeugt, die durch Selektion, Mutation und Rekombination in eine neue Population überführt wird. Diese Nachfolgeneration sollte wegen des Selektionsdrucks bessere Lösungen enthalten. Das Verfahren läuft über mehrere Generationen, bis ein zufriedenstellendes Ergebnis gefunden wurde oder keine Verbesserungen mehr erzielt werden. Die Selektion, also das Bewerten einer Lösung, erfolgt anhand von

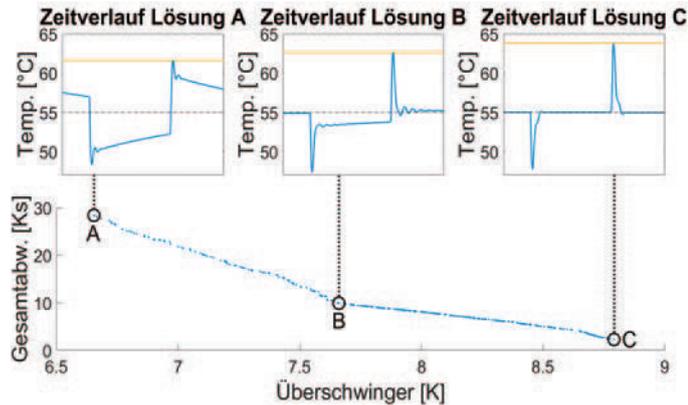


Abbildung 51: Alternative Regelungen eines Durchlauferhitzers und ein guter Kompromiss zwischen schneller und komfortabler Wassererwärmung.

Simulationsmodellen. Dabei wird das Verhalten des jeweiligen Geräts unter repräsentativen Lastzyklen vorhergesagt. Auf diese Weise erhält man eine Approximation der sogenannten Pareto Front. Die Pareto Front enthält alle Lösungen, die sich nicht gleichzeitig bezüglich beider Ziele verbessern lassen. D.h. eine Verbesserung in einem Ziel bewirkt zwangsläufig eine Verschlechterung im anderen Ziel.

Ein Kompromiss führt zur Lösung

Abbildung 51 zeigt das Ergebnis einer solchen Optimierung basierend auf einem generischen Simulationsmodell. Im unteren Teil der Grafik ist die approximierten Pareto Front dargestellt. Jeder Punkt entspricht einer Lösung. Für jede Lösung werden der Wert der Gesamtabweichung und der Wert des Überschwingers gegenübergestellt. Die Gesamtabweichung ist die über die Zeit integrierte Abweichung von der Wunschtemperatur, und der Überschwinger ist die größte Abweichung von der Wunschtemperatur. Beides soll möglichst klein sein: Sowohl die Gesamtabweichung von der Wunschtemperatur als auch der Überschwinger. Man erkennt, dass jedoch nicht beides gleichzeitig erreicht werden kann – die perfekte Lösung ist nicht möglich und wir müssen einen geeigneten Kompromiss finden. Genauer gibt es hier einen Trade-Off zwischen beiden Optimierungszielen: Will man die Gesamtabweichung reduzieren, so geht das nur auf Kosten eines zunehmenden Überschwingers. Exemplarisch sind für drei Lösungen (A, B und C) auch die zugehörigen Temperaturverläufe mit jeweils zwei Lastwechseln dargestellt, die die jeweiligen Vor- bzw. Nachteile zeigen.

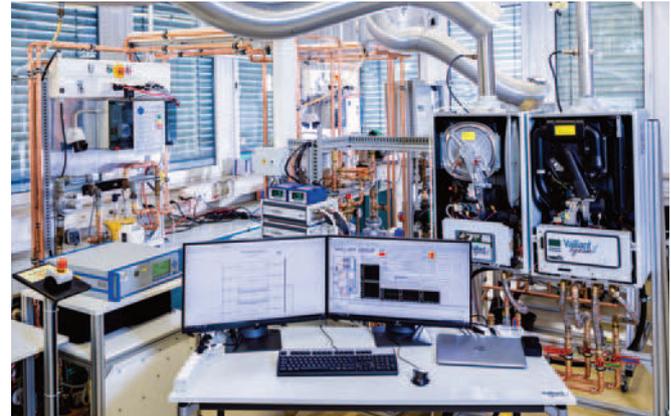


Abbildung 52: Ein Messstand im Labor der Vaillant GmbH (Foto: Mike König, Remscheid).

Optimierungsprobleme treten natürlich auch in vielen anderen Bereichen der Heizgeräteentwicklung auf. Im Rahmen von gemeinsam betreuten Abschlussarbeiten und Promotionen wurden und werden aktuelle Fragestellungen mit modernen und an die jeweilige Problemstellung angepassten Methoden behandelt. So wurden u.a. im Rahmen von Masterarbeiten die Parameter von Simulationsmodellen optimiert festgesetzt, um die Qualität und Vorhersagekraft der Simulationen möglichst gut an die im Labor gemessenen Daten (Abbildung 52) anzupassen. Zukünftig werden auch zunehmend KI-Methoden zum Einsatz kommen, z.B. wenn es um die optimierte Reihenfolgeplanung von Messzyklen geht.



Welcher Windturbinentyp liefert an welchem Standort die meiste Energie? Dazu müssen Sie wissen, wie hoch die Chancen für alle Windgeschwindigkeiten an diesem Ort sind. Die Weibullfunktion liefert diese Chance.

Statistik bestimmt die Wahl der Windturbine

Nichts ist so variabel wie der Wind. Manchmal stürmt es, manchmal ist es windstill. Manchmal weht der Wind aus dem Westen, manchmal aus dem Osten. Aber wo immer man die Windgeschwindigkeit auf der Erde misst, ist die Chance auf eine bestimmte Windgeschwindigkeit durch eine mathematische Funktion gegeben, die Weibull-Funktion. Der schwedische Ingenieur und Wissenschaftler Waloddi Weibull beschrieb diese Funktion erstmals 1939.

Die Weibull-Funktion unterscheidet sich von der bekanntesten Verteilung aus der Statistik, der Standard- oder Normalverteilung. Die Standardverteilung ist symmetrisch; die Weibull-Funktion ist asymmetrisch. Es handelt sich eigentlich um eine Standardverteilung, die bei positiven Windgeschwindigkeiten abgetastet wird. Für die Weibull-

Funktion zählt nur die Größe der Geschwindigkeit und nicht die Richtung. Wie bei einer Windkraftanlage spielt auch hier die Windrichtung keine Rolle, denn die Anlage dreht sich mit dem Wind.

In der Weibull-Funktion ist die Wahrscheinlichkeit der Windgeschwindigkeit Null sehr klein. Mit zunehmender Windgeschwindigkeit steigt die Wahrscheinlichkeit zunächst bis zu einem bestimmten Maximum. Danach nimmt die Chance, die Windgeschwindigkeit zu erhöhen, langsam ab. Nur zwei Faktoren bestimmen, wie gestreckt oder spitz die Wahrscheinlichkeitsverteilung ist: ein Skalierungsfaktor und ein Formfaktor.

In den Niederlanden gibt es viele verschiedene Windgeschwindigkeiten, was bedeutet, dass die Weibull-Funktion weit über das ganze Land ausgedehnt ist. Auf der anderen Seite gibt es am Äquator relativ wenige unterschiedliche Windgeschwindigkeiten: Dort weht er entweder hart oder nicht. Dies führt zu einer stark ausgeprägten Weibull-Funktion. Bei der Installation einer Windkraftanlage oder eines Windparks ist es unerlässlich, die Weibull-Funktion am vorgesehenen Standort genau zu kennen. Denn damit ist es möglich, den zu erwartenden Energieertrag zu berechnen.

Kommunizierende Schiffe

„Um die Weibull-Funktion an einem bestimmten Standort zu bestimmen, muss man den Wind an diesem Standort mindestens ein Jahr lang kontinuierlich messen“, sagt der Professor für Windenergie Gerard van Bussel von der Technischen Universität Delft. „Wenigstens haben Sie alle Jahreszeiten einmal erlebt. Sie müssen dann die Informationen aus der gemessenen Weibull-Funktion mit den von den Turbinenherstellern gelieferten Informationen über die Beziehung zwischen Turbinenleistung und Windgeschwindigkeit für einen bestimmten Typ von Windturbinen kombinieren. Daraus können Sie berechnen, welche Windkraftanlage an diesem Standort am besten passt und wie hoch der geschätzte Energieertrag ist.“

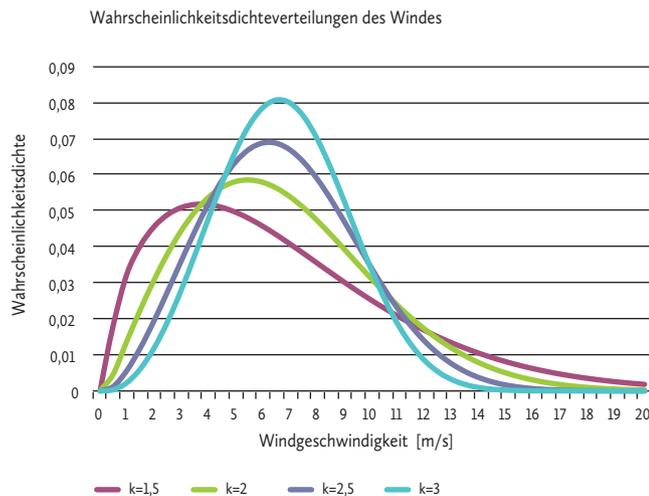


Abbildung 53: Beispiel für Weibull-Funktionen bei verschiedenen Formfaktoren (k).

Zuerst könnte man meinen, dass man sich für eine Windkraftanlage entscheiden sollte, die so viel Energie wie möglich liefert. In der Praxis sind die Energieunternehmen jedoch eher an einer möglichst zuverlässigen Versorgung mit einer bestimmten Energie interessiert. „Leider sind eine maximale Energieausbeute und maximale Zuverlässigkeit zwei kommunizierende Gefäße“, sagt van Bussel, „und so muss man sich zwischen beiden entscheiden. Dann mag es sein, dass Sie mit einer Fünf-Megawatt-Anlage die meiste Kilowattstunde produzieren, aber mit einer Zweieinhalb-Megawatt-Anlage liefern Sie die zuverlässigste Energie, einfach weil diese zweieinhalb Megawatt früher erreicht werden.“

Im Jahr 2014 werden die Niederlande 4,5% ihrer Energie aus Windkraft gewinnen, das sind 2.300 Megawatt. Der Plan für 2020 sieht vor, dass nicht weniger als zwölftausend Megawatt aus Wind kommen sollen, davon die Hälfte aus Onshore-Windparks und die andere Hälfte aus Offshore-Windparks.

Windmühlen auf See

Die Niederlande verfügen nun über zwei Offshore-Windparks: North Sea Wind in Egmond aan Zee und den Offshore-Windpark Prinses Amalia in IJmuiden. Bis 2020 sollen noch drei weitere hinzukommen. „Wir sind noch dabei zu lernen, wie man Windparks auf See baut“, sagt van Bussel. „Es ist nicht klug, landgestützte Turbinen auch auf See einzusetzen. Auf See verschleißten Windturbinen viel schneller, auch weil die Luft so salzig ist und der Wind stärker weht. Hinzu kommt, dass Reparaturen auf See teuer und bei weitem nicht jeden Tag möglich sind. Das erfordert robuste Turbinen.“

Viele Menschen unterschätzen die Anforderungen, die an Windkraftanlagen gestellt werden, sagt van Bussel. „Die größte Turbine der Welt hat Rotorblätter mit einem Durchmesser von hundertachtzig Metern. Das ist zweieinhalb Mal so lang wie das größte Passagierflugzeug der Welt, der Airbus A380. Und dieser Rotordurchmesser geht sogar bis zu zweihundertfünfzig Meter.“ Hinzu kommt, dass eine Windkraftanlage zwanzig Jahre lang ununterbrochen unter Bedingungen arbeiten muss, die mit den Start- und Landebedingungen eines Flugzeugs vergleichbar sind, also genau in den Momenten, in denen ein Flugzeug am meisten leidet. Van Bussel: „In zwanzig Jahren werden die Windturbinen gleich aussehen, aber sie werden alle clevere Tricks haben, so wie sich auch Autos und Flugzeuge in den letzten Jahrzehnten von Low-Tech zu High-Tech entwickelt haben. Aber die Weibull-Funktion bleibt unverändert.“

Dr. Michael Meister



Eine kurze Zusammenfassung Ihrer Karriere

Der Mathematik bin ich bereits seit meiner Schul- und Studienzeit bis in mein Berufs- und Politikerleben hinein sehr verbunden. In der Schule hatte ich großartige Lehrer, die uns Schülern die Naturwissenschaften mit großem Enthusiasmus nahebrachten. Als erstes Prüfungsfach in der Oberstufe wählte ich dann Mathematik; so war es eigentlich selbstverständlich, im Studium an der damaligen TH Darmstadt das Fach Mathematik mit Nebenfach Informatik zu belegen. Meine Doktorarbeit befasst sich mit der „Untersuchung des lokalen Abschneidefehlers für ein linearisiertes Einschrittverfahren zur numerischen Lösung von Systemen quasilinearer parabolischer Differentialgleichungen“.

Nach meinem Wehrdienst war ich dann fünf Jahre als Flugdynamik-Experte am Operationszentrum der Europäischen Raumfahrtbehörde (ESA/ESOC) in Darmstadt tätig. Dort war ich verantwortlich für die Bestimmung der Umlaufbahnen und die Lagekontrolle von Forschungssatelliten. Mit der Missionsplanung werden geeignete Zeitfenster für den Einsatz von wissenschaftlichen Instrumenten definiert. Im Jahr 1994 wurde ich das erste Mal in den Deutschen Bundestag gewählt. Dort war ich u.a. Vorsitzender der Arbeitsgruppe Finanzen und später lange Zeit stellvertretender Vorsitzender der CDU/CSU-Fraktion für die Bereiche Wirtschaft, Mittelstand, Finanzen und Haushalt. Von 2014 bis 2018 war ich Parlamentarischer Staatssekretär im Bundesministerium der Finanzen, seit 2018 bin ich in derselben Funktion im Bundesministerium für Bildung und Forschung (BMBF) tätig.

Was hat Mathematik für Ihre Karriere bedeutet?

Die Mathematik hat mich vor allem gelehrt, wie wichtig die Problemanalyse und das strukturierte Herangehen an die Erarbeitung von Lösungen ist. In der Mathematik muss man kurze, präzise Aussagen treffen und immer den gesamten Lösungsraum betrachten. Und auf Basis dieser Analyse muss dann der am besten geeignete

Lösungsweg zur Erreichung des Ziels gewählt werden. Das heißt, als Mathematiker schließt man keine Lösung von vornherein aus. Ich denke deshalb kommunizieren Mathematiker auch anders als Nicht-Mathematiker.

Meine Kenntnisse in der Mathematik waren für meine Tätigkeit bei der ESOC unabdingbar. Dort musste ich die Mathematik auf konkrete Anwendungsfelder in der Praxis anwenden. Und auch in meinem Politikerleben haben mir die mathematischen Arbeitstechniken und eine logische Herangehensweise an Probleme sehr geholfen. Gerade in der Finanzpolitik, mit der ich mich mehr als 15 Jahre beschäftigt habe, war mir mein mathematisches Grundverständnis sehr hilfreich. So konnte ich problemlos haushalts- oder steuerrechtliche Fragen bearbeiten und komplexe Vorgänge wie z.B. Tragfähigkeitsanalysen der öffentlichen Finanzen verstehen.

Welche Rolle spielt Mathematik in Ihrem Ministerium? Und allgemeiner in der Branche, in der Sie tätig sind?

Das BMBF misst der Mathematik eine zentrale Rolle in Wissenschaft und Forschung bei. Um die wichtigsten gesellschaftlichen Fragen unserer Zeit zu beantworten – also Fragen zu Klimaschutz, Digitalisierung, Nachhaltigkeit oder auch den Ursprüngen unseres Universums – braucht es Mathematik. Mathematische Lösungskonzepte, Modelle und Verfahren ermöglichen es, quantitative Vorhersagen zu machen und belastbare Antworten zu finden. Die Mathematik ist die Grundlage für alle Naturwissenschaften und Technologien und damit auch für die Entwicklung von Innovationen. Das BMBF beteiligt sich an der Finanzierung zahlreicher Großgeräte der naturwissenschaftlichen Grundlagenforschung, die modernste Technologien erfordern und eine oft beeindruckende Komplexität und Größe erreichen. Beispiele dafür sind die Beschleunigeranlage FAIR, das Deutsche Elektronen-Synchrotron DESY oder die europäische Röntgenlaser-Forschungseinrichtung XFEL.

Auch in der Bildung – dem zweiten Politikfeld, für das das BMBF zuständig ist – nimmt die Mathematik im Kanon mit den anderen Naturwissenschaften eine wichtige Rolle ein. Um auch weiterhin im globalen wirtschaftlichen Wettbewerb ganz vorne mit dabei zu sein, benötigen wir gut ausgebildete Fachkräfte. Deshalb ist die Förderung des MINT-Nachwuchses dem BMBF – aber auch mir persönlich – sehr wichtig. Im MINT-Aktionsplan hat das BMBF seine zahlreichen Förderprogramme zur Stärkung der MINT-Bildung entlang der Bildungskette gebündelt. Eine zentrale Maßnahme

ist zum Beispiel der bundesweite Ausbau der außerschulischen MINT-Angebote für Kinder und Jugendliche. Neben den bekannten Schülerwettbewerben wie den Mathe-Olympiaden oder Jugend forscht leisten diese MINT-Cluster einen Beitrag dazu, Jugendliche im entscheidenden Alter für Mathematik, Informatik und Naturwissenschaften zu begeistern und ihr Interesse an einem MINT-Beruf zu wecken.

Wie sieht ein typischer Arbeitstag aus?

Mir gefällt an meinem Beruf, dass jeder Arbeitstag unterschiedlich ist, je nachdem, ob ich in den Sitzungswochen des Bundestags in Berlin bin oder in meinem hessischen Wahlkreis Bergstraße Termine wahrnehme.

In Berlin ist meine Arbeitswoche sehr durchgetaktet: Montag und Dienstag treffen sich verschiedene Parteigremien, Mittwochvormittag nehme ich an den Sitzungen des Bundestagsausschusses für Bildung, Forschung und Technikfolgenabschätzung teil und den Rest der Woche vertritt das BMBF in den Plenarsitzungen des Bundestags. Außerdem bearbeite ich in meinem Büro Vorgänge, die parlamentarischen Bezug haben, halte Reden auf Konferenzen oder telefoniere mit Abgeordneten zu bildungs- und forschungspolitischen Themen. Eine Arbeitswoche hat da durchaus an die 70 Arbeitsstunden!

In meinem Wahlkreis führe ich Gespräche mit Bürgern, Vertretern der Kommunalpolitik, Unternehmen oder Vereinen und kümmere mich um deren Anliegen. Und ich genieße die Zeit mit meiner Familie!

Eines aber ist sicher: Ich treffe dank meiner Abgeordnetentätigkeit sehr viele interessante Menschen und befasse mich mit immer wieder neuen Aufgaben und Fragestellungen. Und ich kann wirklich etwas bewirken, d.h. Bürgern bezüglich ihrer Anliegen helfen und wichtige gesellschaftliche Themen voranbringen.

Sind, aus Ihrer Sicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Entscheidend sind aus meiner Sicht die interdisziplinäre Zusammenarbeit und der Transfer von der Wissenschaft in die Anwendung. Nur so kann das Potential der Mathematik voll ausgeschöpft werden. Das BMBF unterstützt dies z.B. mit dem Förderschwerpunkt „Mathematik für Innovationen“. Mit rund 5 Millionen Euro pro Jahr werden

anwendungsbezogene mathematische Forschungsarbeiten gefördert, die zu Lebensqualität und Wertschöpfung in Deutschland beitragen. Derzeit laufen hier Projekte zu den Themenkreisen Gesundheit, Energiewende und Digitalisierung. Die Themenvielfalt ist dabei sehr groß. Beispielsweise arbeiten Mathematiker an Ansätzen, um die medizinische Versorgung auf dem Land zu verbessern, die Einspeisung regenerativer Energien ins Stromnetz unter Wahrung der Netzstabilität zu erleichtern oder Lösungen für das Small-Data-Problem bei der Tumordiagnostik zu finden.

Wichtig ist aber auch die Kommunikation mit der Gesellschaft. Mathematiker sollten sichtbar sein und die Beiträge der Mathematik zum „Verstehen und Gestalten der Welt“ verständlich machen. Hier fällt mir z.B. Prof. Beutelsbacher ein, der mit dem Mathematikum im Gießen das erste mathematische Mitmach-Museum der Welt geschaffen hat. Auch Koryphäen wie Prof. Scholze, der im Jahr 2018 mit gerade einmal 30 Jahren die wichtigste Auszeichnung für Mathematiker, die Fields-Medaille, gewonnen hat, tragen dazu bei, die Faszination der Mathematik einer breiten Öffentlichkeit nahe zu bringen.

Was raten Sie jungen Leute hinsichtlich Mathematik?

Mit offenen Augen durch die Welt zu gehen, denn die Mathematik ist überall in unserem täglichen Leben präsent. Beim Einparken in eine Parklücke, beim Preisvergleich beim Lebensmitteleinkauf, bei der Nutzung eines Navigationsgeräts oder beim Falten eines Papierfliegers.

Neugier ist eine wichtige Antriebskraft für uns Menschen und ganz besonders für Kinder und Jugendliche. Neugier bedeutet offen sein für neues Wissen und neue Erfahrungen. Den Phänomenen der realen Welt und den unterschiedlichen Facetten eines Anwendungsproblems mit Neugier zu begegnen, ist der erste Schritt auf der Suche nach Lösungen.

Junge Menschen sollten die vielfältigen Möglichkeiten, die die Mathematik bietet, nutzen. Das abstrakte Denken, das man in der Mathematik lernt, ist in vielen Berufen relevant. Deshalb haben Mathematiker exzellente Berufsaussichten. Das wird auch in den kommenden Jahren so sein. Für junge Menschen bietet die Mathematik vielfältige Möglichkeiten, sich mit spannenden Fragestellungen zu befassen und an der Lösung drängender gesellschaftlicher Herausforderungen mitzuarbeiten.

Wie gelingt es Ihrer Meinung nach am besten, dass Ideen der Grundlagenforschung den Sprung in die Anwendung schaffen?

Wichtig ist es, die unterschiedlichen Akteure zusammenzubringen. Die Mathematik ist ein sehr schlagkräftiges Werkzeug – mit enormer Anwendungsvielfalt. Das Bewusstsein für diese Stärke ist eine der Voraussetzungen, um überhaupt Lösungen in der Mathematik zu suchen. Auf der anderen Seite fordern Anwendungsprobleme auch die Grundlagenforschung heraus und geben Impulse zu ihrer Weiterentwicklung.

Im bereits erwähnten BMBF-Förderschwerpunkt „Mathematik für Innovationen“ arbeiten Forschende gemeinsam mit der Wirtschaft an neuen Methoden für Herausforderungen aus der Anwendung. Diesen Austausch zwischen Wissenschaft und Wirtschaft unterstützen wir auch durch zahlreiche Veranstaltungsformate. Ein Aspekt der nicht zu unterschätzen ist, ist dabei der „Transfer über Köpfe“. Indem sich junge Wissenschaftler in den Projekten an Anwendungsproblemen probieren und zu neuen Lösungen beitragen, stärken sie gleichzeitig ihr Kompetenzprofil. Es ist kein Geheimnis, dass Absolventen mit praktischen Erfahrungen auf dem Arbeitsmarkt sehr begehrt sind – und mit ihnen gelangen auch neue mathematische Ideen in die Unternehmen.

Die Mathematik und ihre Bedeutung für unsere Gesellschaft haben sich in den letzten Jahren insbesondere durch Digitalisierung und Automatisierung verändert. Was bedeutet das für junge Menschen?

Die Mathematik existiert seit vielen tausenden Jahren, doch sie entwickelt sich immer noch weiter. Durch die Digitalisierung und Künstliche Intelligenz dringt Mathematik in immer mehr Arbeits- und Lebensbereiche vor. Der Bedarf an mathematischen Methoden und Instrumenten und ihr Anwendungsspektrum ist dabei gewaltig. Es reicht von der Internet-Suchmaschine und Social Media über Wetter- und Klimavorhersagen und die virtuelle Produktentwicklung bis hin zur effektiven Steuerung der Energienetze, bildgebenden medizinischen Verfahren und der individualisierten Medizin. Mathematik ermöglicht es, komplexe Daten zuverlässig und schnell auszuwerten und neue Erkenntnisse zu gewinnen. Die Digitalisierung kann dabei helfen. In den nächsten Jahren werden in diesen Bereichen zahlreiche Fachkräfte und Experten dringend gebraucht, um diesen Prozess der Veränderung mitzugestalten.



Zuwanderung und Integration in sich verändernden Gesellschaften

Ein mathematisches Netzwerkmodell simuliert die Immigration- und Koexistenzprozesse in der Europäischen Union.

Dr. Yao-Li Chuang und Prof. Dr. Maria R. D'Orsogna, Department of Mathematics, California State University, Los Angeles, fassen ihre Arbeiten zur Modellierung von Immigration zusammen:

Immigration gab es schon immer

Menschliche Migration war schon immer ein Teil der Menschheitsgeschichte. Unsere frühesten Vorfahren verließen Afrika vor über 70.000 Jahren und siedelten sich in Europa, Asien und Amerika an, auf der Suche nach neuen Jagdgründen. Bis zum Jahr 10.000 v. Chr. war der größte Teil der Landmasse der Erde besiedelt, und die Menschheit passte sich an neue Umgebungen an, lernte zu kooperieren, entwickelte neue Werkzeuge und erfand die Landwirtschaft.

In jüngerer Zeit verlassen Arbeiter ihre Heimat auf der Suche nach Arbeit oder besseren wirtschaftlichen Möglichkeiten, Flüchtlinge fliehen vor Bürgerkriegen, brutalen Diktaturen oder religiöser Verfolgung. Andere migrieren nach Naturkatastrophen, als Reaktion auf Hunger,

Dürren oder den Verlust von Ackerland. Nach Angaben der Internationalen Organisation für Migration (IOM) machen internationale Migranten im Jahr 2019 Migranten 3,5% der Weltbevölkerung ausmachen, ein Rekordhoch. Es wird erwartet, dass der Klimawandel weltweit noch viel mehr Menschen vertreiben wird.

Die Vereinigten Staaten haben seit ihrer Gründung viele Wellen der Einwanderung erlebt. Pilger und Puritaner waren unter den ersten, die sich im 17. Jahrhundert an der Ostküste niederließen, auf der Suche nach religiöser Freiheit. Versklavte Westafrikaner kamen gegen ihren Willen für mehr als hundert Jahre bis 1865, bis der 13. Zusatzartikel die Sklaverei abschaffte. Im 19. Jahrhundert trieb eine massive Hungersnot viele verarmte Iren an die Ostküste. Eine große Anzahl von Deutschen kam in der gleichen Zeit an, um der Armut und den Unruhen nach der Märzrevolution und den Aufständen von 1848-1849 zu entkommen. Sie ließen sich im Mittleren Westen nieder und gründeten blühende Gemeinden. Bei der Volkszählung im Jahr 2000



waren die Amerikaner deutscher Abstammung zahlreicher als jede andere Gruppe. Der kalifornische Goldrausch zog asiatische Einwanderer an, vor allem aus China, bis der 'Chinese Exclusion Act' im Jahr 1882 unterzeichnet wurde. Schließlich kam es zu Beginn des 20. Jahrhunderts zu einer groß angelegten Einwanderung aus Süd- und Mitteleuropa, hauptsächlich mittellose Italiener und verfolgte osteuropäische Juden. Die Weltwirtschaftskrise von 1929 und der Zweite Weltkrieg stoppten den Zustrom von Migranten in die Vereinigten Staaten, der während des Kalten Krieges mit der Einwanderung von politischen Flüchtlingen aus der UdSSR und Kuba wieder einsetzte. Im Jahr 1965 wurde der 'Immigration and Nationality Act' verabschiedet, der Quoten auf Basis der Nationalität abschaffte und die Familienzusammenführung ermöglichte. Dieses Gesetz, zusammen mit deutlich verbesserten wirtschaftlichen Bedingungen in Europa, veränderte die Migrationsmuster in die Vereinigten Staaten, so dass heute die Mehrheit der Neuankömmlinge aus Asien und Lateinamerika stammt. Europa, seit mehr als drei Jahrhunderten

eine Quelle von Migranten in Richtung USA, hat in den letzten Jahren selbst begonnen, Migranten aus Afrika, Asien und dem Nahen Osten anzuziehen.

Intrakontinentale Migration zwischen den Kontinenten ist ebenfalls auf dem Vormarsch, was durch die EU-Gesetzgebung begünstigt wird. Ein besonders anschauliches Beispiel für eine groß angelegte Migration nach Europa ist die Ankunft von mehr als einer Million Menschen in Deutschland und auf dem Rest des Kontinents von mehr als einer Million Menschen, zumeist Flüchtlingen aus Syrien, ab dem Sommer 2015 und in Folge des 2011 in Damaskus ausgebrochenen Bürgerkriegs. Über diesen schicksalhaften Sommer ist viel geschrieben worden: die bemerkenswerte Großzügigkeit der deutschen Bevölkerung, die politischen Differenzen auf nationaler und europäischer Ebene über die richtige Flüchtlingspolitik, die Belastung der EU-Grenzländer wie Italien, Griechenland und Ungarn, die nicht die Ressourcen oder den Willen hatten, eine so große Zahl von Neuankömmlingen zu bewältigen, die Ausbeutung der Migranten durch Schlepper und die allmähliche Zunahme der Unzufriedenheit in weiten Teilen der Gesellschaft über tatsächliche oder vermeintliche Fragen der Sicherheit, der Identität, des Säkularismus und des Multikulturalismus.

Unabhängig vom größeren gesellschaftspolitischen Diskurs bleibt eine Wahrheit bestehen: Viele Migranten haben sich in den letzten Jahren in europäischen Städten niedergelassen und werden dies auch weiterhin tun, ähnlich wie in den Vereinigten Staaten, die eine längere Einwanderungstradition und eine längere Geschichte von Zurückweisungen durch Einheimische, Erfolgsgeschichten und gesellschaftlichen Veränderungen haben, die durch den heiklen Übergang der Migranten von Außenseitern zu Insidern hervorgerufen wurden.

Wie wird Immigration eine Erfolgsgeschichte

Es stellt sich also eine zentrale Frage: Wie können wir sicherstellen, dass die Ankunft von Neuankömmlingen sowohl für die Migranten als auch für die langfristig Ansässigen von Vorteil ist? Die Integration von Zuwanderern in die breitere Gesellschaft ist wesentlich für die Schaffung lebendiger und zusammenhängender Gemeinschaften. Neuankömmlinge, die sich aufgrund der Umstände, der Abneigung von Einheimischen, mangelnder Motivation und/oder Ressourcen nicht gut integrieren, neigen dazu, sich selbst zu segregieren und Inselgruppen zu bilden. Diese Enklaven bieten den Zuwanderern zwar Vorteile und ein Gefühl der Zugehörigkeit, können sie sie aber auch

an der vollen bürgerlichen Teilhabe hindern und zu dem Risiko führen, dass Parallelgesellschaften entstehen.

Viele mathematische Modelle wurden eingeführt, um die Migration zu untersuchen, die oft als ein 'Push-Pull-Phänomen' beschrieben wird: Der Push entsteht durch die schlechten sozioökonomischen Verhältnisse in den Heimatländern der Migranten, der Pull resultiert aus den erstrebenswerteren Bedingungen in den Zielländern. Einer der ersten Gelehrten, der Migration im Sinne von Push-Pull-Begriffen beschrieb, war der deutsch-britische Geograph Ernst Ravenstein, der in den späten 1800er Jahren eine Reihe von „Gesetzen der Migration“ festlegte, auf denen die moderne Migrationstheorie aufbaut. Später wurden diese Gesetze mathematisiert, was zu verschiedenen Ausdrücken für M_{ij} , die Anzahl der Personen, die von Ort i nach Ort j ziehen, führte. Eine allgemeine Form ist von der Art

$$M_{ij} = k \frac{R_i + A_j}{d_{ij}}$$

wobei k eine Proportionalitätskonstante ist, d_{ij} die Entfernung zwischen Ursprung i und Ziel j ist und R_i, A_j die „abstoßenden“ und „anziehenden“ Eigenschaften des Ursprungs i bzw. des Ziels j sind. Beispiele für R_i, A_j können Arbeitslosenquoten, Löhne, Wohnkosten oder Personen auf dem Arbeitsmarkt sein, die um dieselben Arbeitsplätze konkurrieren. Eine alternative Form ist vom Typ

$$M_{ij} = k \frac{P_i^\alpha + P_j^\beta}{d_{ij}^\gamma}$$

wobei P_i, P_j die Populationen in den Herkunftsgebieten i und Zielgebieten j sind und α, β, γ Parameter, die an die Daten angepasst werden müssen. Der Wanderungssaldo aus der Region i ist gegeben durch $\sum_j (M_{ij} - M_{ji})$, wobei sich die Summe über alle möglichen Hoheitsgebiete j ergibt, in die oder aus denen Einwanderungs- oder Auswanderungsströme mit Verbindung zum Hoheitsgebiet i existieren. Manchmal werden die bekannten Migrationsströme verwendet, um Schätzungen für die Push- und Pull-Faktoren R_i, A_j zu rekonstruieren, ein anderes Mal werden die Daten verwendet, um die Werte von α, β, γ zu schätzen. Diese Modelle wurden über mehrere Jahre hinweg zur Untersuchung der internationalen und internen Migration angewandt, u.a. in und innerhalb der USA, Groß-

britanniens, Neuseelands, Italiens, Vietnams, zwischen der EU und den Nachbarländern, und wurden weiter angepasst, um Niederschlags- und Temperaturänderungen einzubeziehen, um die Auswirkungen des Klimawandels auf die Migration zu untersuchen.

Ein mathematisches Modell der Immigration

Um Probleme im Zusammenhang mit der Integration von Zuwanderern besser zu verstehen, kann ein mathematisches Modell, das die gesellschaftliche Dynamik unter Einbeziehung von Elementen aus der Netzwerktheorie und der Spieltheorie untersucht, verwendet werden kann. Es wird angenommen, dass die Gesellschaft aus zwei Gruppen besteht, die aus N_H „Gastgebern“ und N_G „Gästen“ bestehen und als Knoten dargestellt werden, die in einem zeitlich veränderlichen sozialen Netzwerk interagieren. Jeder Knoten repräsentiert ein Individuum oder eine kleine Gruppe von gleichgesinnten Individuen, entweder Gastgeber oder Gast. Beide Typen interagieren auf der Grundlage rationaler Entscheidungen miteinander, da sie versuchen, ihren sozioökonomischen Status zu verbessern, was zu Veränderungen im sozialen Netzwerk führt. Jeder Knoten i trägt eine zeitabhängige Einstellung x_i^t gegenüber anderen und erhält eine zu maximierende Nutzenfunktion U_i^t , die von seinen m_i^t Verbindungen abhängt. Im Laufe der Zeit passen die Knoten, um ihren Nutzen zu erhöhen, ihre Einstellungen an und schaffen neue oder kappen alte Verbindungen; als Ergebnis entwickelt sich das Netzwerk in Richtung Integration oder Segregation zwischen Gastgebern und Gästen. Während die Nutzenfunktion spieltheoretischen Regeln folgt, wird angenommen, dass sich die Einstellungen durch Meinungsdynamik entwickeln. Die beiden beeinflussen sich gegenseitig auf synergetische Weise. Eine Veranschaulichung finden Sie in Abbildung 54.

Die Einstellungen x_i^t variieren zwischen $0 \leq x_i^t \leq 1$ für Gastgeber und zwischen $-1 \leq x_i^t \leq 0$ für Gäste; der Betrag $|x_i^t|$ gibt den Grad der Feindseligkeit gegenüber der anderen Gruppe an. Somit charakterisiert $x_i^t \rightarrow 0^\pm$ die aufgeschlossenen Gäste bzw. die gastfreundlichsten Gastgeber, während $x_i^t = \pm 1$ die höchste Stufe der Fremdenfeindlichkeit darstellt. Der Nutzen U_i^t ist gegeben durch eine paarweise Belohnung, zu der jeder mit i verbundene Knoten j beiträgt, und durch eine Kostenfunktion für die Aufrechterhaltung von m_i^t Verbindungen, so dass

$$U_i^t = \sum_{j \in \Omega_i} A_{ij} \exp\left(-\frac{|x_i^t - x_j^t|^2}{2\sigma^j}\right) - \exp\left(-\frac{m_i^t}{a}\right).$$

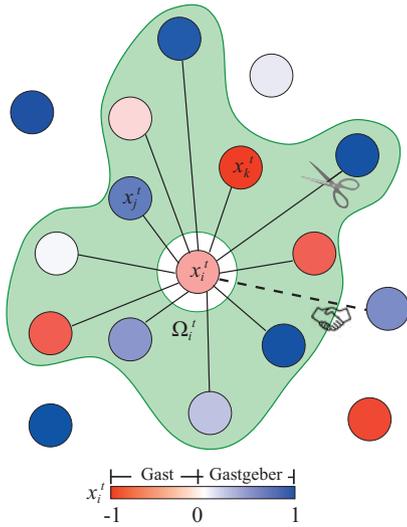


Abbildung 54: Jeder Knoten i ist charakterisiert durch eine variable Einstellung $-1 \leq x_i^t \leq 1$ zum Zeitpunkt t . Negative (rote) Werte stehen für Gäste, positive (blaue) Werte für Gastgeber. Der Betrag x_i^t repräsentiert den Grad der Feindseligkeit des Knotens i gegenüber Mitgliedern der anderen Gruppe. Alle mit Knoten i verbundenen Knoten j, k stellen den grün schattierten sozialen Kreis Ω_i^t von Knoten i zum Zeitpunkt t dar. Der Nutzen U_i^t des Knotens i hängt von seiner Einstellung relativ zu der seinen m_i^t Verbindungen ab. Die Knoten maximieren ihren Nutzen, indem sie ihre Einstellungen x_i^t anpassen und Verbindungen herstellen oder trennen.

Hier ist Ω_i^t die Menge der Knoten, die zum Zeitpunkt t mit i verbunden sind, so dass m_i^t durch seine Kardinalität gegeben ist, $|m_i^t| = |\Omega_i^t|$. Die paarweise Belohnung ist der erste Term auf der rechten Seite der obigen Gleichung und hängt von der Einstellungs-differenz $|x_i^t - x_j^t|$ zwischen den Knoten i und j ab. Je kleiner die Einstellungs-differenz ist, desto höher ist die Belohnung. Wenn also sowohl i als auch j zur gleichen Gruppe gehören, sowohl Gastgeber als auch Gäste, ist die Belohnung für $x_i^t = x_j^t$ maximiert, was zu einem Konsens zwischen ihnen führt. Wenn i und j jedoch aus verschiedenen Gruppen angehören, wird die Belohnung nur dann optimiert, wenn beide Knoten eine kooperativere Haltung einnehmen, $x_i^t \rightarrow 0^-$ und $x_j^t \rightarrow 0^+$. Der Parameter σ steuert die Empfindlichkeit der Belohnung, die Amplitude A_{ij} spezifiziert die maximal erreichbare Belohnung. Der Einfachheit halber kann A_{ij} als Konstante gesetzt werden. Die Kostenfunktion ist der zweite Term auf der rechten Seite von der obigen Gleichung und sinkt mit der Anzahl der

Verbindungen m_i^t , moduliert durch den Skalierungskoeffizienten α . Dieser Term begrenzt die Anzahl der aktiven Verbindungen, die Knoten i aufbauen kann.

Beim Fortschreiten von Zeitschritt t zu Zeitschritt $t + 1$ wird die Konnektivität eines gegebenen Knotens i modifiziert, um den Nutzen U_i^t zu maximieren; seine Einstellung x_i^t wird durch Nachahmung verändert. In der Praxis wird ein zufälliger Knoten j ausgewählt und seine Beziehung zu Knoten i ausgewertet. Wenn die Knoten i und j miteinander verbunden sind und das Trennen der Verbindung zwischen ihnen den Nutzen für den Knoten i erhöht, dann wird die Verbindung getrennt, ansonsten bleiben die beiden Knoten verbunden; wenn die beiden Knoten nicht verbunden sind und das Herstellen einer Verbindung zwischen ihnen den Nutzen für Knoten i erhöht, wird die Verbindung hergestellt, andernfalls bleiben die beiden nicht verbunden. Einstellungen werden geändert, indem ein anderer zufälliger Knoten l ausgewählt wird, der mit i verbunden ist verbunden ist, und durch Verkleinern der Lücke zwischen x_i^t und x_l^t . Die Idee ist, dass, sobald eine Verbindung hergestellt ist, die zwei Knoten dazu neigen, engere Ansichten zu teilen. Wenn Knoten i zum Beispiel ein Gastgeber ist, dann

$$x_i^{t+1} = \max \left(0, x_i^t + \frac{x_l^t - x_i^t}{\kappa} \right) \quad \text{für Gastgeber.}$$

Hier ist die Einstellung des Knotens i zum Zeitpunkt $t + 1$ näher an der Einstellung des Knotens l , den Knoten i zu imitieren versucht, aber wir behalten die untere Grenze bei 0, da aufgrund der Konstruktion alle Gastgeberknoten eine positive Einstellung haben müssen. Analog, wenn der Knoten i ein Gast ist, dann haben wir

$$x_i^{t+1} = \min \left(0, x_i^t + \frac{x_l^t - x_i^t}{\kappa} \right) \quad \text{für Gäste.}$$

Der Parameter κ in beiden Gleichungen regelt die Verhaltensanpassung: Kleine Werte von κ stehen für eine schnelle Nachahmung zwischen den Knoten i und j , während große Werte von κ eine langsamere Nachahmung darstellen.

Schließlich modellieren die Anfangsbedingungen die Art und Weise, wie die Gastgeber anfänglich in die Gemeinschaft integriert werden (oder auch nicht). Ein Extremfall ist der eines perfekt ausgeführten „Willkommens“-Programms, bei dem die Gäste über ausreichende

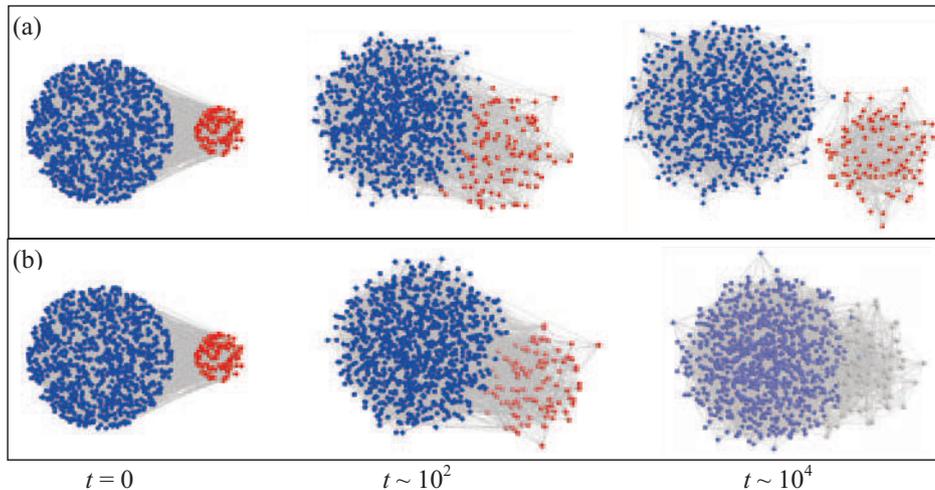


Abbildung 55: Simulierte Dynamik, die zu (a) einer vollständigen Segregation und (b) einer Integration zwischen Gast- (rot) und Gastgeberpopulation (blau) führt. Anfangsbedingungen sind zufällig verbundene Gast- und Gastgeberknoten mit den Einstellungen $x_{i,Gast}^0 = -1$ und $x_{i,Gastgeber}^0 = 1$. Die beiden Abbildungen unterscheiden sich nur durch κ , die Zeitskala der Einstellungsanpassung, mit $\kappa = 1000$ in Abbildung (a), wo segregierte Cluster entstehen, und $\kappa = 100$ in Abbildung (b), wo mit der Zeit ein verbundener Gastgeber-Gast-Cluster entsteht.

soziale Bindungen zu den Gastgebern verfügen und alle Knoten zufällig verbunden sind, unabhängig von ihrer Einstellung und ihrem Nutzen. Dies ist die in Abbildung 55 verwendete Anfangsbedingung. Die andere Anfangsbedingung ist die der nicht existierenden Anfangsressourcen, bei der die Gäste in einer völlig fremden Umgebung ankommen. Die Gastgeber sind natürlich miteinander verbunden, in ihrem eigenen Gleichgewichtszustand, und die Gäste werden ohne jegliche Verbindungen zu Gastgebern oder anderen Gästen eingeführt. In Abbildung 55 zeigen wir zwei repräsentative stationäre Ergebnisse ausgehend von den „perfekten“ Anfangsbedingungen der Vermischung von Gastgebern und Gästen zu Beginn. Der einzige Unterschied zwischen den beiden Tafeln ist der Parameter κ . In der oberen Abbildung, wo der Parameter κ groß ist, werden trotz der günstigsten Anfangsbedingungen Gastgeber und Gäste getrennt und verhalten sich äußerst feindselig. Jegliche anfänglich auferlegte gruppenübergreifende Nützlichkeit führt zu geringen Belohnungen, die im Laufe der Zeit nicht zunehmen so dass alle Verbindungen zwischen Gastgebern und Gästen schließlich gekappt werden. Die Akkumulation von sozioökonomischem Wohlstandes erfolgt eher durch insulare gruppeninterne Verbindungen, so dass schließlich Enklaven entstehen, in denen die beiden getrennten Gruppen einheitliche, aber unterschiedliche Haltungen einnehmen x_i . Im unteren Panel, wo der Wert niedrig ist, entwickeln alle Knoten kooperativere Einstellungen, die die gruppenübergreifenden Belohnungen, so dass Gastgeber und Gäste gemischt bleiben. Schließlich, $x_i \rightarrow 0$ auf allen Knoten.

Bildung von ethnischen Enklaven

Weitere Faktoren, die zur Enklavenbildung führen können, sind ein hohes Gäste-Gastgeber-Verhältnis (N_g/N_h), das die Wahrscheinlichkeit gruppeninterner Verbindungen erhöht und die Kommunikation zwischen der Gast- und der Gastbevölkerung verringert, sowie niedrige Werte der Amplitude A_{ij} , die geringe sozioökonomische Gewinne für Verbindungen zwischen Gastgebern und Gästen ergeben. Ein möglicher Weg, um Segregation zu vermeiden, ist daher die aktive Förderung von Gastgeber-Gast-Interaktionen, z.B. durch Anreize zur Verbesserung der kulturellen Anpassung von Gästen und Gastgebern (Senkung des Faktors in unserem Modell), durch die Sicherstellung, dass lokale Gemeinschaften nicht von Neuankömmlingen überwältigt werden (Senkung des lokalen Gäste-Gastgeber-Verhältnisses N_g/N_h), durch die Bevorzugung von Gästen, die gewünschte Fähigkeiten mitbringen, oder durch die Unterstützung beim Erwerb dieser Fähigkeiten (Erhöhung der Amplitude A_{ij}).

Beachten Sie, dass kulturelle Anpassung nicht zwangsläufig bedeutet, dass beide Seiten ihre Identität aufgeben müssen, sondern dass zur Förderung der Integration die verschiedenen Gruppen einander akzeptieren und versuchen müssen, eine Beziehung aufzubauen und ihre Unterschiede zu überbrücken. Dies ist natürlich die langfristige Herausforderung für die Zukunft. Für weitere Details verweisen wir den interessierten Leser auf Y.-L. Chuang, T. Chou, M.R. D'Orsogna, A network model of immigration: Enclave formation vs. cultural integration, *Networks & Heterogeneous Media* 14(1) (2019), 53-77. DOI: 10.3934/nhm.2019004

Extreme Meinungen verbreiten sich wie ansteckende Krankheiten. Und genau so kann man die Verbreitung und Radikalisierung mathematisch modellieren.

Mathematische Modellierung von Radikalisierungsprozessen

Dr. Yao-Li Chuang und Prof. Dr. Maria R. D’Orsogna, Department of Mathematics, California State University, Los Angeles, fassen ihre Arbeiten zur Modellierung von Radikalisierungsprozessen zusammen:

Die rasche Verbreitung radikaler Ideologien hat in den letzten Jahren zu einer weltweiten Folge von Terroranschlägen geführt. Das Verständnis, wie extremistische Tendenzen entstehen, sich entwickeln und Individuen zur Tat treiben, ist aus kultureller Sicht wichtig, aber auch, um Reaktions- und Präventionsstrategien zu formulieren. Demografische Studien, Interviews mit radikalisierten Personen und die Analyse von Terrordatenbanken zeigen, dass der Weg zur Radikalisierung in mehreren Stufen verläuft, wobei das Alter, der soziale Kontext und der Austausch zwischen Gleichgesinnten eine wichtige Rolle spielen.

Mathematische Modellierer haben Tools aus der statistischen Mechanik und der angewandten Mathematik eingesetzt haben, um beste-

hende und neue Erklärungsmodelle zu beschreiben und neue Strategien zur Terrorismusbekämpfung vorzuschlagen. Hier wollen wir den Ansatz der Kompartimentmodelle für Populationen mit zunehmend extremen Ansichten vorstellen.

Der Keim der Prä-Radikalisierung wurzelt oft in verschiedenen unbewältigten Missständen und persönlichen Frustrationen, wie z.B. Mangel an Beschäftigung und Chancen, rassistische und religiöse Diskriminierung und soziale Ausgrenzung. Individuelles Unbehagen kann auch von sozioökonomischen Ungerechtigkeiten, gegen eine als zu fortschrittlich empfundene Politik, oder umgekehrt aus dem Wunsch nach einer tiefgreifenden gesellschaftlichen Veränderung, kommen. Persönliche Unzufriedenheit führt zu einer Selbstidentifikation, bei der ausgegrenzte oder sich selbst suchende Individuen allmählich beginnen, neue Identitäten und Routinen zu konstruieren, und sich von alten abwenden.

Es werden aktiv Gleichgesinnte gesucht und neue Freundschaften geschlossen. Gegenseitige Ermutigung und das Fehlen einer Gegendialektik ermöglichen es extremistischen Ansichten, sich selbst zu verstärken und zu verfestigen. Hat man erstmal eine extreme Ansicht angenommen, wird es schwierig, die neuen Ideale zu verleugnen; Rechtfertigung und Lob der Gewalt folgen. Eine weitere Aktivismusphase folgt, wenn Radikale sich verpflichten, ihre Überzeugungen militant an andere weiterzugeben, bis äußere Ereignisse wie politische oder gerichtliche Entscheidungen oder einfache Zufälle die Bereitschaft zu gewalttätigem Handeln herauskristallisieren. Eine schematische Darstellung dieses Prozesses ist in Abbildung 56 dargestellt.

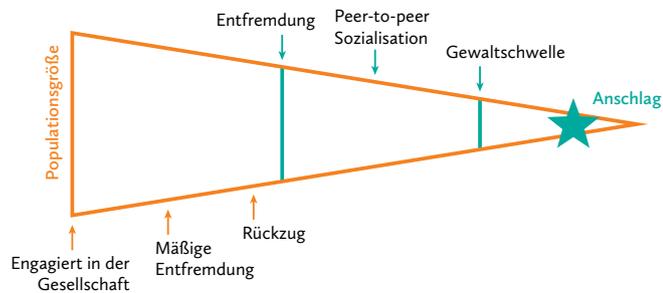


Abbildung 56: Schematische Darstellung des Radikalisierungsprozesses. Engagierte Individuen durchlaufen progressive Phasen des Rückzugs, bis sie sich vom Rest der Gesellschaft abgrenzen. Fortan kultivieren sie eine neue Identität, suchen Gleichgesinnte und bereiten sich auf Gewalt vor. Der Prozess gipfelt in möglichen terroristischen Anschlägen. Nicht alle Individuen durchlaufen die gesamte Hierarchie: Die Zahl der radikalierenden Individuen nimmt mit steigendem Extremismusgrad ab, was zu einem mehrstufigen horizontalen Trichter führt.

Radikalisierung erfolgt typischerweise durch Netzwerke von Gleichaltrigen und kann durch Technologien wie webbasiertem Rekrutierungsmaterial oder Chatrooms erleichtert werden; der IS ist eine bekannte internet-affine Gruppe, die soziale Medien nutzt, um westliche ausländische Kämpfer in Syrien und im Irak zu rekrutieren. Eine RAND-Studie²⁹ zeigt, dass die virtuelle Welt Extremisten die Möglichkeit bietet, ohne physischen Kontakt zu kommunizieren, zusammenzuarbeiten und zu überzeugen, und die Verbindung mit Gleichgesinnten aus der ganzen Welt zu ermöglichen.

Obwohl sie noch in den Kinderschuhen stecken und meist experimentell sind, werden weltweit mit großen Erwartungen De-Radikalisierungsprogramme entwickelt und umgesetzt. Sie umfassen

Bildung, Kontakt zu Literatur, Sport und Kunst, psychologische Beratung, Jobtraining und informelle Einzelgespräche. Das Ziel ist es, Radikalisierung zu verhindern, bekannte gewalttätige Extremisten zu entlassen und ehemalige verurteilte Extremisten wieder in die Gemeinschaft zu integrieren.

Radikalisierung in Meinungsdynamikmodellen

Wir diskutieren hier Radikalisierungsmodelle, die in den Kontext der Meinungsbildung eingebettet sind, ein sehr natürlicher Ausgangspunkt, da Radikale als Individuen mit extremen „Meinungen“ betrachtet werden können. Hier ändern sich die Überzeugungen eines Individuums – und der Gesellschaft insgesamt – durch persönlichen Kontakt, den Einfluss der Medien oder aktuelle Ereignisse. Die radikale Meinung und ihre Verbreitung wird dabei ähnlich wie eine Epidemie modelliert (vgl. Kapitel zur COVID-19 Modellierung).

Mean-Field-Modelle gehen davon aus, dass sich Individuen mit einer bestimmten Überzeugung gleichförmig verhalten und eine homogene Population definieren; es wird angenommen, dass jede dieser Populationen mit anderen interagiert, die ihre Ansichten auf reversible oder irreversible Weise beeinflussen, sodass im Laufe der Zeit die Größe einer bestimmten Population zunehmen oder abnehmen kann. Diese Dynamiken werden in der Regel durch Systeme von gewöhnlichen Differentialgleichungen (DGLn) beschrieben. Diese vereinfachten Modelle berücksichtigen nicht die Mechanismen, wie einzelne Individuen zum Fanatismus verleitet werden, jedoch ermöglichen sie die Identifizierung von Parametern oder Mechanismen, die beobachtete Trends antreiben. Das Modell von Carlos Castillo-Chavez und Baojun Song³⁰ ist eines der ersten Modelle, die Populationsdynamik in die Untersuchung der Übertragung fanatischer Verhaltensweisen einführt.

Konzepte wie die (effektive) Reproduktionszahlen R_0 , die verwendet werden, um vorherzusagen, wann eine Krankheit endemisch wird ($R_0 > 1$) und wann nicht ($R_0 < 1$), werden an die Verbreitung radikaler Ideologien angepasst. Insbesondere wird eine so genannte „Fanatiker-Hierarchie“ eingeführt mit der Gesamtpopulation $T(t)$, die in eine Untergruppe $G(t)$, die keine Neigung zur Radikalisierung hat, und drei Untergruppen von Individuen die sich in verschiedenen Stadien ihres Engagements für die extreme Ideologie befinden. Letztere sind $S(t)$, eine Gruppe von Anfälligen, die noch nicht radikalisiert sind, aber anfällig und offen für die jeweilige radikale Ideologie sind; die Population $E(t)$, Individuen, die sich vor kurzem in Fanatiker ver-

wandelt haben („Halbfanatiker“); und schließlich $F(t)$ Individuen, die sich vollständig radikalen Ansichten verschrieben haben. Die Summe der verschiedenen Teilpopulationen ergibt die Gesamtpopulation: $T = G + S + E + F$. Das Modell von Castillo-Chavez und Song lautet

$$\begin{aligned} \frac{dG}{dt} &= \Lambda - \beta_1 \frac{GC}{T} + \gamma_1 S + \gamma_2 E + \gamma_3 F - \mu G, \\ \frac{dS}{dt} &= \beta_1 \frac{GC}{T} - \beta_2 \frac{S(E+F)}{C} - \gamma_1 S - \mu S, \\ \frac{dE}{dt} &= \beta_2 \frac{S(E+F)}{C} - \beta_3 \frac{EF}{C} - \gamma_2 E - \mu E, \\ \frac{dF}{dt} &= \beta_3 \frac{EF}{C} - \gamma_3 F - \mu F, \end{aligned}$$

wobei $C(t) = T(t) - G(t)$ diejenigen Individuen darstellt, die anfällig sind, teilweise oder vollständig radikalisiert sind („Kern-Population“). Das System beinhaltet eine Eintrittsrate, den sog. „Geburtsterm“, Λ in die nicht-anfällige Population G und eine universelle Austrittsrate („Sterberate“) μ für alle Teilpopulationen, so dass nach Summation aller vier Gleichungen $dT/dt = \Lambda - \mu T$ gilt, was $\lim_{t \rightarrow \infty} T(t) = \Lambda/\mu$ liefert. Alle anderen Terme sind mit spezifischen Übergängen zwischen den verschiedenen Teilpopulationen verbunden, und zwar mit Rekrutierungsraten β_i und Rückföhraten γ_i . Wir ersetzen T durch den Grenzwert Λ/μ und $G(t)$ durch $\Lambda/\mu - C(t)$ in dem obigen DGL-System für die Zustandsvariablen S , E und F . Das reduzierte System ist „dynamisch“ äquivalent.

Ein zweiter globaler Schwellenwert $R_3 = \beta_3/\gamma_3$ steuert die Etablierung der fanatischen Bevölkerung $F(t)$ und folglich die Beharrlichkeit der fanatischen Ideologie. Wenn $R_3 \leq 1$ gilt, dann stirbt die fanatische Population im Grenzwert aus: $\lim_{t \rightarrow \infty} F(t) = 0$. Für diesen Fall $R_3 \leq 1$ benutzt man diesen Grenzwert und die Dimension des Modells kann weiter reduziert werden. Man erhält das folgende zweidimensionale System:

$$\begin{aligned} \frac{dS}{dt} &= \beta_1(1 - C)C - \beta_2 \frac{SE}{C} - \gamma_1 S, \\ \frac{dE}{dt} &= \beta_2 \frac{SE}{C} - \gamma_2 E, \\ C &= S + E. \end{aligned}$$

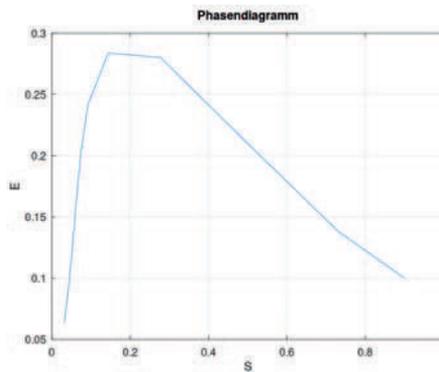
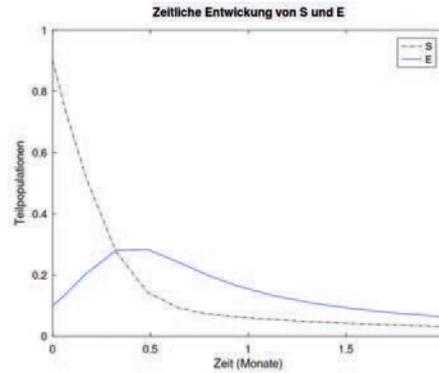


Abbildung 57: Simulation des reduzierten zweidimensionalen Systems.

In der Abbildung 57 sieht man eine Simulation dieses Systems mit den Parametern $S(0) = 0.9$, $E(0) = 0.1$, $\beta_1 = 3$, $\gamma_1 = 2$, $\beta_2 = 10$ und $\gamma_2 = 4$. Man sieht, wie die halbfanatistische Teilbevölkerung $E(t)$ zunächst auf 0.28 ansteigt und dann wieder abfällt.

Zum Beispiel können Individuen der nicht-empfänglichen Population G für radikale Ideologien empfänglich werden, wenn sie ihnen ausgesetzt sind; der Push der radikalen Botschaft wird über das Populationsverhältnis $C/T < 1$ und der zugehörigen Rate β_1 modelliert, so dass der Nettofluss $G \rightarrow S$ durch $\beta_1 GC/T$ ausgedrückt wird. In ähnlicher Weise hängt der Fluss vom anfälligen Zustand in die ersten Stufen des Fanatismus $S \rightarrow E$ von der Indoktrination und dem Vorbild von den radikaleren Individuen ab und ist gegeben durch $(E + F)/C$ moduliert durch die Rate β_2 . Der Übergang vom

radikalen Ausgangszustand zum voll engagierten Zustand $E \rightarrow F$ wird stattdessen allein von der radikalen Gruppe über F/C angetrieben und durch die Rate β_3 moduliert.

Das DGL-System definiert eine Hierarchie in dem Sinne, dass Übergänge durch die $G \rightarrow S \rightarrow E \rightarrow F$ -Stadien durch radikalere Populationen vorangetrieben werden, die die schwächeren beeinflussen. Die drei Rückführaten γ_i , für $i = 1, 2, 3$, repräsentieren die De-Radikalisierung aus den Untergruppen der Anfälligen, der kürzlich Konvertierten bzw. der vollständig Fanatisierten. Im Modell gibt es keine umgekehrten Zwischenübergänge: Radikale Individuen F können z.B. in die nicht-anfällige Population G zurückkehren, aber sie kehren nicht zu ihrer ersten Radikalisierungsstufe E zurück.

Trotz seiner Einfachheit bietet das Modell nützliche Einblicke in gesellschaftliche Ergebnisse, in Form von Attraktorpunkten und Schwellenwerten. Natürlich ist die wichtigste Frage, unter welchen Bedingungen eine endliche fanatische Bevölkerung F entsteht. Die Analyse des DGL-Systems zeigt, dass für $\gamma_1 > \beta_1$ kein Radikalisierungsgrad aufrechterhalten wird ($S = E = F = 0$), was bedeutet, dass eine Möglichkeit, einen radikalen Diskurs in einer Gesellschaft zu vermeiden, darin besteht, den Beginn des radikalen Prozesses in der frühen $G \rightarrow S$ Phase zu verhindern, wenn Individuen anfällig für Extremismus werden.

Abhängig von anderen Parameterkombinationen können Attraktorpunkte ohne fanatische Populationen ($S^* \neq 0, E = F = 0$) oder ($S^* \neq 0, E^* \neq 0, F = 0$), und mit fanatischen Populationen ($S^* \neq 0, E^* \neq 0, F^* \neq 0$) identifiziert werden. Ebenso wichtig sind die Anfangsbedingungen und Zeitskalen: Selbst wenn z.B. für den Gleichgewichtszustand $F = 0$ vorhergesagt wird, kann eine anfänglich kleine Gruppe von Extremisten erfolgreich in die Bevölkerung eindringen und zu einer großen fanatischen Population führen, bevor diese zu zerfallen beginnt.

Auf diesem DGL-System mit vier Größen basieren zahlreiche weitere Studien. Zum Beispiel die Möglichkeit von zwei konkurrierenden radikalen Gruppen, die von derselben allgemeinen, nicht anfälligen Population G ausgehen, wurde modelliert, indem man zwei Rekrutierungsraten $q\beta_1$ und $(1-q)\beta_1$ einbezog, die in zwei unterschiedliche anfällige Gruppen münden. Jede dieser beiden Kerngruppen folgt demselben Radikalisierungsprozess, der in dem DGL-System dargestellt ist, ohne zu interagieren. Beide Untergruppen haben auch

einen Kreuzkontakt über die $S \rightarrow E \rightarrow F$ -Hierarchie. Zwischen den beiden radikalen Zweigen gibt es einen Wettbewerb bei der Rekrutierung und den Bindungsbemühungen, so dass Fanatiker neue Anhänger aus Konvertiten der entgegengesetzten Ideologie gewinnen. Es wird insbesondere gezeigt, dass die beiden Gruppen nicht beide im Gleichgewicht ohne gegenseitige Interaktionen koexistieren können: kompetitive Rekrutierung ist notwendig für die Entstehung von zwei endlichen fanatischen Populationen F_1, F_2 .

Kompartiment-Modelle wie dieses wurden auch auf aktuelle extremistische Bewegungen angewendet, zum Beispiel auf den Einfluss separatistischer Gruppen im spanischen Baskenland³¹. Es wurden Teilpopulationen erstellt und Übergangparameter wurden mit Hilfe von demographischen und Wahldaten geschätzt; Immigrations- und Emigrationsdaten wurden ebenfalls einbezogen. Die ideologische Entwicklung der Gesellschaft wurde so über einen Zeitraum von 35 Jahren analysiert, einschließlich Projektionen in die Zukunft. Ebenso wurde dieses DGL-System für die Untersuchung der Radikalisierung rechtsextremer Bewegungen in Deutschland angepasst³², wo Umfragen auf der *Likert-Skala*, die ab 2002 alle zwei Jahre wiederholt wurden, zur Kalibrierung verwendet wurden. Es zeigte sich, dass die Nutzer mehrere Ebenen des Engagements durchlaufen, bis sie zu Anwerbern oder realen terroristischen Akteuren werden. Alternative Kompartiment-Modelle definieren z.B. Teilpopulationen über den kulturellen Hintergrund oder etwa einfach über das Lebensalter.

Für weitere Details verweisen wir auf die umfangreiche Arbeit von Yao-Li Chuang und Maria R. D'Orsogna³³, in der auch alternative mathematische Modelle von Radikalisierungsprozessen vorgestellt werden.

²⁹ I. von Behr, A. Reding, C. Edwards, L. Gibbon, Radicalisation in the digital era: The use of the internet in 15 cases of terrorism and extremism, RAND Europe (2013).

³⁰ C. Castillo-Chavez, B. Song, Models for the transmission dynamics of fanatic behaviors, Bioterrorism-Mathematical Modeling Applications in Homeland Security, Philadelphia: SIAM 28 (2003), 155-172.

³¹ F.J. Santonja, A.C. Tarazona, R.J. Villanueva, A mathematical model of the ressure of an extreme ideology on a society, Computers and Mathematics with Applications 56 (2008), 836–846.

³² T. Deutsch, Mathematische Modellierung von Radikalisierungsprozessen am Beispiel von rechtsradikalen Gruppierungen in Deutschland, Bachelorarbeit, Bergische Universität Wuppertal, November 2014.

³³ Y.L. Chuang, M.R. D'Orsogna, Mathematical models of radicalization and terrorism, arXiv preprint (2019):1903.08485.



Wie findet man eigentlich gute Zuordnungen? Welche Schülerin erhält einen Platz an ihrer Wahlschule? Welcher Schüler erhält seinen Wahlkurs? Welche Mannschaften sollen in Turnieren gegeneinander antreten? Und wann ist eine Zuordnung eigentlich gut?

Faire Schulplatzvergabe: von Boston in die ganze Welt

Dr. Ágnes Cseh vom Hasso-Plattner-Institut, Potsdam und Prof. Dr. Nicole Megow von der Universität Bremen berichten darüber, wie man gute Zuordnungen findet:

Gute und faire Zuordnungen werden ständig und in den verschiedensten Bereichen des Lebens gebraucht. Welche Schülerin erhält einen Platz an ihrer Wahlschule? Wie werden Schüler auf ihre Kurse verteilt? Welcher Bewerber wird an welcher Universität zugelassen? Welche Mannschaften sollen in Turnieren gegeneinander antreten? Wie werden Wohnheimplätze vergeben? All diese Fragen haben gemein, dass eine beschränkte Menge von Ressourcen den Bewerberinnen und Bewerbern zugeordnet werden soll und hierbei unterschiedliche Präferenzen zu berücksichtigen sind.

Mathematikerinnen und Mathematiker beschäftigen sich mit der Frage, wie man eine gute Zuordnung finden kann. Dafür werden Algorithmen entwickelt, also Verfahren, die solch eine Zuordnung finden können. Doch was ist eigentlich eine gute Zuordnung? Kann man Güte messen und vergleichen?

Betrachten wir als Beispiel die Vergabe von Schulplätzen genauer. Diese funktioniert in verschiedenen Städten und Bundesländern ganz unterschiedlich. Der Trend geht zu zentralisierten Verfahren innerhalb eines Schulbezirks. Hier wird den Familien die Möglichkeit gegeben eine Liste von, sagen wir drei, präferierten Schulen im Schulbezirk anzugeben. Den Schulen und Schultypen sind auch Prioritäten für die Schüler vorgegeben, z.B. nach Leistungsstand oder Länge des Fahrwegs, sowie eine maximale Aufnahmekapazität. Aus diesen Daten soll nun eine möglichst gute Zuordnung von Schülerinnen und Schülern auf Schulen erfolgen. Aber was nun genau gut heißt, ist dabei gar nicht so eindeutig. Es gibt viele mögliche Kriterien, die oft unvereinbar miteinander sind und nicht gleichzeitig perfekt erfüllt werden können. Dementsprechend gibt es für die Schulplatzvergabe mehrere Ansätze.

Boston Algorithmus – die erste Wahl zählt!

Ein einfaches Verfahren sieht wie folgt aus. Man versucht zunächst die Erstwünsche aller Schülerinnen und Schüler zu erfüllen. Dabei werden einige Schulen überangefragt sein und müssen Bewerber abweisen. Haben Bewerber gleich hohe Priorität bei einer Schule so ist ein zufälliges Auswählen nur fair. Was passiert aber mit den Abgewiesenen? Sie werden in einer zweiten Runde möglichst auf ihre Zweitwahl zugewiesen. Ein Teil dieser Schulen ist auch schon voll, und einige Bewerber werden erneut abgewiesen. Nach einer erneuten Runde sind alle Bewerber entweder zugewiesen oder haben ihre Liste von Wahlschulen abgearbeitet und keinen Platz gefunden.

Dieses Verfahren, man nennt es den Boston Algorithmus oder Immediate Acceptance Algorithmus (Algorithmus der sofortigen Annahmefestlegung), wird oder wurde in vielen Schulbezirken mit zentraler Schulplatzvergabe weltweit verwendet. Ist dies ein gutes Verfahren? Das kommt darauf an, was das Ziel ist. In der Optimierung sprechen wir von einer Zielfunktion, welche es zu maximieren gilt. Ist die Zielfunktion die Anzahl der erfüllten Erstwünsche, dann ist dieses Verfahren optimal. Das heißt, für jede Problemstellung (jeden möglichen Fall von Schülerwünschen und Schulen mit Prioritäten und Kapazitäten) wird die maximal mögliche Anzahl an Erstwünschen erfüllt.

Auf der anderen Seite werden Wahlverlierer in Kauf genommen, das heißt, dass es Schülerinnen oder Schüler gibt, die keine ihrer drei Wahlschulen erhalten, obwohl es eine solche Zuordnung gegeben hätte.

Die Schüler und ihre Eltern wissen, dass sie im Boston Algorithmus eventuell keinen Platz an einer ihrer drei Wahlschulen bekommen. Darüber hinaus ist ihnen klar, dass die Reihenfolge der Schulen extrem wichtig ist. Es kann passieren, dass die Schülerin Lilli die höchste Priorität an ihrer zweitplatzierten Schule A hat, sie wird aber nicht angenommen, weil Schule A ihre Kapazität bereits in der ersten Runde gefüllt hat – eventuell mit Schülern, die an der Schule eine geringere Priorität haben als Lilli. Sie ist berechtigterweise frustriert. Hätte Lilli Schule A auf Platz 1 ihrer Wahlliste gesetzt, wäre sie dort angenommen worden.

Solche Situationen lassen sich durch taktisches Wählen vermeiden. Die Familien sind vom Algorithmus gezwungen, die drei Schulen sehr vorsichtig zu wählen. Viele geben nicht ihre Wunschschule, sondern eine akzeptable, aber eventuell weniger beliebte Schule auf Platz 1 an, um die Chancen zu erhöhen, sich einen Platz gleich in der ersten Runde zu sichern. Das führt dazu, dass der Boston Algorithmus zwar beweisbar die maximal mögliche Anzahl von Schülerinnen und Schülern ihren Erstwahlschulen zuweist, aber diese Erstwahlen gar nicht die tatsächlichen Erstwahlen sind!

Diese Behauptung stammt von Atila Abdulkadiroğlu und Tayfun Sönmez, die 2003 einen vielbeachteten wissenschaftlichen Artikel geschrieben haben, in dem die Stadt Boston als Beispiel erwähnt wird. Noch im selben Jahr bekamen die zwei Wissenschaftler den Auftrag der Stadt Boston, die Schulplatzvergabe dort mathematisch zu analysieren und zu optimieren. Boston hat nach den Vorschlägen der Wissenschaftler den Algorithmus gewechselt, und die Schülerinnen und Schüler können im neuen System ihre wahren Schulpräferenzen ohne taktische Bedenken angeben. In vielen Städten weltweit wird dieses alternative Verfahren erfolgreich angewendet. Der einzige Verlierer ist der Ruf der Stadt: Der alte Algorithmus wurde nach Boston benannt, der Stadt, die ihn als Erste abgewählt hat.

Ein Verfahren gegen den Frust

Wie unterscheidet sich das neue Verfahren vom Boston Algorithmus? Auch hier versucht man Schülerinnen und Schüler jeweils in mehreren Runden auf ihre favorisierte, oder eben Zweit- oder Drittwahl-

Schule zuzuweisen. Der entscheidende Unterschied ist, dass hier Entscheidungen nicht sofort und endgültig getroffen werden sondern über die Runden hinweg noch angepasst werden. Deshalb wird der Algorithmus auch Deferred Acceptance Algorithmus (Algorithmus der verzögerten Annahmeentscheidung) genannt. Wenn Lilli in der ersten Runde ihre Erstwahlschule A nicht bekommen konnte und ihre Zweitwahl B schon voll ist, dann hat sie in der zweiten Runde trotzdem noch die Chance, einen Platz an Schule B zu erhalten, falls dort ein Schüler vorläufig zugeordnet wurde, der die Schulkriterien weniger erfüllt als Lilli.

Dieses Verfahren garantiert, dass es keinen berechtigten Frust gibt; Frust darüber, dass man einen Schulplatz nicht erhält der einem laut Schulprioritäten eigentlich zustünde. Das Ergebnis nennt man auch eine stabile Zuordnung. Tatsächlich werden hier sogar alle Schülerinnen und Schüler auf die für sie beste Schule zugewiesen, solange hierfür keine bessere Schülerin abgewiesen werden muss. Ganz konkret für Lilli bedeutet es, dass sie von Schule A auf keinen Fall eine Absage bekommen kann, nur weil Schule A ihre Kapazität bereits gefüllt hat. Die Absage in einer stabilen Zuordnung ist nur dann berechtigt, wenn Schule A ihre Kapazität mit Schülerinnen und Schülern aufgefüllt hat, die alle mindestens so gut den Schulprioritäten entsprechen wie Lilli.

Stabile Zuordnungen und der Deferred Acceptance Algorithmus wurden zum ersten Mal 1962 von David Gale und Lloyd Shapley in einem wissenschaftlichen Artikel beschrieben. Seitdem spielen sie eine wichtige Rolle bei zahlreichen Anwendungen. Das wurde auch von dem Nobelpreiskomitee so gesehen. Alvin Roth und Lloyd Shapley erhielten im Jahre 2012 den Alfred-Nobel-Gedächtnispreis für Wirtschaftswissenschaften, umgangssprachlich Wirtschaftsnobelpreis genannt, für die von ihnen entwickelte Theorie zu stabilen Zuordnungsverfahren und deren Anwendungen im Design von Marktmechanismen.

Die Rolle der Mathematik

Ist der Deferred Acceptance Algorithmus nun tatsächlich besser als der Boston Algorithmus? Das kommt auf die Ziele an. Wenn Lilli taktik-frei nach ihren wirklichen Wünschen wählen können soll und wenn uns Fairness in dem Sinne wichtig ist, dass es keinen berechtigten Frust geben soll, dann JA.

Dennoch, bei der Untersuchung des Problems der Schulplatzvergabe tauchen viele weitere Fragen auf, die abgewogen werden müssen. Darf es x berechtigt frustrierte Schüler geben, wenn dafür y Wahlverlierer vermieden werden können, oder wenn dadurch z Schülerinnen mehr ihre Erstwahl bekommen? Wie viele Wahlverlierer darf ein zusätzlich erfüllter Erstwunsch kosten?

Dies sind Entscheidungen, die nicht wir Mathematikerinnen und Mathematiker treffen, sondern Behörden, Politikerinnen, Elternvertreter, usw. Wir bieten dabei Entscheidungsunterstützung an und helfen beim Abwägen von gegensätzlichen Zielen und der Auswahl passender Verfahren. Dies tun wir, indem wir Probleme analysieren, Verfahren und mögliche Schwächen identifizieren und diese, wenn möglich, beheben indem wir neue Verfahren entwickeln. Außerdem garantieren wir als Mathematiker durch formale Beweise, dass die Eigenschaften der Algorithmen tatsächlich zutreffen. Damit fair auch wirklich fair ist.

Ein mathematisches Modell kann nützlich sein, um den potenziell schädlichen Folgen von gefälschten und Anti-Wissenschafts-Narrativen entgegenzuwirken und die hypothetischen Bedingungen zu untersuchen, unter denen eine nachhaltige Konspiration möglich sein könnte.



FAKE NEWS

FAKE NEWS

FAKE NEWS

FAKE NEWS

Ausfallwahrscheinlichkeiten für Verschwörungstheorien

Bei Verschwörungstheorien glauben einzelne Individuen, dass Ereignisse und Machtverhältnisse heimlich von bestimmten geheimen Gruppen und Organisationen manipuliert werden. Viele dieser vermeintlich erklärenden Vermutungen sind nicht falsifizierbar, nicht beweisbar oder nachweislich falsch, dennoch bleibt die öffentliche Akzeptanz hoch. Die Bemühungen, die Öffentlichkeit von der Gültigkeit medizinischer und wissenschaftlicher Erkenntnisse zu überzeugen, können durch solche Narrative behindert werden, die den Eindruck von Zweifeln oder Unstimmigkeiten in Bereichen erwecken, in denen die Wissenschaft gut etabliert ist. Umgekehrt gibt es historische Beispiele für aufgedeckte Verschwörungen und es kann für die Menschen schwierig sein, zwischen vernünftigen und zweifelhaften Behauptungen zu unterscheiden.

Dr. David Robert Grimes von der Universität Oxford hat in seiner Arbeit³⁴ ein einfaches mathematisches Modell für Verschwörungen mit mehreren Akteuren entwickelt, das die Wahrscheinlichkeit des Scheiterns für jede gegebene Verschwörung liefert. Die Parameter für das Modell werden anhand von Literaturbeispielen bekannter Skandale geschätzt, und die Faktoren, die den Erfolg und Misserfolg

von Verschwörungen beeinflussen, werden untersucht. Das Modell von Grimes wird auch verwendet, um die Wahrscheinlichkeit von Behauptungen aus einigen weit verbreiteten verschwörerischen Überzeugungen zu schätzen; diese sind nämlich, dass die Mondlandungen gefälscht wurden, dass der Klimawandel ein Schwindel ist, dass Impfungen gefährlich sind und dass ein Heilmittel für Krebs von Interessengruppen unterdrückt wird.

Anti-Wissenschafts-Verschwörungsnarrative – ein kurzer Überblick

Verschwörungstheorien, die ein schändliches, hinterhältiges Handeln von Wissenschaftlern unterstellen, sind allgegenwärtig. Dr. David Robert Grimes beschränkt sich in seiner Arbeit auf vier prominente Verschwörungstheorien:

NASA-Mondlandung. Die erfolgreiche Apollo-11-Mission im Jahr 1969 brachte erstmals Menschen auf den Mond, eine bahnbrechende Errungenschaft der Menschheitsgeschichte. Es gibt eine Randgruppe, die glaubt, dass die Mondlandungen zu Propagandazwecken vorgetäuscht wurde. Sie berufen sich dabei auf angeblichen Unge-

reimtheiten in Bildern die auf der Mondoerfläche aufgenommen wurden.

Klimawandel. Trotz der überwältigenden Stärke der Beweise, die den wissenschaftlichen Konsens der anthropogenen globalen Erwärmung unterstützen, gibt es viele, die diesen Konsens ablehnen. Viele von ihnen behaupten, der Klimawandel sei ein von Wissenschaftlern und Umweltschützern inszenierter Schwindel, angeblich um Forschungseinnahmen zu erzielen.

Impfungen. Verschwörerische Überzeugungen über Impfungen sind in der Anti-Impf-Bewegung verbreitet, so wird. z.B. geglaubt, dass es einen Zusammenhang zwischen Autismus und dem MMR-Impfstoff gibt. Dieser Glaube hat in mehreren Ländern die Inanspruchnahme wichtiger Impfungen verringert und hat zu einem Wiederauftreten von Krankheiten wie z.B. Masern geführt hat.

Krebsheilmittel. Der Glaube, dass ein Heilmittel für Krebs von Interessengruppen zurückgehalten wird, ist weit verbreitet. Er wird oft von Befürworter einer alternativen angeblichen Heilung verwendet, und die Behauptung der Verschwörungstheorie fungiert als Erklärungsansatz, um den völligen Mangel an klinischen Beweisen für solche Behauptungen zu erklären.

COVID-19. Die Coronavirus-Pandemie hat zu einem deutlichen Anstieg der medizinischen Desinformation und COVID-Verschwörungsnarrativen in den sozialen Medien geführt, wie z.B. dass COVID ein Hoax ist oder absichtlich hergestellt wurde, dass 5G-Frequenz-Strahlung das Coronavirus verursacht hat und dass die Pandemie eine List von großen Pharmaunternehmen ist, um von einem Impfstoff zu profitieren³⁵.

Herleitung des Modells

Es wird beim Modell von Grimes angenommen, dass bei einer gegebenen Verschwörung die Verschwörer im Allgemeinen größtenteils darauf bedacht sind, ihre Aktivitäten zu verbergen. Weiterhin wird angenommen, dass ein (absichtliches oder versehentliches) Informationsleck von einem beliebigen Verschwörer ausreicht, um die Verschwörung aufzudecken und sie überflüssig zu machen. Diese Aufdeckung einer Verschwörung ist ein relativ seltenes und unabhängiges Ereignis. Wir können dann die Poisson-Statistik anwenden und die Wahrscheinlichkeit von mindestens ein Leck, das zum Scheitern der Verschwörung führt, als

$$L(t) = 1 - e^{-\int_0^t \Phi(\tau) d\tau}$$

ausdrücken, wobei Φ die mittlere Anzahl der pro Zeiteinheit erwarteten Ausfälle ist. Diese ist wiederum eine Funktion der Anzahl der Verschwörer $N(t)$ und p , der intrinsischen Ausfallwahrscheinlichkeit pro Person pro Jahr. Dann kann man Φ angeben durch

$$\Phi = 1 - (1 - p)^{N(t)}$$

und wenn man der Kürze halber $\psi = 1 - p$ schreibt, kann die Wahrscheinlichkeit eines Verschwörungsausfalls als Funktion der Zeit umgeschrieben werden

$$L(t, N(t)) = 1 - e^{-\int_0^t (1 - \psi^{N(\tau)}) d\tau}$$

Wenn $\Phi(t)$ eine Konstante ist (homogener Poisson-Prozess), vereinfacht sich L zu

$$L(t, N(t)) = 1 - e^{-t (1 - \psi^{N(t)})}$$

Es gibt mehrere Möglichkeiten für die Anzahl der Verschwörer $N(t)$; die geeignete Auswahl hängt von der Art der Verschwörung ab.

Wenn eine Verschwörung eine konstante Aufrechterhaltung erfordert, dann ist die Verschwöreranzahl, die erforderlich ist, um die Fiktion aufrechtzuerhalten, ungefähr konstant mit der Zeit. Dies bezieht sich auf Situationen, in denen ein aktiver Beitrag zur Vertuschung eines Ereignisses oder zur Aufrechterhaltung einer Täuschung unerlässlich ist. In diesem Fall nimmt die Zahl der beteiligten Verschwörer eine einfache Form $N(t) = N_0$ an, wobei N_0 die anfängliche Anzahl der Verschwörer ist.

Wenn es sich bei der Verschwörung stattdessen um ein einzelnes Ereignis handelt, nach dem keine neuen Verschwörer benötigt werden, sterben die Beteiligten mit der Zeit aus, wodurch die Wahrscheinlichkeit der Aufdeckung sinkt. Wenn dies der Fall ist, kann eine *Gompertz'sche Überlebensfunktion* für die Funktion $N(t)$ verwendet werden. Wenn das Durchschnittsalter der beteiligten Personen zum Zeitpunkt des Ereignisses t_e ist, dann ist

$$N(t) = N_0 e^{\frac{\alpha}{\beta} (1 - e^{\beta(t+t_e)})}$$

wobei N_0 die anfängliche Anzahl der beteiligten Personen ist und α und β Konstanten für die Gompertz'sche Kurve sind. Für

Menschen können wir $\alpha = 10^{-4}$ und $\beta = 0,085$ wählen, um die menschliche Sterblichkeit zu beschreiben.

Wenn Verschwörer aufgrund interner Streitigkeiten oder anderweitig schnell entfernt werden, (eine Aktion, die selbst wohl ein meta-verschwörerisches Ereignis ist), kann es Umstände geben, wo wir $N(t)$ als einen *exponentiellen Zerfall* modellieren können. Wenn Mitglieder schnell entfernt werden und nach einer Periode t_2 nur noch die Hälfte übrig ist, dann ist die Zerfallskonstante $\lambda = \ln 2/t_2$ und die Anzahl der Verschwörer zu einem bestimmten Zeitpunkt ist

$$N(t) = N_0 e^{\lambda t}.$$

Diese Gleichung beruht auf der Annahme, dass ein schnelles Entfernen von Mitwissern die Aufdeckungswahrscheinlichkeit pro Mitwisser nicht verändert.

Wir sehen, dass eine Erhöhung von $N(t)$ immer zu einer Erhöhung von $L(t)$ führt, egal welche Form für die Verschwörerdichte gewählt wird. Die zeitliche Ausfallrate ist etwas komplizierter; für den konstanten Fall wird L monoton mit der Zeit anwachsen. Wenn stattdessen nicht-konstante Formen verwendet werden, ist L nicht-linear mit der Zeit. Der Zeitpunkt t_m , an dem L in diesen Fällen ein Maximum ist, ergibt sich durch Lösen von $\partial L/\partial t = 0$, das die folgende Gleichung liefert

$$1 - \psi^N(t_m) \left(1 + t_m \ln(\psi) \frac{\partial N}{\partial t} \Big|_{t_m} \right).$$

Diese Gleichung kann man mit numerischen Verfahren lösen. Die maximale Ausfallwahrscheinlichkeit ist dann $L(t_m)$ und die Form von $N(t)$ prägt deutlich die Dynamik des Problems, wie in Abbildung 58 dargestellt.

Parameterabschätzung

Um das Gompertz-Modell zu verwenden, müssen die Parameter geschätzt werden. Insbesondere der Parameter p , die Wahrscheinlichkeit eines intrinsischen Lecks oder Fehlers, ist extrem wichtig. Per Definition sind Details der Verschwörung selten bekannt, aber wir können die Parameter sehr konservativ schätzen, indem wir Daten von exponierten Beispielen verwenden. Zur Schätzung betrachten wir aufgedeckte Verschwörungen, bei denen ausreichende Daten über die Dauer und die Anzahl der Verschwörer öffentlich verfügbar sind. Als Beispiel wollen wir hier die PRISM-Affäre der NSA (National Security Agency) betrachten, also die Spionage von zivilen Internet-

nutzer, die das Anzapfen von Glasfaserkabel, Abhören von Telefongesprächen usw. umfasste. Mit den verfügbaren Daten kann man p konservativ schätzen. Dabei nehmen wir an, dass nach der Dauer t , wenn Verschwörungen aufgedeckt werden, die Wahrscheinlichkeit ihres Scheiterns bei $L \geq 0,5$ liegt. Eine untere Schranke für p ist dann

$$p \geq 1 - \sqrt[N(t)]{1 - \frac{\ln 2}{t}}.$$

Bei einigen dieser Schätzungen besteht eine beträchtliche und unvermeidliche Unsicherheit, insbesondere bei der Anzahl der Personen, die das Ereignis vollständig kennen. Im Fall von PRISM ist die Zahl 30.000 die Gesamtzahl der NSA-Mitarbeitern. In Wirklichkeit wäre der Anteil der Beschäftigten, die Kenntnis von diesem Programm hätten, wahrscheinlich viel geringer, aber wir nehmen die obere Zahl, um die Schätzung von p zu minimieren. Angesichts des kurzen Zeitraums nehmen wir weiterhin an, dass die Anzahl der Verschwörer über den Zeitraum vor der Aufdeckung des Ereignisses ungefähr konstant geblieben ist. Hinzu kommt, dass die Lebensdauer der Verschwörung nicht immer klar ist – im Fall der NSA liegen die Schätzung bei ca. 6 Jahren. Weiterhin verwenden wir die geschätzten Werte $p = 4,09 \cdot 10^{-6}$ und $\psi = 0,99999591$.

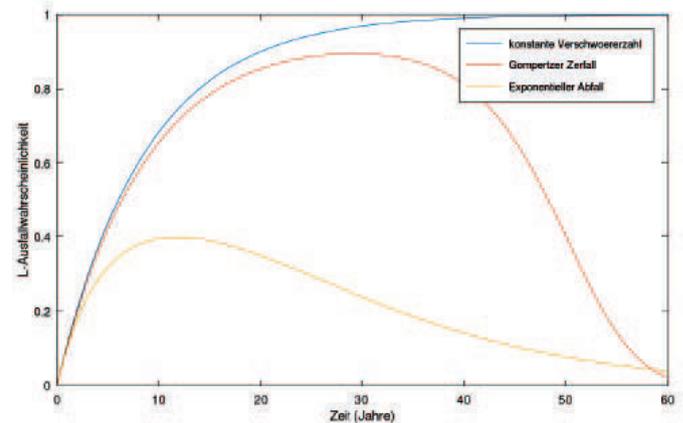


Abbildung 58: Projizierte Ausfallwahrscheinlichkeit L für eine Verschwörung mit 30.000 anfänglichen Verschwörern und $p = 4,09 \cdot 10^{-6}$ (Beispiel PRISM) und Halbwertszeit $t_2 = 8$.

³⁴ D.R. Grimes, On the Viability of Conspiratorial Beliefs, PLoS ONE 11(1) (2016), e0147905. (+ Korrektur)

³⁵ D.R. Grimes, Medical disinformation and the unviable nature of COVID-19 conspiracy theories, PLoS ONE 16(3) (2021), e0245900.

Die Zahl π beherrscht die moderne Satelliten-Kommunikation.

Schnelle Satelliten-Kommunikation mit π

„Superschnelles Internet für alle!“. Dieser Werbeslogan ist heutzutage überall zu finden. Aber ist das wirklich für jeden möglich? Internet-Provider mögen die Idee nicht wirklich, ein teures Internet-Kabel zu einem kleinen Dorf in den Bergen Sibiriens zu verlegen. Diese Menschen müssen sich mit einer Satellitenverbindung begnügen und diese ist nicht immer sehr schnell. Das liegt daran, dass mehr Menschen in entlegenen Gebieten der Welt die gleiche Verbindung nutzen wollen. Außerdem wird die Verbindung in Sibirien auch dann langsamer, wenn der Satellit in diesem Moment gerade über dem Südpol fliegt!

Menschen an abgelegenen Orten auf der ganzen Welt können sich seit kurzem glücklich schätzen - dank der Firma SpaceX und ihrem ehrgeizigen Projekt namens Starlink. Dieses Unternehmen ist dabei, ein großes Netzwerk aus mehreren Satelliten aufzubauen, um eine schnelle und konstante Verbindung für jeden Ort der Erde zu bieten. Um überall auf der Welt eine schnelle Verbindung zu garantieren, will SpaceX irgendwann 42 Tausend Satelliten ins All schießen! In der Animation kann man sehen, wie das aussehen wird.

Zum Zeitpunkt der Erstellung dieses Artikels (Februar 2021) befinden sich etwa 1000 Satelliten in einer Umlaufbahn um die Erde. Diese Umlaufbahn um die Erde wird auch Orbit genannt.

Um Unfälle zu vermeiden, muss der Start von 42 Tausend Satelliten gut durchdacht sein. Deshalb ist es wichtig, genau berechnen zu können, wie die Umlaufbahn des Satelliten aussehen wird. Und da kommt die Konstante π ins Spiel! In der untenstehenden Formel sieht man, dass mit Hilfe von π die Periode (p) eines Orbits berechnet wird:

$$p = \frac{2\pi r \sqrt{r}}{\sqrt{Gm}}$$

Mit anderen Worten: Diese Formel berechnet, wie lange ein Satellit für einen Umlauf um die Erde benötigt. Diese Formel kann auch in eine neue Formel umgeschrieben werden, mit der man den Radius (r) der Bahn des Satelliten bestimmen kann:

$$r = \left(\frac{p\sqrt{Gm}}{2\pi} \right)^{\frac{2}{3}}$$

Mit diesem Radius ist man in der Lage, die Höhe des Satelliten zu bestimmen. Wenn die Periode (p) und der Radius (r) berechnet sind, ist es möglich, die Geschwindigkeit (v) des Satelliten zu bestimmen. Auch hier spielt π eine Rolle:

$$v = \frac{2\pi r}{p}$$

Die Konstante π wird also in den Berechnungen verwendet, um Satelliten sauber in ihrer Umlaufbahn um die Erde zu halten. Das Starlink-Netzwerk von SpaceX ist eine der vielen Anwendungen, bei denen Satelliten eine Rolle spielen. Auch die Navigation im Auto oder auf dem Telefon verdanken wir der Technik hinter den Satelliten.

Verkehrsknoten von digitalen Datenpaketen

In Frankfurt am Main steht der größte Internet-Knotenpunkt Deutschlands, der Deutsche Commercial Internet Exchange (DE-CIX). DE-CIX trägt zu einem schnellen, zuverlässigen und kostengünstigen Internet bei.

Wenn Sie als Verbraucher einen Film von einem digitalen Medienunternehmen wie Netflix ansehen, erhalten Sie diesen Film von Ihrem eigenen Internet-Provider. Aber wie bekommt der Internet-Provider den Film von Netflix? Das geschieht beim Frankfurter Deutsche Commercial Internet Exchange, kurz DE-CIX (www.de-cix.net). Viele Großstädte in der Welt haben ihren eigenen Internet Exchange. Diese Internet Exchanges sind so konzipiert, dass angeschlossene Kunden digitale Daten kostengünstig, effizient und zuverlässig miteinander austauschen können. Kunde A tut dies nicht direkt und getrennt mit den Parteien B, C und D, sondern über einen solchen Verkehrsknoten gleichzeitig mit B, C und D.

Der DE-CIX ist der weltweit größte kommerzielle Internet-Knoten. Er wurde 1995 offiziell gegründet und befand sich im ehemaligen

Telegramm-Postamt, das schon ans Glasfasernetz angeschlossen war. DE-CIX ist nach angeschlossenen Internetdiensteanbietern der viert- und nach Verkehrsaufkommen der größte Internetknoten der Welt (Stand: April 2019).

Was der Hamburger Hafen für die Schifffahrt und der Flughafen Frankfurt für die Luftfahrt ist, das ist DE-CIX für den Internetverkehr.

Ethernetswitch

Bei DE-CIX müssen Netzwerkingenieure den Datenfluss zwischen allen Kunden optimieren. Sie verwenden Geräte, die Daten von einem Sender auf der Vorderseite empfangen, z.B. Netflix, und die gleichen Daten an den richtigen Empfänger auf der Rückseite senden, z.B. Ihren Internet-Provider.

Diese Geräte werden als *Ethernet-Switches* bezeichnet: große Schränke voller Kabel. DE-CIX verfügt über Zugangs-Ethernet-Switches, die direkt mit einem Kunden verbunden sind, und Kern-Ethernet-Switches, die die Zugangs-Switches miteinander verbinden. Die Verwendung von zwei Arten von Switches hat den Vorteil, dass DE-CIX auf diese Weise mehrere Möglichkeiten schafft, Daten von A nach B zu senden. Dies verringert das Risiko von Staus bei hohem Verkehrsaufkommen und erhöht die Zuverlässigkeit, falls eine der Straßen unerwartet ausfällt. Dabei muss das Problem gelöst werden, wie man die Kunden so über den ersten Typ von Ethernet-Switch aufteilen kann, dass die Kosten so niedrig wie möglich sind, während die Geschwindigkeit und Zuverlässigkeit des Datenverkehrs so hoch wie möglich sein muss.

Um dieses Problem zu lösen, kann man mathematische Methoden der *Kombinatorik* verwenden. Die Berechnung aller möglichen Verbindungen zwischen allen angeschlossenen Kunden würde eine unpraktische Menge an Rechenzeit in Anspruch nehmen. Die Kombinatorik hilft, schnell gute Lösungen zu finden. Dabei machen viele Kombinationen, die in der Theorie möglich sind, in der Praxis keinen Sinn. So wäre es beispielsweise sehr teuer, nur einen großen Kunden an eine Zugangsvermittlungsstelle anzuschließen, weil der gesamte Verkehr dieses Kunden dann von DE-CIX zu anderen Vermittlungsstellen transportiert werden müsste. Die Algorithmen, die in der Software verwendet werden, können dies berücksichtigen. Eine einfache Sortierung, die das Problem sehr viel einfacher macht, besteht darin, Kunden, die viele Daten miteinander austauschen, so weit wie möglich an denselben Ethernet-Switch anzuschließen.

Graphentheorie

Ein zweites Problem, das hier mathematisch gelöst wird, hat mit der proaktiven Prüfung zu tun, ob Verbindungen zwischen Ethernet-Switches noch ordnungsgemäß funktionieren. Zu diesem Zweck wird die *Graphentheorie* verwendet, die mathematische Theorie der Netzwerke, die Knoten verbinden. So kann festgestellt werden, wie mit möglichst wenigen Testgeräten noch alle Verbindungen überprüft werden können. Der Algorithmus, der dafür verwendet wird, ist eine Variante des berühmten *Dijkstra-Algorithmus*, der zum Beispiel auch in der Navigationssoftware in Autos verwendet wird.

Die Menge des über den DE-CIX abgewickelten Internetverkehrs wächst jedes Jahr deutlich an, so dass jedes Jahr mehr Ethernet-Switches installiert werden müssen. Dies geschieht gewöhnlich in

den ruhigen Sommermonaten. Im Jahr 2019 beträgt die Spitzenkapazität des DE-CIX nicht weniger als acht Terabit pro Sekunde - in der Größenordnung von zehntausend Mal schneller als das, was Sie als Verbraucher zu Hause haben. Und dank einer deutlichen Überkapazität und der Verteilung von Ethernet-Switches über unzählige physische Standorte in Frankfurt und vielen anderen Austauschpunkten in Deutschland und weltweit wird eine schnelle und robuste Internet-Infrastruktur geschaffen. Durch die Lockdown-Maßnahmen während der COVID-19-Pandemie stieg der Datenverkehr am DE-CIX seit Anfang/Mitte März 2020 weiter an, speziell im Bereich Videokommunikation und im Bereich Gaming.



Abbildung 59: DE-CIX funktioniert wie ein Kreisverkehrsplatz für den Internetverkehr: Über DE-CIX kann Kunde A alle anderen Kunden gleichzeitig erreichen.

Prof. Dr.-Ing. Stefan Kurz

Eine kurze Zusammenfassung Ihrer Karriere

Prof. Dr.-Ing. Stefan Kurz (1966), Studium der Elektrotechnik an der Universität Stuttgart, Diplom 1992, Promotion 1998. Verschiedene berufliche Stationen bei der Robert Bosch GmbH und im akademischen Umfeld, unter anderem Professor für Theoretische Elektrotechnik und numerische Feldberechnung an der (damaligen) Universität der Bundeswehr Hamburg. „Finland Distinguished Professor“ an der (damaligen) Technischen Universität Tampere (2010-2014).

Heute Tätigkeit als Hauptreferent (Senior Chief Expert) für „Hybride Modellierung“ (Physik und Daten) am „Bosch Center for Artificial Intelligence“ (BCAI). Gleichzeitig Kooperationsprofessor für mathematische Modellierung elektromagnetischer Felder an der Technischen Universität Darmstadt.

Was hat Mathematik für Ihre Karriere bedeutet?

Mathematik hat sich als roter Faden durch alle Stationen meiner Karriere gezogen, was zum Beispiel an den Ausrichtungen meiner Professuren erkennbar ist.

Bereits als Schüler faszinierte mich Mathematik, insbesondere die Erkenntnis, dass Mathematik eine Art „lingua franca“ der physikalischen Welt ist. Aus diesem Grund wählte ich im Gymnasium die klassischen Leistungskurse Mathematik und Physik. Ich entschied mich für das Studium der Elektrotechnik, weil ich neugierig auf deren Anwendungen war. Aber bereits während des Grundstudiums war ich mehrfach versucht zur Mathematik zu wechseln – bis im Hauptstudium dann die Theorie Elektromagnetischer Felder und Wellen hinzukam. Dieses mathematisch geprägte Thema begleitet mich bis heute, beispielsweise für die Auslegung supra-leitender Magnete des „Large Hadron Collider“ (LHC), des großen Teilchenbeschleunigers der europäischen Organisation für Kernforschung CERN in Genf.



Neben der akademischen Seite spielt die Anwendung der Mathematik auch in meinem beruflichen Umfeld eine entscheidende Rolle. Heute bin ich im „Bosch Center for Artificial Intelligence“ tätig, wo ich mich neben strategischen Fragen um „Hybride Modellierung“ kümmerge. Im Kern geht es um die Kopplung von physikalisch basierten mit datengetriebenen Modellen, was für Bosch von grundlegender Bedeutung ist. Das gesamte vorhandene Wissen um die Produkte und ihre Wirkzusammenhänge aus mehr als 130 Jahren Firmengeschichte stammt überwiegend aus der Physik. Dieses Wissen wird durch „Hybride Modellierung“ in die Welt des maschinellen Lernens und der künstlichen Intelligenz integriert. Außerdem helfen solche „Hybriden Modelle“, die künstliche Intelligenz sicher, robust und erklärbar zu machen. Dies ist eine Grundvoraussetzung, um künstliche Intelligenz im industriellen Kontext anwenden zu können.

Die „Hybriden Modelle“ sind gleichzusetzen mit mathematischen Modellen, aus verschiedenen mathematischen Bereichen. Beispielsweise werden stochastische Prozesse aus der Wahrscheinlichkeitstheorie mit Differentialgleichungen gekoppelt.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihrer Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Mathematik spielt eine zentrale Rolle bei der Produktentwicklung bei Bosch. Als international führendes Technologie- und Dienstleistungsunternehmen ist Bosch in unterschiedlichen Branchen vertreten. Die Aktivitäten gliedern sich in die vier Unternehmensbereiche Mobility Solutions, Industrial Technology, Consumer Goods sowie Energy and Building Technology. Ich möchte dazu zwei Beispiele geben.

Permanente Innovation bei den Entwicklungsmethoden ist ebenso wichtig wie Produktinnovation. Ein Schlüsselement in der heutigen Produktentwicklung ist die Simulation. Dies erfordert physikalische und mathematische Modelle aus der realen Welt. Ebenso notwendig sind effiziente numerische Lösungsverfahren, die es sozusagen ermöglichen, Fragen an die Modelle auf dem Computer stellen zu können.

Vielleicht haben Sie schon einmal vom sogenannten Moore'schen Gesetz gehört, nach Gordon Moore, dem Mitbegründer der Firma Intel. Es besagt, dass sich die Leistungsfähigkeit von Computerhardware empirisch etwa alle 18 Monate verdoppelt. Was weniger bekannt ist: eine ähnliche exponentielle Entwicklung findet man auf dem Gebiet mathematischer Algorithmen, zum Beispiel für die Lösung der stationären Wärmeleitungsgleichung. Vom Gaußschen

Eliminationsverfahren, das Sie vielleicht aus der Schule kennen, bis zur modernen Mehrgittermethode wurde über knapp 30 Jahre hinweg ein Leistungsgewinn um den Faktor Hunderttausend erzielt³⁶. Ist das nicht beeindruckend? Das ist es, was mich an der angewandten Mathematik so fasziniert.

Derartige Leistungssteigerungen sind essentiell im Zeitalter der digitalen Transformation, wo wir eine schnelle Zunahme von vernetzten und automatisierten Systemen sehen. Im Hinblick auf diese zunehmende Komplexität können die Produkte von morgen nicht mit den Methoden von gestern entwickelt werden.

Mein zweites Beispiel stammt aus dem Bereich der Entwicklung von Antriebsmotoren für Elektrofahrzeuge. Dort ist Mathematik der Schlüssel zu Produktivität und funktionalen Verbesserungen in der Entwicklung. Mit der sogenannten Mehrzieloptimierung werden konstruktive Parameter systematisch variiert, um zehntausende ihrer Varianten auf einem Hochleistungs-Computer prozessieren zu können. So entwerfen wir Komponenten an ihren physikalischen Grenzen, berücksichtigen dabei die Markterfordernisse und lösen Widersprüche bei den Entwurfszielen auf.

Sind, in Ihrer Hinsicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

In den letzten Jahren sehe ich diverse Initiativen, um die wichtige Rolle der Mathematik in den Schulen und der Gesellschaft zu stärken, worüber ich mich sehr freue. Mit meiner Kooperationsprofessur an der Technischen Universität Darmstadt bin ich ja selbst als Brückenbauer unterwegs. Durch spannende Anwendungsbeispiele aus Industrie und Forschung gelingt es, die theoretischen Grundlagen nachvollziehbar zu illustrieren. So entsteht ein fruchtbarer Dialog für beide Seiten. Im gemeinsamen Diskurs entstehen neue Denkmuster für die Anwendung in Forschung und Industrie. Mit der Zeit ist zudem ein wertvolles Netzwerk gewachsen, das auch für interessierte Studierende eine tolle Gelegenheit bietet, sich beruflich zu orientieren.

Den Austausch zwischen den Generationen empfinde ich darüber hinaus als eine echte Bereicherung. Für mich ist eine Lehrveranstaltung gelungen, wenn ich mit den Studierenden eine Art Lerngemeinschaft bilden kann, aus der alle Beteiligten reicher hinaus als hineingehen. Dadurch gelingt es mir, am Puls der Zeit zu bleiben.

³⁶ OECD Global Science Forum, Report on Mathematics in Industry 2008, S. 8



Die Mathematik des Drängelns

Die Dynamik von Fußgängerströmen ist ein junges und wachsendes Forschungsfeld in dem unterschiedliche wissenschaftliche Disziplinen arbeiten. Die Mathematik hat in diesem interdisziplinären Feld eine wichtige integrierende Rolle. Die Anwendungen reichen von der Organisation von Großveranstaltungen, über die Optimierung von Verkehrsanlagen bis zur Gestaltung von Fluchwegen für Schulen oder Museen.

Abbildung 60: Ein Experiment mit 700 Probanden in dem Stauungen an Kreuzungen untersucht wurden. Die Experimente fanden 2013 im Rahmen eines Forschungsprojektes zu Großveranstaltungen in den Messehallen in Düsseldorf statt. (Quelle: Forschungszentrum Jülich / Foto: Marc Strunz-Michels)

Armin Seyfried beschäftigt sich seit 15 Jahren mit der Dynamik von Personenströmen. Mira Küpper arbeitet in Ihrer Promotion an der Analyse von Laufwegen und Warteverhalten an Bahnsteigen. Hierfür kooperieren sie mit Wissenschaftler*innen aus dem Ingenieurwesen, der Mathematik, den Computerwissenschaften und der Sozialpsychologie. Die Zusammenarbeit mit Mathematiker*innen ist wichtig, da sie zu neuen Impulsen in der Datenanalyse und Modellierung führen.

Obwohl der Zusammenhang zwischen Fußgängern und der Mathematik nicht direkt ersichtlich ist, gibt es viele mathematische Methoden die in der Forschung zur Fußgängerdynamik eine zentrale Rolle spielen und aus denen sich eine Vielzahl von praktischen Anwendungsbereichen ergeben. Wie bekommt man mehrere tausend Besucher in ein Fußballstadion und auch wieder heraus? Wie kann man berechnen wie breit die Fluchtwege aus einem Gebäude sein müssen, sodass im Falle einer Evakuierung möglichst viele Menschen in kurzer Zeit das Gebäude verlassen können? Wo kommt es dabei zur Entstehung von Stau und wie kann man dies verhindern?

Simulation von Fußgängerverkehr

Die Schwierigkeit bei solchen Fragestellungen besteht darin, dass sich im Gegensatz zum Autoverkehr, bei dem es konkrete Verkehrsregeln gibt (Fahrspuren, Rechts-vor-Links usw.) Fußgänger kreuz und quer bewegen können. Um die Bewegungen von Fußgängern detailliert untersuchen zu können, werden häufig Fußgänger-Experimente unter „Laborbedingungen“ durchgeführt, siehe Abbildung 60. Im Rahmen von verschiedenen Forschungsprojekten des Lehrstuhls „Computersimulation für Brandschutz und Fußgängerverkehr“ der Bergischen Universität und dem „Institute for Advanced Simulation“ des Forschungszentrum Jülich wurden bereits zahlreiche solcher Experimente durchgeführt. Im Kontext der Fußgänger-Dynamik bedeutet dies, mehrere hundert freiwillige Probanden in einer großen Halle durch Korridore und Engstellen drängeln zu lassen, um die Bewegungen von Personen wissenschaftlich beobachten zu können. Die Probanden tragen bei den Experimenten farbige Mützen und werden von Kameras aus der Vogelperspektive gefilmt, sodass im Nachhinein die Laufwege der Personen nachvollzogen werden können. Mit verschiedenen Experimentaufbauten lässt sich dabei untersuchen, wann die Probanden eine geordnete Warteschlange bilden und welche Faktoren dazu führen, dass gedrängt wird.

Experimente für die Sicherheit im Alltag

Analysiert werden die durch Experimente erhobenen Daten in Hinblick auf die Dichte; das heißt wie viele Personen sich auf einem Quadratmeter Fläche aufgehalten haben; die Geschwindigkeit der Probanden, sowie den Fluss, welcher die Anzahl der Personen beschreibt, die in einem festen Zeitintervall eine vorgegebene Linie zum Beispiel eine Engstelle durchquert haben. Durch den Vergleich der Ergebnisse von unterschiedlichen Versuchsanordnungen kann analysiert werden wie sich eine Änderung des Aufbaus (z.B. Korridore mit verschiedenen Breiten) oder der Anweisung der Probanden (z.B.

„Nur die Schnellsten bekommen einen guten Sitzplatz“) auswirkt.

Da Experimente immer nur in einem überschaubaren Rahmen und damit für vereinfachte Fragestellungen durchgeführt werden können und es zum Beispiel viel zu aufwändig wäre in einem Experiment zu prüfen wie lange es dauert ein ganzes Fußballstadion zu evakuieren, werden solche komplexen Fragestellungen über Simulationen untersucht. Als Simulation wird ein Computermodell bezeichnet mit dem versucht wird die Realität nachzubilden.

Mit guten Computermodellen zur Bewegung von Personenströmen kann die Sicherheit von Personen in Alltagssituationen erhöht werden. Daher handelt es sich um ein Forschungsgebiet mit Alltagsbezug zu dem jeder Einzelne eigene Erfahrungen beisteuern kann.

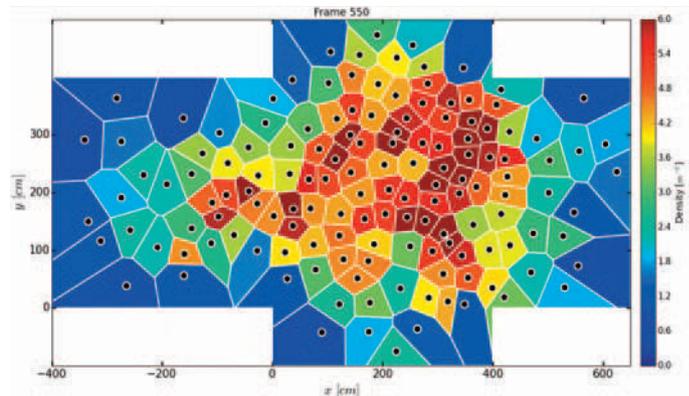


Abbildung 61: Voronoi-Zerlegung des Raumes für die Bestimmung individueller Dichten. Bei den Rohdaten, schwarze Punkte, handelt es sich um Positionen der Fußgänger. Sie wurden in einem ähnlichen Experiment gemessen, wie es Abbildung 60 zeigt. Die Werte zeigen, dass die Dichte mehr als 6 Personen auf einem Quadratmeter betragen kann.

Mit mathematischer Optimierung zu sicheren und klimafreundlichen Flügen

Eine spannende Herausforderung für die mathematische Theorie, Forschung und Praxis der optimalen Steuerung angewandt auf ein Problem von hoher gesellschaftlicher, ökologischer und ökonomischer Relevanz.

Abbildung 62: Ein ruhiger Nachmittag in Heidelberg.

Ein Bericht aus der wissenschaftlichen Praxis von Prof. Dr. Ekaterina Kostina (Institut für angewandte Mathematik und Interdisziplinäres Zentrum für Wissenschaftliches Rechnen der Ruprecht-Karls-Universität Heidelberg).

Wenn Ihr zum Himmel hinaufschaut und seltsame Muster von Kondensationsspuren seht (siehe Abbildung 62), habt Ihr Euch dann schon einmal gefragt, wie viele Flüge es so jeden Tag gibt?

Sehen wir uns die Zahlen an, die im Jahresbericht³⁷ von EUROCONTROL, der Organisation für die zentrale Koordination der Flugsicherheit in Europa, veröffentlicht werden: Es ist kaum zu glauben, aber im Jahr 2019, also noch vor der COVID-19-Pandemie, hat die Zahl der Flüge in Europa 11,1 Millionen erreicht. Allein in Deutschland starteten, landeten 2019 an jedem Tag mehr als 9.000 Flugzeuge in Deutschland oder flogen darüber. Im Jahr darauf waren es dann wegen der Pandemie in Europa zwar weniger als 5 Millionen Flüge, aber Experten gehen davon aus, dass im Falle einer erfolgreichen Impfkampagne schon 2024 der Flugverkehr das Niveau von 2019 wieder erreichen könnte. Langfristig wird sogar mit weiterem starkem Wachstum gerechnet. Selbst in Pandemiezeiten kann man nicht abstreiten, dass Flugreisen zu einer Selbstverständlichkeit in unserer Gesellschaft geworden sind und die Menschen mobiler denn je machen. Der wachsende Flugverkehr bringt allerdings viele Herausforderungen mit sich, nicht zuletzt auch für das Flugverkehrsmanagement.

Das oberste Ziel des Flugverkehrsmanagements ist die Flugsicherheit. Das bedeutet, dass die Fluglotsen bei jedem Flug dafür sorgen müssen, dass alle Flugzeuge jederzeit ausreichend Abstand zueinander halten. Das ist aber nicht die einzige Schwierigkeit. Das Flugverkehrsmanagement muss sich auch mit den Verspätungen im System auseinandersetzen. Laut EUROCONTROL Jahresbericht 2019 waren bereits ein Zehntel aller Flüge in Europa verspätet. Das klingt vielleicht wenig, bewirkt aber erhebliche Störungen der gesamten Flugpläne, erschwert die Flugsicherheit und führt zu Staus in der Luft und auf den Flughäfen. Verspätungen sind nicht nur unangenehm für die Passagiere, sie führen auch zu längeren Flugstrecken, einer Verringerung der operativen Effizienz, einem Anstieg des Treibstoffverbrauchs und damit zu einem Anstieg der CO₂-Emissionen.

Und damit sind wir bei den Umweltproblemen, die mit dem Flugverkehr und den ehrgeizigen Klimazielen der EU verbunden sind. Ihr wisst vielleicht, dass die Europäische Union zur Bekämpfung des Klimawan-

dels und der durch CO₂-Emissionen verursachten globalen Erwärmung strenge Nachhaltigkeitsziele für eine klimaneutrale und emissionsfreie Mobilität bis 2050 vereinbart hat. Laut „European Aviation Environmental Report 2019“⁴⁸ ist der Luftverkehr derzeit für 3 % der weltweiten Treibhausgasemissionen verantwortlich, aber langfristige Prognosen sehen den relativen Anteil des Luftverkehrs weiter steigen, da er auch mittelfristig auf fossile Brennstoffen angewiesen ist.

Ein Beispiel: ein typisches Flugzeug mit zwei Düsentriebwerken verbraucht während eines einstündigen Fluges mit 150 Passagieren 2.700 kg Kerosin, was zu einer Emission von etwa 8,5 Tonnen Kohlendioxid führt. Für Flüge in Europa summiert sich das dann z.B. 2016 auf insgesamt 171 Millionen Tonnen CO₂, fast doppelt so viel wie 1990.

Zu den Strategien zur Emissionsreduzierung gehören die Erneuerung der bestehenden Flugzeugflotte durch leichtere und treibstoffeffizientere Flugzeuge, die Entwicklung effizienterer Düsentriebwerke, kohlenstoffneutrale Biokraftstoffe, Wasserstoff- oder Elektroantriebe oder die Reduzierung des Luftwiderstands. Aber auch effiziente Flugverkehrsmanagementsysteme und die Planung treibstoffeffizienterer Flugrouten können eine wichtige Rolle bei der Emissionsreduzierung spielen, da das Flugverkehrsmanagement etwa 6 % des gesamten „Gate-to-Gate-Treibstoffverbrauchs“ beeinflussen kann.

Besonders während der verbrauchsintensiven Start- und Landephase können Fluglotsen und Flugplaner einen wichtigen Beitrag zur Reduzierung der Belastung leisten. Um all diese Herausforderungen zu bewältigen, benötigen Fluglotsen hochwertige Entscheidungsunterstützungswerkzeuge, um den anstehenden Verkehrsfluss effizient zu bewältigen. Die Entwicklung eines solchen Werkzeugs ist ohne moderne mathematische Methoden der optimalen Steuerung kaum möglich.

Historischer Exkurs

Moderne Methoden der optimalen Steuerung haben ihre Wurzeln in der Theorie der Variationsrechnung, die mit dem sogenannten „Brachistochron“-Problem (griechisch: brachistos=kürzeste, chronos=Zeit) begann. Das Problem wurde 1696 von Johann Bernoulli formuliert: Wie lautet die Kurve der kürzesten Fallzeit eines Massenpunktes, der unter dem Einfluss der Schwerkraft reibungsfrei von einem Punkt A zu einem Punkt B rollt? Die Lösung des Problems (das übrigens weder die direkte Verbindungsgerade der Punkte A und B, noch ein Kreisbogen ist) durch Johann Bernoulli und seinen

Bruder Jacob, und später durch so eminente Wissenschaftler wie Gottfried Wilhelm Leibniz und Isaac Newton, Leonhard Euler und Joseph-Louis Lagrange zu den Arbeiten von Adrien-Marie Legendre und Karl Weierstraß und schließlich zur modernen Optimierungstheorie. Die Notwendigkeit, reale angewandte Probleme, z.B. der Raumfahrt, zu lösen, führte in den frühen 1950er Jahren zur Entstehung des sogenannten Maximumprinzips von Pontryagin in der Theorie der optimalen Steuerung. Das Maximumprinzip beschreibt die Bedingungen, die von der optimalen Lösung erfüllt werden müssen und hat sich in der Anwendung als äußerst wertvoll erwiesen.

Bereits in den 1980er Jahren wurden diese Methoden zur Optimierung nichtlinearer Prozesse eingesetzt, teilweise mit beeindruckenden Ergebnissen. Allerdings muss die optimale Steuerung „direkt“ bestimmt werden, aus einem sehr schwierig aufzustellenden und zu lösenden Randwertproblem mit adjungierten Gleichungen. Mit zunehmender Komplexität der praktischen Anwendungen erwies sich dieser Ansatz als nicht mehr geeignet, die Probleme zu lösen. Daher wurde er durch Methoden ersetzt, bei denen optimale Steuerungen „direkt“ als Entscheidungsvariablen berechnet werden. Moderne Verfahren basieren auf der „indirekten Mehrzielmethode“, einer Methode, die in den 1980er Jahren von dem Heidelberger Mathematiker Hans Georg Bock eingeführt wurde. Am Interdisziplinären Zentrum für Wissenschaftliches Rechnen der Universität Heidelberg werden von Ekaterina Kostina und anderen Arbeitsgruppen hocheffiziente Verfahren dieser Art für viele verschiedene Problemklassen und Komplexitäten entwickelt und verbessert.

Wie können wir die Methoden der optimalen Steuerung auf Probleme des Flugverkehrsmanagements anwenden?

Wir diskutieren ein Optimalsteuerungsproblem zur Bestimmung verbrauchsoptimaler und kollisionsfreier Flugbahnen für mehrere Flugzeuge im Reiseflug und im Landeanflug. Wir untersuchen also einen Luftraumsektor S in der Luft während eines bestimmten Zeitraums und zielen darauf ab, den optimalen Weg zu finden, alle Flugzeuge durch S zu navigieren unter Einhaltung von Sicherheitsabständen. Zunächst beginnen wir mit der mathematischen Modellierung der Bewegungen aller Flugzeuge in dem gegebenen Sektor. In unserer Anwendung besteht das Modell aus einem nichtlinearen System von Differentialgleichungen der Flugdynamik, das beschreibt, wie sich die „Zustände“, d.h. Flugrichtung, Geschwindigkeit und die räumliche Position in der Zeit entwickeln.

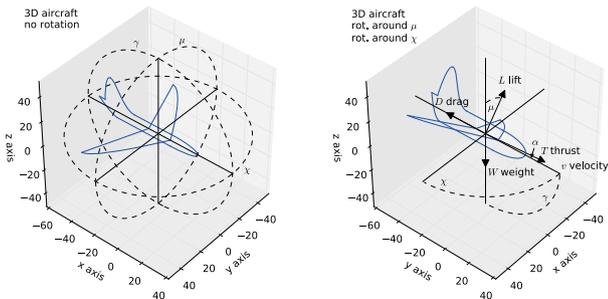


Abbildung 63: Darstellung der Zustände und auf das Flugzeug wirkenden Kräfte.

Die dreidimensionale Flugbahn lässt sich durch sieben Zustände und drei Steuerungen beschreiben und repräsentiert die Translationsdynamik des Flugzeugs. Die sieben Zustände in Abbildung 63 sind die (x, y) -Position, die Höhe z , die Geschwindigkeit v , der Kurswinkel χ , der Steigwinkel γ und die Flugzeugmasse m . Die drei Steuerungen sind der Anstellwinkel α , der Querneigungswinkel μ und die Schubhebelstellung δ . Der Anstellwinkel α ist der Winkel zwischen den Tragflächen und der ankommenden Luft. Der Anstellwinkel wird durch die Klappe an den Flügeln gesteuert und hat die Aufgabe, die aerodynamischen Kräfte und Momente um den Schwerpunkt des Flugzeugs auszugleichen. Der Querneigungswinkel μ wird durch die Querruder bewirkt und hat die Aufgabe, die Richtung des Auftriebsvektors in Bezug auf den Windgeschwindigkeitsvektor bei einem gegebenen Anstellwinkel zu ändern. In unserem Modell wird der Schub durch die Schubhebelposition δ gesteuert. Die Schubhebelposition hat einen Wert zwischen dem Null- und Maximalschub, der von den Triebwerken erzeugt wird.

Als Ergebnis erhalten wir die folgenden Differentialgleichungen

$$\begin{aligned}\dot{x}(t) &= v(t) \cos \chi(t) \cos \gamma(t), \\ \dot{y}(t) &= v(t) \sin \chi(t) \cos \gamma(t), \\ \dot{z}(t) &= v(t) \sin \gamma(t), \\ \dot{v}(t) &= \frac{1}{m(t)} \left(T_{\max} \delta(t) - D(v(t), \alpha(t)) \right) - g \sin \gamma(t), \\ \dot{\chi}(t) &= \frac{1}{m(t)v(t)\gamma(t)} L(v(t), \alpha(t)) \sin \mu(t), \\ \dot{\gamma}(t) &= \frac{1}{m(t)v(t)} L(v(t), \alpha(t)) \cos \mu(t) - g \sin \gamma(t), \\ \dot{m}(t) &= -f_r T_{\max} \delta(t).\end{aligned}$$

Die Konstanten T_{\max} - der maximale Schub, und f_r - die Treibstoffverbrauchsrate, hängen vom Flugzeugmodell ab. Für die aerodynamischen Auftriebs- und Widerstandskräfte gibt es spezielle Modelle (Gleichungen), die vom Flugzeugmodell, den atmosphärischen Bedingungen, der effektiven Flügelfläche und der Luftdichte abhängen.

Die Trajektorien der einzelnen Flugzeuge müssen mehrere Nebenbedingungen erfüllen, wie z.B. Anfangs- und Endbedingungen. Weitere Ungleichungsbedingungen definieren die Leistungsgrenzen des Flugzeugs, wie z.B. die maximale Geschwindigkeit, sorgen aber auch für einen komfortablen Flug. Für die Modellierung des Optimierungsproblems sind viele weitere Faktoren wichtig. Die notwendige Gesamtzeit sollte eingehalten werden. Außerdem wollen wir, dass das Flugzeug möglichst unnötige Kurven vermeidet. Ebenso möchten wir, dass die Flughöhe innerhalb vorgeschriebener Grenzen bleibt. Darüber hinaus wollen wir den Treibstoffverbrauch so gering wie möglich halten.

Das optimale Steuerungsproblem für jedes Flugzeug kann also wie folgt formuliert werden: Minimierung der Zielfunktion unter Berücksichtigung der Gleichungen der Flugzeugdynamik sowie aller Nebenbedingungen.

Um Konflikte zwischen mehreren Flugzeugen im Sektor S zu lösen, führen wir die Differentialgleichungen, Zielfunktionen und Nebenbedingungen aller Flugzeuge zu einem gesamten optimalen Steuerungsproblem zusammen. Für das Gesamtproblem führen wir die paarweisen räumlichen Restriktionen ein, die sicherstellen, dass der Sicherheitsabstand zwischen allen Flugzeugen zu jeder Zeit eingehalten wird. Diese paarweisen Einschränkungen kann man sich wie Sicherheitsellipsoide um jedes Flugzeug vorstellen. Auch das Einfädeln mehrerer Flugzeuge beim Landeanflug führt auf weitere räumliche und zeitliche Abstandsbeschränkungen.

Das Gesamtproblem ist aufgrund seiner hohen Dimension und der einzuhaltenden räumlichen und zeitlichen Beschränkungen sehr schwierig zu lösen, kann aber mit an der Universität Heidelberg entwickelten Methoden der optimalen Steuerung, die die speziellen Strukturen der zugrundeliegenden Probleme ausnutzen, sehr effizient gelöst werden.^{39 40 41}

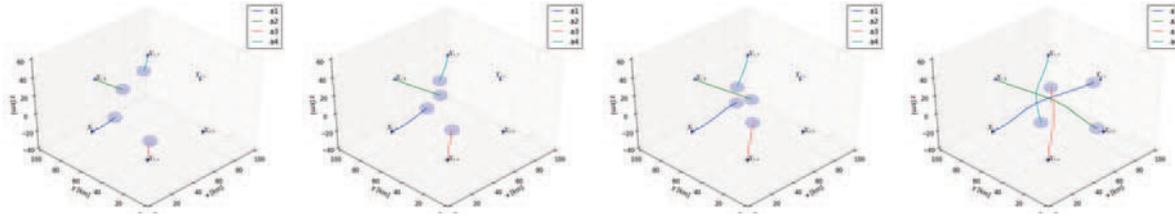


Abbildung 64: Kollisionsfreie und CO₂-effiziente Flugbahnen in einem Sektor S.

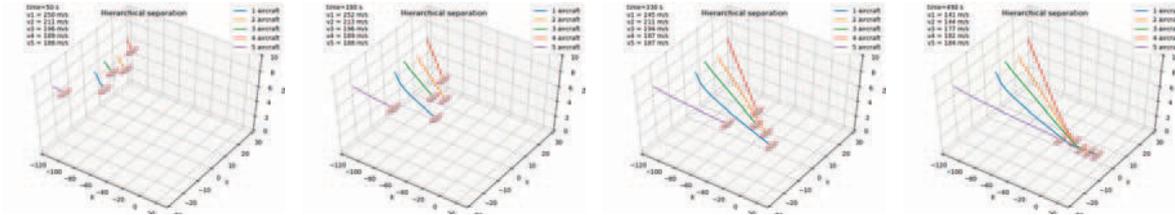


Abbildung 65: Kollisionsfreie und CO₂-effiziente Flugbahnen beim Landeanflug

Die mit dem Computer erhaltenen Lösungen haben ein sehr hohes Potential für die Berechnung von kollisionsfreien und CO₂-effizienten Flugbahnen, siehe Abbildungen 64 und 65

Optimale Lösungen reduzieren Verspätungen im System und deren Dominoeffekte. Flugumwege werden reduziert, was eine bessere Ausnutzung des Luftraums und weitere Treibstoffeinsparungen ermöglicht.

Aktuelle und zukünftige Arbeiten

Für den praktischen Einsatz ist das oben beschriebene Modell leider noch zu einfach und muss erheblich erweitert werden. Unterschiedliche Flugzeugtypen haben spezifische Verbrauchswerte und unterscheiden sich oft deutlich in ihrem aerodynamischen Flug- und Steuerungsverhalten. Außerdem müssen äußere Einflüsse, die auf das Flugverhalten wirken wie Wind- und Wetterverhältnisse modelliert und berücksichtigt werden. Solche externen Einflüsse sind Störungen und mögliche Fehlerquellen, die nicht oder nicht genau vorhersehbar sind. Wind- und Wetterlage können sich lokal schnell ändern und die Flugdynamik stark beeinflussen (Turbulenzen, Seitenwinde, Scherwinde). Um diese Faktoren zu berücksichtigen, muss die aktuelle Situation sehr schnell und zuverlässig bewertet und quantifiziert und die optimale Lösung mit Echtzeitmethoden an die

Störung angepasst werden („Feedback“). Bei der Entwicklung solcher Methoden haben wir in den letzten Jahren bereits enorme Fortschritte gemacht.

Unser Forschungsziel ist es, in Kooperation mit den Einrichtungen der Flugsicherung und den Fluggesellschaften, leistungsfähige Entscheidungsunterstützungssysteme zu entwickeln, die die vielfältigen Herausforderungen an den Flugverkehr der Zukunft auf mathematisch optimale Weise bewältigen helfen. Dafür brauchen wir hochmotivierten wissenschaftlichen Nachwuchs, der sich für die mathematische Optimierungstheorie ebenso wie für die Modellierung der physikalisch-technischen Prozesse spannender Anwendungen begeistert!

³⁷ <https://www.eurocontrol.int/publication/performance-review-report-prr-2019>

³⁸ <https://www.eurocontrol.int/publication/european-aviation-environmental-report-2019>

³⁹ M. Schubert, Optimal Control of 3D Separation Management, Master thesis, Ruprecht-Karls-Universität Heidelberg, 2017

⁴⁰ V. Semenov, Multi-Stage Strategy to Numerical Trajectory Optimization for Multiple Aircraft Approaching an Airport, Master thesis, Ruprecht-Karls-Universität Heidelberg, 2019.

⁴¹ A. Raptakis, Optimization of Flight Trajectories of an Aircraft with Respect to Fuel Consumption, Master thesis, Ruprecht-Karls-Universität Heidelberg, 2020.



Weniger Flugzeugabstürze aufgrund besserer Wahrscheinlichkeitsmodelle

Die meisten Flugunfälle werden durch eine Kombination von technischen, menschlichen und Umweltfaktoren verursacht. Bessere mathematische Wahrscheinlichkeitsmodelle verringern die Wahrscheinlichkeit eines Unfalls.

In den letzten fünfzig Jahren ist das Risiko eines tödlichen Flugunfalls in der Zivilluftfahrt spektakulär gesunken: von rund 30 tödlichen Unfällen bei einer Million Flügen im Jahr 1959 auf weniger als 0,5 bei einer Million Flügen im Jahr 2012. Weil jeder Unfall natürlich immer noch einer zu viel ist, suchen Flugzeughersteller und Luftfahrtbehörden ständig nach Möglichkeiten, die Flugsicherheit weiter zu verbessern.

Der Sicherheitsforscher Alfred Roelen vom National Aerospace Laboratory (NLR) in Amsterdam verwendet die Mathematik als Hilfsmittel bei der Modellierung der Wahrscheinlichkeit eines Flugzeugabsturzes. „Wir suchen nach Faktoren, die die größte Rolle bei der Entstehung

von Unfällen spielen, und reduzieren sie dann“, sagt Roelen. „Einerseits tun wir dies, indem wir große Mengen von Flugdaten sammeln und statistisch analysieren. Und auf der anderen Seite, indem man bessere Modelle der Wahrscheinlichkeit erstellt, dass während eines Fluges etwas schief geht.“

Es kann ein technisches Problem geben, zum Beispiel mit einem Flügel, einem Motor oder dem Autopiloten. Aber auch auf menschlicher Ebene kann etwas schief gehen, zum Beispiel beim Piloten oder beim Fluglotsen. Darüber hinaus können wetterbedingte Probleme auftreten. Fast immer ist ein Flugunfall auf eine Kettenreaktion von Faktoren zurückzuführen.

Eis auf dem Flügel

Roelen gibt ein Beispiel für eine Kettenreaktion, die er mathematisch modellieren kann: „Angenommen, es schneit auf einem Flughafen. Nach dem üblichen Verfahren werden Schnee und Eis vor dem Abflug aus einem Flugzeug entfernt. Es besteht jedoch immer die Möglichkeit, dass dies aus welchen Gründen auch immer nicht ordnungsgemäß geschieht. Wie groß ist die Chance, dass ein Flugzeug mit Schnee oder Eis auf einem seiner Flügel abhebt? Wie groß ist dann die Chance, dass sich – einmal in der Luft – die Strömung um diesen Flügel löst und das Flugzeug abrupt absinkt? Und schließlich, wie groß ist die Chance, dass der Pilot dies rechtzeitig bemerkt und angemessene Maßnahmen ergreift?“

Die europäischen und amerikanischen Luftfahrtbehörden stellen strenge Anforderungen an die Ausfallwahrscheinlichkeit eines Flugzeugs. Unter anderem müssen die Flugzeughersteller durch mathematische Analysen nachweisen, dass die Wahrscheinlichkeit eines katastrophalen Unfalls weniger als 1 in einer Milliarde Flugstunden beträgt (mit anderen Worten: etwa 1 Unfall in mehr als hunderttausend Jahren). Roelen: „Diese Wahrscheinlichkeit ist so gering, dass man das Flugzeug nicht in der Praxis testen kann. Wir müssen uns also etwas anderes einfallen lassen. Wir tun dies, indem wir das Flugzeug in einzelne Komponenten zerlegen und die Ausfallwahrscheinlichkeit für jede Komponente messen oder abschätzen.“

Die Ausfallhäufigkeiten für die technischen Komponenten sind am einfachsten zu bestimmen, in der Regel durch Versuche im Labor. Die Bestimmung von Ausfallhäufigkeiten für menschliches Verhalten ist viel schwieriger. Wenn ein Fluglotse während eines Landeverfahrens zum Piloten „tausend Fuß tief“ sagt und er meinte „zweitausend Fuß tief“, kann das tödlich sein. Mit psychologischen Tests wird ermittelt, wie oft Menschen in Abhängigkeit von der Hektik in ihren Köpfen Fehler machen.

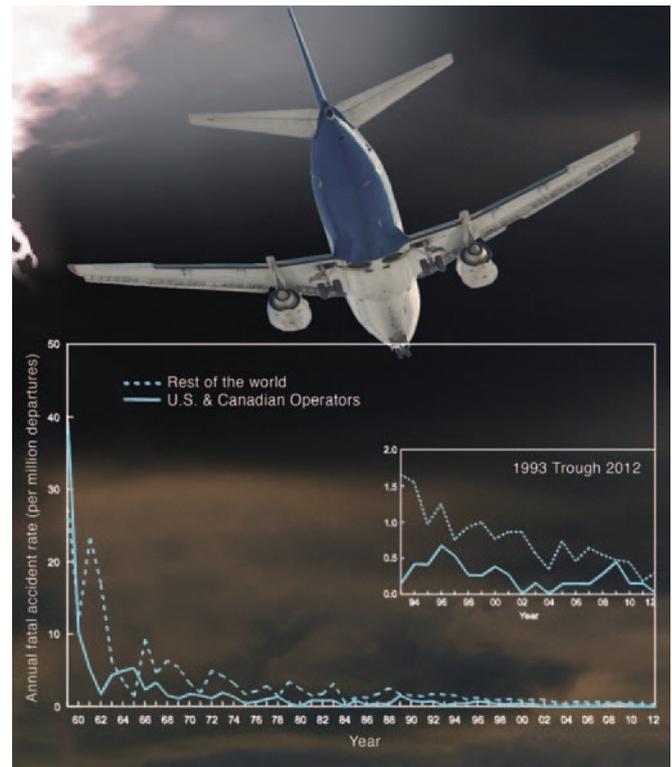
Roelen: „Das einfachste Modell bestimmt die Wahrscheinlichkeit für menschliches Versagen in drei Situationen: eine Person ist sehr beschäftigt, mittelmäßig beschäftigt oder ruhig. Es ist logisch, dass ein Mensch die meisten Fehler macht, wenn er sehr beschäftigt ist. Aber wenn er sehr ruhig ist, ist seine Aufmerksamkeit nicht so konzentriert und er macht auch mehr Fehler als wenn er im Durchschnitt beschäftigt ist.“

Killer

Vor zwanzig Jahren war einer der größten „Killer“ in der Zivilluftfahrt eine Art Unfall, der als kontrollierter Flug ins Gelände bezeichnet wird: Ein Flugzeug ist normalerweise auf dem Weg zu einem Flughafen. Mit dem Flugzeug ist alles in Ordnung, aber unerwartet fliegt es unweit des Flughafens in einen Berg. Dank einer detaillierten statistischen Analyse dieser Art von Unfällen konnten Roelen und seine Kollegen zeigen, was die Hauptursache war.

Roelen: „Auf Flughäfen, die keine Funksignale zur Führung eines Flugzeugs bei der Landung verwendeten, war die Wahrscheinlichkeit eines solchen Unfalls fünfmal höher. Es ist zum Teil unserer Arbeit zu verdanken, dass mehr Flughäfen die notwendige Funkausrüstung angeschafft haben. Heutzutage werden diese Geräte zunehmend durch GPS-Geräte ersetzt. Auch hier werden mathematische Modelle verwendet, um die Wahrscheinlichkeit eines Misserfolgs abzuschätzen.“

Die Zahl der tödlichen Unfälle pro Million Flüge zwischen 1959 und 2012 (Quelle: Boeing)



Digitale Schutzmechanismen knacken

um sie sicherer zu machen

Heutzutage geht es bei der Kryptographie nicht mehr um Geheimnisse wie noch vor einem Jahrhundert, sondern hauptsächlich um Mathematik. Und dass die moderne Kryptographie immer mehr alltägliche Anwendungen absichert.

Im Jahr 2008 knackten Forscher der Radboud Universität Nijmegen die OV-Chipkarte. Sie zeigten, wie sie damit kostenlos reisen konnten. Aber das Problem war viel größer als die OV-Chipkarte allein. Der Chip in dieser niederländischen Reisekarte für den ÖPNV, der Mifare Classic, wurde in mehr als einem Jahrzehnt weltweit zu mehr als einer Milliarde Karten verarbeitet. Und dazu gehören auch Zugangsausweise zu Regierungsgebäuden und militärischen Einrichtungen.

Es ist eine gute Praxis für Wissenschaftler, den Hersteller nach der Entdeckung einer Sicherheitsverletzung zu informieren und ihm Zeit für die Reparatur des Produkts zu geben: sechs Monate für die Anpassung der Hardware und sechs Wochen für die Software. Die

Radboud-Wissenschaftler warnten 2008 die niederländische Regierung, den internen Sicherheitsdienst und den Hersteller der OV-Chipkarte, die niederländische Firma NXP. Die Panik war groß. NXP versuchte, eine wissenschaftliche Veröffentlichung der Forscher über den Sicherheitsbruch zu verhindern, aber der Richter entschied, dass die Wahrheit von öffentlichem Interesse sei und erlaubte die Veröffentlichung.

Geheimer Schlüssel

Roel Verdult, einer der damaligen Hacker und heute Doktorand an der Radboud Universität, sieht es als soziale Aufgabe von Forschern der digitalen Sicherheit, die Sicherheit kritisch zu untersuchen. „Es ist

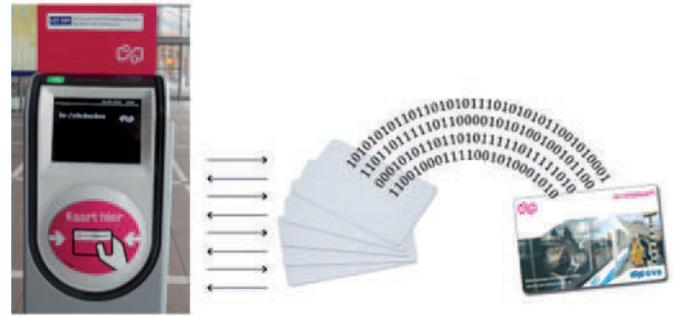
sehr schwierig, den theoretischen Beweis zu erbringen, dass die digitale Sicherheit so und so stark ist“, sagt Verdult. „Deshalb wird in der Praxis ein pragmatischer Ansatz verfolgt. Das bedeutet, dass es Teil des Aufbaus eines guten Kryptosystems ist, auch zu versuchen, es zu knacken. Auf diese Weise können die Wissenschaftler aus allen Arten der vorgeschlagenen Schutzmaßnahmen die besten auswählen.“

Sowohl das Erstellen als auch das Brechen kryptographischer Schutzvorrichtungen basiert auf Mathematik. Im Mittelpunkt steht ein kryptographischer Algorithmus: ein Berechnungsrezept, das geheime digitale Schlüssel erzeugt. Eine häufig verwendete Methode ist die Durchführung einer mathematischen Operation mit drei Zahlen: zuerst eine bekannte Zahl, zum Beispiel die Identifikationsnummer einer Karte, dann ein geheimer Wert, der als Schlüssel dient, und schließlich eine Zufallszahl, die an Ort und Stelle generiert wird. Je schwieriger der Schlüssel und je zufälliger die Zahl, desto schwieriger ist er zu knacken und desto besser die Sicherheit.

Im Beispiel der OV-Chipkarte funktioniert es wie folgt. Sobald sich eine Chipkarte einem Lesegerät nähert, sendet sie eine eindeutige Identifikationsnummer an das Lesegerät. Mit dieser Zahl erzeugt der Leser eine Reihe von kryptographischen Schlüsseln. Die Karte und das Lesegerät prüfen mit Hilfe elektromagnetischer Signale schnell, ob sie den geheimen Schlüssel kennen. Wenn sie dies tun, wird der Reisende ein- oder ausgecheckt.

„Der Algorithmus, der bei der Mifare Classic und damit auch bei der OV-Chipkarte verwendet wurde, war bereits in der Entwurfsphase unsicher“, erklärt Verdult. „Aber da der Algorithmus geheim gehalten wurde, dauerte es einige Zeit, bis wir die Konstruktionsfehler entdecken konnten. Unmittelbar danach warnten wir alle vor den bestehenden Schwächen. Obwohl wir die Ersten waren, die offen über die Probleme sprachen, ist es nicht unwahrscheinlich, dass diese Schwächen bereits insgeheim von anderen ausgenutzt wurden.“

Eigentlich sollte der Ausgangspunkt sein, dass die Sicherheit eines Kryptosystems nur vom Schlüssel und nicht vom Algorithmus profitieren sollte, meint Verdult. „Die Zufalls- und Geheimzahlen werden immer wieder neu berechnet, während der Algorithmus derselbe bleibt. Ein guter Algorithmus wird nicht unsicher, wenn er öffentlich ist. Tatsächlich kann dann jeder überprüfen, ob es sich um einen guten



Algorithmus handelt. Und jeder kann Verbesserungen vorschlagen, um es noch stärker zu machen.“

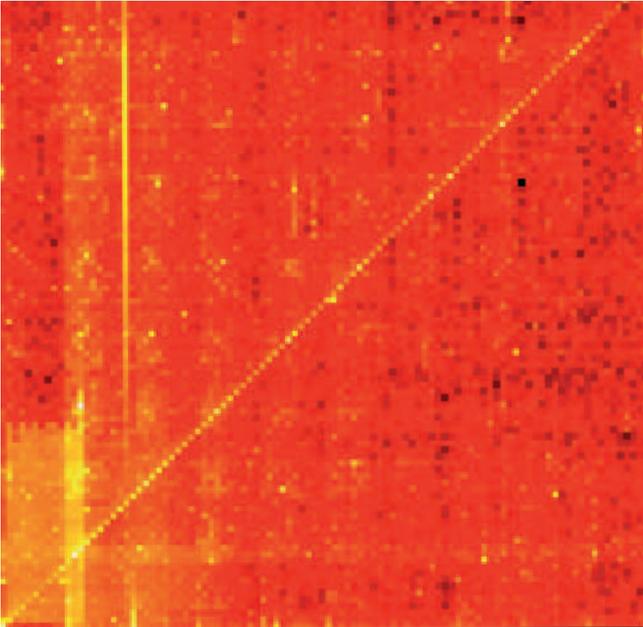
Neuer Pass

Die Empfehlungen der Forscher für ein stark verbessertes Kryptosystem wurden von der niederländischen Regierung sofort beherzigt. Die Regierung ist nun dabei, einen sicheren nationalen Pass für den Zugang zu Ministerien und anderen wichtigen Gebäuden einzuführen. Ein gutes Beispiel für die Nützlichkeit ihrer Arbeit, so Verdult. Aber bei der OV-Chipkarte sieht die Sache ganz anders aus. „Eigentlich gibt es dort nur wenige Stellen zur Verbesserung der Sicherheit. Aber im Kern gibt es immer noch den gleichen unsicheren Mifare Classic-Algorithmus. Wirklich seltsam, denn selbst der Hersteller des Mifare Classic rät vom Kauf des Chips ab.“

Nach dem Hacken der OV-Chipkarte im Jahr 2008 hörte man einige Leute als Entschuldigung sagen, dass jedes System gehackt werden kann. Aber Verdult betont, dass Sie diese Aussage von einem Kryptoforscher niemals hören werden: „Es gibt viele Algorithmen, die sicher sind. Ein bekanntes Beispiel ist der *Advanced Encryption Standard* (AES), der 1998 eingeführt wurde und nach mehr als fünfzehn Jahren immer noch nicht zu knacken ist.“ Es ist die Aufgabe der Kryptoforscher, herauszufinden, welche Algorithmen sicher sind und in welchen Anwendungen sie am besten eingesetzt werden können.

An der Bergischen Universität Wuppertal gibt es Lernstationen zur Kryptographie, die SpionCamps, siehe <https://ddi.uni-wuppertal.de/www-madin/material/spion-camp.html>

Muster in PIN-Codes



**Entdecken Sie die Muster,
und wählen Sie von nun an
einen besseren Code!**

Was tun Sie, wenn Sie den PIN-Code für Ihre Bankkarte wählen dürfen? Oder wenn Sie einen vierstelligen Code für das Schließfach in einem Hotelzimmer festlegen können? Sie sind wahrscheinlich berechenbarer als Sie denken.

Der Datenanalytiker Nick Berry sammelte Millionen von Daten aus durchgesickerten Passwörtern und PIN-Code-Datenbanken. Letztendlich hatte er eine Sammlung von Millionen von Benutzern, die einen vierstelligen Code wählten. Insgesamt traten mindestens einmal zehntausend mögliche Kombinationen auf, aber einige waren viel häufiger als andere: zum Beispiel wählte eine von zehn Personen 1234 als supergeheimen Code. Ein Dieb betritt einen von fünf Safes, indem er nur die fünf Codes 1234, 1111, 0000, 1212 und 7777 ausprobieren.

Berry machte dieses faszinierende Bild, das zeigt, wie oft die Kombinationen auftreten. Auf der horizontalen Achse befinden sich die ersten beiden Ziffern des PIN-Codes von 00 bis 99. Auf der vertikalen Achse befinden sich die letzten beiden Ziffern des PIN-Codes, wiederum von 00 bis 99. Jedes Quadrat repräsentiert also genau einen PIN-Code. Je öfter der PIN-Code gewählt wird, desto leichter ist er. 1234 ist also weiß und die am wenigsten gebräuchliche PIN 8068 ist schwarz. Dieses Bild zeigt allerlei Interessantes darüber, wie Menschen ihre PIN-Codes wählen.

Das Benfordsche Gesetz – Wie Mathematik Zahlenfälscher entlarvt



Ein einfaches
mathematisches Gesetz
hilft Fälschungen in
verschiedenen Bereichen
aufzudecken.

Die Geschichte von Benfords Gesetz beginnt mit einem Artikel von Frank Albert Benford⁴² aus dem Jahre 1938 und zeigt wie auch theoretische Ergebnisse konkrete praktische Anwendungen, wie etwa Aufdeckung von Steuerbetrug, haben können.

Frank Benford interessierte sich für Muster in Zahlendaten und beobachtete eine Ziffernverzerrung bei der Verteilung der ersten oder der führenden Ziffern und wollte dies mathematisch erklären. Gerade in einem Zeitalter, in dem wir ständig mit riesigen Datenmengen bombardiert werden, ist diese Entdeckung sehr wichtig. Man nennt diese Gesetzmäßigkeit auch „Newcomb-Benford Gesetz“ (NBG), da es genau genommen bereits 1881 vom Mathematiker Simon Newcomb entdeckt wurde⁴³ nachdem er bemerkt hatte, dass die in Logarithmentabellen diejenigen Seiten mit der Eins als erster Ziffer

schmutziger als andere Seiten waren, also häufiger benutzt worden waren. Und so schrieb Benford in seinem Artikel:

There may be, in the relative cleanliness of the pages of a logarithm table, data on how we think and how we react when dealing with things that can be described by means of numbers.

Frank Benford, 1938

Die Verteilung von Ziffern in Zahlen

Bevor wir das Gesetz beschreiben können, müssen wir zunächst die Notation festlegen. Jede positive Zahl x kann im Dezimalsystem als $M(x) \cdot 10^E$ geschrieben werden, wobei $M(x) \in [1, 10)$ die sogenannte Mantisse und E der Exponent ist. Der ganzzahlige Anteil der Mantisse (die Zahl vor dem Komma) wird als führende Stelle oder erste Stelle bezeichnet. Die Zahl 14568,79 wird in wissenschaftlicher Notation als $1.456879 \cdot 10^4$ geschrieben. Die Mantisse ist 1.456879, der Exponent ist 4, die führende Ziffer ist 1.

Die naheliegendste Vermutung wäre zu behaupten, dass bei einem allgemeinen Datensatz alle Zahlen mit gleicher Wahrscheinlichkeit die führende Ziffer sind. Wir würden dann also annehmen, dass wir in etwa 11 % der Fälle eine führende Ziffer von 1,2,...,9 beobachten sollten. Hierbei sind wir davon ausgegangen, dass jede Zahl $1/9$ der Zeit und nicht $1/10$ der Zeit vorkommt, da 0 nur die führende Ziffer für die Zahl 0 ist.). Der Inhalt des Benfordschen Gesetzes ist, dass dies häufig nicht der Fall ist; insbesondere erwarten wir in vielen Situationen, dass die *führende Ziffer* d mit einer Wahrscheinlichkeit von etwa $P(d) = \log_{10} \left(\frac{d+1}{d} \right)$ bzw. $\log_{10}(d+1) - \log_{10}(d)$ ist, was bedeutet, dass die Wahrscheinlichkeit einer ersten Ziffer von 1 etwa $\log_{10} 2 \approx 30,1\%$ beträgt, während eine erste Ziffer von 9 etwa $\log_{10} \left(\frac{10}{9} \right) \approx 4,58\%$ der Zeit vorkommt.

Eine Menge von Zahlen erfüllt das Benfordsche Gesetz für die führende Ziffer, wenn die Wahrscheinlichkeit eine erste Ziffer von d zu beobachten ungefähr gleich $P(d) = \log_{10} \left(\frac{d+1}{d} \right)$ ist. Schließlich könnten wir, anstatt nur die führende Ziffer zu untersuchen, die gesamte Mantisse untersuchen. Anstatt also nach der Wahrscheinlichkeit für eine erste Ziffer von 1 oder 2 oder 3 zu fragen, fragen wir nun nach der Wahrscheinlichkeit für die Beobachtung einer Mantisse z.B. zwischen 1 und 3. Diese Verallgemeinerung wird häufig als das Starke Benfordsche Gesetz bezeichnet: Wir sagen, dass ein Daten-

satz das Starke Benfordsche Gesetz erfüllt, wenn die Wahrscheinlichkeit von einer Mantisse in $[1, m)$ zu beobachten, $\log_{10} m$ ist. Beachten Sie, dass das starke Benfordsche Gesetz das Benfordsche Gesetz impliziert: die Wahrscheinlichkeit einer ersten Stelle von d ist einfach die Wahrscheinlichkeit, dass die Mantisse in $[d, d+1)$ liegt, also $\log_{10}(d+1) - \log_{10}(d)$.

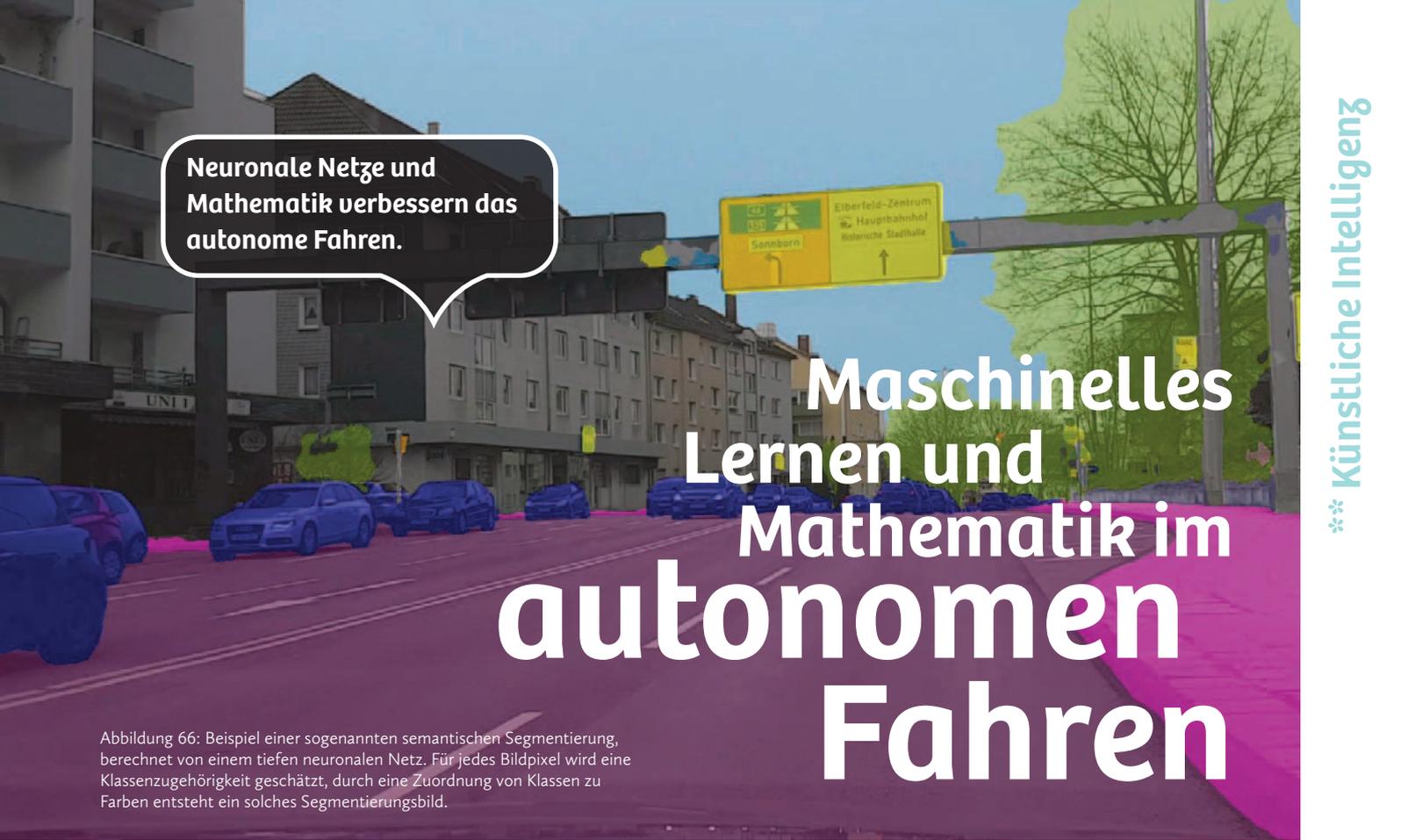
Aufdeckung von Betrug

Das Benfordsche Gesetz kann man vielfältig einsetzen, etwa von Wirtschaftsprüfern bei der Aufdeckung von Steuerbetrug, Fälschung von Bilanzen oder generell bei dem Erkennen von Unregelmäßigkeiten im Rechnungswesen. Ebenso kann es in der Wissenschaft dazu dienen Datenfälschungen aufzudecken. Bekannt geworden sind auch Untersuchungen von signifikanten Unregelmäßigkeiten bei Wahlen, z.B. bei den Präsidentschaftswahlen 2009 im Iran.



⁴² Frank Benford, The law of anomalous numbers, Proceedings of the American Philosophical Society, 78, No. 4 (1938), 551–572.

⁴³ Simon Newcomb, Note on the frequency of use of the different digits in natural numbers, American Journal of Mathematics, 4 (1881), 39–40.



Neuronale Netze und
Mathematik verbessern das
autonome Fahren.

Maschinelles Lernen und Mathematik im autonomen Fahren

Abbildung 66: Beispiel einer sogenannten semantischen Segmentierung, berechnet von einem tiefen neuronalen Netz. Für jedes Bildpixel wird eine Klassenzugehörigkeit geschätzt, durch eine Zuordnung von Klassen zu Farben entsteht ein solches Segmentierungsbild.

Maschinelles Lernen, insbesondere die Verfahren des Deep Learning basierend auf sogenannten tiefen neuronalen Netzen, werden im autonomen Fahren und seinen Vorstufen hauptsächlich in drei Anwendungsfeldern verwendet:

- 1 Bei der Umfeldwahrnehmung mittels der Verarbeitung von Kamera und Sensordaten
- 2 Zum Erlernen und Umsetzen einer Fahrstrategie
- 3 Zur Überwachung des Innenraums

In diesen Anwendungsfeldern haben die neuen Deep-Learning-Algorithmen, entwickelt in den 2010er Jahren, Vorhersagegenauigkeiten entwickelt die mit vorherigen Methoden unerreichbar schienen. Im Folgenden werden die 3 Anwendungsfelder des maschinellen Lernens genauer beschrieben:

Umfeldwahrnehmung durch die Verarbeitung von Sensordaten

Bei der Umfeldwahrnehmung spielen hauptsächlich drei Sensoren eine große Rolle: Kamera-, LiDAR- und Radar-Sensoren. Exotischere Varianten wie Wärmesensoren sind aktuell weniger verbreitet. Auf Basis der Sensoren werden andere Verkehrsteilnehmer, die Straße, Hindernisse etc. detektiert. Die Gesamtheit der Detektionen wird in ein 3D Umfeldmodell überführt und liefert die Grundlage für die nachfolgende Fahrplanung und Fahrstrategie.

Kamera. Kameradaten werden durch 3 Farbwerte in jedem Pixel repräsentiert, wodurch ein HD Bild mehrere Millionen numerische Werte beinhaltet. Anhand dieser Pixelwerte möchte man Objekte im Bild detektieren und einer Klasse (PKW, Lastwagen, Bus, Mensch,

Fahrrad, etc.) zuordnen. Mathematisch gesehen möchten wir eine Funktion finden, die den gegebenen Pixelwerten Positionen und Klassen aller Objekte im Bild zuordnet. Eine (klassische) händische Modellierung basierend auf selbst erdachten Regeln erreicht hier nicht die gewünschte Detektionsgenauigkeit. Die Methoden des Deep Learnings haben diese Anwendung revolutioniert.

Beim Deep Learning werden im übertragenen Sinne Schablonen von hierarchisch organisierten Funktionen mit einer großen Zahl von freien (lernbaren) Parametern verwendet. Die Parameter werden in einer mathematischen Optimierung daraufhin angepasst, dass sie einen vorgegebenen funktionalen Zusammenhang erlernen (mehr Information dazu befindet sich weiter unten in diesem Abschnitt). So kann ein tiefes neuronales Netz beispielsweise Lernen, jedem Bildpixel seine Klassenzugehörigkeit zuzuordnen.

Dies ist in der Abbildung 66 dargestellt. Dabei wird zur Illustration jeder Klasse eine Farbe zugeordnet. Um eine solche Funktion zu erlernen, bedarf es große Mengen sogenannter Labels. Aktuell werden neuronale Netze für diese Aufgabe mit Größenordnungen von 10.000en von Bildern trainiert, für jedes dieser Bilder haben Menschen die Objekte im Bild mit Kantenzügen umrandet. Die Akquise solcher Labels ist zeitaufwändig und kostspielig, jedoch für eine akkurate Umfeldwahrnehmung derzeit noch unumgänglich.

Radar. Ein weiterer, in der Automobilindustrie weit verbreiteter Sensor ist Radar (**RA**dio **D**etection **A**nd **R**anging). Obwohl der Radar keine Bilder im klassischen Sinne liefert, hat der Sensor mehrere Vorteile: die Qualität der gemessenen Daten hängt nicht von der Sichtbarkeit ab, und so liefert Radar nutzbare Daten auch in Szenarien wo Kamera oft Probleme hat, wie etwa unter schlechter Beleuchtung, Nebel oder im schlechten Wetter. Als weiterer Vorteil gilt, dass man durch Radar nicht nur die Position, sondern auch die Geschwindigkeit des Objekts direkt messen kann. Der Radar funktioniert auf Basis von elektromagnetischen Wellen: ein Signal wird ausgesendet und wenn dieser auf ein Hindernis trifft wird er zurückreflektiert. Dieses Echo wird dann mit dem gesendeten Signal verglichen, um den Zeitversatz zwischen beiden Signalen zu messen. Der Zeitversatz gibt Informationen über die Distanz des reflektierenden Hindernisses.

Mehrere Messungen erlauben sogar, die Distanzänderung über die Zeit, d.h. die Geschwindigkeit zu messen. Durch den Einsatz meh-

rerer Radar-Antennen nebeneinander ist es zusätzlich möglich, Informationen über den Beobachtungswinkel und Höhe eines Reflektionsobjektes zu extrahieren. Geeignete mathematische Verfahren überführen die empfangenen Zeitsingale in die Positionsdaten der reflektierenden Ziele, sogenannte Detektionspunkte, die eine Grundlage für Erkennung der Objekten (Fußgänger, Autos) bilden. Die Verarbeitung der Radardaten war über die Jahre hinweg durch klassische Signalverarbeitungsmethoden dominiert, doch in letzten Jahren setzten sich die Machine Learning Techniken auch hier durch. Da die Radardaten aber sehr unterschiedlich zu von Kameras aufgenommenen Bildern sind, muss man dafür auch die Deep Learning Technologie anpassen eine Herausforderung an der auch in der Wuppertaler Niederlassung von Aptiv gearbeitet wird. Für die Entwicklung geeigneter Deep Learning so wie auch der klassischen Signalverarbeitung Methoden ist das Verständnis der Mathematik hinter diesen Technologien essenziell.

Lidar. Neben Kamera und Radar ist beim autonomen Fahren auch ein Lidar (**L**ight **D**etection **A**nd **R**anging) Sensor von großer Bedeutung. Ähnlich wie beim Radar, misst auch Lidar zurückreflektierte Signale: in diesem Fall geht es aber statt einer elektromagnetischer Welle um einen Lichtstrahl, der mit entsprechender Verzögerung reflektiert wird. Diese Zeitverzögerung wird ebenfalls in die Ortsinformation des Reflektionspunktes überführt. Durch mehrere ausgesendete Lichtstrahlen kann eine Vielzahl von Reflektionspunkten generiert werden, die eine drei-dimensionale Punktwolke bilden, in der man (je nach Qualität des Lidars) die Umgebung gut erkennen kann. Die Vielzahl der Reflektionspunkte und die Genauigkeit der Positionsmessung ist einer der klaren Vorteile der Lidarsensoren. Auf anderer Seite, repräsentiert die Verteilung der Punktdaten im dreidimensionalen Raum eine Herausforderung für viele der Deep Learning Methoden, die oft spezifisch für zweidimensionale Bilder entwickelt wurden. Bei der Entwicklung neuer Algorithmen ist die Laufzeit der Algorithmen genauso zu berücksichtigen wie die Qualität der Erkennung: wenn der Algorithmus zu lange braucht, würden die Ergebnisse zu spät vorliegen, um in der Situation angemessen zu reagieren.

Ob Kamera, Radar, oder Lidar: jeder dieser Sensoren bringt mit sich neben den Vorteilen auch einige Nachteile. Gerade das macht ein weiteres Forschungsgebiet wichtig: die Sensorfusion. Mit mehreren unterschiedlichen Sensoren lässt sich das Risiko, dass alle Sensoren in demselben Moment fehlschlagen minimieren.

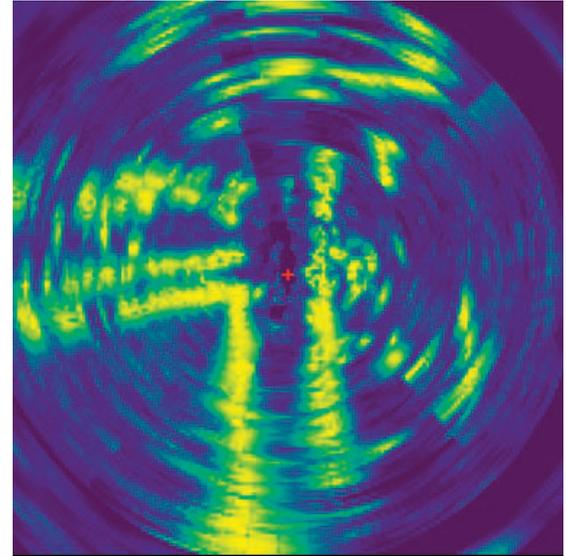
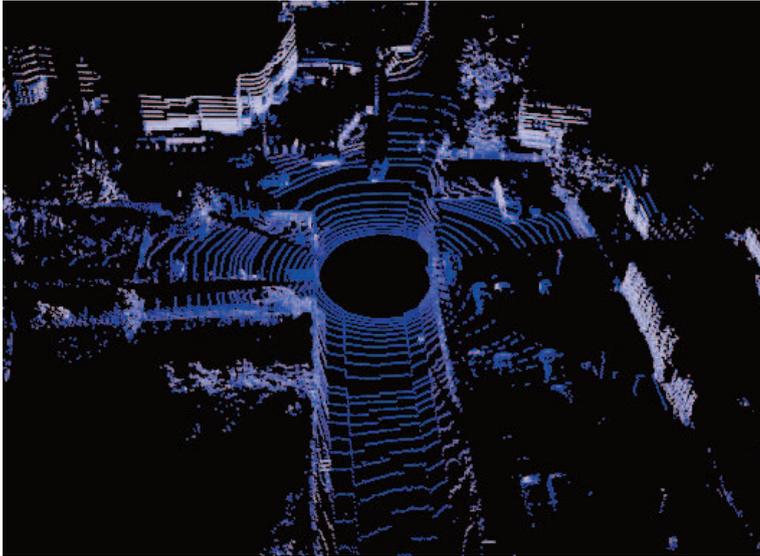


Abbildung 67: Beispiel der Daten aus Lidar (links) und der Daten aus Radar (rechts)

Fahrer-Assistenz und automatisierte Fahrzeug-Steuerung

Anhand der aus Sensordaten extrahierten Informationen über andere Verkehrsteilnehmer (wie z.B. Fußgänger, Fahrräder, Autos) und über die Umgebung, kann das Fahrzeug automatisiert Maßnahmen treffen. Dabei kann es je nach Level der Automatisierung um z.B. eine Warnung an den Fahrer, automatisches Notbremsen oder sogar um volle Steuerung des Fahrzeugs gehen. Vor allem bei komplexeren Aufgaben die sich voller Automatisierung nähern, kann es schwierig sein, das Verhalten des Autos nur durch Regeln programmieren, und ein weiteres Teil von Maschinellern Lernen kommt zum Einsatz: Reinforcement Learning. Bei dieser Methode geht es darum, anhand aller zu dem Zeitpunkt verfügbaren Informationen (Umgebung, Positionen und Bewegungsrichtungen anderer Verkehrsteilnehmer, aber auch aller solchen verfügbaren Daten aus der Vergangenheit) optimale Entscheidung zu treffen. Dabei müssen nicht nur die sofortige, sondern auch die weiterfolgende Konsequenzen dieser Entscheidung in Betracht genommen werden.

Um die Entscheidung als eine mathematische Funktion von allen verfügbaren Informationen zu repräsentieren, greift man wieder oft auf die tiefe Netzwerke zu. Die Konsequenzen dieser Entscheidung

werden dann durch einen positiven numerischen Wert bewertet falls die erwünscht sind (wie z.B. das Ankommen ins Ziel), und mit einem negativen Wert falls unerwünscht (wie z.B. verursachen eines Unfalls). Mit Hinsicht auf die Sicherheit trainiert man dieses Netzwerk in der Regel nur in einer Simulation. Während des Trainings wird das Netzwerk mit einer Verkehrssituation konfrontiert und gibt anhand der verfügbaren Informationen eine Empfehlung. Die empfohlene Handlung kann dann in der Simulation ausgeführt werden, muss aber nicht notwendig: mit gewisser Wahrscheinlichkeit trifft das Programm eine andere Entscheidung um die unbekannte Folgen zu untersuchen. Anhand der Folgen werden danach die Parameter des neuronalen Netzwerks so geändert, dass das Netzwerk künftig in ähnlichen Situationen Empfehlungen mit möglichst positiv-bewerteten Konsequenzen liefert.

Eine andere Methode die optimale Steuerung zu lernen heißt Imitation Learning. Bei dieser Methode werden als Daten aufgenommene (nicht-autonome) Fahrten benutzt: während des Trainings bekommt dann das neuronale Netzwerk die aufgenommene Daten aus einem Szenario als Eingang und versucht anhand dieser die aufgenommene Entscheidung des menschlichen Fahrers zu imitieren.

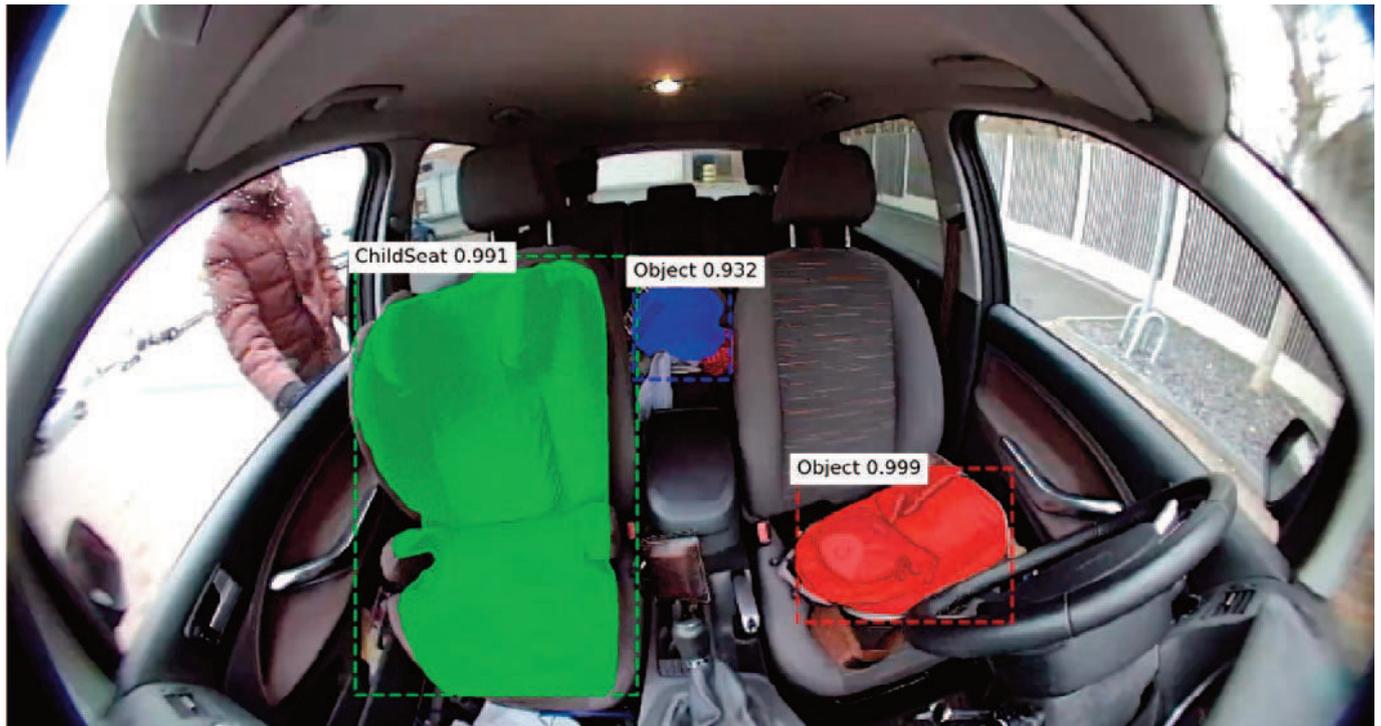
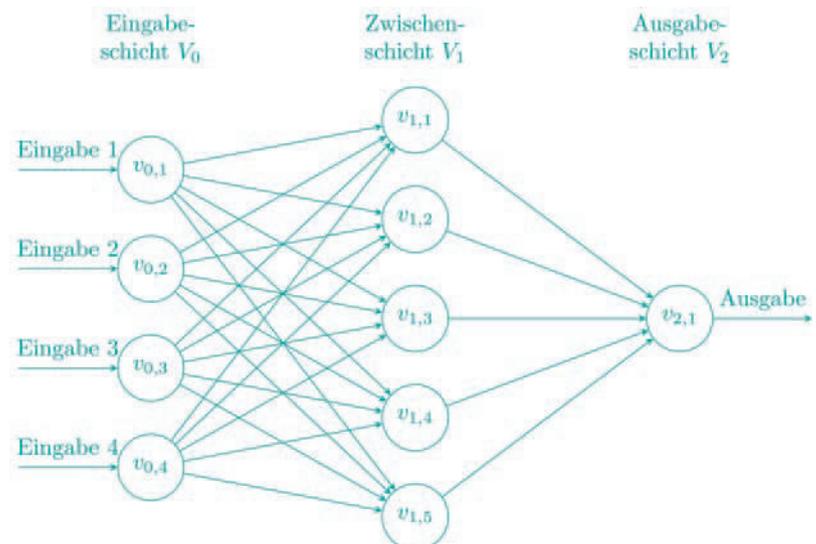


Abbildung 68: Beispiel einer Instanzsegmentierung im Innenraum eines Autos. Zuerst werden Objekte lokalisiert und von Rechtecken umrandet. Anschließend wird die Instanz innerhalb des Rechtecks segmentiert.

Abbildung 69: Skizze eines Feedforward-Netzes mit 3 Lagen. Tiefe neuronale Netze in der Bildverarbeitung sind im allgemeinen um viele Größenordnungen komplexer.



Überwachung des Innenraums

Ähnlich wie bei der Erkennung des Umfeldes mittels Kameras, ist die Erkennung der Personen und Objekte im Innenraum ebenfalls von großer Bedeutung, sowohl für das automatisierte Fahren als auch für Assistenzsysteme. Dabei kommen im Innenraum ebenfalls Kameras und Deep Learning zum Einsatz, siehe Abbildung 68. Die Detektion von Personen, die Erkennung ihrer Verfassung, die Detektion von Kindersitzen, Kindern und Haustieren sind beispielsweise von höherem Interesse. In Vorstufen des automatisierten Fahrens muss z.B. ermittelt werden, ob der Fahrgast im Fahrersitz in der Verfassung ist, das Steuer zu übernehmen. Auch könnten akut auftretende Symptome einer Erkrankung detektiert werden. Das Fahrzeug kann außerdem darauf Hinweisen, wenn ein Kind oder ein Haustier im Auto zurückgelassen wird. In Mobilitätslösungen, die den ÖPNV durch kleinere autonome Gruppentaxis voraussichtlich in Zukunft ergänzen werden, spielt es auch eine Rolle, die Fahrgäste zu überwachen um beispielsweise Übergriffe zu erkennen.

Erlernen funktionaler Zusammenhänge mittels Deep Learning

Nachdem wir nun viele Beispiele für die Anwendung des Deep Learning im automatisierten Fahren gesehen haben, beleuchten wir noch einmal die Rolle der Mathematik im Deep Learning, denn diese geht über die bloße Approximation funktionaler Zusammenhänge hinaus. In einem neuronalen Netz, der Funktionsschablone des Deep Learnings, finden hauptsächlich arithmetische Operationen (Additionen und Multiplikationen statt) statt. Neuronen bilden die kleinsten Einheiten der neuronalen Netze und erhalten als Eingabe ein Signal in Form einer Zahl von anderen Neuronen, zu denen eine Verbindung in Form einer sogenannten Kante (als Pfeil in der Abbildung 69) besteht. Auf jeder Kante lebt ein sogenanntes Gewicht, ein freier Parameter der neuronalen Netzes, welcher gelernt wird. Das Gewicht wird mit dem Signal multipliziert (also gewichtet) und alle gewichteten Signale werden im Neuron summiert und anschließend wird eine Schwellwertfunktion auf die gewichtete Summe angewandt. Wird ein bestimmte Gesamtsumme überschritten, wird das Neuron aktiviert und leitet ein Signal an die anderen mit ihm verbundenen Neuronen weiter. Diese modellieren ist biologisch inspiriert. Neuronalen Netzen für die beschriebenen Anwendungen sind meistens sogenannte Feedforward-Netze. Bei diesen sind die Neuronen gruppenweise in Lagen organisiert und eine Lage leitet Signale nur an nachfolgende Lagen (also vorwärts) weiter (vgl. Abbildung 69).

Abgesehen davon, dass neuronale Netze nichts weiter als Mathematik sind, spielen verschiedene klassische Disziplinen der Mathematik bei folgenden Fragestellungen eine gewichtige Rolle, hier nennen wir die prominentesten Beispiele:

- 1 Wie viele lernbare Gewichte werden benötigt, damit ein neuronales Netz theoretisch einen gegebenen funktionalen Zusammenhang zu einer gewünschten Genauigkeit approximieren kann? Die grobe Richtung lautet: Mehr Gewichte entsprechen mehr Kapazität und erlauben, komplexe Zusammenhänge darzustellen. Um diese Frage jedoch genauer und auch quantitativ zu beantworten (welche Anzahl an Neuronen wird benötigt?), wird immer genauere Theorie entwickelt. Da jedoch auch immer wieder neue Architekturen von neuronalen Netzen entwickelt werden, wird auch neue Theorie nötig.
- 2 Wenn das neuronale Netz nun die Kapazität in Form von freien Gewichten besitzt, wie findet man die richtigen Gewichte? Optimale Gewichte zu finden ist ein Problem, das nicht in annähernd in akzeptabler Rechenzeit, auch nicht mit den größten Computern der Welt, zu lösen ist. Effiziente Algorithmen (sogenannte Heuristiken), die möglichst gute Lösungen für das Problem der Gewichtsfindung liefern, sind Gegenstand aktueller Forschung.
- 3 Wie viele Daten braucht man, damit man einen gegebenen funktionalen Zusammenhang zu einer gewünschten Genauigkeit approximieren kann? Grob gilt: mehr Daten erlauben besseres Lernen. Auch hier sind quantitative Antworten von Interesse. Diese Fragestellung und das Finden immer genauerer Antworten ist Gegenstand der Forschung in der Stochastik / Statistik.
- 4 Wie werden die arithmetischen Operationen in neuronalen Netzen möglichst effizient in Computern umgesetzt? Ingenieure und Informatiker entwickeln dazu geeignete Hardware, Mathematiker und Informatiker entwickeln effiziente Implementierungen (Programme) von neuronalen Netzen und außerdem immer wieder neue Netzwerkarchitekturen, die genauere Vorhersagen machen und/oder weniger Rechenoperationen dafür benötigen.

Mathematische Architekturen für Neuronale Netze

Ohne Mathematik geht es auch bei der KI nicht. Optimale Netzwerkarchitekturen für komplexe Anwendungen sind am besten, wenn sie die zugrundeliegende mathematische Theorie abbilden. Dies gilt insbesondere für bildgebende Verfahren der Medizin.





KI, das ist doch Informatik, oder? So denken wohl die meisten, wobei es ohne Mathematik auch in diesem neuen, faszinierenden Forschungsfeld nicht geht. Ein besonders anspruchsvolles Feld der KI-Anwendungen sind sogenannte inverse Probleme, hierzu zählen die meisten bildgebenden Verfahren der Medizin ebenso wie die Parametrisierung von mathematischen Finanzmodellen oder Techniken der Überwachung und Steuerung komplexer Produktionsprozesse. Mit derartigen Problemen beschäftigten sich Prof. Dr. Peter Maass und seine Mitarbeiter Sören Dittmer, Tobias Kluth, Johannes Leuschner und Maximilian Schmidt an der Universität Bremen.

Aber was genau sind nun inverse Probleme? Das sind Probleme bei denen die gesuchte physikalische oder technische Größe nicht direkt gemessen, sondern indirekt bestimmt wird. Ein einfaches Beispiel ist die Messung der Höhe einer Quecksilbersäule (Messung) aus der dann die aktuelle Raumtemperatur (gesuchter Parameter) 'berechnet' wird. Dieses Beispiel ist allerdings doch zu einfach, im engeren Sinn spricht man von inversen Problemen erst, wenn zwei Schwierigkeiten hinzukommen. Als erstes werden nur Prozesse betrachtet, bei der beliebig kleine Änderungen der Messung zu beliebig großen Abweichungen des gesuchten Parameters führen können. Lange Zeit wurden derartige Probleme als „schlecht gestellt“ bezeichnet, in der Annahme, dass hier die physikalische Modellierung unvollständig oder ungenau sei. Erst Mitte der 1960er Jahre konnte mit Methoden der Funktionalanalysis gezeigt werden, dass diese Instabilität leider bei vielen Problemen inhärent ist und nicht durch bessere Messtechniken oder exaktere Modellierung behoben werden kann. Diese Instabilität wäre an sich kein Problem, wenn exakte Messdaten zur Verfügung stehen würden. Hier kommt allerdings die zweite Schwierigkeit

ins Spiel: Messdaten sind nie exakt, sondern bestenfalls im Rahmen einer vorgegebenen Messgenauigkeit bekannt.

Abbildung 70: Drei CT-Rekonstruktionen, die aus demselben Datensatz erzeugt wurden. Die linke Rekonstruktion ist offensichtlich wertlos (noisy), obwohl sie die Messdaten genauso gut wiedergibt wie die beiden anderen. Ob die mittlere oder rechte Rekonstruktion bevorzugt wird, hängt davon ab, was der medizinische Experte damit machen will. Eine Diagnose, die auf der Gewebetextur der Organe beruht würde die mittlere bevorzugen, eine Diagnose, die rein auf Form und Morphologie der Organe abzielt die rechte.

Abbildung 70 zeigt drei Rekonstruktionen aus CT-Daten, die alle die vorgegebenen Messdaten im Rahmen der Messgenauigkeit des Tomographen wiedergeben. Welche ist nun die richtige Rekonstruktion? Eine der grundlegenden Erkenntnisse der Theorie inverser Probleme besagt, dass dies bei derartigen instabilen Problemen auf die jeweilige Anwendung ankommt. In dem vorliegenden Beispiel würde ein Arzt, der Diagnosen an Hand der Gewebestruktur stellt, die mittlere bevorzugen, ein Arzt, der rein an der Form der Organe interessiert ist, wohl eher die rechte auswählen.

Es gibt also keine „richtigen“ Lösungen, sondern bestenfalls „gute“ Lösungen, die über ein Kriterium, das von Experten des Anwendungsgebiets individuell formuliert wird, ausgewählt werden.

Derartige inverse Probleme zählen zu den komplexesten Herausforderungen der angewandten Mathematik und ihre Lösung ist zudem äußerst rechenintensiv. Erste Versuche, dies mit Hilfe neuronaler Netze oder anderer Verfahren der KI effizienter zu lösen, sind allerdings gescheitert oder führten zu suboptimalen Ergebnissen.

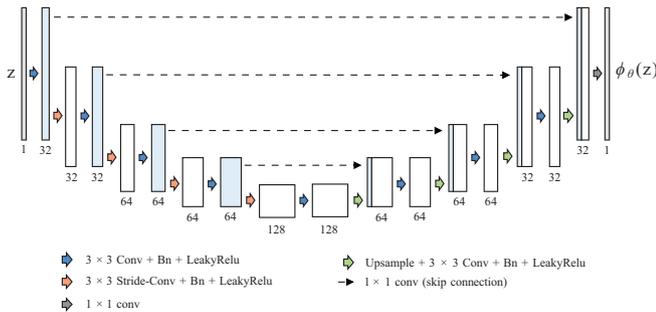
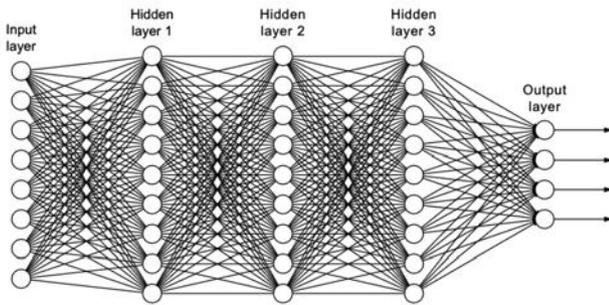
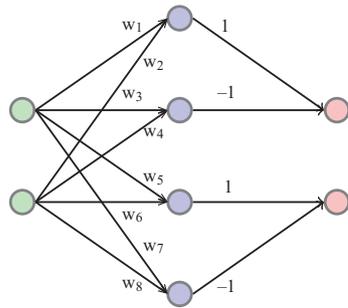


Abbildung 71: Oben ist ein einfaches neuronales Netz zu sehen, das als Eingabe zwei Zahlen erhält und ebenso zwei Werte ausgibt, das mittlere Bild ist ein „fully connected feedforward“ Netz und das untere Bild stellt das häufig verwendete „U-Net“ mit ca. 500.000 freien Parametern dar.

Eine direkte Anwendung der KI-Verfahren, die im Bereich der Computer Vision und Bildverarbeitung so erfolgreich sind, scheitert an der unvermeidbaren Instabilität derartiger inverser Probleme.

Bevor wir über konkrete Anwendungen neuronaler Netze für inverse Probleme reden, lassen Sie uns erst einmal klären, was neuronale Netze überhaupt sind und wie diese mathematisch beschrieben werden. Neuronale Netze sind dem menschlichen Nervensystem nachempfunden. Beim Menschen werden Reize über Nervenbahnen weitergeleitet und dabei entweder verstärkt oder reduziert. In den sogenannten Neuronen werden dann die ankommenden Reize verschiedener Nervenbahnen gebündelt und über einen bio-chemischen Prozess in einen ausgehenden Reiz zur Weiterleitung transformiert. Anschaulich kann man sich das Nervensystem als ein Gewirr von Nervenbahnen vorstellen, die die eingehenden Reize an eine erste Schicht von Neuronen weiterleiten. Von dort aus werden dann analog die ausgehenden Reize dieser ersten Neuronenschicht gewichtet und an eine zweite Schicht weitergeleitet usw.

Mathematisch gesprochen sind die Reize einfach Zahlen, die mit Gewichten multipliziert (Nervenbahn) und danach zusammengezählt werden. Die Aktion der Neuronen entspricht der Auswertung einer typischerweise nicht-linearen Funktion, diese wird Aktivierungsfunktion genannt. Das Ergebnis dieser Auswertung wird dann an die nachfolgenden Neuronenschichten weitergegeben.

Letztendlich entsteht eine Rechenvorschrift, die eine Menge von Zahlen als Eingabe bekommt und über die iterative Anwendung von Gewichtung (lineare Transformation, entspricht Weiterleitung von Reizen entlang von Nervenbahnen) und Aktivierungsfunktion (entspricht der Aktionen, die in einem Neuron von statten gehen) eine festgelegte Anzahl von Ausgabewerten erzeugt. Ein einfaches neuronales Netz ist in Abbildung 71 dargestellt. Die Festlegung der Anzahl der Schichten, der Anzahl der Neuronen je Schicht sowie die Auswahl der Aktivierungsfunktion bestimmen die Architektur des Netzwerks.

So wie der Mensch aus Erfahrung lernt kann nun ein neuronales Netz trainiert werden, in dem ihm Trainingsdaten mit bekanntem Ergebnis vorgelegt werden. Die Gewichte des Netzwerks werden dann so berechnet, dass das Netzwerk dieses Ergebnis bestmöglich reproduziert.

Um zu verstehen, was dies mit KI zu tun hat, muss man sich erst vergegenwärtigen, dass neuronale Netze, genauer gesagt „fully connected feed forward“ Netzwerke mit L Schichten als Realisierung von L Schritten eines Iterationsverfahrens interpretiert werden können. Ein derartiges neuronales Netz hat eine Vektor x_0 als Eingabegröße. Die Werte des Vektors werden von links nach rechts, Schicht für Schicht weitertransportiert. Zwischen den Schichten, werden die einzelnen Werte mit Gewichten w multipliziert, das entspricht der Weiterleitung inkl. Dämpfung oder Verstärkung von Reizen entlang der Nervenbahnen. In den „Neuronen“ der einzelnen Schichten werden dann die einzelnen ankommenden Werte zusammengefasst und komponentenweise nichtlineare Transformation unterworfen, das entspricht den biochemischen Reaktionen, die in den einzelnen Neuronen des Körpers z.B. entscheiden, ob der Reiz überhaupt weitergeleitet werden soll.

Mathematische Grundlagen

Der Weg zu einem erfolgreichen Einsatz neuronaler Netze bei Inversen Problemen führt über Netzwerkarchitekturen, die aus tiefliegenden Forschungsergebnissen zur Analysis Inverser Probleme abgeleitet sind.

Hierzu braucht man Einblicke in die mathematischen Grundlagen, sowohl von inversen Problemen als auch von neuronalen Netzen.

Bei einem inversen Problem wird – in einer allgemeinen Formulierung – ein unbekannter Parameter x , für den lediglich Messwerte y^δ bekannt sind, die über einen Messprozess F mit einer Messgenauigkeit δ gemessenen Daten wurden. In dieser Formulierung sind alle Parameter x , die

$$\| F(x) - y^\delta \| \leq \delta$$

erfüllen, potentielle Lösungen des Problems. Unter diesen wird nun derjenige Parameter ausgewählt, der ein zusätzliches Kriterium R_{expert} am besten erfüllt. Die genaue Festlegung von R kann recht allgemein sein, hier können die Anwendungsexperten Eigenschaften kodieren, wie z.B. dass die Kanteninformation in den Rekonstruktionen besonders wichtig ist. Mit Hilfe der Lagrange-Multiplikatoren führt dies zu einem Minimierungsproblem

$$\min_x \| F(x) - y^\delta \|^2 + \alpha R_{\text{expert}}(x).$$

Derartige Minimierungsaufgaben können für ein nicht-lineares F und nicht-differenzierbare Funktionale R_{expert} nicht direkt gelöst

werden. Ableiten nach x und Nullsetzen erlaubt aber die Konstruktion von Iterationsverfahren, die ausgehend von einem geeignet gewählten x_0 Schritt für Schritt die Lösung annähern. Derartige Gradientenabstiegsverfahren führen z.B. auf

$$x_{k+1} = \text{Prox}_R \left(x_k - \lambda F'(x_k)^* (F(x_k) - y^\delta) \right).$$

Um zu verstehen, was dies mit KI zu tun hat, muss man sich erst vergegenwärtigen, dass neuronale Netze, genauer gesagt „fully connected feed forward“ Netzwerke mit L Schichten als Realisierung von L Schritten eines Iterationsverfahrens interpretiert werden können. Ein derartiges neuronales Netz hat eine Vektor x_0 als Eingabegröße. Die Werte des Vektors werden von links nach rechts, Schicht für Schicht weitertransportiert. Zwischen den Schichten, werden die einzelnen Werte mit Gewichten w multipliziert, das entspricht der Weiterleitung inkl. Dämpfung oder Verstärkung von Reizen entlang der Nervenbahnen. In den „Neuronen“ der einzelnen Schichten werden dann die einzelnen ankommenden Werte zusammengefasst und komponentenweise weiterverarbeitet, das entspricht den biochemischen Reaktionen, die in den einzelnen Neuronen des Körpers z.B. entscheiden, ob der Reiz überhaupt und wenn ja wie stark weitergeleitet werden soll.

Mathematisch formuliert lässt sich ein fully connected feedforward Netzwerk mit L Schichten zur Lösung des inversen Problems also ebenfalls als iterative Rechenvorschrift schreiben. Zu gegebenem y^δ und mit beliebig gewähltem x_0 wird $b = -\lambda F'(x_0)^* y^\delta$ gesetzt. Dann entspricht die Anwendung des neuronalen Netzes der Iterationsvorschrift

$$x_{k+1} = \varphi (Wx_k + b).$$

Für die spezielle Wahl $W = I - \lambda F'(x_k)^* F(x_k)$, $b = -\lambda F'(x_k)^* y^\delta$ und $\varphi = \text{Prox}_R$ sieht man die Analogie zu dem analytischen Gradientenabstiegsverfahren. Entscheidend ist aber, dass die Aktivierungsfunktion φ des Netzwerks entsprechend der mathematischen Theorie als proximal Mapping Prox_R der Experteninformation R gewählt wird. Darauf wäre man ohne die mathematische Theorie im Hinterkopf nicht gekommen. In diesem Fall entspricht jedenfalls der Output des Netzwerks, der mit $\Phi(y^\delta, W)$ bezeichnet wird, exakt der L -ten Iterierten des analytischen Gradientenabstiegsverfahrens.

Der Vorteil der neuronalen Netze, ist jedoch, dass die Matrix W und der Bias-Vektor b an Hand von experimentell gemessenen Daten optimal an die jeweilige Anwendung angepasst werden: Hierzu werden

zu mehreren bekannten Parametern $x_i, i = 1, \dots, N$ die Messdaten y_i erhoben. Danach werden die Gewichte W über die Minimierung der sogenannten Loss-Funktion

$$\sum_{i=1}^N \|\Phi(y_i^\delta, W) - x_i(W)\|^2$$

bestimmt. So kann das Netzwerk optimal an eine gegebenen Datenstruktur angepasst werden und selbst feinste Modellierungsdetails des Messoperators sind in den Messdaten enthalten. Damit wird das Netzwerk dahin getrimmt, Outputs zu erzeugen, die ähnlich denen der Testdaten sind.

Anwendung

Als Anwendungsbeispiel wollen wir ein besonders anspruchsvolles und relativ neues tomographisches Verfahren wählen. Bei dem sogenannten „magnetic particle imaging“ (MPI) werden kleinste Nanopartikel in die Blutbahn injiziert. Diese Nanopartikel haben einen magnetisierbaren metallischen Kern und eine bioaktive Hülle. Diese Partikel werden nun mit dem Blutfluss transportiert und erlauben – sofern man ihre aktuelle Position nachverfolgen kann – detaillierten Aufschluss über Blockaden in Blutgefäßen und über die Herzfunktion. Prinzipiell kann die bioaktive Hülle auch so konstruiert werden, dass sich die Nanopartikel an bestimmten metabolisch charakteristischen Regionen ansiedeln und so detailliert funktionale Erkenntnisse liefern. Diese Nanopartikel werden danach über den natürlichen Kreislauf wieder ausgeschieden und sind so potentiell eine harmlose Alternative zu radioaktiven Kontrastmitteln.

Dabei entsteht allerdings die Frage, wie kann die Position der Nanopartikel nachverfolgen nachdem sie injiziert wurden? Hierzu wird ein dynamisches Magnetfeld außerhalb des Körpers angelegt, um gezielt Nanopartikel in bestimmten Bereichen zu magnetisieren. Die dynamische Änderung dieses äußeren Magnetfelds führt nun dazu, dass Nanopartikel anfangen zu schwingen und ein eigenes elektromagnetisches Feld zu erzeugen. Dieses wird wiederum durch Spulen außerhalb des Körpers gemessen. In einem einfachen linearen Modell wird das elektromagnetische Feld, das von einem einzelnen Nanopartikel in Position x erzeugt wird, mit $s(x, t)$ bezeichnet. Ist nun $c(x)$ die Anzahl oder Konzentration der Nanopartikel an der Position x so entsteht das Gesamtsignal durch Summation (Integration) der Beiträge an allen Positionen im Körper Ω

$$y(t) = \int_{\Omega} c(x)s(x,t) dx.$$

Das ist allerdings ein stark vereinfachtes Modell das z.B. sowohl die komplexe Messcharakteristik der messenden Spulen und die Partikel-Partikel Interaktionen ebenso wie den Einfluss der Größenverteilung der Nanopartikel außer Acht lässt. Dies lässt sich alles nur sehr schwer modellieren. Aber, wie gesagt, alle diese Modellierungseinheiten sind in hinreichend guten Testdatensätzen enthalten. Dies macht den Einsatz daten-getriebener Verfahren in diesem Bereich so attraktiv.



	$\tau=0$	$\tau=1$	$\tau=3$	$\tau=5$
DIP $D\ell^1 + P$ -				
KACZ $D\ell^2 + P\ell^2$				
KACZ $D\ell^2 + P\ell^1$				

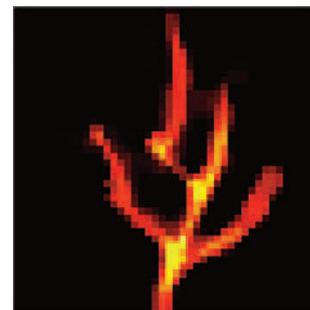


Abbildung 72: Links ist ein MPI-Gerät des UKE (Universitätsklinikum Hamburg Eppendorf) zu sehen, die mittleren Grafiken sind Rekonstruktionen eines real gemessenen Phantoms, das einmal herkömmlich (KACZ) und einmal über neuronale Netze (DIP) rekonstruiert wurde, das rechte Bild zeigt eine KI-Rekonstruktion eines Adernbaums.

Wie soziale Netzwerke bei der Behebung von Softwarefehlern helfen

* Künstliche Intelligenz

Mathematiker und Informatiker nutzen Methoden der künstlichen Intelligenz, um automatisch große Mengen von Fehlerberichten in großen Softwaresystemen zu kategorisieren. Graphentheorie, wie sie auch in sozialen Medien zum Einsatz kommt, hilft ihnen dabei, diese Aufgabe besonders effizient zu erledigen.

Bugtracking in Open Source Software Projekten

Software kommt in unserer modernen Gesellschaft fast überall zum Einsatz, sei es in unseren Smartphones, im Auto, bei Onlinediensten, oder in Produktionsanlagen von Unternehmen. Ihre Programmierung ist dabei meist so komplex, dass diese ein Team von Softwareentwicklerinnen und -entwicklern erfordert. Häufig handelt es sich hierbei um bezahlte Mitarbeiterinnen und Mitarbeiter einer Firma, welche die Software herstellt, testet und verkauft. Immer häufiger wird Software aber auch auf Grundlage einer sogenannten Open Source Lizenz von einer Community erstellt, die aus vielen Freiwilligen besteht, die sich über Online-Plattformen austauschen und mittels Internet-Technologien global zusammenarbeiten. Populäre Beispiele für Software, die so entstanden ist, sind der Browser Mozilla

Firefox, das mobile Betriebssystem Android oder auch die Software, welche die Grundlage für die Online-Enzyklopädie Wikipedia bildet.

Da es in diesen sogenannten Open Source Software (OSS) Communities häufig keine bezahlten Testnutzer gibt, ist das Feedback der Nutzer für die Qualität des Endprodukts von besonderer Bedeutung. Hierzu betreiben fast alle Projekte sogenannte „Bug tracking“-Plattformen. Dieser Name beruht auf der Tatsache, dass man einen Fehler in einer Software in der Informatik auch als „Bug“, (englisch für „Insekt“, oder Käfer) nennt ein Begriff, der aus einer Zeit stammt, als Computerprozessoren noch so groß waren, dass ein darin umherkrabbelnder Käfer tatsächlich einen Fehler auslösen konnte. Hat man als Nutzer einer Open Source Software nun ein Problem, welches man für einen Fehler hält, kann man einen sogenannten „Bugreport“, d.h. einen Fehlerbericht, schreiben. Darin beschreibt man möglichst genau, welche Version der Software man verwendet, wie sich der Fehler ausdrückt, wann und unter welchen Umständen dieser auftritt und welches Ergebnis man eigentlich erwartet hätte. Nun hofft man darauf, dass die Entwickler sich das Problem anschauen und möglichst bald eine Lösung anbieten.

Leider stellt die schiere Anzahl solcher Fehlerberichte die Entwickler besonders erfolgreicher Open Source Projekte vor große Herausforderungen. So mussten die Entwickler von Mozilla Firefox über einen Zeitraum von zehn Jahren mehr als 112.000 Fehlerberichte von Nutzern abarbeiten. Wenn die Bearbeitung jeder dieser Berichte nur 10 Minuten dauert, würde eine einzelne Person hierzu mehr als 2.300 Arbeitstage benötigen. Interessanterweise zeigt eine Analyse aber auch, dass mehr als 80 % der Fehlerberichte sich nicht auf einen Fehler in der Software zurückführen lassen, sondern bspw. auf einem Missverständnis der Nutzer oder einer Fehlbedienung beruhen. Nur ein kleiner Teil dieser großen Anzahl von Fehlerberichten hilft also tatsächlich bei der Verbesserung der Software. Wenn es uns nun also gelingt möglichst schnell genau diese hilfreichen Fehlerberichte zu identifizieren, hilft das dabei Softwarefehler möglichst schnell und effizient zu beheben. Können wir Methoden der Mathematik einsetzen, um dieses Problem zu lösen?

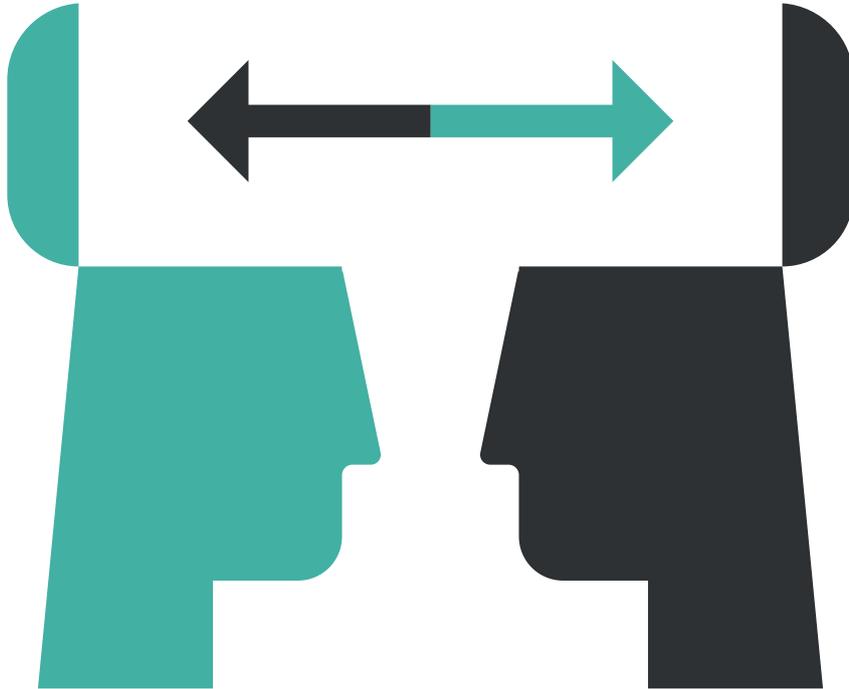
Soziale Netzwerke in Open Source Projekten

Heutzutage kennt bereits jedes Kind soziale Medien wie TikTok, Twitter, Instagram oder Facebook. Auch wenn sich diese Plattformen in vielen Details unterscheiden, liegt doch allen ein gemeinsames

Konzept zugrunde, welches in der Mathematik bereits vor fast 300 Jahren entwickelt wurde: ein soziales Netzwerk von Nutzern, welches in Form eines sogenannten Graphen dargestellt werden kann. Einen solchen Graphen kann man sich in etwa vorstellen wie Malen nach Zahlen, d.h. er besteht aus Punkten, den sogenannten Knoten, die über Linien, den sogenannten Kanten, paarweise miteinander verbunden sein können. Graphen aus Knoten und Kanten werden in vielen Bereichen der Wissenschaft eingesetzt, um komplexe Beziehungen zwischen verschiedenen Dingen darzustellen, bspw. Bücher in einer Bibliothek, die über Querverweise miteinander verbunden sind, Tierarten in einem Ökosystem, deren Überleben voneinander abhängt, oder Menschen, die in einem sozialen Netzwerk miteinander kommunizieren, bekannt oder befreundet sind. Die mathematische Beschreibung und Analyse solcher komplexen Graphen ist die Grundlage vieler aus dem Alltag bekannter Technologien. Beispielsweise liegt der Suche mit Suchmaschinen wie Google oder Bing eine Analyse von Verbindungen, sogenannter Hyperlinks, zwischen Milliarden von Webseiten zugrunde. Freundschaftsempfehlungen in sozialen Medien wie Twitter oder Instagram werden auf Grundlage der Verbindungen zu unseren Freunden, sowie deren Verbindungen untereinander berechnet. Und wenn wir Online einkaufen, beruhen die von Einkaufsplattformen vorgeschlagenen Produkte häufig auf einem Graphen, in dem diejenigen Produkte miteinander verbunden werden, die häufig zusammen bestellt werden.

Doch wie kann die mathematische Analyse solcher komplexen Graphen nun dabei helfen, die Flut von Fehlerberichten in Open Source Software Projekten zu beherrschen? Hierzu haben sich Forscherinnen und Forscher einen Trick ausgedacht, bei dem die Mitglieder einer Open Source Community in Form eines sozialen Netzwerks dargestellt werden. Wann immer zwei Personen über die jeweilige Kollaborationsplattform miteinander kommunizieren werden diese über eine Kante miteinander verbunden. Über die Zeit entsteht so ein großer komplexer Graph, der die Community von Freiwilligen eines Open Source Projekts darstellt und der es uns erlaubt diese mit mathematischen Methoden zu analysieren.

Wann immer ein Nutzer nun einen Fehlerbericht einsendet steht nun nicht nur dieser Fehlerbericht zur Verfügung, sondern wir kennen auch den Knoten im sozialen Netzwerk, der zum jeweiligen Nutzer gehört. Wie stark dieser Knoten mit anderen Nutzern vernetzt ist, zu welchen anderen Knoten diese Verbindungen bestehen, oder an welcher Position im Netzwerk sich diese befinden kann nun – neben



den Daten des Fehlerberichtes – von einem maschinellen Lernverfahren genutzt werden, um automatisch diejenigen Fehlerberichte zu identifizieren, die tatsächlich auf einen wichtigen Fehler in der Software hindeuten. Die Idee dahinter ist, dass die Position einer Person innerhalb einer Open Source Community Rückschlüsse darüber zulässt, wie gut sich die jeweilige Person mit den Details der Software auskennt und wieviel Erfahrung diese darin hat, ein Problem so zu beschreiben, dass die Entwicklerinnen und Entwickler den zugrundeliegenden Fehler im Code finden können. Eine künstliche Intelligenz kann nun also lernen, aus welchen „Regionen“ des sozialen Netzwerks der Open Source Community besonders hilfreiche Fehlerberichte kommen.

Ein Test dieses Verfahrens in einem Datensatz mit Millionen von Fehlerberichten zu verschiedenen Open Source Projekten zeigt, dass diese Methode in mehr als 90 % aller Fälle richtige Vorhersagen macht, d.h. von 100 Fehlerberichten, die die Methode als potenziell

hilfreich identifiziert hat resultieren mehr als 90 tatsächlich aus einem Fehler in der Software. Das Verfahren kann daher genutzt werden, um schnell diejenigen Fehlerberichte zu identifizieren, welche mit großer Wahrscheinlichkeit tatsächlich auf einen Fehler in der Software zurückzuführen sind. Diese können von den Entwicklern dann zuerst angeschaut werden, um Probleme möglichst schnell zu beheben. Dieses Verfahren, welches auf mathematischen Konzepten beruht, die mehrer Jahrhunderte alt sind, trägt also in unserer digitalen Gesellschaft dazu bei, Softwarefehler möglichst schnell zu finden und zu beheben.



Die Mathematik hinter den Empfehlungen von Netflix und Amazon Prime

Ein Algorithmus schlägt Ihnen auf subtile Weise vor, welches Buch Sie als nächstes lesen oder welchen Film Sie sich ansehen sollen. Dahinter steckt lineare Algebra.

Internetunternehmen, wie z.B. Google, Amazon, Netflix, wissen am besten, wie man zu gesuchten Informationen, wie etwa ein Buch oder ein Film, kommt. Dazu verwenden sie Methoden der linearen Algebra, also die Mathematik der Vektoren und Matrizen, genauegenommen die Matrixvervollständigung und die *Singulärwertzerlegung*⁴⁴.

Wenn man einen Film bei einem Streamingdienst gesehen hat, hat man im Anschluß die Möglichkeit ihn mit einem bis fünf Sternen zu bewerten. Auf Basis dieser persönlichen Bewertungen und ihrer Auswahl von Filmen wird bestimmt, welche weiteren Filme Ihnen gefallen könnten. Anschließend bekommt man entsprechende Empfehlungen zugeschickt.

Ein einfacher Weg, um herauszufinden, ob Ihnen z.B. „Das Boot“ gefallen würde, wäre der Durchschnitt aller Bewertungen, die der Film von anderen Nutzern erhalten hat. Dies ist jedoch kein sehr guter Ansatz, da er ignoriert, ob Sie generell U-Boot Filme mögen oder nicht.

Aber wie kann man entscheiden, welche Benutzer wie Sie sind? An dieser Stelle kommt die lineare Algebra ins Spiel. Stellen Sie sich eine riesige Tabellenkalkulation mit Spalten vor, die nach jedem der Zehntausenden von Filmen benannt sind, die auf Netflix verfügbar sind. Jeder der mehreren Millionen Abonnenten entspricht einer Zeile in dieser Tabelle. Wenn ein Benutzer einen Film bewertet, wird ein Eintrag ausgefüllt. Wenn ich zum Beispiel „Avatar – Aufbruch nach Pandora“ drei Sterne gebe, wird die Spalte, die diesem Film entspricht, in der Zeile mit meinem Namen aktualisiert.

Nur wenige bewerten mehr als hundert Filme, so dass nur ein winziger Bruchteil dieser riesigen Tabelle tatsächlich ausgefüllt wird. Die Aufgabe des Empfehlungsalgorithmus ist es, die gesamte Tabelle auf der Grundlage dieser sehr spärlichen Informationen auszufüllen. Wie kann das effizient gemacht werden?

Der Netflix-Algorithmus

Vor etwa drei Jahren bot Netflix eine Million Dollar für jeden an, der die Genauigkeit der Empfehlungen um 10 % verbessert⁴⁵. Teams von Software-Ingenieuren und Enthusiasten aus der ganzen Welt haben an dieser Herausforderung teilgenommen. Die erfolgreichsten Ansätze basieren auf linearer Algebra: Während die Tabelle der Benutzerpräferenzen riesig ist, gibt es vielleicht ein paar Dutzend stereotypische Bewertungsprofile. Der Geschmack eines jeden Menschen wird als eine Mischung aus diesen Profilen beschrieben. Zum Beispiel mögen Sie vielleicht Krimis und gelegentlich eine Dokumentation. Lineare Algebra kann diese stereotypen Profile identifizieren und die magische Mischung liefern, die Ihre Vorlieben beschreibt.

Verbesserungen von 8 bis 9 % gegenüber dem bestehenden Netflix-Algorithmus wurden schnell nach Bekanntgabe des Preises erreicht. Das verlangte 10 %-Ziel erwies sich jedoch als schwer zu erreichen. Nach drei Jahren hat es ein Team von Programmierern aus aller Welt im Juli 2009 endlich geknackt. Sie erreichten dies, indem sie berücksichtigten, dass sich unsere Vorlieben im Laufe der Zeit stark verändern.

Yehuda Koren von Yahoo! Research Israel war einer der ersten Teilnehmer, die diese 10 %-Marke überschreiten konnten, indem er den Bewertungen eine zeitliche Komponente gab⁴⁶. Dazu hatte er ein Modell entwickelt, das die zeitliche Dynamiken in den kollaborativen Filtern oder Empfehlungssystemen sensibel genug beschrieb⁴⁷. Obwohl aktuelle Daten mehr über die aktuellen Präferenzen eines Benutzers verraten, gehen durch die Untergewichtung älterer Bewertungen zu viele wertvolle Informationen verloren. Der Trick von Koren bestand darin, alles beizubehalten, was das langfristige Verhalten des Nutzers vorhersagt, und gleichzeitig das temporäre Rauschen, die „Launen“ herauszufiltern. Zu diesem Zweck hat Korens Modell eine Komponente für die Aufnahme des dauerhaften Signals und eine andere für die Erkennung von Anzeichen von vorübergehendem Rauschen, wie z.B. bestimmte abrupte Änderungen in den Bewertungen.

Eine Unterhaltung mit Maschinen

Mächtige Algorithmen formen, was Sie in Ihrem Webbrowser sehen. Sie können Ihnen auch auf subtile Weise vorschlagen, welches Buch Sie als nächstes lesen und welchen Film Sie sich ansehen sollten. Wenn Sie auf diese Vorschläge mit Ihrem Feedback antworten, beeinflussen Sie nicht nur Ihre zukünftigen Empfehlungen, sondern auch die aller anderen Nutzer desselben Dienstes. Sie nehmen an einem Gespräch zwischen Maschinen und Menschen teil.

⁴⁴ T. Hastie, R. Mazumder, J.D. Lee, R. Zadeh, Matrix completion and low-rank SVD via fast alternating least squares, *The Journal of Machine Learning Research*, 16(1) (2015), 3367-3402.

⁴⁵ www.netflixprize.com

⁴⁶ www.wired.com/2009/09/bellkors-pragmatic-chaos-wins-1-million-netflix-prize/

⁴⁷ Y. Koren, The BellKor Solution to the Netflix Grand Prize, (2009). www.asc.ohio-state.edu/statistics/dmsl/GrandPrize2009_BPC_BellKor.pdf

Sprachen durch Zählen von Wörtern bändigen

Dank einer mathematischen Gesetzmäßigkeit in der Sprache kann Google im Bruchteil einer Sekunde eine Suchanfrage beantworten oder einen Text automatisch übersetzen.



Welches ist das am häufigsten verwendete Wort im Niederländischen? Das hängt davon ab, um welche Art von Sprache es sich handelt. Wenn es um das geschriebene Niederländisch in Zeitungen und Zeitschriften geht, ist „das“ das am häufigsten verwendete Wort. Nimmt man jedoch das gesprochene Niederländisch, erscheint oben das Wort „Ja“. Und auf Twitter ist „Ich“ der Spitzenreiter.

Doch das geschriebene Niederländisch, das gesprochene Niederländisch und das Twitter-Holländisch haben eines gemeinsam: Das häufigste Wort innerhalb einer solchen Domäne ist doppelt so häufig wie Nummer zwei auf der Rangliste, dreimal so häufig wie Nummer drei und so weiter. Wenn wir die Frequenz des häufigsten Wortes auf 1 setzen, dann bilden die Worthäufigkeiten die Reihe 1, 1/2, 1/3, 1/4, ...

Dieses in einer mathematischen Formel ausgedrückte Muster wird *Zipfsches Gesetz* genannt, nach dem amerikanischen Linguisten George Zipf, der das Gesetz 1935 entdeckte. „Dieses Gesetz scheint für alle Sprachen und für alle Textsammlungen innerhalb einer Sprache zu gelten, egal ob man in einem chinesischen Rechtsbuch, einer norwegischen Bibel oder in englischsprachigen E-Mails eines großen Unternehmens nachschlägt“, sagt Antal van den Bosch, Professor an der Radboud Universität Nimwegen und Spezialist für Computerlinguistik. „Das Zipfsche Gesetz ist ein empirisches Gesetz, aber es ist ziemlich genau. Nur am Anfang, bei den Top 10 der Wörter, und am Ende, bei den seltenen Wörtern, weicht die Praxis ein wenig von der mathematischen Formel ab.“

Effiziente Suche

Gerade weil das Gesetz von Zipf allgemeingültig ist, kann die Suchmaschine von Google so schnell antworten. Van den Bosch: „Der Trick von Google besteht darin, dass sie einen Wortindex des Webs erstellt haben und diesen ständig aktualisieren. Der Wortindex zeigt an, welches Wort in welchem Dokument enthalten ist. Mit dem Zipf-Gesetz können Sie nun zeigen, dass dieser Wortindex kompakt ist. Und das wiederum bedeutet, dass Sie es kompakt auf Festplatten speichern und problemlos an Datenzentren auf der ganzen Welt verteilen können.“

Warum genau ist dieser Wortindex kompakt? Google hat Zugang zu Dutzenden Milliarden von Webseiten, aber die Anzahl der Wörter pro Sprache liegt „nur“ im Millionenbereich, von denen es in einem offiziellen Wörterbuch normalerweise nur einige Hunderttausend

gibt. Das Zipf-Gesetz lehrt uns nun, dass die Hälfte der Wörter in einer großen Textsammlung nur einmal vorkommt. Dank Zipf wissen wir auch, dass die Top 300 fast alle Funktionswörter (Artikel, Pronomen, Präpositionen ...) und die am häufigsten verwendeten Inhaltswörter (Substantive, Verben, Adverbien, Adjektive) enthalten. Diese beiden Eigenschaften machen den Wortindex kompakt.

Van den Bosch: „Wenn wir einen Suchbegriff eingeben, muss Google nicht in Zehnmilliarden von Dokumenten suchen, sondern in dem viel handlicheren Wortindex. Und wenn jemand vier Schlüsselwörter eingibt, nimmt die Suchmaschine die Überschneidung von vier Sammlungen. Jeder Satz sagt aus, auf welcher Webseite das Wort erscheint. Es ist eine einfache Berechnung, also geht es schnell.“

Automatische Übersetzung

Maschinelle Übersetzungsmaschinen nutzen eine Art abgeleitete Eigenschaft des Zipf-Gesetzes aus, nämlich die Eigenschaft, Wortkombinationen zu verhindern. Google Translate verwendet eine große Datenbank mit bereits vorhandenen Übersetzungen, z.B. offiziell übersetzte Texte des Europäischen Parlaments oder übersetzte Untertitel von Filmen. Um einen neuen Text z.B. vom Niederländischen ins Englische zu übersetzen, sucht die Übersetzungsmaschine nach möglichst langen Wortkombinationen, die in bestehenden Übersetzungen so oft wie möglich auf die gleiche Weise übersetzt wurden.

Die Übersetzungsmaschine sieht zum Beispiel, dass das Shakespeare-Zitat „Juliet is the sun“ immer mit „Julia ist die Sonne“ übersetzt wird. In diesem Fall muss es sich um die korrekte Übersetzung handeln. Wie oft Wortkombinationen in einer bestimmten Reihenfolge vorkommen, wird ebenfalls in Zipf-ähnlicher Weise aufgeteilt: nur eine begrenzte Anzahl von Kombinationen ist sehr häufig. Und so wie der Wortindex kompakt ist, so ist auch der Index der Wortkombinationen kompakt. Deshalb erledigt eine Übersetzungsmaschine ihre Arbeit so schnell.

Diese statistische Übersetzungsmethode eignet sich gut für Texte, die bestehenden Texten sehr ähnlich sind. Aber je einzigartiger und kreativer der Text ist, desto schwieriger ist es für die Übersetzungsmaschine. „Poesie ist notorisch schwierig“, sagt van den Bosch. „Der heilige Gral auf meinem Gebiet lautet daher: Wie können wir sicherstellen, dass Maschinen Sprache wirklich verstehen? Denn Google Translate macht das immer noch nicht, egal wie nützlich es oft ist.“

Prof. Dr. Anita Schöbel

Eine kurze Zusammenfassung Ihrer Karriere

Nach meinem Mathematikstudium an der TU Kaiserslautern promovierte ich im Fachgebiet der diskreten Optimierung. Nach der Promotion arbeitete ich zwei Jahre am Institut für Techno- und Wirtschaftsmathematik, ging dann an die Universität zurück und habilitierte mich dort 2003 in Mathematik. 2004 nahm ich einen Ruf an die Georg-August Universität in Göttingen an, zunächst als W2-Professorin, nach zwei auswärtigen Rufem 2007 als W3-Professorin. Ich war dort Studiendekanin, habe das Simulationswissenschaftliche Zentrum Clausthal-Göttingen mit aufgebaut und einige Jahre geleitet. Ein Forschungssemester verbrachte ich mit meinem Mann und unseren beiden Kindern in Auckland (Neuseeland).

2019 nahm ich einen Ruf an die TU Kaiserslautern und als Leiterin des Fraunhofer- Instituts für Techno- und Wirtschaftsmathematik an. Sechs Jahre war ich im Vorstand der Gesellschaft für Operations Research, 2019/2020 als Präsidentin. In Kaiserslautern bin ich im Hochschulrat, bei Fraunhofer u.a. Sprecherin des Strategischen Forschungsfeldes „Next Generation Computing“ und seit 2020 im wissenschaftlichen Senat der Nationalen Forschungsdateninfrastruktur (NFDI).

Was hat Mathematik für Ihre Karriere bedeutet?

Alles! Mathematik hat mir schon in der Schule Spaß gemacht. Meine Begeisterung entstand, als mich mein Lehrer an der Tafel anhand der quadratischen Ergänzung die Formel zum Auflösen einer quadratischen Gleichung herleiten ließ. Die Entscheidung, Mathematik zu studieren (ich war auch in anderen Fächern gut), ist mir trotzdem nicht leichtgefallen. Aber ich habe sie nie bereut. An der Universität hat mir Mathematik weiterhin Spaß gemacht, deswegen bin ich zur Promotion geblieben. Und zur Habilitation. Ich wurde Professorin und schließlich Leiterin des Fraunhofer-Instituts für Techno- und Wirtschaftsmathematik ITWM.



Das Tolle an der Mathematik sind zwei Dinge:

Erstens: Die Logik und Klarheit der Mathematik. Was bewiesen ist, das stimmt. Mathematik bringt Ordnung durch schöne Strukturen, hilft das Wesentliche zu sehen.

Zweitens: Mathematik ist nützlich. Mathematik hilft, Krebspatienten besser zu behandeln, Energie zu sparen, den Verkehr effizienter zu gestalten, Produktionsprozesse und Materialien zu optimieren. Es gibt so viele Beispiele, die zeigen, wie Mathematik unser Leben verbessert.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihrer Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Das Fraunhofer-Institut für Techno- und Wirtschaftsmathematik (ITWM) in Kaiserslautern zählt zu den größten mathematischen Forschungsinstituten weltweit. Bei unseren fast 250 Mitarbeitenden und mehr als 60 Doktorandinnen und Doktoranden dreht sich fast alles um Mathematik. Wir befassen uns hauptsächlich mit Modellierung, Simulation und Optimierung. Dafür entwickeln wir Mathematik als Schlüsseltechnologie weiter, geben innovative Anstöße, unterstützen Industriepartner bei der Modellierung und Digitalisierung von Prozessen und Produkten sowie der Anwendung und Verbesserung von Hochleistungs-Rechnertechnologien und wir stellen maßgeschneiderter Software- Lösungen bereit. Unsere Kunden finden sich vor allem in den Bereichen Automotive, Maschinenbau, Chemie und Pharma, Textil, Medizintechnik, der Computerindustrie und im Finanzund Versicherungssektor. Auch bei der Bekämpfung der Corona-Krise konnten wir mit mathematischen Methoden unterstützen: Zur Vorhersage der Pandemie, bei der Logistik-Unterstützung rund um das Impfen, bei der Analyse der Durchfeuchtung von Masken oder der Ausbreitung von Aerosolen. Auch die Herstellung von Schutzkleidung lässt sich durch mathematische Methoden optimieren.

Sind, in Ihrer Hinsicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Seit seiner Gründung vor über 25 Jahren hat unser Institut eine Brücke als Bild zur Verdeutlichung genau dieser Vision: Wir bauen Brücken zwischen der realen Welt und der Mathematik. Eine Fragestellung der realen Welt wird durch ein mathematisches Modell in die virtuelle Welt der mathematischen Verfahren gebracht und dort gelöst. Die

Lösung wird dann in der realen Welt interpretiert und dort (wenn alles geklappt hat) erfolgreich eingesetzt. Genau das sehen wir als unsere Aufgabe: Erkenntnisse aus der Universität, neue Forschungsergebnisse und Methoden in die reale Welt zu transferieren, also in Wirtschaft und Gesellschaft anzuwenden und dort Dinge zu verbessern. Unser Motto ist „Mathematik für eine gute Zukunft“.

Aber nicht nur bei uns, auch in vielen anderen Bereichen sehe ich Aktivitäten von Mathematikerinnen und Mathematikern, solche Brücken zu schlagen. Es gibt zwei angewandte mathematische Forschungsgesellschaften in Deutschland, die GAMM (Gesellschaft für angewandte Mathematik und Mechanik) und die GOR (Gesellschaft für Operation Research). Beide treffen sich gemeinsam mit anderen Disziplinen, diskutieren praktische Projekte und Fragestellungen und schlagen somit Brücken zur realen Welt. In der GOR beispielsweise gibt es sehr aktive Arbeitsgruppen, in denen sich Forschende mit Anwendern aus der Industrie und Wirtschaft regelmäßig treffen und sich austauschen. Das schlägt die Brückenpfeiler auf der Seite der Forschung ein. Aber auch auf Seiten der Anwender werden Brückenpfeiler gebaut: Viele Unternehmen haben sehr gut aufgestellte Forschungsabteilungen, in denen Mathematiker und Mathematikerinnen arbeiten, die auch zu Tagungen fahren und gemeinsam mit Kolleginnen an Universitäten Forschungsartikel schreiben. Ich habe das Gefühl, dass hier zurzeit viel zusammenwächst.

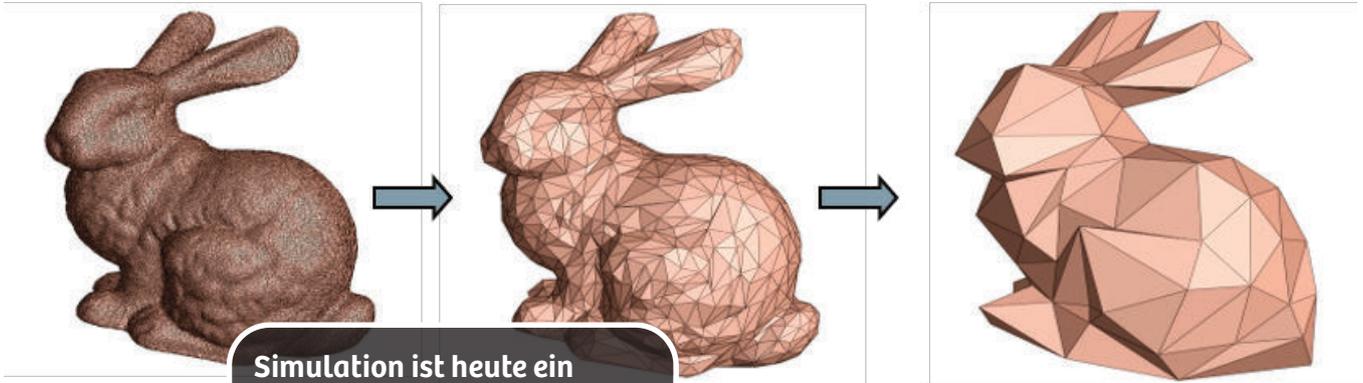
Was raten Sie jungen Leuten hinsichtlich Mathematik?

Hier habe ich gleich drei Ratschläge:

Traut euch, Mathematik zu studieren, wenn es euch Spaß macht! Mathematik ist eine Querschnittswissenschaft, die in so vielen Branchen eine Rolle spielt, dass man ganz unabhängig von eventuellen Wirtschaftsschwankungen immer einen guten Job finden wird.

Arbeitet zusammen mit Kollegen und Kolleginnen! Teamwork ist wichtig in der Mathematik, ganz besonders in der angewandten Mathematik. Gemeinsam ist man kreativer, hat mehr neue Ideen. Das ist wichtig, wenn man in der Mathematik vorankommen möchte. Schaut über den Tellerrand! Mathematik lässt sich in vielen Disziplinen anwenden, wenn man sich darauf einlässt. Gemeinsame interdisziplinäre Arbeit passt manchmal nicht in das normale Schema, ist aber spannend und wenn man etwas hinbekommt, das wirklich eingesetzt wird, ist das ein wunderbares Ergebnis.

Der Digitale Zwilling – Den Betrieb von Anlagen neu denken



Simulation ist heute ein wesentliches Werkzeug in der Entwicklung. Mathematische Innovationen erlauben es, Simulationsmodelle während des Betriebs wieder zu verwenden, und in Realzeit physikalische Vorhersagen zu treffen.

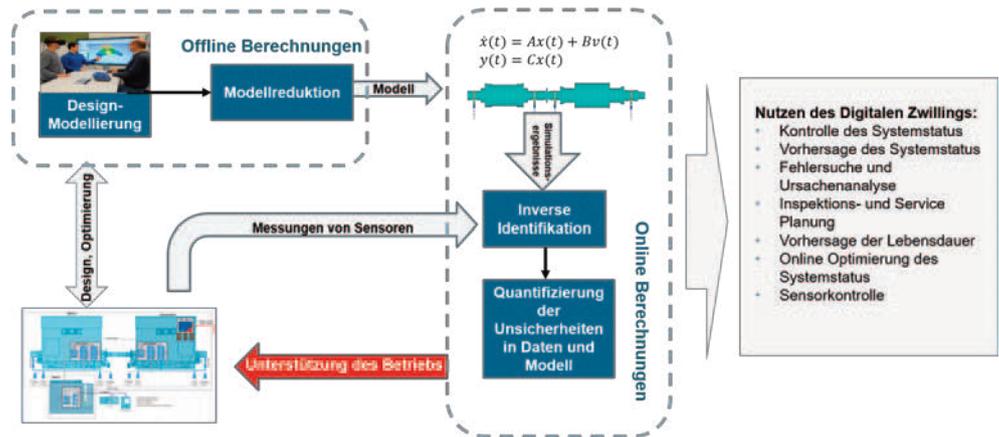
Abbildung 73: Modellreduktion am Beispiel des berühmten Stanford-Hasen (aus einem Vortrag von Wil Schilders, Eindhoven University of Technology)

Dr. Dirk Hartmann und Dr. Utz Wever von der Siemens AG, Technology, München fassen Ihre Erfahrungen zum Digital Twin zusammen:

Was ist ein Digitaler Zwilling?

Digital Zwillinge sind ein neuer Technologietrend, der bereits jetzt einen großen Einfluss verspricht. Sie sind die nächste Welle der Simulationstechnologien, die Simulation mit großen Datenmengen und Technologien der künstlichen Intelligenz verschmilzt. Digitale Zwillinge sind heute so wichtig für die Wirtschaft, dass sie zu den Top 10 der strategischen Technologietrends für 2018 gehören. Es wird erwartet, dass sie zu einer unternehmerischen Notwendigkeit werden, die den gesamten Lebenszyklus einer Anlage oder eines Prozesses abdeckt. Für viele Produkte und Dienstleistungen werden Digitale Zwillinge eine essentielle Grundlage sein. Unternehmen, die nicht reagieren, werden zurückgelassen. Beispielsweise wird prognostiziert, dass Unternehmen, die in die digitale Zwillingstechnologie investieren, die Zykluszeiten kritischer Prozesse um 30 % verbessern werden. Es wird ein Marktpotenzial von fast 50 Mrd. Dollar pro Jahr für entsprechende Angebote prognostiziert.

Abbildung 74: Digitaler Zwilling zur Unterstützung des Betriebs von Anlagen, hier am Beispiel des Antriebsstranges eines elektrischen Motors.



- Nutzen des Digitalen Zwillings:**
- Kontrolle des Systemstatus
 - Vorhersage des Systemstatus
 - Fehlersuche und Ursachenanalyse
 - Inspektions- und Service Planung
 - Vorhersage der Lebensdauer
 - Online Optimierung des Systemstatus
 - Sensorkontrolle

Digitale Zwillinge sammeln alle digitalen Kenntnisse, Modelle und Daten während der gesamten Lebensdauer von Produkten und Systemen – von der Idee bis zum Lebensende. Sie integrieren somit modellbasierte Ansätze, auf denen klassische Simulations- und Optimierungsparadigmen aufbauen, sowie datenanalytische Ansätze. Digitale Zwillinge sind leistungsstarke Wegbereiter für Innovationen und Effizienz. Sie kombinieren ingenieurwissenschaftliches Wissen, das oft in Form von ausführbaren Simulationsmodellen vorliegt, mit verfügbaren Daten und künstlicher Intelligenz. Durch diese Integration ermöglichen Digitale Zwillinge völlig neuartigen Dienstleistungen, wie simulationsbasierter Überwachung und Diagnose oder prädiktiver Wartung, und eröffnen damit neue Geschäftsmöglichkeiten.

Um etwas konkreter zu werden, konzentrieren wir uns auf ein Beispiel, siehe Abbildung 74, welches jedoch sehr oft in der Praxis vorkommt.

Was hat Mathematik mit Digitalen Zwillingen zu tun?

Mathematik ist die Sprache der Digitalen Zwillinge. Die blauen Kästen in Abbildung 74 symbolisieren die mathematischen Disziplinen, die für den Aufbau eines Digitalen Zwillings gebraucht werden. Im Folgenden sollen diese kurz beschrieben werden.

Das Modell Die Grundlage jedes Digitalen Zwillings ist ein Computermodell, welches das reale System beschreibt. Dieses Modell kann zum Beispiel durch ein Neuronales Netz repräsentiert sein, welches durch reale Daten trainiert wurde. Besser ist es, wenn das reale System durch physikalische Gleichungen beschrieben werden kann

(was nicht immer der Fall ist). Für viele Anwendungen existieren bereits ausgereifte kommerzielle Computer-Programme, die auch von Nicht-Spezialisten bedient werden können. Wesentliche Anwendungen sind die Strukturmechanik, der Elektromagnetismus oder die Strömungsmechanik. Oft verlangen die Modellierung der immer komplizierter werdenden realen Systeme ein Kopplung dieser Computer-Programme. Dies führt jedoch zu immer komplexeren und langwierigen Berechnungen.

Modellreduktion Der Kern eines Digitalen Zwillings ist ein Echtzeitfähiges Modell, welches parallel zum Betrieb auf einem Computer läuft. Die oben beschriebenen kommerziellen Computer-Programme sind dafür sicher zu langsam. Deswegen gibt es mathematische Verfahren, die aus einem komplexen Modell ein einfaches Modell generieren. Die wesentliche Eigenschaft des einfachen Modells ist es, bestimmte physikalische Eigenschaften des komplexen Modells zu erhalten. Diesen Eigenschaft kann man gut an dem bekannten *Stanford-Hasen* (Abbildung 73) beobachten. Das rechte Modell wird mit wesentlich weniger Freiheitsgraden beschrieben, ist aber immer noch als Hase zu erkennen.

Inverse Identifikation Die Simulationsergebnisse des reduzierten Modells müssen mit den Sensordaten des realen Systems verglichen werden, was mathematisch zu einem inversen Problem führt. Durch diesen Vergleich wird erstens das mitlaufende Modell fortlaufend synchronisiert und zweitens können durch Abweichungen vielen zusätzliche Information über das reale System gewonnen werden. Ein paar Beispiele sind in Abbildung 74 in der rechten Box aufgezählt.

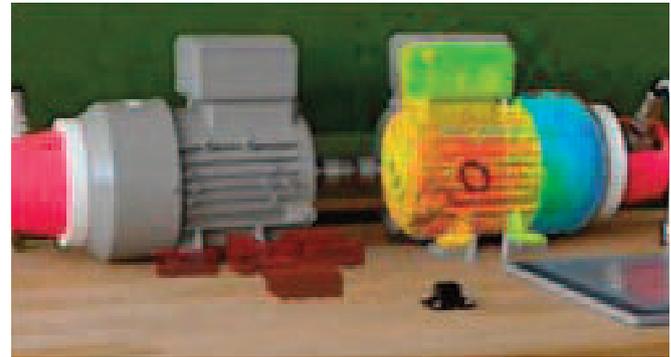
Quantifizierung von Unsicherheiten Alle durch Sensoren gewonnenen Daten und auch das reduzierte Modell sind mit Unsicherheiten behaftet. Der dadurch entstandene Fehler kann nicht reduziert werden (nur unter hohen Kosten), wohl aber quantifiziert. Eine Aussage oder Ergebnis des Digitalen Zwillings kann also immer nur mit einer gewissen Wahrscheinlichkeit getroffen werden. Wird zum Beispiel eine Vorhersage des Systemstatus gewünscht, so kann diese nur innerhalb eines Vertrauensintervalls bereit gestellt werden. Dieses Phänomen kennt man gut von der Wetterkarte.

Beispiel: Ein Digitaler Zwilling für einen Elektromotor

Als Anwendungsfall wird die effiziente thermische Regelung eines asynchronen Elektromotors betrachtet.⁴⁸ Elektromotoren unterliegen einem thermischen Materialverschleiss, welcher eine kontrollierte Temperaturführung erfordert. So werden z.B. große elektrische Asynchronantriebe beim Anfahren durch die Induktionserwärmung stark belastet. Häufige Starts ohne ausreichende Kühlpausen können zu einer Überhitzung der Motoren führen. Es ist jedoch nahezu unmöglich, die Temperatur der tatsächlichen Rotoren zu messen, die sich mit hoher Geschwindigkeit im Motor drehen. Daher basieren die Regelungen oft auf sehr konservativen Heuristiken. Prinzipiell können entsprechende Temperaturen mit Hilfe von 3D-Wärmesimulationen sehr genau berechnet werden. Im Fall der betrachteten Elektromotoren werden lineare Konvektions-Diffusionsmodelle verwendet. Entsprechende 3D-Modelle sind typischerweise aus der Detailkonstruktion verfügbar. Sie sind jedoch rechnerisch zu anspruchsvoll, um im Betrieb eingesetzt zu werden. Durch Modellreduktion können jedoch entsprechende Modelle für die Hintergrundsimulation realisiert werden.

Diese sind sehr effizient auszuwerten und lassen sich nicht nur mit Methoden der Unsicherheitsquantifizierung erweitern, sondern unterstützen auch die kontinuierliche Kalibrierung mit vorhandenen Sensoren auf der Statorseite. So können mit Hilfe von kontinuierlich kalibrierten Hintergrundsimulationsmodellen Temperaturen, die für Sensoren nicht zugänglich sind, virtuell gemessen werden. Mit Hilfe von z.B. Augmented-Reality-Geräten kann der Anwender jederzeit virtuell in das Innere des Motors schauen und Temperaturverteilungen beobachten, vgl. Abbildung 75 (a). Dadurch lassen sich die Kühlzeiten von Elektromotoren deutlich reduzieren, was letztlich die Anlagenverfügbarkeit erhöht. Angereichert mit Methoden zur Unsicherheitsquantifizierung können Konfidenzintervalle für die Rotortemperatur bereitgestellt werden (vgl. Abbildung 75 (b)), die es erlauben, nahe an Betriebsgrenzen zu gehen.

(a)



(b)

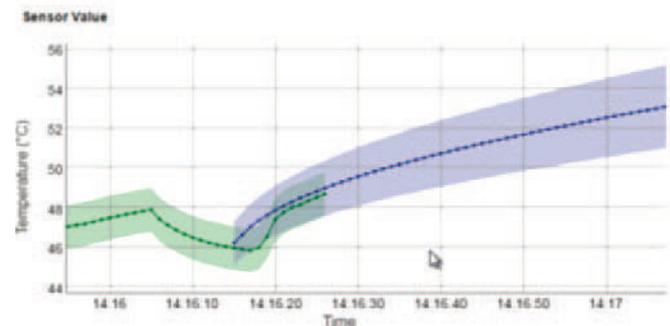


Abbildung 75: (a) Temperaturvisualisierung im Inneren eines Elektromotors mittels Augmented Reality unter Verwendung der Microsoft HoloLens. Screenshot der virtuellen Temperatursensordlösung zur Schätzung der aktuellen Rotortemperatur (grün) sowie zur Vorhersage der Rotortemperatur nach dem Start (blau) an einem bestimmten Punkt. Die eingezeichneten Bänder zeigen die berechneten Unsicherheiten, siehe auch www.youtube.com/watch?v=86vkjykbHRM

⁴⁸ D. Hartmann, M. Herz, M., U. Wever, Model Order Reduction – A Key Technology for Digital Twins, KoMSO Challenge Workshop, Springer (2017).



Mit Mathematik die die Finanz- märkte zähmen

Ein Finanzmarkt ist wie ein lebender Organismus, welcher sich kontinuierlich verändert, nicht vorhersehbar ist, und dennoch dank der Stochastik gezähmt werden kann.

Herr Dr. Martin Friesen von der School of Mathematical Sciences, Dublin City University, Irland, beschreibt seine Erfahrungen zur Finanzmathematik:

Standen noch zu Jahresbeginn 2020 die Kurse an großen Finanzplätzen der Welt auf Rekordniveau, drehte sich bei den Anlegern schlagartig die Stimmung als die Corona Pandemie die Welt in den Würgegriff nahm und unseren Alltag auf den Kopf stellte. Durch coronabedingte Lockdowns wurde unsere Wirtschaft zeitweise lahmgelegt und als Konsequenz kam es im Februar/März 2020 zu einem weltweiten Kurzschluß an den Finanzmärkten. Zwar konnten durch ein, unter anderem, schnelles und beherztes Eingreifen der Zentralbanken die Finanzmärkte wieder stabilisiert werden,⁴⁹ so führt uns aber der *Corona Crash 2020* eindrucksvoll vor Augen, dass nicht vorhersehbare Ereignisse jederzeit auftreten können und mit nicht zu vernachlässigbaren Risiken verbunden sind. Es ist daher unumgänglich, dass sich alle großen und teilweise auch systemrelevanten Akteure an den Finanzmärkten (wie z.B. Banken, Versicherer aber auch Staaten) gegen eine Vielzahl möglicher Risiken absichern.

Um solchen Unsicherheiten zu begegnen, bedarf es der Stochastik als Teilgebiet der Mathematik. Diese ermöglicht es z.B mittels der Wahrscheinlichkeitsrechnung sowie Statistik fortwährend mögliche Zukunftsszenarien anhand von vergangenen Ereignissen zu analysieren und mittels mathematischen Modellen Prognosen für die Zukunft zu erstellen. Das derzeit größte, wenn auch nicht einzige, System zur vollautomatisierten Risikoanalyse heißt Aladdin⁵⁰ und wird vom New Yorker Finanzdienstleister Black Rock betrieben⁵¹. Neben großen Weltkonzernen haben diese Dienstleister auch die europäische Staatengemeinschaft als Kunden.

Verstehen von Risiken

Aber warum sollten Akteure überhaupt an den Finanzmärkte agieren, wenn diese mit Risiken verbunden sind, welche erst mit komplexen mathematischen Modellen gezähmt werden müssen? Diese Frage lässt sich am einfachsten mit dem Sprichwort „*Das Geld muss arbeiten*“ beantworten. Es ist also die Möglichkeit durch Investitionen am Erfolg von Unternehmen teilhaben zu können (z.B. durch Dividendenzahlungen oder steigende Aktienkurse). Der genaue Zusammenhang von Gewinn und Risiko wird dabei vom

Hauptsatz der Finanzmathematik beschrieben und kann mit dem Sprichwort „There is no free lunch“ abgekürzt werden. Ohne Berücksichtigung von Zinsen, kann dieser Zusammenhang in einer einzigen mathematischen Formel wie folgt ausgedrückt werden:

$$E_u[R_{t,u}] = 0, \quad 0 < u < t. \quad (1)$$

Konkret bedeutet es, dass der durchschnittlich erzielte Gewinn oder Verlust zum Zeitpunkt t , sofern man ein Wertpapier zum Zeitpunkt u gekauft hat, stets Null ist. Am einfachsten lässt sich diese Formel mit einem einfachen Münzwurf als Zufallsexperiment verstehen. Wird Kopf geworfen, so steigt der Aktienkurs, wird hingegen Zahl geworfen, so fällt der Kurs. Jeder einzelne Münzwurf ist hierbei unabhängig von den jeweils Anderen (man sagt *die Münze hat kein Gedächtnis*). Da wir im Mittelwert erwarten, dass bei einer fairen Münze Kopf und Zahl gleich oft vorkommen, ist das durchschnittliche Ergebnis Null. Dieses sagt jedoch nichts über den Ausgang einer einzigen Investition aus (ebenso wie man nicht einen einzigen Münzwurf vorhersagen kann). So gibt es durchaus die Möglichkeit einen höheren Gewinn zu erzielen, im Gegenzug muss man jedoch auch bereit sein mögliche Verluste zu akzeptieren. Also kurz gesagt: *Es gibt keine totalsichere Wette!*

Computerhandel: Viele Chancen und neue Risiken

War vor rund 50 Jahren ein Finanzmarkt noch ein rein physischer Marktplatz wo Käufe und Verkäufe von realen Personen getätigt wurden, so hat sich dieser spätestens seit den 2000er Jahren grundlegend gewandelt. Heutzutage sind die Finanzmärkte durch und durch digitalisiert, automatisiert und globalisiert. Tatsächlich wird bereits die überwältigende Mehrheit aller Kaufentscheidungen von ausgeklügelten Algorithmen getroffen, welche in Bruchteilen von Sekunden agieren und reagieren.⁵² Solche computerbetriebenen Systeme basieren zwar zum einen auf mathematischen Modellen, sind jedoch auch zunehmend ausgestattet mit künstlicher Intelligenz, welche sich, so die Erwartungen, automatisch an neue Situationen anpasst und stets die rational beste Entscheidung trifft.

Eine Konsequenz des modernen „Computerhandels“ ist, dass Kursverläufe in kurzen Zeitperioden sehr hohe Ausschläge aufweisen. In der Finanzmathematik spricht man von Volatilität.

Aktuelle Forschungsarbeiten der Mathematik beschäftigen sich daher auch mit der Frage wie solche Effekte möglichst präzise erfasst werden können. Als überaus effektiv hat sich der Ansatz mit soge-

nannten „*Rough Volatility models*“ herausgestellt. Eine wichtige Eigenart dieser Modelle ist die Zunahme früherer Kursschwankungen frei nach dem Motto „starke Ausschläge führen kurzfristig wieder zu starken Ausschlägen,“. Im Hinblick auf den zuvor beschriebenen Münzwurf, würde hier die Münze zusätzlich eine Art „Gedächtnis“ zugewiesen bekommen, sodass man für mögliche Vorhersagen auch stets die vergangenen Kursausschläge berücksichtigen muss. Konkret können bereits heute vorhandene Resultate dazu verwendet werden *faire Preise* für Absicherungen gegen Kursverluste zu errechnen. Auf diese Weise können die Akteure an den Finanzmärkten ihre Risiken zwar nicht gänzlich eliminieren, jedoch zumindest kontrollieren und somit *die unvorhersehbaren Finanzmärkte zähmen*.

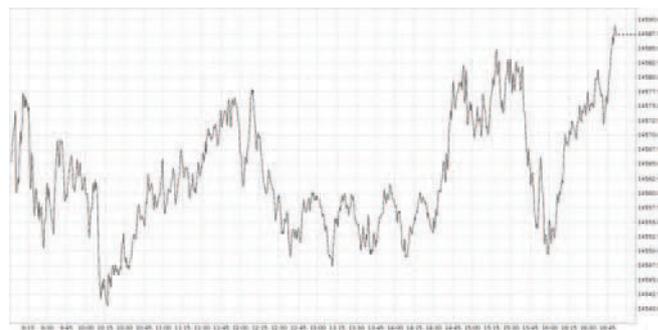


Abbildung 76: DAX 17. März 2021

⁴⁹ www.bundesbank.de/de/aufgaben/geldpolitik/geldpolitische-wertpapierankauefe/pandemic-emergency-purchase-programme-pepp-830356
www.faz.net/aktuell/finanzen/finanzmarkt/ezb-stockt-krisenprogramm-pepp-um-500-milliarden-euro-auf-17095348.html

⁵⁰ www.it-times.de/news/aladdin-der-super-algorithmus-von-blackrock-135395/

⁵¹ www.faz.net/aktuell/finanzen/fonds-mehr/vermoegensverwalter-blackrock-der-schwarze-riese-12057048.html

⁵² www.heise.de/newsticker/meldung/Hibernia-Express-Erstes-neues-Transatlantik-Kabel-in-12-Jahren-2778847.html



Modellierung Negativer Zinsen – Nur Bares ist Wahres?

Mathematische Modelle helfen aktuelle Entwicklungen an den Finanzmärkten zu modellieren und können helfen mittels ökonomischen Modellen Erkenntnisse zu gewinnen.

Jörg Kienitz von Acadiasoft Inc., ist Privatdozent an der Bergischen Universität Wuppertal, Adjunct Associate Professor an der Universität Kapstadt und Betreiber der Seite finciraptor.de. Er beschreibt am Beispiel negativer Zinsen wie Mathematik hilft, aktuelle Finanzmarktphänomene zu modellieren.

Senkung des Leitzinses

Negative Zinsen sind uns allen mittlerweile wohlbekannt. Letztlich erfahren wir dies, wenn wir Zinsen für Tagesgeld vergleichen oder über „Verwahrtgelte“ für Einlagen lesen. Wie kam es zu dieser Situation? Im Zuge der Finanzkrise 2007 musste die Europäische Zentralbank die Stabilität in der Eurozone erhalten. Dabei setzte sie auf die

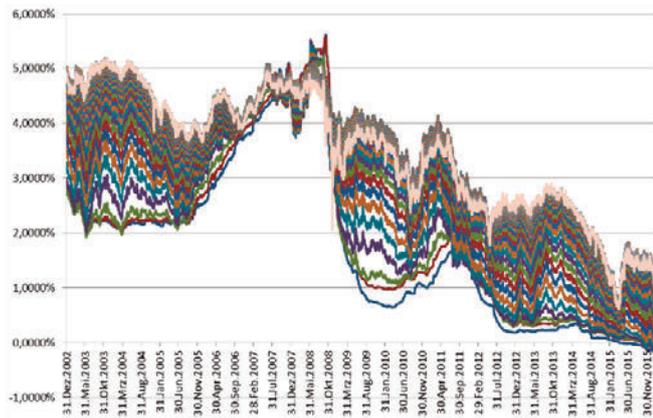


Abbildung 77: Entwicklung unterschiedlicher Zinsraten (Swapraten), EUR im Zeitraum von 31.12.2002 bis 10.11.2015.

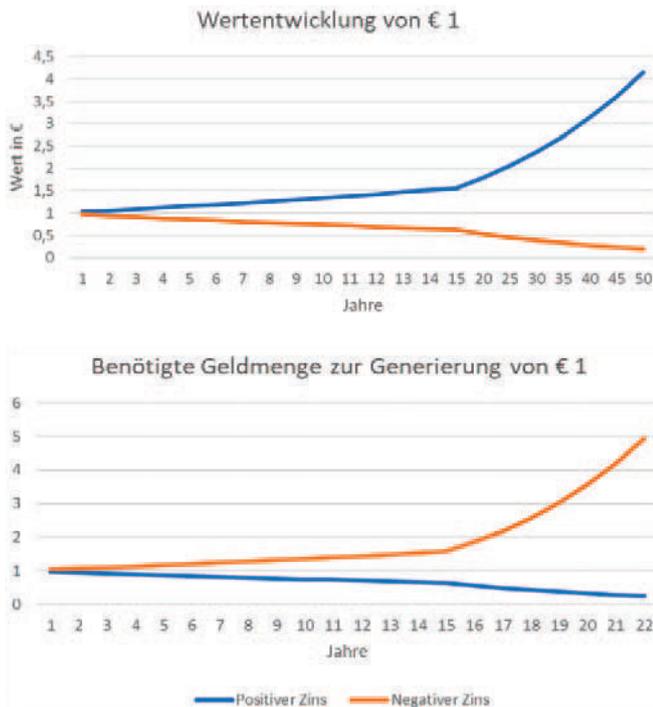


Abbildung 78: oben: Zinskurve für eine konstante Zinsrate von 3 % (blau) und -3 % (orange); unten: Diskontierungsfaktor für eine konstante Zinsrate von 3 % (blau) und -3 % (orange).

Erhöhung der Geldbasis durch den Kauf von Staatsanleihen („Quantitative Easing“) und die Senkung des Leitzinses. Letzterer besagt, zu welchem Zinssatz sich Geschäftsbanken Geld leihen können. Die eingeleitete schrittweise Senkung zeigte allerdings nicht die erhoffte Wirkung, was zu weiteren Senkungen führte und schließlich dazu, dass dieser Zins negativ wurde. Derzeit⁵³ steht er bei -0,5 %.

Länder wie die Schweiz, Schweden, Dänemark und Japan haben negative Zinsen bereits angewendet. Im Falle der Schweiz resultierte dieser Schritt aus starken Geldzuflüssen, die mangels Alternativen in schweizerische Zinsanleihen investiert wurden. Die Auswirkungen für den Export oder die auch die Verteuerung des Konsums waren erheblich. Dem wirkte die schweizerische Nationalbank entgegen, indem sie den Leitzins unter 0 % senkte.

Modellierung von Zinssätzen

In der mathematischen Modellierung betrachten wir Zinssätze, welche an eine Laufzeit geknüpft sind. Am Markt notieren Zinssätze zu unterschiedlichen Laufzeiten, siehe Abbildung 77. Ein Zinssatz beschreibt die Entwicklung einer Geldmenge, wenn diese für die entsprechende Laufzeit angelegt wird. Aus solchen Zinssätzen werden Zinskurven konstruiert. So sind in Abbildung 78 Zinskurven für konstante Zinsraten von 3 % und -3 % dargestellt. Außerdem wird die sogenannte Diskontkurve gezeigt. Während man aus der Zinskurve ablesen kann wie sich 1 € für unterschiedliche Laufzeiten entwickelt, so lässt sich aus der Diskontkurve ablesen welche Geldmenge notwendig ist um 1 € zu einer Laufzeit zu erhalten, siehe Abbildung 78. Für positive Zinsen ist der Wert kleiner als 1 für negative jedoch größer.

Für die Bestimmung des Wertes eines Zinsgeschäfts unterscheidet man Produkte, die einzig und allein durch die aktuelle Zinskurve bestimmt werden. So kann der Wert eines Instruments, welches jedes Jahr, z.B. am 11.06., einen Zinssatz von 3 % und nach 10 Jahren außerdem den Nennwert zahlt, z.B. 100 €, den Wert zu errechnen. Hierzu liest man in der Diskontkurve den entsprechenden Wert ab, multipliziert diesen mit dem Zins (Nennwert 3 % = 3 €) und addiert diese Werte auf. Final wird der diskontierte Rückzahlungswert dazu gerechnet. Für die Diskontkurven aus Abbildung 78 ergeben sich genau 100 €, während im Fall mit negativen Zinsen 171,37 € berechnet werden. Solche Instrumente nennt man auch linear. Darüber hinaus gibt es auch nicht-lineare Instrumente wozu unter anderem Zinsbegrenzer gehören. Die Auszahlungen solcher Instrumente sind

nicht-lineare Funktionen eines Zinssatzes, z.B. das Maximum oder Minimum bzgl. eines Begrenzungswertes.

Es ist nicht möglich – auch nicht im linearen Fall – diesen Wert für alle zukünftigen Zeitpunkte zu berechnen. Die Zinskurve ist ein dynamisches Objekt, welches durch Handel an Geld- und Kapitalmärkten sowie Entscheidungen der Geschäfts- und Zentralbanken oder Investitionen von Pensionskassen und Fonds getrieben wird. Dies wird in Abbildung 77 gezeigt, wo Zinssätze für verschiedene Laufzeiten in der Zeit vom 31.12.2002 bis 10.11.2015 gezeigt werden. Sowohl die Höhe als auch das Verhältnis der Zinsraten zueinander schwankt.

Stochastische Zinsmodelle

Daher werden in der Finanzmathematik dynamische, stochastische Modelle verwendet, die Entwicklungen von Zinsraten beschreiben. Preise von Zinsprodukten werden dann als Erwartungswerte bzgl. von Wahrscheinlichkeitsverteilungen berechnet. Für Zinsprodukte wird gewährleistet, dass die modellierten Zinsraten genau getroffen werden. Jetzt sehen wir auch warum Produkte, die nur die aktuellen Zinsraten benötigen, linear heißen. Dies ergibt sich aus der Linearität des Erwartungswertes. Für nicht-lineare Funktionen ist die Berechnung nicht so einfach. Hierfür wird die Kenntnis der Verteilung benötigt um den Wert bestimmen zu können. Hier spielen dann auch weitere Parameter wie Volatilität, Schiefe oder Wölbung eine Rolle, die die Verteilung bestimmen.

Diese stochastische Entwicklung der Zinsraten ist sehr kompliziert. Es ist möglich eine Zinsrate, mehrere Zinsraten oder die gesamte Zinskurve, also unendlich viele Zinsraten gleichzeitig zu modellieren. Im einfachsten Fall modellieren wir die stochastische Entwicklung eines stochastischen Faktors und betrachten die Kurve als Funktion dieses Faktors. Die Entwicklung wird durch einen deterministischen Teil, den sog. Drift und einen stochastischen Teil beschrieben. Letzterer wird häufig als Brownsche Bewegung angenommen. In diesem Fall spricht man beim stochastischen Teil von der Diffusion.

Die Dynamik kann durch geschickte Wahl der Diffusionskomponente in verschiedener Weise gesteuert werden⁵⁴. Zum Beispiel kann diese Komponente als Multiplikation mit einer konstante und der zufälligen Realisierung der Brownschen Bewegung aber einer Funktion des aktuellen Zustands und der Zeit modelliert werden. Möchte man eine Normalverteilung nutzen, so wählt man den ersten Ansatz. Man sagt auch die stochastische Bewegung reagiert absolut auf die

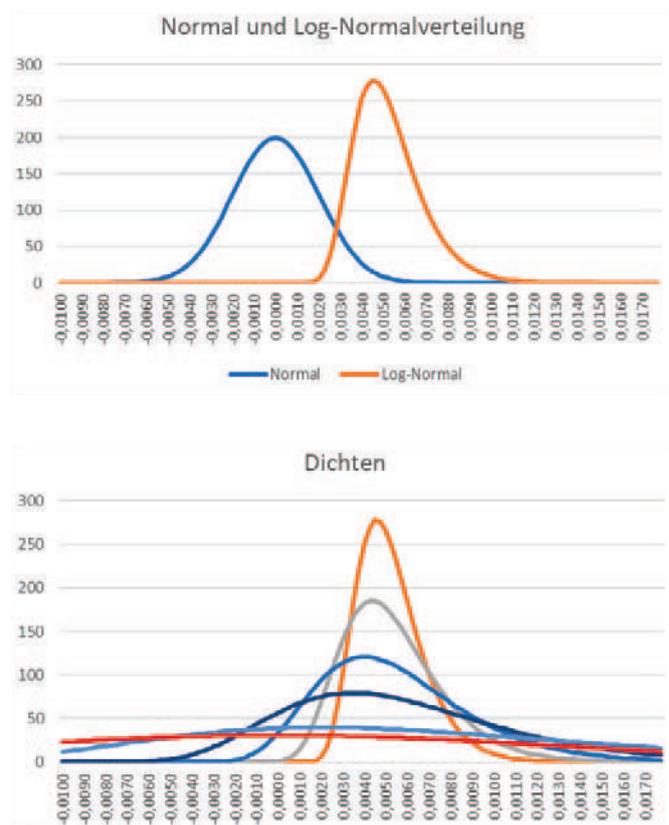


Abbildung 79: links: Normale und Logarithmisch-Normal Verteilung; rechts: Verteilung zu Logarithmisch-Normalen (orange), Vershobenen Diffusionen mit verschiedenen Shiftparametern und Normaler Verteilung (rot).

zufällige Bewegung der treibenden Brownschen Bewegung. Wählt man hingegen den zweiten Ansatz so erhält man eine Logarithmisch-Normale Verteilung, siehe auch Abbildung 79 (linke Seite). Interessant ist, dass diese Wahl auch den Definitionsbereich der stochastischen Dynamik einschränkt. Im ersten Fall kann jede reelle Zahl realisiert werden. Wählt man die zweite Möglichkeit, so ist das nicht mehr möglich und es können nur positive reelle Zahlen realisiert werden.

Negative Zinsen

Wie können wir nun den Bogen zu der Modellierung von Zinsraten schlagen? Zu Beginn deren Modellierung wurden einfache Modelle basierend auf Normalverteilungen genutzt. Zu dieser Zeit aber waren negative Zinsen ein Artefakt und resultierten aus der Einfachheit der Modelle. Jedoch achtete man darauf möglichst Modellparameter zu verwenden, die nur eine sehr geringe Wahrscheinlichkeit für die Realisierung negativer Zinsen erlaubten. Später wurden diese Modelle zur Vermeidung negativer Zinsen weiterentwickelt. Gleichzeitig wurden mehr, bzw. unendlich dimensionale Modelle betrachtet, um Verhältnisse von Zinsraten zueinander abzubilden. Das Augenmerk lag vielmehr darin andere Marktgegebenheiten abzubilden, z.B. den sogenannten Skew. Dies beschreibt den Umstand, dass zur Berechnung nicht-linearer Produkte die Eigenschaften wie Schiefe oder Wölbung mit betrachtet werden müssen, siehe Abbildung 79 (rechte Seite). Hierzu bot es sich an Mischungen von Normalverteilung und Logarithmischen-Normalverteilung zu nutzen. Dies führt zum Beispiel zu sogenannten verschobenen Diffusionen. Diese Hilfsmittel erlaubten eine bessere Modellierung des Skew, resultierte aber wieder in der Möglichkeit der Realisierung negativer Zinsraten.

In der aktuellen Marktsituation greift man wieder auf solche Modelle zurück, die explizit negative Zinsen zulassen. Sicherlich kann der Zinssatz nicht beliebig kleine Werte annehmen, was die Verwendung der Normalverteilung fraglich erscheinen lässt. Letztlich kann man sich allerdings wieder mit der Wahl der Parameter aus der Schlinge ziehen und versuchen, Wahrscheinlichkeiten unterhalb eines Schwellenwertes sehr klein werden zu lassen. Die verschobenen Diffusionen erlauben hier eine explizite, direkte Modellierung des Schwellenwertes. In diesen Fällen können negative Werte bis hin zum Schwellenwert modelliert werden, siehe Abbildung 79 (rechte Seite).

Es sei bemerkt, dass die verwendeten stochastischen Modelle und numerischen Methoden unproblematisch hinsichtlich negativer Zinsen sind. Es ergeben sich aber erhebliche Probleme bei der ökonomischen Modellbildung. Zum Beispiel sei hier der Schwellenwert genannt. Eine theoretische untere Grenze kann durch die Existenz von Bargeld bestimmt werden, die dann von Abhebegebühren und Lagerkosten abhängt. Neben dieser physischen Untergrenze versucht man auch, mittels komplexerer Modelle ökonomisch plausible Werte zu etablieren. Für den Bargeldfall nimmt man Abhebegebühren von 2 % und Lagerungskosten von 1 % pro Jahr an, beträgt die untere Grenze $-2,98 \%$ pro Jahr, $-(100 \% - 2 \%) (100 \% - 1 \%) = -0,98 \cdot 0,99 = -2,98 \%$ pro Jahr.

Ökonomische Schwierigkeiten und Implikationen negativer Zinsen können modellhaft mathematisch betrachtet werden und (komplexe) numerische und stochastische Methoden stehen zur Verfügung, um diese Modelle umzusetzen. Allerdings beruhen diese Modelle unter anderem auf ökonomischen und politischen Annahmen, die natürlich nicht mathematisch verifiziert werden können. Hier ist ein enges Zusammenspiel der unterschiedlichen Disziplinen notwendig. Da die Thematik Auswirkungen auf Gesellschaften und schließlich Menschen hat sollte mit großer Sorgfalt und offener Kommunikation und Darstellung von Annahmen umgegangen werden. Letztlich muss jeder die im Titel gestellte Frage selbst beantworten. Allerdings kann mittels des Bargeldes eine untere Schranke für negative Zinsraten etabliert werden, die ohne „Bares“ so nicht hält.

⁵³ Stand 14. April 2021

⁵⁴ Der Drift wird aus finanzmathematischen Gründen durch die Diffusionskomponente und die aktuelle Zinskurve bestimmt.

Beipackzettel für Finanzinstrumente: PRIIPs!

Wie kann man die Ertragsaussichten verschiedener Finanzinstrumente vergleichen? In einer idealen Welt für InvestorInnen nimmt das Vermögen stetig, ohne große Aufregungen, zu. Tatsächlich sind verschiedene Anlageinstrumente aber mit höchst unterschiedlichen Laufzeiten, Risiko- und Ertragsprofilen ausgestattet und daher *ex ante*, also vor einer Anlageentscheidung, sehr schwierig zu vergleichen.

Für die sogenannten *Packaged Retail and Insurance-based Investment Products (PRIIPs)* schreibt die Europäische Verordnung (EU) Nr. 1286/2014 vor, dass jedes solche PRIIP mit einem Basisinformationsblatt (englisch: Key Information Document, KID) auszustatten ist, in dem die Herstellerin (im Regelfall eine Bank, Versicherung oder Kapitalanlagegesellschaft) das Risiko, den möglichen Ertrag und die Kosten in Verbindung mit dem Produkt in einer standardisierten Weise auf höchstens drei A4-Seiten quantifiziert. So sollen unterschiedliche PRIIPs vergleichbar gemacht werden.

Was sind Beispiele von PRIIPs? Was sind Beispiele, die keine PRIIPs sind? Die Schlüsselworte sind „Packaged“ und „Retail“: Wenn ein (beispielsweise: Floating to Fix-) Swap-Geschäft⁵⁵ mit einer institu-

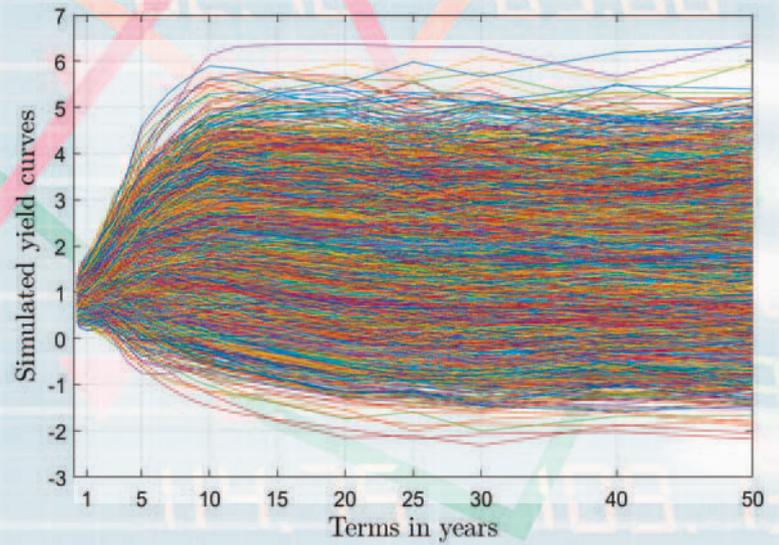


Abbildung 80: Für das Jahr 2031 (zehn Jahre ab jetzt) ergeben sich die 10.000 simulierten Zinskurven der Abbildung. Jede einzelne Kurve zeigt dabei auf der y-Achse den Zinssatz für die Laufzeit (in Jahren) auf der x-Achse.

tionellen Investorin abgeschlossen wird, dann ist es kein PRIIP, wenn ein Swap mit genau den gleichen Bedingungen an eine Endkundin (Retail) verkauft wird, dann schon.

Ebenso: Wenn ein Instrument nicht verpackt ist, ist es kein PRIIP. Eine Einzelaktie von Apple ist kein PRIIP, eine deutsche Bundesanleihe ist kein PRIIP, eine Euribor verzinste Emission der Commerzbank ist kein PRIIP.

Im Gegensatz dazu (immer unter der Annahme der Retail-Kundin): Ein Exchange Traded Fund auf den EuroStoxx50 ist ein PRIIP (verpackt), eine kündbare Fixzinsanleihe der Credit Agricole ist ein PRIIP (Anleihe verpackt mit Kündigungsrecht), ein Floater mit Cap und Floor der Hessischen Landesbank ist ein PRIIP (Floater verpackt mit Caplets und Floorlets).

Ebenso sind PRIIPs (exemplarische, in keiner Weise vollständige Aufzählung): Swaps, Optionen, Termingeschäfte, Wertpapierfonds, Wandelanleihen, strukturierte Aktienanleihen, strukturierte Zinsinstrumente, fondsgebundene Lebensversicherungen.

Wo versteckt sich die Mathematik?

Ergänzend zur PRIIPS-Verordnung, schreibt die Delegierte Verordnung (EU) 2017/653 vor, wie die Basisinformationsblätter genau auszugestalten sind und welche Berechnungen dafür anzustellen sind. Das Aufwändige dabei sind für den Großteil der PRIIPs die Berechnungen, die für die Risikoquantifizierung und die Performance-Szenarien (Ertragsaussichten) anzustellen sind. Solche Bewertungen berechnen den fairen Wert eines Finanzinstrumentes auf der Basis der aktuell vorliegenden Marktdaten, die aber für PRIIPs, die noch nicht abgelaufen sind, in der Zukunft liegen. Im Mathematik-Studium werden solche Bewertungsverfahren im Themenbereich „Finanzmathematik“ behandelt.

Andreas Binder von UnRisk arbeitet mit seinem Team von Mathematikerinnen, Physikerinnen, Informatikerinnen seit mehr als 20 Jahren an mathematischer Software für die Bewertung und Risikoanalyse von Finanzinstrumenten. Zu PRIIPs merkt er an: „Die Idee, dass verschiedene Instrumente vergleichbar gemacht werden, ist aus meiner Sicht ein richtiger und wichtiger Schritt in Richtung Anlegerschutz und Kostentransparenz. Im Detail kann aber die Erstellung eines Basisinformationsblattes enorm aufwändig sein. Für langlaufende (> 5 Jahre) strukturierte Zinsinstrumente müssen gemäß der

Delegierten Verordnung 70.000 Bewertungen durchgeführt werden. Dabei bedeutet Bewertung das Lösen einer partiellen oder einer stochastischen Differentialgleichung mit möglicherweise Zehntausenden Unbekannten auf dem Zeit-Zins-Gitter. Wenn eine einzelne Bewertung nur eine Sekunde dauert, reden wir von Rechenzeiten von einem Tag auf der Single-CPU.“

Gemeinsam mit der TU Berlin forscht das österreichische Unternehmen in einem Teilprojekt von ROMSOC (www.romsoc.eu) an Modellreduktionstechniken für die Risikoanalyse.

⁵⁵ Ein solcher Swap tauscht für einen festgesetzten Nominalbetrag fixe und variable Zinszahlungen. So kann beispielsweise für eine Hausfinanzierung der Zinssatz eines Kredits von fix in variabel (oder umgekehrt) umgewandelt werden, ohne dass der Basiskredit mit möglicherweise hohen Kosten gekündigt werden muss.

Prof. Dr. Helga Rübsamen-Schaeff



Eine kurze Zusammenfassung Ihrer Karriere

Die promovierte Chemikerin leitete zunächst das Chemotherapeutische Forschungsinstitut Georg-Speyer-Haus in Frankfurt und baute es zu einer erfolgreichen akademischen Einrichtung der medizinischen Grundlagenforschung auf (1987-1993, Schwerpunkte: HIV, Onkologie). 1994-2006 zunächst Leiterin der Virusforschung der Bayer AG, ab 2001 Leitung der gesamten Infektionsforschung (Medikamente gegen Viren und multiresistente Bakterien). 2006 Gründung der Firma AiCuris, Geschäftsführung bis 2015, Weiterführung der Arbeiten an Medikamenten gegen Infektionskrankheiten. 2017 die Zulassung des ersten Medikaments, Previmis, - ein außergewöhnlicher Erfolg für eine junge Biopharma-Firma. Prof. Rübsamen-Schaeff ist aktuell Aufsichtsrätin der AiCuris, Mitglied des Aufsichtsrats und des Gesellschafterrats der Merck KGaA bzw. E. Merck KG, ferner Aufsichtsrätin der 4SC.

Sie erhielt von der IUPAC (International Union of Pure and Applied Chemistry) den Titel „Distinguished Woman in Chemistry and Chemical Engineering“, die Ehren-Mitgliedschaft der Gesellschaft Deutscher Chemiker, sie ist Mitglied der Nationalen Akademie der Wissenschaft Leopoldina und wurde 2018 mit dem Deutschen Zukunftspreis, verliehen durch den Bundespräsidenten der Bundesrepublik Deutschland ausgezeichnet. 2019 erhielt sie den Innovationspreis des Landes Nordrhein-Westfalen. Im März 2020 wurde sie mit der Löffler-Frosch Medaille der Gesellschaft für Virologie ausgezeichnet.

Was hat Mathematik für Ihre Karriere bedeutet?

In der Medikamenten-Entwicklung muß man viel modellieren, z.B. das Anfluten und den Abbau von Spiegeln der Medikamente im Körper, man muß aber bei der Erfindung neuer Medikamente auch Computer-Programme benutzen die berechnen, wie genau ein Wirkstoff in die zu hemmende Struktur passt, bzw. wie man ihn optimieren kann. Und schließlich braucht man auch eine Menge Statistik, um Forschungsergebnisse, aber auch die Ergebnisse klinischer Studien zu verifizieren, um nur einige Anwendungen zu nennen.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihre Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Wie bereits ausgeführt, spielt die Mathematik in allen Schritten, die zu einem neuen Medikament führen, eine herausragende Rolle.

Wie sieht ein typischer Arbeitstag aus?

Nachdem ich nicht mehr im Labor stehe, besteht ein Großteil der Arbeit (neben den klassischen Management-Aufgaben) in der Analyse der Daten, entweder solcher, die von anderen publiziert werden oder der eigenen.

Sind, aus Ihrer Sicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Ich fürchte, dass dies nicht der Fall ist, wenngleich die Mathematik für so viele Bereiche des Lebens, der Wirtschaft und der Forschung von grundlegender Bedeutung ist.

Was raten Sie jungen Leuten hinsichtlich Mathematik?

Dasselbe, das ich für die Chemie rate: „Lassen Sie sich auf das Studium ein. Chemie ist eine solch wichtige Grundlagen-Wissenschaft, von der aus Sie in die verschiedensten Anwendungen gehen können“. Ebenso ist Mathematik eine solch wichtige Grundlage zur Beschreibung aller Prozesse, dass auch nach dem Studium der Mathematik für das spätere berufliche Leben die unterschiedlichsten Karrieren und Spezialisierungen möglich sind.

Unbemannte Flugzeuge verwenden Kamerabilder zur Navigation. Aber in diesen Bildern gibt es alle Arten von Verzerrungen: Winkel ändern sich, Objekte verformen sich. Es braucht eine Menge Tricks, um zu berechnen, wie die Realität aussieht.

Berechnung von Kamerabildern

Kamerabilder finden an den unterschiedlichsten Orten Anwendung. Jaap van de Loosdrecht und Klaas Dijkstra vom Computer Vision Knowledge Centre (NHL University of Applied Sciences) arbeiten an unbemannten Flugzeugen, zum Beispiel an einem Projekt zur Überwachung der Bewegungen von Kühen in einem Stall. In diesem Stall befinden sich mehrere Kameras, und es ist beabsichtigt, auf einer Karte anzugeben, wo jede Kuh zu welchem Zeitpunkt war. Das ist schwieriger, als es scheint. Die Kameras sind alle in einem anderen Winkel aufgehängt, und eine Kuh kann auf dem einen Bild viel größer aussehen als auf dem anderen. Die verwendeten Weitwinkelobjektive ergeben zudem eine tonnenförmige Verzerrung: alles im Bild scheint an den Rand gezogen zu sein. Alles in allem sind die Bilder auf alle möglichen Arten verzerrt.

Um diese Verzerrung abzubilden, erstellen die Forscher vor der Kamera ein Muster, von dem alle Entfernungen bekannt sind. Sie verwenden ein Muster mit zehn Reihen zu je vierzehn Punkten. Alle Punkte sind schwarz, mit Ausnahme der ersten beiden in der oberen Reihe, so dass Sie immer sehen können, wie die linke obere Ecke des Bildes aussehen sollte. Aus der Verzerrung dieser ordentlich angeordneten Punkte lässt sich die genaue Bildverzerrung ableiten. Dies wird als Kamerakalibrierung bezeichnet. Der Zweck der Kalibrierung besteht darin, letztendlich von einem Kamerabild zurück zu den ursprünglichen dreidimensionalen Koordinaten in der Realität zu berechnen.

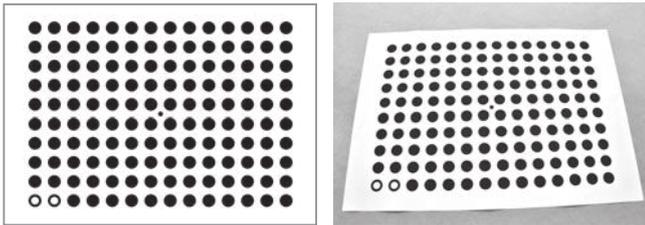


Abbildung 81: Links das Punktmuster, rechts das Foto, das es erzeugt. Die Linien sind nicht mehr gerade, und einige Punkte erscheinen größer als andere.

Punktewolke

Bei der Kalibrierung müssen fünfzehn Unbekannte geschätzt werden. Dies sind z.B. die Verzerrung des Objektivs in drei verschiedenen Richtungen, sondern auch der Winkel, in dem die Kamera hängt. Wir verwenden dieses Muster von 14×10 Punkten. Man erhält also eine Wolke von 140 Punkten in einem 15-dimensionalen Raum und suchen eine nichtlineare Beziehung zwischen diesen Punkte.

Die Forscher beschreiben ihre Sicht der Mathematik als praktisch: sie sind besonders an Lösungen interessiert, die gut funktionieren. Deshalb suchen sie nicht nach der genauen Lösung, sondern nach ein Ansatz, der gut genug für ihre Anwendungen ist. Dass sie mit sogenannten *genetischen Algorithmen* tun: ein Computerprogramm, das die Evolution nachahmt. Sie beginnen mit einem Anzahl der Zufallsschätzungen für die 15 Unbekannten. Dann berechnen sie für jede dieser Schätzungen den entsprechenden Fehler: Je besser die gefundene nichtlineare Beziehung die Punktewolke im 15-dimensionalen Raum beschreibt, desto kleiner ist der Fehler.

Als nächstes wird den besten Schätzungen erlaubt, sich zu reproduzieren (es ist einfach das Überleben des Stärkeren, „survival of the fittest“). Zwei Schätzungen werden zu einer neuen Schätzung kombiniert. Wie in der Natur ist ihr „Nachwuchs“ nicht nur eine Mischung seiner beiden Elternteile, sondern es finden auch einige wenige zufällige Veränderungen statt. Dies schafft eine neue Generation. Dieser Prozess wiederholt sich immer wieder. Jedes Mal wird den besten Schätzungen erlaubt, sich zu reproduzieren, bis eine Lösung gefunden wird, bei der der Fehler klein genug ist. Wie klein der Fehler sein kann, hängt von der Anwendung ab.

Flugzeuge

Unbemannte Flugzeuge müssen in der Lage sein, aus ihren Kamerabildern pixelgenaue Entfernungen und Winkel abzuschätzen. In Zukunft werden diese kleinen Flugzeuge Inspektionen von Windparks durchführen müssen, anstatt der Menschen, die dies jetzt unter Einsatz ihres eigenen Lebens tun. Flugzeuge, die jetzt eingesetzt werden, haben oft zusätzliche Kameras und brauchen extra Computer um vor Ort zurechnen. Es wird angestrebt, Kameras und Rechentechnik im Flugzeug selbst einzubauen, denn nur dann kann diese Technologie wirklich überall eingesetzt werden. Aber dazu müssen wir in der Lage sein, diese Kamerabilder sehr gut zu analysieren.

Softwarefehler können Schäden in Höhe von Hunderten Millionen Euro verursachen und sogar Menschenleben kosten. Die Mathematik hilft zu beweisen, dass die Software absolut fehlerfrei ist.

Mathematik beweist, dass Software keine Fehler hat

Da Computersoftware immer komplexer wird, wird es immer wichtiger, die Wahrscheinlichkeit von Fehlern so weit wie möglich zu reduzieren. Dies gilt sicherlich für lebenskritische Softwareanwendungen in Autos, Flugzeugen und Krankenhäusern, aber zunehmend auch für die Software von Unternehmen. Das niederländische Unternehmen ASML ist der weltweit größte Hersteller von Maschinen, die Computerchips auf Siliziumscheiben drucken. Große Chiphersteller wie Intel, Samsung und TSMC verwenden ASML-Maschinen, um ihre eigenen Computerchips auf Siliziumscheiben zu drucken. Diese Chips finden sich zum Beispiel in den neuesten iPhones und iPads.

Jede ASML-Maschine wird von einem kolossalen Softwareprogramm gesteuert. Die Grundlage für dieses Computerprogramm wurde vor 25 Jahren gelegt und seither kontinuierlich erweitert und verbessert. Das Programm verfügt jetzt über mehr als dreißig Millionen Codezeilen, und niemand kann alle Details übersehen. ASML beschäftigt 900 Mitarbeiter für die Wartung, Verbesserung und Erweiterung der Software. Softwarefehler können Schäden in Höhe von Hunderten Millionen Euro verursachen und sogar Menschenleben kosten. Die Mathematik hilft zu beweisen, dass ein Stück Software absolut fehlerfrei ist.

Kostspielige Fehler

Programmieren ist eine akribische Arbeit und eine goldene Regel in der Software-Welt besagt, dass im Durchschnitt 10 Fehler in 1.000 Zeilen Computercode vorkommen. Für die ASML-Maschine bedeutet dies, dass es bis zu 300.000 Fehler in der Software geben kann. In der Praxis wird ein Kunde viele dieser Fehler nicht bemerken, aber einige

Fehler können das Gerät stundenlang abschalten. Die Maschine kostet 40 Millionen Euro, und für jede Stunde, in der die Maschine angehalten wird, verlieren die ASML-Kunden, die die Maschine benutzen, schnell Hunderttausende von Euro an Einnahmen.

Traditionell werden Fehler durch das Testen von Software entdeckt. Das Problem beim Testen besteht darin, dass man zwar das Vorhandensein von Fehlern nachweisen kann, nicht aber, dass es keine Fehler enthält. Weil Software-Fehler so teuer sind, verwendet ASML seit einigen Jahren mathematische Beweis-Techniken, die nachweisen können, dass Teile der Software keine Fehler enthalten.

Jedes Stück Software besteht im Wesentlichen aus einer Abfolge von Entscheidungen: wenn A wahr ist, führen Sie B aus; wenn A nicht wahr ist, führen Sie C aus. Angenommen, ein Programm enthält eine dieser Arten von Entscheidungen, dann kann es in 2 möglichen Zuständen sein. Bei 10 Entscheidungen sind dies bereits $2^{10} = 1024$ Zustände und bei 1000 sogar bis zu 2^{1000} Zustände. Wenn man garantieren will, dass es keine Fehler in der Software gibt, dann muss man alle möglichen Kombinationen von Entscheidungen ausprobieren. Selbst mit hundert Kombinationen pro Minute ist es für ein großes System wie das von ASML nicht machbar, dies in einem vernünftigen Zeitrahmen zu erreichen.

Mathematische Tricks

Der Trick besteht darin, die Anzahl der möglichen Zustände, in denen sich eine Software befinden kann, zu reduzieren. Angenommen, das Programm muss die Aufgaben A, B und C erledigen und die Reihenfolge spielt keine Rolle. So enden alle sechs Kombinationen ABC, ACB, BAC, BCA, CAB und CBA alle im gleichen Zustand Q.

Beim klassischen Testen der Software müssen Sie alle sechs Kombinationen ausprobieren. Die mathematische Beweismethode sieht, dass alle Kombinationen zu Q führen. Bei dieser Beweismethode müssen Sie nur über eine viel kleinere Zahl als die Gesamtzahl der



Abbildung 82: Beispiel eines ASML-Lithographiegerätes. Einige große Herausforderungen für ASML-Lithographiemaschinen: sehr genaues, ultraschnelles und ultrakleines Linienschreiben auf Siliziumscheiben, die sich in Computerchips verwandeln sollen. Quelle: ASML

möglichen Zustände nachdenken. Dann ist es möglich, die Fehlerfreiheit in Teilen der gesamten Software nachzuweisen.

Ein weiterer Trick ist die *Symmetrie-Reduktion*. Angenommen, die ASML-Maschine kann 3 Produkte gleichzeitig verarbeiten, während sich gleichzeitig 6 Produkte in der Maschine befinden können. Die mathematische Beweismethode sieht dann sozusagen, dass die Verarbeitung von Produkt 1, 2 und 3 gleichbedeutend ist mit der Verarbeitung von Produkt 2, 3 und 4 und so weiter. Allerdings muss der Ingenieur dieses Wissen vermitteln, indem er der mathematischen Beweismethode mitteilt, dass die Produkte sauber in der Reihenfolge verarbeitet werden. Mit solchen mathematischen Tricks kann man letztlich beweisen, dass die Roboterarme, die die Siliziumscheiben bewegen, niemals kollidieren oder dass die Reihenfolge, in der man die Messungen vornimmt, immer korrekt ist.

Verbrechen durch Mathematik verstehen

Warum begehen Menschen Verbrechen? Warum sind manche unbeeindruckt von den Konsequenzen, während andere niemals das Gesetz brechen würden, selbst wenn niemand hinsieht? Gibt es umweltbedingte, sozioökonomische oder sogar biologische Faktoren, die zu kriminellem Verhalten führen? Gibt es mögliche Abschreckungsmaßnahmen? Was ist die angemessene Strafe? Wie erkennen wir überhaupt ein Fehlverhalten? Lassen sich all diese Fragen in mathematische Begriffe fassen?

Dr. Yao-Li Chuang und Prof. Dr. Maria R. D'Orsogna, Department of Mathematics, California State University, Los Angeles, fassen ihre Arbeiten zur Modellierung von Verbrechen zusammen:

Manche Verbrechen geschehen bekanntlich in der Hitze des Gefechts. Sie werden getrieben von plötzlichen und starken Impulsen, ausgelöst durch Eifersucht, Wut oder als Reaktion auf eine Provokation. Diese „Verbrechen aus Leidenschaft“ werden oft als Fälle von vorübergehender Unzurechnungsfähigkeit erklärt. Vorvermittelte Verbrechen hingegen beinhalten Planung und sorgfältige Recherche im Vorfeld. Emotionen wie Rache, Gier oder Neid können das Verbrechen immer noch motivieren, aber es wird im Voraus geplant, wie sich die Dynamik entfalten wird, oft mit der Absicht, sich der Strafverfolgung zu entziehen. Schließlich entstehen Gelegenheitsverbrechen, wenn ein Individuum erkennt, dass die Umstände optimal sind, um eine illegale Handlung zu begehen, und beschließt, die Gelegenheit zu ergreifen, ohne provoziert zu werden oder ohne Vorsätze zu haben.

Wohnungseinbrüche sind repräsentative Beispiele für Gelegenheitsverbrechen: Sofern keine besondere Absicht besteht, einen Ort zu überfallen, an dem wertvolle Gegenstände wie Geld, Kunst oder Schmuck aufbewahrt werden, sind Wohnungseinbrüche das Ergebnis dreier konvergierender Faktoren: ein leichtes/wertvolles Ziel, die Abwesenheit eines Wächters und ein motivierter Täter. Einfach ausgedrückt, ist ein eingebrochenes Haus zur falschen Zeit am falschen Ort, wenn der Einbrecher zuschlägt. Wäre der Eigentümer anwesend gewesen oder hätte das Sicherheitssystem sicherer gewirkt, hätte der Einbrecher einen anderen Ort oder vielleicht gar keinen gewählt. Die erstmals 1979 von Lawrence Cohen und Marcus Felson vorgestellte Routine-Aktivitäts-Theorie formalisiert die Entstehung von opportunistischer Kriminalität in soziologischer Hinsicht und wurde auf Einbrüche, sexuelle Übergriffe, Taschendiebstähle, Fahrzeugdiebstähle und Raubüberfälle angewendet.

Es gibt Hinweise darauf, dass wiederholte oder nahezu wiederholte Straftaten bei Gelegenheitsdelikten typisch sind. Bei Einbrüchen bedeutet dies, dass in dasselbe Haus (oder in benachbarte Häuser) wiederholt eingebrochen wird. Obwohl es kontraintuitiv klingen mag, finden wiederholte und nahezu wiederholte Straftaten statt, weil der Täter nach der ersten Tat mit dem Ort, den Routinen und der Sicherheit vertraut ist und die Rückkehr zum selben Haus (oder zu nahe gelegenen Häusern) weniger Unsicherheit und mehr Effizienz ermöglicht. Der ursprüngliche Täter hat möglicherweise auch Informationen an andere weitergegeben, so dass die Nachahmung durch einen anderen Täter ebenfalls eine Rolle spielen kann. Wiederholte und nahezu wiederholte Einbrüche wurden in vielen städtischen Zentren weltweit gemeldet, wobei sich das Einbruchrisiko über einige Wochen und

mehrere Stadtblöcke nach einem ersten Einbruch erhöht. Einige statistische Messungen zeigen, dass frühere Straftaten der beste Prädiktor für zukünftige Straftaten ist.

Zerbrochene Fensterscheiben (engl. 'broken windows') sind eine verwandte und bekannte Theorie in der Kriminologie, die postuliert, dass sichtbare Zeichen selbst kleinerer Verbrechen, wenn sie unbeachtet bleiben, weitere und schwerere Verbrechen fördern. Die Theorie wurde 1982 von James Wilson und George Kelling entwickelt³⁶ und der Titel bezieht sich auf den Anblick zerbrochener Fensterscheiben, die in New York City in den späten 1970er Jahren so allgegenwärtig waren, einem Ort und einer Zeit, die durch hohe Kriminalitätsraten gekennzeichnet waren. Zeichen der Unordnung in der Umgebung (zerbrochene Fensterscheiben, aber auch Graffiti, Müll, geparkte Autos ohne Nummernschilder) senden die Botschaft, dass die Gegend nicht gepflegt oder gut bewacht ist und dass illegale Aktivitäten toleriert werden. Diese Botschaft zieht mehr und schlimmere Kriminalität an. Bald können aus einem zerbrochenen Fenster oder einem Graffiti viele werden, die den städtischen Verfall signalisieren. Seit den späten 1980er Jahren hat diese Theorie die Polizeiarbeit in New York City geprägt, die eine Null-Toleranz-Strategie selbst für kleinere Verbrechen als Abschreckung für größere Verbrechen entwickelte. Die Erfahrung anderer Gemeinden in Massachusetts und in den Niederlanden zeigt, dass die Erhöhung des Ordnungsgefühls in der Nachbarschaft hilft, die Kriminalität einzudämmen. Allerdings ist die Broken-Windows-Theorie in den letzten Jahren umstrittener geworden, da sie zu exzessiven präventiven Eingriffen und zu übermäßiger Polizeiarbeit führen kann, wie z.B. bei der Anwendung von „stop and frisk“ in New York City, bei der Personen, die aufgrund ihres Aussehens oder Verhaltens verdächtig werden, Kriminelle zu sein, kurzzeitig von der Polizei angehalten werden, ohne dass ein wirklicher Beweis für eine Beteiligung an einem Verbrechen vorliegt.

Mathematische Modellierung von Verbrechen

Gelegenheitsverbrechen, die zu weiteren Verbrechen anregen, lassen sich am besten mathematisch behandeln, da sie keine individuellen, sehr persönlichen Motivatoren wie bei Rachemorden oder Dreiecksbeziehungen beinhalten. Wohnungseinbrüche, bei denen die Täter immer wieder dieselben Orte aufsuchen und „Hotspots“ der Kriminalität schaffen, sind ein natürlicher Ausgangspunkt, da das Ziel, ein Haus, feststeht und nur die Bewegung des Täters beschrieben werden muss. Das Ziel ist es, einen mathematischen Rahmen zu schaffen, der Elemente aus den oben beschriebenen soziologischen und kri-

minologischen Theorien enthält, so dass man Einbrüche (und andere opportunistische Straftaten) mit quantitativen Werkzeugen untersuchen kann.

Wir beginnen mit einem agentenbasierten Modell, in dem ein Einbrecher durch eine virtuelle Stadt, z.B. auf einem Gitter, wandert und einen zufälligen Spaziergang (engl. 'random walk') durchführt, der durch ein sich dynamisch entwickelndes „Attraktivitätsfeld“ beeinflusst wird. Jedes Mal, wenn ein Verbrechen begangen wird, wird der Einbrecher entfernt und taucht später an einem anderen Ort wieder auf, bereit, erneut zuzuschlagen. Sobald ein Verbrechen begangen wurde, nimmt die Attraktivität des Ortes zu. Die Idee ist, dass der Täter nach einer Straftat zu seinem Heimatort zurückkehrt und dass die erhöhte Attraktivität die Mobilität zukünftiger Täter durch eine positive Rückkopplung in Richtung bereits eingebrochener Orte lenkt, gemäß den Prinzipien der wiederholten und fast wiederholten Straftat. Die Attraktivität wird auch auf natürliche Weise zu einem intrinsischen Basiswert abnehmen, wenn keine Straftaten begangen werden.

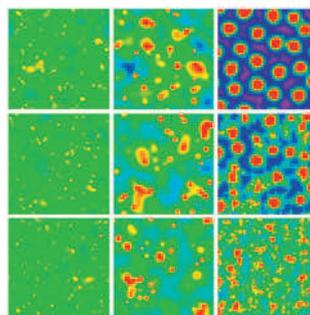


Abbildung 83: Verbrecheninduzierte Hotspots im Attraktivitätsfeld, ausgehend von einem einheitlichen Wert: Anfangs sind alle Standorte in der virtuellen „Stadt“ gleich attraktiv. Wenn es viele Kriminelle gibt, aber das Attraktivitätsfeld nach einem Verbrechen nicht signifikant ansteigt, treten Hotspots nie auf (obere Reihe); wenn es viele Kriminelle gibt und das Attraktivitätsfeld nach einem Verbrechen signifikant ansteigt, treten Hotspots auf und sind stationär (untere Reihe). Wenn es wenige

Kriminelle gibt, erscheinen Hotspots nur, wenn die Änderungen im Attraktivitätsfeld groß sind, aber sie sind vorübergehend (mittlere Reihe). Die Farben folgen dem Regenbogenspektrum von Violett (Minimum) bis Rot (Maximum), wobei Grün den Mittelpunkt darstellt.

Unter bestimmten Bedingungen lassen sich diese Regeln in ein kompaktes System so genannter Diffusions-Reaktions-Differentialgleichungen gießen, die wie folgt koppeln und A , die Dichte der Verbrecher und das Anziehungsfeld, berücksichtigen

$$\frac{\partial A}{\partial t} = \eta \nabla^2 A - (A - A_0) + \rho A,$$

$$\frac{\partial \rho}{\partial t} = \vec{\nabla} \cdot \left[\vec{\nabla} \rho - \frac{2\rho}{A} \vec{\nabla} A \right] - \rho A + \Gamma.$$

Dabei stellt der Ausbreitungsparameter η dar, wie stark sich die Attraktivität auf nahe A_0 ist der Basiswert der Attraktivität der virtuellen Stadt, zu der sich A bei Abwesenheit von Kriminalität natürlich hinbewegt. A_0 ist der Basiswert der Attraktivität der virtuellen Stadt, zu der sich A bei Abwesenheit von Kriminalität natürlich hinbewegt, und Gamma ist die Rate der Ankunft von Kriminellen in der Stadt. Hotspots entstehen, wenn sich Verbrechen in bestimmten Gebieten aufgrund eines starken Anstiegs des Attraktivitätsfeldes, das Kriminelle nach einem Einbruch anzieht, häufen und durch eine große Anzahl von Kriminellen, die diese Gebiete besuchen, aufrechterhalten werden. Wenn jedoch die kriminelle Aktivität zu stark ist, z.B. wenn es zu viele Kriminelle gibt und das Attraktivitätsfeld überall einheitlich hoch wird, gehen die Hotspots ineinander über und schließlich wird die gesamte Stadt zu einer einzigen Verbrechenswüste. Die Ergebnisse für das Attraktivitätsfeld A für verschiedene Parameter sind in Abbildung 83 dargestellt und zeigen, wo sich Kriminelle tendenziell häufen und wo Einbrüche wahrscheinlicher sind. Dieses Modell wurde erweitert, um die räumliche Unordnung in der Stadt mit einzubeziehen (durch Berücksichtigung ungleichmäßiger Werte von A_0) und, was am wichtigsten ist, um Methoden zur Ausmerzungen von Hotspots und zur Bestimmung der strategischen Lage von Strafverfolgungsbehörden zu untersuchen.

Die oben erwähnten Phänomene der Häufung von Straftaten und der Beinahe-Wiederholung von Ereignissen werden auch bei Bandenriminalität, Terroranschlägen und Protesten beobachtet, und zwar aufgrund anderer Formen der Ansteckung, bei denen ein anfängliches zufälliges Ereignis andere Ereignisse in der Nähe auslösen kann, entweder als „Nachahmer“ oder als Vergeltungsmaßnahmen. Die zeitlichen Muster erinnern an die Aktivität von Erdbeben, bei denen ein erstes Ereignis viele Nachbeben in der Nähe auslöst. Infolgedessen wurden mathematische Techniken, die ursprünglich zur Untersuchung tektonischer Aktivität entwickelt wurden, angepasst, um Beinahe-Wiederholungsphänomene in soziologischen und Verhaltenskontexten zu untersuchen. Eine der wichtigsten dieser Methoden ist der selbsterregende Punktprozess, auch *Hawkes-Prozess* genannt. Wenn es keine vorherigen Ereignisse gibt (keine Erdbeben oder keine Verbrechen), ist die Anzahl der neuen Ereignisse, die pro Zeiteinheit auftreten sollen, bezeichnet mit $\lambda(t)$, durch eine Konstante $\lambda(t) = \mu$ gegeben. Dies zeigt an, dass zu Beginn des Prozesses Ereignisse zu jeder Zeit mit der gleichen Wahrscheinlichkeit auftreten werden. Wenn es jedoch ein anfängliches Ereignis zum Zeitpunkt t_1 gibt, ein erstes Erdbeben oder ein erstes Verbre-

chen, ist die Anzahl der zukünftigen Ereignisse pro Zeiteinheit vorübergehend größer und gegeben durch

$$\lambda(t) = \mu + \alpha e^{-\beta(t-t_1)}.$$

Beachten Sie, dass unmittelbar nach dem ersten Ereignis $\lambda(t_1) = \mu + \alpha > \mu$, und dass nach einer Zeit $1/\beta$ nach t_1 , $\lambda(t)$ zu seinem Anfangswert μ zurückkehrt. Das bedeutet, dass das erste Ereignis als Auslöser für weitere Ereignisse pro Zeiteinheit innerhalb der Zeitspanne $1/\beta$ gewirkt hat. Natürlich kennen wir die Werte von α , β nicht, aber wenn man eine Reihe von geografisch geclusterten Ereignissen hat, die zu Zeiten $t_1, t_2, \dots, t_i, \dots$ auftreten, die Teil desselben Prozesses sind und die $\lambda(t)$ auf dieselbe Weise beeinflussen, können diese Ereignisse an die folgende Form angepasst werden

$$\lambda(t) = \mu + \sum_{t_i < t} \alpha e^{-\beta(t-t_i)},$$

aus der α , β abgeleitet werden können. Große Werte von α und kleine Werte von β bedeuten, dass die Beinahe-Wiederholungseffekte groß sind und länger andauern. Man kann auch die Anzahl der „Kind“-Ereignisse abschätzen, die durch ein einzelnes „Eltern“-Ereignis ausgelöst werden, und zeigen, dass diese durch $\gamma = \alpha/\beta$ gegeben ist. Die Modellierung selbsterregender Punktprozesse wurde auf Wohnungseinbrüche, gewaltsame zivile Todesfälle im Nahen Osten, Bandengewalt, Terroranschläge und Proteste im Zuge von Abriegelungsmaßnahmen zur Bekämpfung von COVID-19 in verschiedenen Ländern angewendet.

Strategien zur Verringerung von Kriminalität

Ein letztes Thema ist das der Rehabilitation. Lange Zeit, vor allem in den Vereinigten Staaten, bestand der Hauptzweck der Inhaftierung darin, Straftäter zu bestrafen. Harte Strafen wurden als Abschreckung für zukünftige Straftaten gesehen, nicht nur für den Täter selbst, sondern auch als Beispiel für andere, die von kriminellen Aktivitäten Abstand nehmen sollten, da sie im Falle einer Verhaftung lange Zeit im Gefängnis verbringen könnten. Die Ergebnisse waren jedoch nicht ermutigend, da die Kriminalitätsrate in den letzten Jahrzehnten in den gesamten Vereinigten Staaten gestiegen ist. Vielleicht ist ein anderer Weg möglich? In der Tat kann die Rehabilitation von Straftätern ein nützlicheres Mittel sein, um die Kriminalität zu reduzieren, indem man den Menschen hilft, ihr Verhalten zu ändern, und zwar durch eine Berufsausbildung, um sich auf Jobs außerhalb der Gesellschaft vorzubereiten, durch Beratung bei

Drogenmissbrauch, durch Hilfe bei der Bewältigung von Wut und Gewaltimpulsen und durch den Aufbau elterlicher Fähigkeiten. Es stellt sich also die Frage „Peitsche oder Zuckerbrot“: Ist es besser, Straftäter zu bestrafen (die Peitsche) oder ihnen zu helfen, sich zu rehabilitieren (das Zuckerbrot), um die Kriminalität zu reduzieren?

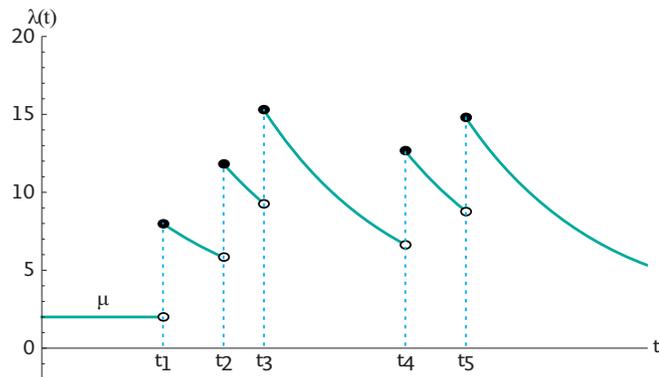


Abbildung 84: Dynamik eines selbsterregenden Prozesses. Bis zum ersten Ereignis (ein Erdbeben oder ein Verbrechen) zum Zeitpunkt t_1 ist die Anzahl der erwarteten Ereignisse pro Zeiteinheit eine Konstante, $\lambda(t) = \mu$. Nach dem ersten Ereignis zum Zeitpunkt t_1 steigt $\lambda(t \rightarrow t_1^+)$ auf $\mu + \alpha$, wie durch den ersten Sprung bei t_1 angezeigt. Die Größe $\lambda(t)$ beginnt dann zu sinken, so dass sie, wenn es keine weiteren Ereignisse gäbe, nach einer Zeit $1/\beta$ wieder in Richtung des Anfangswert μ zurückkehren würde. Ein neues Ereignis zum Zeitpunkt t_2 bewirkt jedoch einen zweiten Sprung in $\lambda(t)$, so dass $\lambda(t \rightarrow t_2^+) = \mu + \alpha e^{-\beta(t-t_1)} + \alpha$. Der Prozess setzt sich auf diese Weise fort. In diesem Bild ist $\mu = 2$, $\alpha = 6$, $\beta = 0.15$ und die Reihe der Ereignisse findet zu den Zeiten $\{t_1, t_2, t_3, t_4, t_5\} = \{6, 9, 11, 18, 21\}$ statt. Alle Zeiten werden in dieser Darstellung in willkürlichen Einheiten angenommen.

Eine Möglichkeit, diese Frage anzugehen, ist die evolutionäre Spieltheorie, bei der angenommen wird, dass Individuen Entscheidungen treffen, in diesem Fall, ein Verbrechen zu begehen oder nicht, in Abhängigkeit von drei Faktoren: ihre frühere Verhaftungsgeschichte – hohe frühere Strafen verringern die Wahrscheinlichkeit einer Rückfälligkeit – das soziale Umfeld – je mehr Kriminelle anwesend sind, desto größer ist die Wahrscheinlichkeit, ein Verbrechen zu begehen – und die Ressourcen für die Rehabilitation, die Personen mit früheren Verhaftungen angeboten werden. Eine weitere Möglichkeit ist, dass Personen mit einer kriminellen Vergangenheit ihr kriminelles Leben (und das evolutionäre Spiel) als reformierte Personen dauerhaft ver-

lassen. Schließlich wird angenommen, dass das Individuum ab einer bestimmten Anzahl von Straftaten sein Verhalten nicht mehr ändert und ebenfalls das Spiel verlässt, allerdings als Rückfalltäter. Die soziologischen Faktoren, die bestimmen, welche Entscheidungen getroffen werden (z.B. die Ressourcen für die Rehabilitation oder die Höhe der Strafe), werden durch mathematische Parameter dargestellt. Ihre Änderung führt zu unterschiedlichen Ergebnissen, mit unterschiedlichen Anteilen von Individuen, die nie Straftaten begehen, die sich bessern oder die bis zum Ende des Spiels unverbesserlich bleiben.

Eine wichtige mathematische Größe ist das finale Verhältnis zwischen denjenigen, die nie Straftaten begangen haben oder die sich gebessert haben, P , den „Paladinen“ der Gesellschaft und den Rückfälligen U , den „Unverbesserlichen“. Große Werte von P/U repräsentieren „gute“ Gesellschaften mit viel mehr tugendhaften Menschen als hartgesottenen Kriminellen, und kleine Werte von P/U repräsentieren „schlechte“ Gesellschaften mit vielen Kriminellen und viel Kriminalität. In einer idealen Welt hätte man alle möglichen Ressourcen um Kriminelle in angemessener Weise zu bestrafen und zu rehabilitieren. In der realen Welt können die Gesamtressourcen jedoch begrenzt sein, so dass, wenn die Gesellschaft beschließt, die mit der Bestrafung verbundenen Kosten zu erhöhen (z.B. durch eine Verlängerung der Strafen oder durch die lange Aufbewahrung von Kriminellen unter hohen Sicherheitsvorkehrungen), zwangsläufig die mit der Rehabilitation verbundenen Kosten gesenkt werden.

Wir stellen fest, dass die erfolgreichste Strategie zur Erhöhung des Verhältnisses P/U und zur Verringerung der Anzahl der Kriminellen (und der Verbrechen) darin besteht, die Ressourcen optimal so zu verteilen, dass Kriminelle nach ihrer Bestrafung wirkungsvolle Interventionsprogramme erfahren, insbesondere in den ersten Phasen ihrer Rückkehr in die Gesellschaft. Übermäßig harte oder nachsichtige Strafen sind weniger wirksam. Das Dilemma „Peitsche versus Zuckerbrot“, kann also durch eine „Peitsche und Zuckerbrot“-Strategie gelöst werden: Der beste Weg, die Kriminalität zu verringern, besteht darin, sicherzustellen, dass die Strafe hart genug ist, um Kriminelle davon abzuhalten, erneut straffällig zu werden. Gleichzeitig muss sichergestellt werden, dass genügend Ressourcen für die Rehabilitation zur Verfügung stehen.

⁵⁶ 56J.Q. Wilson, G.E. Kelling, Broken Windows. The Police and Neighborhood Safety, in: The Atlantic Monthly. März 1982.



Das Haus des Täters: Was ist Geographic Profiling?

Bei einem Serienmörder ist es schwierig vorherzusagen, wer das nächste Opfer sein wird, aber die Arithmetik kann sagen, wo der Täter wohnt.

In der allerersten Episode der Krimiserie *Numb3rs* widersetzt sich ein Serienmörder der Polizei. Die Detektive suchen nach einem Muster an den Fundorten der Leichen. Auf diese Weise hoffen sie vorherzusagen, wo der Mörder beim nächsten Mal zuschlagen wird. Der Bruder eines Polizisten, der zufällig Mathematiker ist, erklärt, dass dies Unsinn sei: „Wenn man einen Gartenregner nimmt und sich das Muster der Tropfen ansieht, ist es unmöglich vorherzusagen, wo der nächste Tropfen landen wird. Aber das Muster der Tropfen wird Ihnen sagen, wo der Sprinkler ist.“

Dann erstellt der Mathematiker in Übereinstimmung mit dieser Gleichung ein Modell, das den Wohnort des Mörders vorhersagt. Die Polizei durchsucht das fragliche Viertel, aber DNA-Untersuchungen zeigen, dass der Täter dort nicht wohnt. Der Mathematiker passt sein Modell an, um den Unterschied zwischen Wohnen und Arbeiten zu berücksichtigen, und in dem nun extra ausgewiesenen Bereich weiß die Polizei, wo der Mörder zu finden ist. Dann stellt sich heraus, dass er gerade umgezogen ist und dass das erste Modell seine alte Straße ordentlich als verdächtig bezeichnet hatte.

Die Fernsehmacher brauchten sich für diese Geschichte nicht viel einfallen zu lassen, denn so ist es im wirklichen Leben fast genau so passiert. Nur war es in diesem Fall nicht der Bruder eines Polizisten, der die Lösung brachte. Während seiner Beförderung 1991 entwickelte der Kriminologe Kim Rossmo eine Formel, um zu bestimmen, wo der

Täter wahrscheinlich von den Tatorten aus wohnt. Die Formel besteht aus zwei Termen. Ein Term beschreibt, dass der Täter nicht zu nahe an seinem eigenen Haus zuschlagen wird. Der andere Term besagt, dass nach dieser Pufferzone die Wahrscheinlichkeit einer Straftat langsam abnimmt. Spätere Verfeinerungen der Formel berücksichtigen auch die Schwere des Verbrechens. Je weiter weg von zu Hause, desto gewalttätiger ist die Tat. Das Ergebnis der Formel ist genau wie in *Numb3rs*: eine farbige Karte, die pro Bezirk anzeigt, wie wahrscheinlich es ist, dass der Täter dort wohnt.

Genau wie im Film

Kim Rossmo ist jetzt Professor an der Texas State University und leitet das *Center for Geospatial Intelligence and Investigation*. Die Episode von *Numb3rs* sieht verdächtig nach der Suche nach einem Vergewaltiger aus, der zehn Jahre lang in Louisiana aktiv war. 1998 bat der damals ziemlich verzweifelte Polizeibeamte Kim Rossmo um Hilfe. Rossmo sammelte einige Tage lang Daten und erstellte dann farbige Karten, die zeigen, wo der Täter wahrscheinlich lebte. Damit blieb ein Gebiet von etwa einem Quadratkilometer zu untersuchen: die Gefahrenzone. Aber alle Männer in der Nachbarschaft stellten sich nach einem DNA-Test als unschuldig heraus.

Dann erhielt die Polizei einen Hinweis auf einen neuen Verdächtigen, der, wie sich herausstellte, gar nicht in dem von Rossmo bezeichneten Gebiet wohnte. Ein kluger Detektiv entdeckte jedoch, dass der Mann gerade umgezogen war und mitten in der Gefahrenzone wohnte. Genau wie in *Numb3rs*. Die DNA auf einer weggeworfenen Zigarette beweist, dass dieser Verdächtige tatsächlich der Täter ist.

Das einzige deduktive System

Kim Rossmo hofft, mit seiner Arbeit die Wissenschaft des *Geographic Profiling* zu präzisieren. Viele forensische Wissenschaft basiert auf Ableitungen statt auf logischem Denken: Sie ist induktiv statt deduktiv. Deduktive Argumentation basiert rein auf Fakten und Logik. Denken Sie darüber nach:

***Alle Menschen sind sterblich.
Sokrates ist ein menschliches Wesen.
Sokrates ist also sterblich.***

Die induktive Argumentation basiert auf Beobachtungen, aus denen Sie möglichst plausible Schlussfolgerungen ziehen. Denken Sie darüber nach:

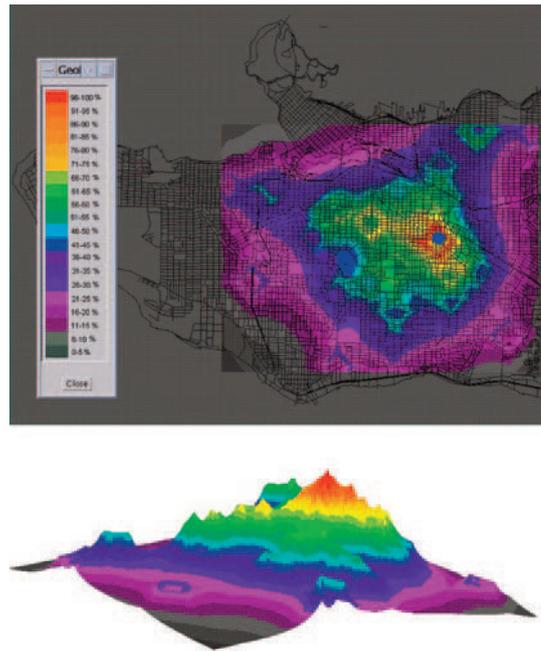


Abbildung 85: Bilder eines Geoprofils und einer Gefährdungsfläche für eine Serie von Raubüberfällen in Vancouver, BC, Kanada. Der blaue Punkt auf dem Geoprofil zeigt an, wo die Täter wohnten.

In der vergangenen Woche ließ mein Nachbar jeden Morgen um acht Uhr seinen Hund raus. So wird der Nachbar morgen früh um 8:00 Uhr mit seinem Hund Gassi gehen.

Induktive Argumentation funktioniert in der Praxis oft gut, aber sie ist nicht wasserdicht. Im obigen Beispiel kann es durchaus sein, dass der Nachbar verschläft. Rossmo: „Die meiste Wissenschaft ist Induktion: Man zeichnet seine Beobachtungen auf und macht daraus Verallgemeinerungen. Das einzige wirkliche deduktive System ist die Mathematik.“

In der Zwischenzeit hat das System von Rossmo zur Lösung von Hunderten von Fällen beigetragen. Rossmo betont in Interviews, dass es ihm nicht um Vorhersagen geht. Im Gegenteil, sein Modell geht auf den Ausgangspunkt zurück. Um dies zu erklären, verwendet er die Metapher des Gartensprengers. Er scherzt, er sei zu 99 % sicher, dass die Autoren von *Numb3rs* diese Idee aus einem Interview haben. Die Schöpfer mussten sich wirklich nicht viel selbst einfallen lassen.

Wie gross sind die Chancen, dass die **DNA-Spur** des Verdächtigen gefunden wird?

Verdächtige eines Verbrechens hinterlassen oft Spuren: von Schuhabdrücken bis hin zu Kleidungsfasern und DNA-Spuren. Die forensischen Statistiken werden bei der Berechnung der Beweiskraft solcher Spuren immer genauer.



Seit 1988 wird die DNA-Analyse in Deutschland genutzt, um Straftäter eine Tat nachzuweisen. Die häufigste Vorgehensweise ist der Vergleich einer am Tatort gefundenen DNA-Spur mit der DNA von Tatverdächtigen oder mit der DNA-Analyse-Datei (DAD), in der u.a. die DNA-Muster von bereits in Erscheinung getretenen Tatverdächtigen und von anderen Tatorten gespeichert werden.

Die DNA, die der Mensch fast in jeder Zelle seines Körpers trägt, besteht aus mehreren Milliarden Einzelteilen, die bei jedem Menschen in Doppelsträngen unterschiedlich angeordnet sind. Dabei gibt es vier Grundbausteine, die sich in bestimmten Abschnitten unterschiedlich oft wiederholen. Durch diese unterschiedliche Anzahl an Wiederholungen in genau definierbaren Bereichen lassen sich sogenannte DNA-Muster feststellen, die für jeden Menschen individuell sind, siehe Abbildung 86. Es kann durchaus sein, dass bei verschiedenen Menschen die Anzahl der Wiederholungen in bestimmten Abschnitten gleich ist. Durch Populationsstudien wurde ermittelt, wie hoch die Wahrscheinlichkeit bei einem zufälligen Menschen der Bevölkerung für eine bestimmte Anzahl von Wiederholungen in jedem Abschnitt ist.



Abbildung 86: Der Wissenschaftler analysiert das in Genetik, Forensik, Arzneimittelforschung, Biologie und Medizin verwendete DNA-Gel.
© gopixa / shutterstock"

Während das DNA-Identifizierungsmuster anfangs auf fünf Abschnitte beschränkt war, werden in der EU inzwischen mindestens zwölf Vergleichsabschnitte erfasst. Zudem kann inzwischen aus kleinsten Spuren schon mehr analysiert werden als früher. Je mehr Abschnitte aus einer aufgefundenen Spur aufbereitet werden können, desto zuverlässiger können Aussagen darüber getroffen werden, ob die Vergleichsperson auch die Spur verursacht hat. „Auch wenn eine DNA-Spur am Tatort für sich genommen meist noch keine Täterschaft beweisen muss, ist sie dennoch ein wichtiger Beweis, bzw. Ansatz für weitere Ermittlungen“, sagt der junge Kriminalkommissar Befeldt.

Ermittlung eines Mörders in den Niederlanden durch DNA-Analyse

1999 wurde das damals sechzehnjährige Mädchen Marianne Vaatstra vergewaltigt und ermordet, als sie mit dem Fahrrad nach Hause fuhr. Auf ihrem Körper befand sich eine DNA-Spur, die höchstwahrscheinlich vom Täter stammte. In den folgenden Jahren verhaftete die Staatsanwaltschaft mindestens zwölf Verdächtige, doch der Täter wurde nicht gefunden. Im September 2012 begann eine Untersuchung der DNA-Übereinstimmung. Die Idee war, dass der Täter vielleicht über Familienangehörige ausfindig gemacht werden könnte. DNA-Analyse-Techniken können auf der Grundlage der hinterlassenen DNA-Spur Hinweise darauf geben, ob der mögliche Täter mit einem oder mehreren Teilnehmern am DNA-Populationsscreening genetisch verwandt ist. 7.200 Männer, die in einem Umkreis von fünf Kilometern um den Tatort leben, gaben freiwillig ihre DNA ab. Das Niederländische Forensische Institut (NFI) verglich die DNA-Profile dieser Freiwilligen mit der DNA-Spur des Täters. Unerwartet fand die NFI eine direkte Übereinstimmung mit der DNA eines Jasper S. Er hatte freiwillig am Bevölkerungsscreening teilgenommen und musste nicht einmal über Familienmitglieder zurückverfolgt werden. Im Dezember 2012 gestand Jasper S. den Mord. Er wurde zu achtzehn Jahren Gefängnis verurteilt.

Aufklärung eines Mordes nach 25 Jahren

Auch in Deutschland werden sog. Cold Cases durch DNA-Analysen aufgeklärt. So konnte 2020 der Mord an Brigitta J. in Sindelfingen aufgeklärt werden. Nachdem die Ermittlungen 1995 ohne Erfolg eingestellt worden waren, führten neue Abgleiche der damals am Opfer aufgefundenen DNA-Spuren zum Tatverdächtigen Hartmut M., dessen DNA-Spuren in der DAD vorhanden waren, da er 2007 bereits wegen eines Tötungsdeliktes verurteilt worden war, siehe

Abbildung 87. Kriminaldirektor Mathias Bölle, der Leiter der Kriminalpolizeidirektion Böblingen sagt zu diesem Erfolg: „Dieser Fall zeigt uns einmal mehr, zu welch herausragenden Ergebnissen die hartnäckige und intensive Arbeit unserer Kolleginnen und Kollegen im Zusammenspiel mit modernen kriminaltechnischen Methoden führen kann.“

Blutgruppe

Die Aufklärungen dieser Fälle sind nur zwei der vielen Beispiele für eine erfolgreiche Anwendung der forensischen DNA-Analyse auf z.B. Speichel, Haare, Blut oder Sperma, die zurückgelassen wurden. „In den letzten Jahren hat sich die forensische Forschung zunehmend von einem Handwerk zu einer Wissenschaft entwickelt“, sagt der Statistiker Marjan Sjerps von der NFI, ebenfalls Teilzeitprofessor für forensische Statistik an der Universität Amsterdam. Die Mathematik, insbesondere die Wahrscheinlichkeitstheorie und die Statistik, spielen in dieser Wissenschaft eine wichtige Rolle. Sjerps: „Vor dreissig Jahren war die genaueste Aussage, die wir machen konnten, dass die Blutgruppe einer Spur mit der eines Verdächtigen übereinstimmt und dass die Wahrscheinlichkeit eines Zufalls zum Beispiel zehn Prozent betrug. Bei den heutigen DNA-Profilen liegt diese Wahrscheinlichkeit bei weniger als 1 zu 1 Milliarde.“ Diese Wahrscheinlichkeitsbestimmung wird in einem bestimmten Typ von Wahrscheinlichkeitsmodell verwendet: das *Bayes'sche Wahrscheinlichkeitsmodell*. Dieses Modell hat die wichtige Eigenschaft, dass die Wahrscheinlichkeit einer Hypothese auf natürliche Weise angepasst wird, wenn im Laufe der Zeit zusätzliche Beweise verfügbar werden.

Wenn am Tatort eine DNA-Spur gefunden wird, bittet die Staatsanwaltschaft die NFI, die Spur zu analysieren und festzustellen, wie stark die Beweise dafür sind, dass ein Teil oder die gesamte DNA vom Verdächtigen stammt. Die Stärke der Evidenz wird durch das so genannte *Likelihood Ratio (LR)* gemessen. Je größer die LR, desto stärker ist der Beweis, dass die Spur wirklich von dem Verdächtigen stammt. Die LR ist das Verhältnis zwischen zwei Wahrscheinlichkeiten. Die Wahrscheinlichkeit im Zähler ist die Wahrscheinlichkeit, dass es eine Übereinstimmung gibt, wenn die Spur mit Gewissheit vom Verdächtigen kommt. Diese Chance ist fast 1, wenn es sich um eine sehr schöne unvermischte Spur handelt und wenn die Analyse perfekt durchgeführt wird. Die Wahrscheinlichkeit im Nenner ist die Wahrscheinlichkeit, dass es eine Übereinstimmung gibt, wenn die Spur mit Sicherheit nicht von dem Verdächtigen stammt.



Abbildung 87: Polizeibeamter, der Kriminelle in Handschellen am Schreibtisch im Innenbereich befragt. © New Africa / shutterstock

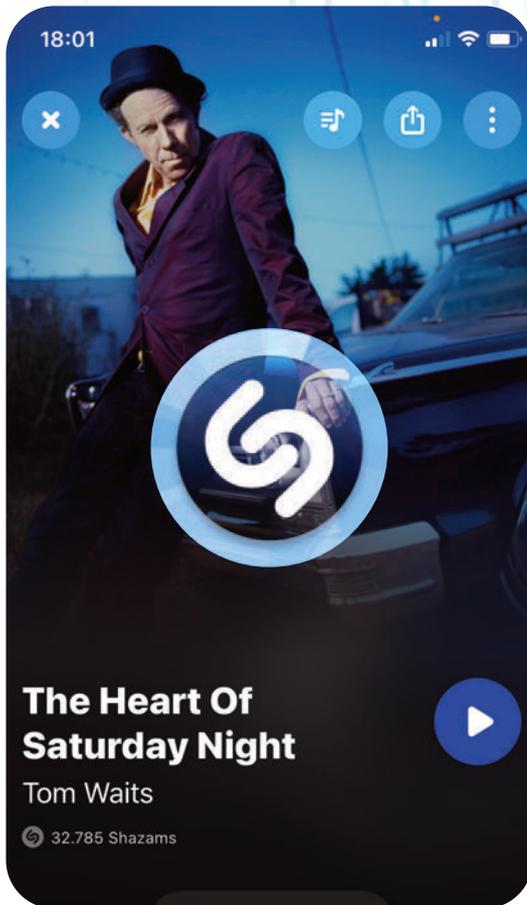
Je seltener das DNA-Profil, desto geringer ist diese zweite Chance. Sjerps: „Die zuverlässige Bestimmung dieser beiden Wahrscheinlichkeiten für alle Arten von Spuren ist eine große Herausforderung für die forensische Statistik.“

Identifizierung des Opfers

Die gleiche Art der Wahrscheinlichkeitsberechnung wie im Fall Vaatstra wurde von der NFI erfolgreich für die Identifizierung der Opfer des Flugzeugunglücks in den libyschen Tripolis im Jahr 2010 eingesetzt. Dies führte zum Tod von 71 Niederländern. Nur eine Person überlebte den Unfall, der neunjährige niederländische Junge Ruben. Die NFI verwendete DNA von Familienmitgliedern wie Eltern, Brüdern und Schwestern, um die Opfer zu identifizieren. Sjerps meint, dass die forensische Statistik in den kommenden Jahren den Beweiswert von immer mehr verschiedenen Spuren immer genauer bestim-

men wird: „Dank neuer technischer Entwicklungen werden wir zum Beispiel die DNA oder die chemische Zusammensetzung von immer kleineren Spuren bestimmen können. Wir können auch gemischte Spuren – wie zum Beispiel eine Zigarette, die sich zwei Raucher teilen – immer genauer analysieren. Darüber hinaus erstellt die NFI nun getrennte Berichte über die Beweiskraft beispielsweise von Fingerabdrücken, DNA-Spuren und Kleidungsfasern vom selben Tatort. Künftig werden wir auch den kombinierten Beweiswert dieser Spuren berechnen.“

Jeder Song hat einen einzigartigen digitalen Fingerabdruck. Shazam kennt die Formel für diesen Fingerabdruck und erkennt damit Lieder.



Wie erkennt Shazam ein Lied so schnell?

Ich höre ein wunderschönes Lied aus den Lautsprechern widerhallen und richte das Mikrofon meines Mobiltelefons darauf. Ich drücke 'Shazam', sehe, wie die App den Ton aufnimmt und sogar auf die verschiedenen Tonhöhen reagiert, und nach ein paar Sekunden erscheint „The Heart Of Saturday Night – Tom Waits“ auf dem Bildschirm. Sogar das Albumcover, das den Song enthält, erscheint auf dem Bildschirm. Es ist jedes Mal ein magischer Moment, wenn Shazam Ihnen sagt, wer das Lied singt und wie der Titel lautet.

Shazam ist ein ursprünglich amerikanisches Unternehmen, das 1999 gegründet wurde. Seit 2008 bietet das Unternehmen seinen Musikliebhabern eine kostenlose App für Mobiltelefone an. Shazam verfügt nun über eine Datenbank mit mehr als elf Millionen Liedern, deren digitaler Fingerabdruck ermittelt wurde. Nur Lieder aus dieser ständig wachsenden Datei können von Shazam erkannt werden. Jeden Tag wird weltweit bis zu fünfzehn Millionen Mal auf Shazam zugegriffen. Selbst die Hintergrundgeräusche in einem Café oder ein bisschen Klangverzerrung verwirren den Musikliebhaber nicht. Was ist das Geheimnis?

Digitaler Fingerabdruck

Das Geheimnis von Shazam ist der Algorithmus, mit dem es einen digitalen Fingerabdruck eines Liedes oder Musikstücks erstellt. Dieser digitale Fingerabdruck ist ein einzigartiges Klangmuster eines Liedes, genau wie ein Fingerabdruck ein einzigartiges Linienmuster auf den Fingern jeder Person ist.

Sie können von jeder Tonaufnahme ein Spektrogramm erstellen: ein dreidimensionales Diagramm, in dem auf der x-Achse die Zeit, auf der y-Achse die Frequenz des Tons und auf der z-Achse die Stärke des Tons angegeben ist. Jeder Punkt in der Grafik sagt also aus, welche Frequenz und welche Amplitude der Schall zu welchem Zeitpunkt hat. Obwohl dieses Spektrogramm alle Charakteristika des Tons enthält, enthält es so viele Informationen, dass es unbequem ist, alles digital zu speichern, um später darauf zurückgreifen zu können.

Shazam konzentriert sich daher auf die intensiven Fragmente aus einem Lied, sagen wir die Spitzenwerte im Spektrogramm. Der Algorithmus bestimmt die Zeit dieser Peaks und die mit ihnen verbundenen Frequenzen. Das Ergebnis ist eine zweidimensionale Grafik von Punkten, die im Spektrogramm hervorstechen, so als ob Sie nur die Gipfel einer Berglandschaft auf einer Karte anzeigen würden. Diese Informationen reichen in der Regel nicht aus, um ein Lied zu identifizieren. Deshalb haben die Designer von Shazam einen cleveren Trick angewandt. Sie berechnen, wie sich jeder Peak zu anderen nahegelegenen Peaks verhält. Diese wird durch die Zeit zwischen den Peaks und die Frequenzunterschiede zwischen ihnen bestimmt.

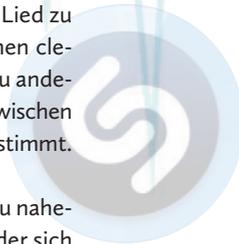
Die Informationen über einzelne Peaks und ihre Beziehung zu nahegelegenen Peaks bestimmen den digitalen Fingerabdruck, der sich in der Praxis als (nahezu) einzigartig erweist. Shazam vergleicht den digitalen Fingerabdruck eines Liedes, das Sie als Benutzer aufnehmen, mit allen Fingerabdrücken, die in ihrer Datenbank vorhanden sind. Wenn es eine Übereinstimmung gibt, erscheinen Titel und Künstler auf dem Bildschirm.

Treffer-Vorhersage

Obwohl Shazam für Popmusik gut funktioniert, hat der Musikkenner Probleme mit klassischer Musik. Shazam auf meinem Smartphone erkennt ein Stück aus Mahlers erster Sinfonie, aber wenn ich ihm ein Stück aus Sibelius' erster Sinfonie zeige – ebenfalls ein kraftvolles, charakteristisches Stück – hat er keine Ahnung. Das ist nicht ver-

wunderlich, denn viele klassische Musikstücke dauern nicht wie bei der Popmusik ein paar Minuten, sondern manchmal eine Stunde. Wenn Sie nur zehn Sekunden aus einem einstündigen Musikstück hören, ist es viel schwieriger zu bestimmen, aus welchem Stück es stammt. Außerdem macht es einen großen Unterschied, wie ein Stück aufgeführt wird. Das kann von Orchester zu Orchester und von Dirigent zu Dirigent sehr unterschiedlich sein.

Eine der neuesten Anwendungen von Shazam ist die Vorhersage von Treffern. Dazu kombiniert das Unternehmen Online-Liedrezensionen mit der Frage, inwieweit Shazam für ein bestimmtes Lied verwendet wird: Wenn Hörer ein Stück eines Liedes zu Shazam hochladen, um zu wissen, wer es singt, tun sie dies in der Regel, weil sie es für ein gutes Lied halten. Und wenn die Zahl der Menschen, die dasselbe Lied wiedererkennen wollen, innerhalb weniger Tage dramatisch ansteigt, könnte das Lied ein Hit werden.



THE WINNER IS

Beim Eurovision Song Contest flammt die Diskussion über Vetternwirtschaft jedes Jahr wieder auf. Eine Analyse zeigt, wie fair die Vergabe der Punkte tatsächlich ist.

Wie fair ist der Gesangswettbewerb?

Jedes Jahr schicken Dutzende von Ländern einen Song zum Eurovision Song Contest (ESC). Die Länder geben sich gegenseitig Punkte und am Ende des Abends gibt es einen Gewinner. Danach wird oft über diese Punkte gemurrt, besonders wenn das eigene Land nicht gewonnen hat. Das ist nicht fair: Die Ostblockländer stimmen nur füreinander, genau wie die Skandinavier. Das ist alles Vetternwirtschaft.

Berüchtigt sind auch Griechenland und Zypern, die sich gegenseitig immer viele Punkte geben. Im Jahr 2003 zum Beispiel belegte Griechenland mit 'Never let you go' (wer erinnert sich noch?) den 17. von 26 Plätzen. Die Ballade erhielt insgesamt fünfundzwanzig Punkte, aber Zypern gab, wie jedes Jahr, das Maximum von zwölf Punkten an seinen Nachbarn. Das muss doch abgekartet sein?

Was ist fair?

Der Mathematiker Michel Vellekoop und die Ökonometrikerin Laura Spierdijk entdeckten an der Kaffeemaschine der Universität Twente, dass sie beide das Liederfestival fanatisch verfolgten. Sie kamen auf die sogenannte ungerechte Verteilung der Punkte zu sprechen. Sie fragten sich, ob die diesbezüglichen Beschwerden wahr seien. Sie beschlossen, auf saubere, wissenschaftliche Weise herauszufinden, ob die Punktevergabe fair war⁵⁷.

Ein Problem war, dass die Qualität eines eingereichten Liedes nicht bekannt ist. Vellekoop und Spierdijk lösten dieses Problem, indem sie die durchschnittliche Punktzahl eines Liedes als Maßstab für die Qualität heranzogen. Dann untersuchten sie, wie einzelne Länder von diesem Durchschnitt abweichen, mit anderen Worten, wie stark ihre

Präferenzen sind. Vellekoop: „Die Niederlande geben im Durchschnitt 2,5 Punkte zu viel an Belgien und 1,6 Punkte zu wenig an Malta. Natürlich ist es schön zu sehen, wer sich gegenseitig viele Punkte gibt, aber um etwas über die zugrunde liegende Struktur zu sagen, muss man mehr tun und ein Modell formulieren.“

Viele vorhandene Artikel betrachten nur die aggregierten Daten. Aber wenn man die Statistiken aller Länder zusammen betrachtet, sieht man manchmal Dinge, die auch durch andere Variablen erklärt werden können. Zum Beispiel können zwei Nachbarländer einander viele Punkte geben, weil sie den gleichen Musikgeschmack haben. Vellekoop und Spierdijk beschlossen daher, die verschiedenen Länder getrennt zu betrachten und zu bestimmen, welche Variablen ihre Präferenzen am besten erklären. „Was wir zu erwähnen hatten, ergab sich ganz direkt aus der Folklore. Geben sich Nachbarländer gegenseitig mehr Punkte? Spielt Religion eine Rolle? Wie wichtig ist Sprache?“

Diese Variablen genau zu definieren, erwies sich als eine Menge Arbeit. Wie messen Sie, wie nahe zwei Länder einander stehen? Und gelten Länder mit einem Meer dazwischen als Nachbarn? Wenn Großbritannien ein Nachbar der Niederlande ist, dann ist es auch ein Nachbar Italiens. Die Meeresnachbarn fielen schließlich in das Modell und für die Entfernung zwischen den Ländern wurde die Entfernung zwischen ihren Hauptstädten genommen. Es war auch komisch, dass es bereits viele Definitionen für andere Dinge gab. Es stellte sich zum Beispiel heraus, dass es Messsysteme gab, die anzeigen, wie viele zwei Sprachen einander ähneln.

Eine Aussagekräftige Theorie

Vellekoop war überrascht, als er entdeckte, wie sie all diese Daten kombinieren konnten: „Ich hatte immer das Gefühl, dass man mit Abhängigkeiten vorsichtig sein muss und dass es ein Problem ist, dass wir alle Arten von Varianzen und Skalen austauschbar verwenden. Aber dann stellte sich heraus, dass es ein Theorem gibt, das besagt, dass man durch ein paar sehr schwache Forderungen immer noch eine Annäherung an den guten Parameter erreicht und eine ziemlich konsistente Schätzung all dieser Abweichungen vornehmen kann. In der Ökonometrie ist dies alles sehr gut bekannt, aber als Mathematiker ist es sehr schön zu sehen, wie mächtig die Theorie sich in einem anderen Bereich erweist.“

Ihr Modell zeigte, welche Variablen in jedem Land einen großen Einfluss haben. Zum Beispiel geben die Niederlande einem Solosänger, der auf Französisch singt, viele Punkte. Vellekoop bezeichnet die Vorliebe für eine Sprache nicht als unfair: „Man kann nichts dagegen tun, wenn man Französisch schöner findet.“ Übrigens erwies sich diese Sprachpräferenz in Griechenland und Zypern als die stärkste. Sie bestrafen Länder, die in einer Sprache singen, die weit von ihrer eigenen entfernt ist. Dass diese Länder sich gegenseitig so viele Punkte geben, liegt nicht nur daran, dass sie Nachbarn sind. Die Sprache ist eine stärkere erklärende Variable.

Im Allgemeinen scheinen Nachbarn sich gegenseitig etwas mehr Punkte zu geben, aber die Variable „Nachbarland“ stach in den meisten Ländern nicht enorm hervor. In den skandinavischen Ländern scheinen sich die gegenseitigen Punkte besser durch ähnlichen Musikgeschmack als durch Vetternwirtschaft erklären zu lassen. Große Ausnahmen sind Estland, Lettland und Litauen. Vellekoop: „Sie lieben ihre Nachbarn sehr: Im Durchschnitt bekommen sie fast vier Punkte zu viel. Unsere allgemeine Schlussfolgerung war, dass die baltischen Staaten etwas über ihr Abstimmungsverhalten zu erklären haben.“

Analyse von Sozialen Netzwerken

Das US-amerikanische Filmdrama „The Social Network“ von 2010 beschreibt die Entstehungsgeschichte des sozialen Netzwerks Facebook. Die Analyse von Sozialen Netzwerken untersucht Verbundenheit von Gruppenmitgliedern mit gewissen Gemeinsamkeiten. Die Links beinhalten verwandtschaftliche Beziehungen, informationelle Verbindungen, oder physische Annäherungen. Das Ziel hierbei ist z.B. die Identifikation des Führers des Netzwerks oder die Rolle von Untergruppen (sog. Cluster) im Netzwerk. In der Realität könnten diese Netzwerke Gangstergruppen, Ko-Autorenschaft (Citation index)⁵⁸, Online-Netzwerke (Instagram, Facebook), Email-Kommunikation, Zuspield-Struktur einer Fußballmannschaft⁵⁹, usw. sein.

Um Netzwerke mathematisch zu analysieren, verwendet man *Graphentheorie* und *Zentralitätsmaße*. Ein Graph besteht aus einer Menge von Punkten (sog. *Knoten*), meist Personen. Jeweils zwei Knoten sind verbunden (oder auch nicht) durch eine *Kante*. Die Kanten zeigen den Informationsfluss innerhalb der Gruppe. Hier, für den Gesangswettbewerb ESC, betrachten wir gerichtete, einfache, vollständige Graphen. Man benötigt spezielle Größen, um die Personen im sozialen Netzwerk bewerten zu können.

Zentralitätsmaße

Die Analyse eines Netzwerks basiert auf Zentralitätsmaßen; sie sagen etwas aus über die 'Wichtigkeit' von Akteuren im Netz, wie z.B. das Ausmaß, in dem die Akteure Beziehungen auf sich konzentrieren oder die Möglichkeit für Akteure, die Beziehungen anderer durch ihre Lage im Netz stören zu können. Aus den Zentralitätsmaßen der Akteure kann man auf die Machtverteilung im Netz schließen. Die wichtigsten drei Maße sind

- Grad-Zentralität: direkte Verbindungen im Netz
- Closeness-Zentralität misst, wie schnell ein Knoten einen anderen erreichen kann
- Betweenness-Zentralität: Welches Potential hat ein Akteur zur Kontrolle des Interaktionsflusses innerhalb des Netzes?

Grad-Zentralität

Die Grad-Zentralität ist die Summe der direkten Beziehungen eines Akteurs. Diese Anzahl der direkten Verbindungen eines Akteurs mit anderen gilt als Maßzahl für die Aktivität eines Akteurs im Netz: je mehr Beziehungen ein Akteur hat, desto zentraler ist er. Allerdings sind direkte Verbindungen alleine nicht entscheidend. Ein Pfad ist ein Weg über mehrere Kanten, der 2 Ecken verbindet und die Pfadlänge ist die Anzahl der enthaltenen Kanten. Die Distanz zwischen A und B, $d(A,B)$, ist die Länge eines kürzestmöglichen Pfades von A nach B, ein solcher Pfad heißt *geodätisch*.

Closeness-Zentralität

Die Closeness-Zentralität ist die 'Autonomie im Netz'; sie misst die Schnelligkeit der Interaktion im Netz. Hierbei werden indirekte Verbindungen über Dritte mit einbezogen. Die Closeness-Zentralität beschreibt die Schnelligkeit, mit der es einem Akteur möglich ist, jeden anderen Akteur zu erreichen: je weniger Umwege er über Dritte machen muss, desto zentraler ist der Akteur. Je näher ein Punkt zu allen übrigen steht, desto effektiver und unabhängiger kann er sich im Netz bewegen, er ist weniger angewiesen auf die Bereitschaft der anderen zur Kooperation. Dieser *Closeness-Score* von Knoten C berechnet sich als $C=1/d(C,A) + 1/d(C,B) + \dots$, d.h. je kleiner die Distanz zu C, je größer der Kehrwert. Hier werden nicht nur unmittelbar benachbarte Knoten gewertet.

Betweenness-Zentralität

Zentral sind die Akteure, über die die meisten Verbindungen laufen, und dies soll mit der Betweenness-Zentralität gemessen werden. Wichtig sind hier die indirekten Beziehungen über Dritte, die ein



Abbildung 88: ESC 2010: Gerichteter Graph, der das Abstimmverhalten modelliert.

Akteur auf sich vereinigt. Ebenso steht die Möglichkeit der Kontrolle des Interaktionsflusses eines Akteurs als beteiligter Dritter im Vordergrund. Akteure sind umso mächtiger, je mehr kürzeste Verbindungswege sie zwischen anderen Punkten unterbrechen können. Der Betweenness-Score von C als Link zwischen A und B ist die Zahl der geodätischen Pfade von A nach B die durch C gehen geteilt durch die Zahl aller geodätischen Pfade von A nach B. Dann ist der gesamte Betweenness-Score von C die Summe aller Einzelberechnungen aller möglichen Knoten A, B.

Eurovision Song Contest (ESC)

Der ESC ist ein Netzwerk von Ländern mit geographischen/historischen Beziehungen, das untereinander Punkte austauscht. Man kann mit einem dynamischen Netzwerk: das Abstimmverhalten beim Song Contest über mehrere Jahre analysieren. Allerdings gibt es zwei Probleme: ständige Regeländerungen und sich ändernde Finalteilnehmer (bis auf die 'Big Five': Deutschland, Frankreich, Spanien, Großbritannien und Italien).

Im grundlegenden Bewertungssystem (seit 1975) vergibt jedes Land Punkte zu 10 anderen Ländern, z.B. Land A gibt die Punkte $\{1, 2, 3, 4, 5, 6, 7, 8, 10, 12\}$ an die 10 (besten?) Länder $\{B, C, D, E, F, G, H, I, J, K\}$. Die Länder dürfen sich nicht selbst wählen. Die Darstellung geschieht durch gerichtete, gewichtete, dynamische Graphen,

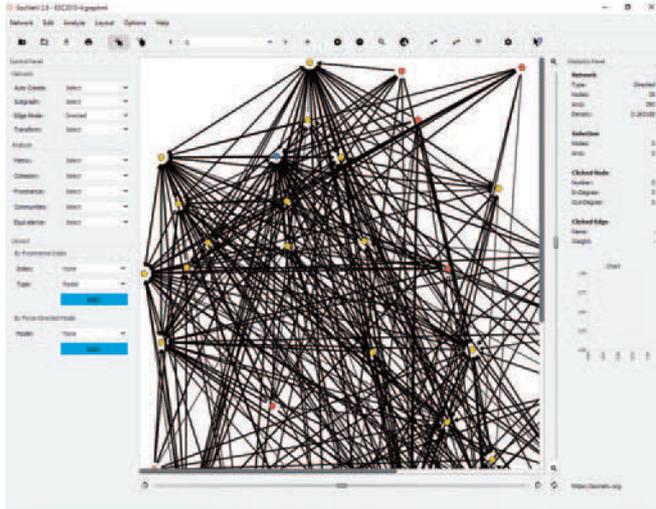


Abbildung 89: Software SocNetV mit dem Graphen des ESC 2010.

und zwar mit gerichteten Kanten für die Punktvergabe und gewichtet mit der jeweiligen Punktzahl.

Cliquenbildung beim Eurovision Song Contest

Meist haben in sozialen Netzwerken zwei Knoten mit einem gemeinsamen dritten Knoten eine höhere Wahrscheinlichkeit miteinander verbunden zu sein. Falls es Cliquen im ESC gibt, wird der beobachtete Clustering Koeffizient höher sein als der Koeffizient von einem 'Zufalls-Contest'. Falls zwei Länder beide für ein drittes Land stimmen, dann ist die Wahrscheinlichkeit höher, dass sie auch für sich gegenseitig stimmen. Der sog. *Clustering Koeffizient* C eines Knotens v ist die Wahrscheinlichkeit, dass zwei Nachbarn eines gegebenen Knotens v selbst Nachbarn sind. Zur Bestimmung von C zählt man Kanten zwischen den Nachbarn eines gegebenen Knotens v und teilt dies durch $d(v)!/2!(d(v)-2)!$. Das Ausrufezeichen hier bezeichnet die Fakultät. Die Mittelung von C über alle Knoten liefert den Clustering Koeffizienten des Netzwerks.

Analyse des Eurovision Song Contest 2010

Schliesslich wollen wir den ESC 2010 analysieren. Die Frage ist, ob Lena Meyer-Landrut durch Cliquen-Bildung gewonnen hat. Dazu benutzten wir die freie Software SocNetV, Version 2.9 und haben das gesamte Abstimmverhalten in einem Graphen visualisiert, siehe Abbildung 88.

Deutschland ist der blaue Knoten mit der Nummer 22. Die roten Knoten haben Deutschland Punkte gegeben, die gelben haben Punkte bekommen und/oder gegeben. Man kann sich nun mit der Software SocNetV die vielen Zentralitätsmaße einfach ausgeben lassen, z.B. „Prominence/Betweenness Centrality“ oder „Communities/Cliques“, siehe Abbildung 89. Bei letzteren hat Lena eine Cliquengröße 34, was hier ein unauffälliger Wert ist: sie hat den ESC somit fair gewonnen! Übrigens, zum Experimentieren wird das verwendete Datenfile ESC2010 für SocNetV auf der Webseite des Buches zur Verfügung gestellt.

-
- ⁵⁷ L. Spierdijk, M. Vellekoop, The structure of bias in peer voting systems: lessons from the Eurovision Song Contest, *Empirical Economics*, 36(2) (2009), 403-425.
- ⁵⁸ A.L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, T. Vicsek, Evolution of the social network of scientific collaborations, *Physica A: Statistical Mechanics and its Applications*, 311(3-4) (2002), 590- 614.
- ⁵⁹ C. Cotta, A.M. Mora, J.J. Merelo, C. Merelo-Molina, A network analysis of the 2010 FIFA world cup champion team play, *Journal of Systems Science and Complexity*, 26(1) (2013), 21-42.

Prof. Dr. Volker Mehrmann



Eine kurze Zusammenfassung Ihrer Karriere

Volker Mehrmann schloss sein Diplom in Mathematik 1979, sein Staatsexamen Mathematik/Physik 1980, seine Dissertation in 1982 und seine Habilitation 1987 an der Universität Bielefeld ab. Er verbrachte Forschungsjahre an der Kent State University 1979-1980, University of Wisconsin 1984-1985, und dem IBM Forschungszentrum Heidelberg 1988-1989. Von 1990-1992 vertrat er eine Professur an der RWTH Aachen und er hatte den Lehrstuhl für Numerische Algebra an der TU Chemnitz von 1993-2000. Seit 2000 ist er Universitätprofessor für Numerische Mathematik an der TU Berlin. Von 2008 bis 2016 war er Sprecher des DFG-Forschungszentrums MATHEON und des Einsteinzentrums ECMath. Er war Präsident der Gesellschaft für Angewandte Mathematik und Mechanik (GAMM) 2011-2013 und ist seit 2020 Präsident der European Mathematical Society (EMS).

Seine Forschungsinteressen sind Numerische Mathematik, Wissenschaftliches Rechnen, Angewandte und Numerische Lineare Algebra, Steuerungstheorie und Dynamische (insbesondere differentiell-algebraische) Systeme, sowie in den letzten Jahren die energiebasierte mathematische Modellierung insbesondere mit Hilfe Port-Hamiltonscher Systeme.

Was hat Mathematik für Ihre Karriere bedeutet?

Mathematische Strukturen und deren Nutzung in der Entwicklung, Analysierbarkeit und Implementierbarkeit von effizienten numerischen Methoden für die industrielle Praxis, war und ist immer noch während meiner ganzen Karriere der Antrieb für meine Forschung und Lehre. Es hat sich über die Jahre ergeben, dass dies nicht nur eine sehr erfolgreiche Strategie für die Konstruktion von effizienten und robusten numerischen Methoden, sondern auch eine wunderbare Motivation für mathematische Grundlagenforschung. Als Beispiel seien die Entwicklungen zu strukturhaltenden Linearisierungen von Matrixpolynomen, geboren aus einem Industrieproblem hat diese Frage eine ganz neue Theorie initiiert mit hunderten mathematischer Publikationen. Einen weitgehenden Beleg findet man in den Büchern:

- M. Grötschel, K. Lucas und V. Mehrmann (Hrsg.), Produktionsfaktor Mathematik. Wie Mathematik Technik und Wirtschaft bewegt, acatech DISKUTIERT, acatech-Deutsche Akademie der Technikwissenschaften, Springer Verlag, Berlin, Heidelberg, 2008.
- P. Deuffhard, M. Grötschel, D. Hömberg, U. Horst, J. Kramer, V. Mehrmann, K. Polthier, F. Schmidt, C. Schütte, M. Skutella, and J. Sprekels (Editors), Matheon, Mathematics for Key Technologies Series in Industrial and Applied Mathematics 1, EMS Publishing House, Zürich, Switzerland, 2014.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihrer Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Da ich als Hochschullehrer in der Forschung und Lehre tätig bin, ist Mathematik ein zentraler Teil meines Leben.

Wie sieht ein typischer Arbeitstag aus?

Vorlesungen und Seminare vorbereiten und halten, Bachelor, Master und Doktorarbeiten betreuen, akademische Selbstverwaltung, Bürokratiekram mit der Uni-Verwaltung regeln, Forschungsprojekte beantragen und durchführen, mit Kolleg:innen Forschungsfragen diskutieren, aktiv an mathematischen Fragestellungen arbeiten, Ergebnisse für die Publikation aufbereiten, Publikationen schreiben, Vorträge vorbereiten und halten, Publikationen, Forschungsprojekte und Abschlussarbeiten begutachten, in Aufsichtsgremien von Forschungsinstituten Rat geben und vieles mehr.

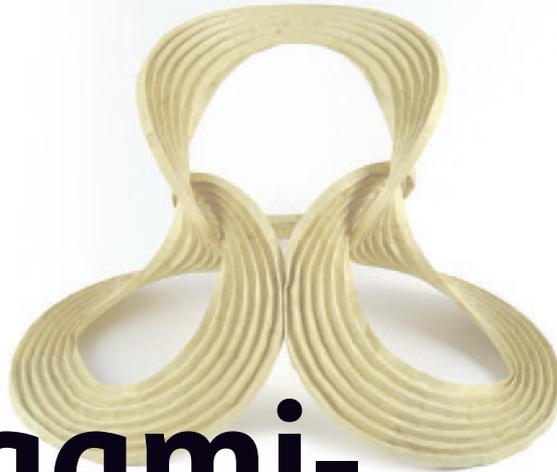
Sind, in Ihrer Hinsicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Hier ist erheblicher Nachholbedarf. Viele Mathematiker:innen wollen oder können eine Anwendung ihrer Forschungsergebnisse auf Wirtschaft oder Gesellschaft nicht angehen. Das ist auf jeden Fall legitim, denn die theoretische Grundlagenforschung ist sehr wichtig, da sich die Mathematik ohne diese nicht weiterentwickelt und dann langfristig auch die Anwendungen nicht gemacht werden können. Allerdings fehlt oft die Kommunikation über die erzielten Ergebnisse, um die Brücken zu schlagen und oft erlebe ich auch eine gewisse Überheblichkeit gegenüber den Kolleg:innen die diese Brücken schlagen. Diese ist in keinem Fall zu rechtfertigen.

Was raten Sie jungen Leuten hinsichtlich Mathematik?

Wenn man/frau Mathematik und intellektuelle Herausforderungen mag, ist es das tollste was man/frau machen kann. Mathematik durchzieht unsere Gesellschaft in allen Bereichen, Mathematiker:innen werden überall gebraucht, die Karriere-Chancen sind hervorragend. Aber es sollte einem Freude machen, dann macht es auch glücklich.

Die Mathematik hinter der Faltung von Origami-Mustern inspiriert nicht nur schöne Kunst, sondern führt auch zu einem besseren Verständnis der Faltung von Proteinen, den Arbeitspferden im menschlichen Körper.



Origami- Mathematik im Museum für Moderne Kunst

Abbildung 90: „Computational Origami“ von Erik und Martin Demaine.
Diese drei Origami-Figuren sind Teil der ständigen Sammlung des Museum of Modern Art in New York. Quelle: Erik und Martin Demaine.

Origami ist eine alte japanische Kunstform, die auf dem Falten von Papier basiert. Die Origami-Mathematik gibt es erst seit etwa vierzig Jahren. Die Origami-Mathematik beschreibt die Regeln des Origami auf formale Weise und entdeckt, was möglich ist und was nicht. Mit Origami-Mathematik versuchen Mathematiker zu verstehen, wie man aus einem flachen Stück Papier ein beliebiges dreidimensionales Objekt falten kann: einen Frosch, einen Kranich, einen Menschen oder eine abstrakte, geometrische Figur. Es kann alles sein.

Die ultimative Herausforderung besteht darin, eine Berechnungsmethode zu finden, die Ihnen sagt, wie eine beliebige dreidimensionale Struktur am besten gefaltet werden kann. Und „am besten“ bedeutet so etwas wie mit möglichst wenig Falten und einem möglichst kleinen Stück Papier.

Origami-Schleifen

Der amerikanische Mathematiker Erik Demaine (1981) nutzt die Mathematik, um schöne neue Formen aus Papier zu falten. Zusammen mit seinem Vater Martin, der Künstler ist, entwarf und faltete er drei Origami-Formen – „Computational Origami“ genannt -, die man in der ständigen Sammlung des Museum of Modern Art in New York bewundern kann. Jedes Formular verbindet mehrere kreisrunde Papierstücke zu einem faszinierenden *Looping*. Folgen Sie dem *Looping* phantasievoll mit dem Zeigefinger, und Sie werden feststellen, dass der Gesamtdrehwinkel viel größer als 360° ist, normalerweise etwas zwischen 720° und 1080° .

Diese Form des Origami unterscheidet sich vom traditionellen Origami dadurch, dass die Formen tatsächlich aus unendlich vielen unendlich kleinen Faltungen bestehen. Nur so lässt sich eine überall gekrümmte Oberfläche erzeugen. Erik Demaine: „Die zugrunde liegende mathematische Frage ist, wie die Mathematik ein Berechnungsrezept liefern kann, das angibt, welche gekrümmten Formen theoretisch gefaltet werden können.“

Origami-Mathematik hilft nicht nur Künstlern. Es wird in allen Bereichen angewandt, die mit dem Falten zu tun haben: zum Beispiel in der Robotik, aber auch in der Architektur, der bildenden Kunst und der Computergrafik. Zum Beispiel arbeitet der Origami Künstler Robert Lang an einem faltbaren Teleskopobjektiv für den Einsatz im Weltraum. Dort muss man in der Lage sein, die Linse auf eine Ebene von hundert Metern Durchmesser zu falten, aber zusammengeklappt in einem Raumschiff will man ein Paket nicht mehr als zehn

Meter breit halten. Wie macht man das auf die bequemste Art und Weise?

Gefaltete Proteine in Ihrem Körper

Ein weiteres wichtiges praktisches Problem, an dem Demaine mit Origami-Mathematik arbeitet, ist das Problem der Proteinfaltung. Proteine sind die Arbeitspferde des Körpers. Es handelt sich um komplexe, dreidimensionale Moleküle, die auf eine bestimmte Art und Weise zusammengefaltet sind. Diese Faltung bestimmt zu einem großen Teil ihre Funktion. Bei allen Arten von Krankheiten geht bei der Proteinfaltung etwas schief, und Medikamente können das mit speziell entwickelten Proteinen reparieren.

Als Origami-Mathematiker stellt Demaine die Frage, auf welche Art und Weise man eine Proteinform in eine andere falten kann: „Mathematisch gesehen ist das eine Art eindimensionales Origami, denn man kann ein langes Molekül als eine Linie betrachten. Die Frage ist, ob wir eine Berechnungsmethode für den bequemsten Weg der Proteinfaltung finden können. Wir hoffen, dass wir, wenn wir diese mathematische Berechnungsmethode gefunden haben, automatisch Einblick in die natürlichen und chemischen Prinzipien erhalten, die bestimmen, wie sich ein Protein faltet.“

Vater und Sohn Demaine experimentieren viel mit Papier, um ein mathematisches Origami-Problem zu lösen. Sohn Erik: „Das Falten eines Origami-Objektes mit den Händen bildet die Intuition dafür, was möglich ist und was nicht. Es ist, als ob man durch das Zusammenfalten im Kopf eine große experimentelle Datenbank aufbaut. Sie verwenden diese Datenbank, um das mathematische Rätsel zu lösen. Übrigens verwenden alle bekannten Origami-Designer irgendeine Form der Origami-Mathematik. Nur lassen sie nicht einen Computer rechnen, sondern sie rechnen unbewusst im Kopf, auf intuitive Weise.“

E.D. Demaine, J. O'Rourke. *Geometric folding algorithms: linkages, origami, polyhedra*, Cambridge University Press, 2007.

Liebeslyrik: Petrarca und Laura – Chaos in Liebesaffären

Dynamische Systeme
beschreiben Liebesbeziehungen
und helfen so Gedichte aus dem
Mittelalter zu sortieren.

In diesem Kapitel wollen wir diskutieren, inwieweit man Liebe als ein (deterministisches) mathematisches Modell betrachten kann. Obwohl Liebe häufig als etwas spontan gesteuertes, ohne jegliche Struktur beschrieben wird („Liebe auf den ersten Blick“), wollen wir ein Modell aufstellen und es anhand von berühmten Beispielen von Beziehungen wie Romeo und Julia, Petrarca und Laura oder Jack und Rose (Film 'Titanic') kalibrieren. Dabei werden Gefühle wie Angst, Zuneigung, Sehnsucht, Verlangen oder auch Abneigung betrachtet und Liebe als eine Reaktion auf dieses definiert.

Steven Strogatz gilt als einer der Pioniere auf diesem Gebiet: er formulierte ein System von Differentialgleichungen, um die Dynamik der Liebe zu beschreiben. Was zunächst nur als amüsante Motivation für seine Studenten gedacht war, entwickelte sich rasch zu einem eigenen Gebiet der sog. dynamischen Systeme⁶⁰.

Das Grundmodell

Das Modell von Strogatz beschreibt unter Berücksichtigung der Persönlichkeiten die Interaktion zweier Menschen, und zwar im folgenden 2 Größen, die Liebenden. Die eigene Persönlichkeit spielt eine Rolle bei der Reaktion auf die Gefühle des anderen; zum Beispiel kann ein narzisstisch veranlagter Mensch die Gefühle des Partners nicht erwidern. Trotzdem liebt er das Gefühl, selbst geliebt zu werden. Eine selbstbewusste Person würde immer auf Annäherungen des Geliebten mit Annäherungen reagieren.

Auch die eigene Lebensgeschichte kann diese Reaktionen beeinflussen; wurde man von der ersten Liebe enttäuscht oder gekränkt, ist man bei der nächsten Beziehung etwas vorsichtiger und reagiert nicht übermäßig auf die Gefühle des anderen.

In diesem Modell werden zeitnahe Reaktionen angenommen; es können z.B. keine Fernbeziehungen simuliert werden, da hier ein zeitlicher Verlust der Gefühle stattfindet. Liebe ist im Modell die direkte Reaktion auf eine Aktion des Gegenübers, die wieder zu einer Reaktion und zu einer Aktion usw. führt.

Es können auch Gefühlszustände simuliert werden, die eine gewisse Zeit brauchen, um zu entstehen oder zu verschwinden, z.B. Depressionen oder emotionale Taubheit. Die wichtigste Annahme ist, dass die simulierten Personen Gefühle lesen und interpretieren können. Sonst wäre eine Reaktion nicht möglich und das Modell sinnlos. Wir

erwähnen, dass das Modell nicht das allgemeine Bauchgefühl, sondern nur die äußeren Reaktionen berücksichtigt.

Die zwei folgenden Modelle sind linear, also eher einfach gehalten, damit eine weitere Analyse noch möglich ist. Sie haben die Form

$$\dot{X}(t) = \alpha_1 X(t) + \beta_1 Y(t),$$

$$\dot{Y}(t) = \beta_2 X(t) + \alpha_2 Y(t).$$

Dabei beschreiben $X = X(t)$ und $Y = Y(t)$ die Gefühle der zwei liebenden Personen dar: Liebe bei positiven und Abneigung bei negativen Werten. Die Faktoren α_1, α_2 beschreiben den Einfluss der eigenen und β_1, β_2 den der Gefühle des anderen auf die aktuelle Liebe. Der Punkt über X, Y steht für die zeitliche Ableitung.

Romeo und Julia

Dieses Modell wurde von Strogatz zur Motivation seiner Studenten benutzt. Es beschreibt sehr stark vereinfacht den Verlauf der Liebesbeziehung von Romeo (X) und Julia (Y) nach dem Theaterstück von William Shakespeare: Julia wird von Romeo umworben, allerdings sind ihre Familien verfeindet und verbieten diese Liebe. Schließlich brennen beide durch und es kommt zum tragischen Tod beider. Man könnte daher als Parameter nehmen:

$$\alpha_1 = 0.05, \quad \alpha_2 = -0.3, \quad \beta_1 = \beta_2 = 0.1, \quad X(0) = 9, \quad Y(0) = 3.$$

Die positiven β_1, β_2 beschreiben die starke Reaktion auf die Gefühle des anderen. Romeo ist ein junger Mann, der durch die Gefühle beider nur angespornt wird, also ein positives α_1 . Dagegen hat Julia ein negatives α_2 , reagiert also negativ auf die eigenen Gefühle, da sie als junges Mädchen ihre Familie nicht enttäuschen will und somit ihre eigenen Gefühle nicht akzeptiert. Die Startwerte $X(0), Y(0)$ wurden unter der Berücksichtigung der physischen Attraktivität gewählt. Das Zeitintervall ist angelehnt an das Stück 30 Tage.

Man sieht in Abbildung 91, dass sich bei der einfachen Modellierung die Liebe sehr linear entwickelt. Sie steigt nach ihrer ersten Begegnung stetig an und die beiden Liebenden verstärken sich gegenseitig. Es gibt keine Höhen und Tiefen in der gegenseitigen Zuneigung, kein Schwanken einer der Parteien oder sonstige wirkliche Interaktion, wie sie in Shakespeares Stück vorkommen.

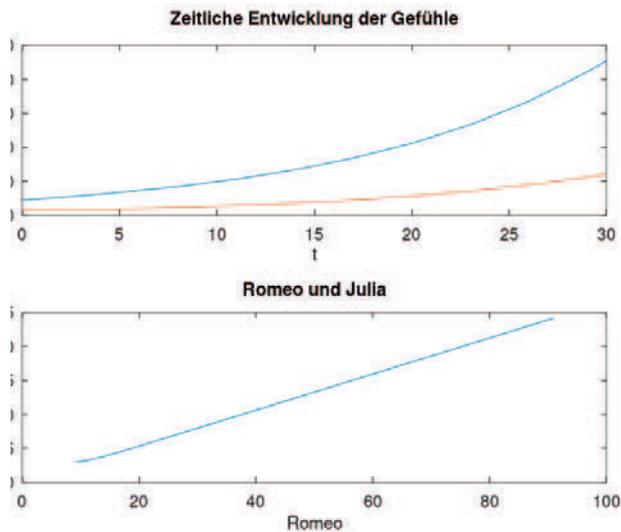


Abbildung 91: Zeitliche Entwicklung der Gefühle (links) und Romeo vs. Julia (rechts).

Um einen näheren Liebesverlauf an das Theaterstück zu haben, kann man das Modell wie folgt erweitern:

$$\dot{X}(t) = \alpha_1 X(t) + \beta_1 Y(t) + r_1 A_2 - bY(t - \tau),$$

$$\dot{Y}(t) = \beta_2 X(t) + \alpha_2 Y(t) + r_2 A_1.$$

Dabei bezeichnen A_1, A_2 Konstanten für die Attraktivität des jeweiligen Partners. Die Faktoren r_1, r_2 beschreiben die jeweilige Reaktion auf die Attraktivität des anderen. Der Term $bY(t - \tau)$ ist ein Verzögerungsterm; er beschreibt eine zeitversetzte Reaktion Romeo auf Julias Gefühle ('Überdenken').

Aus dem Theaterstück entnehmen wir: die Attraktivität von Julia ist höher als die von Romeo ($A_2 > A_1$), die überdachte Reaktion Romeos auf Julias Gefühle ist kleiner als die direkte Reaktion ($b < \beta_1$). Das Modell simuliert auch den Zustand, wenn Romeo und Julia nicht gestorben wären. Sie hätten sich sicherlich nach relativ kurzer Zeit getrennt und das wäre kein 'schönes' Ende für eine der bekanntesten Liebestragödien gewesen.

Petrarca und Laura

Das zweite Modell wurde von Sergio Rinaldi entwickelt⁶¹. Die Motivation dabei war, mithilfe eines Liebesmodells die ungeordneten Gedichte des berühmten Dichters Francesco Petrarca aus dem Mittelalter zeitlich zu ordnen. Diese Gedichte resultieren aus der Inspiration der unerwiderten Liebe zu Laura.

Beschreiben wir zunächst die zwei Personen. Petrarca war ein berühmter italienischer Dichter des 14. Jahrhunderts. Er war 23 Jahre alt und ein wenig attraktiver Mann, als er Laura zum ersten Mal traf. Wie viele Dichter zu der Zeit, war er gesellschaftlich nicht hoch angesehen. Dagegen war die 16-jährige Laura eine sehr attraktive, jungverheiratete Frau und hatte einen hohen gesellschaftlichen Stand.

Aus ihrer ersten Begegnung entstand eine 27-jährige unerwiderte Liebe von Petrarca zu Laura, die die literarische Inspiration von Petrarca lieferte und er begann Liebesgedichte zu schreiben. Trotz des Todes von Laura nach 27 Jahren, war Petrarca noch darüber hinaus mehrere Jahre inspiriert und schrieb insgesamt rund 200 Gedichte an sie.

Das resultierende Modell von Rinaldi basiert auf dem Grundmodell, das um eine dritte Gleichung ergänzt wurde:

$$\dot{L}(t) = -\alpha_1 X(t) + \beta_1 A_P + R_L P(t),$$

$$\dot{P}(t) = -\alpha_2 P(t) + \beta_2 \frac{A_L}{1 + \delta Z(t)} + \beta_2 L(t),$$

$$\dot{Z}(t) = -\alpha_3 Z(t) + \beta_3 P(t).$$

Dabei stehen L und P für die Liebe von Petrarca und Laura. Die Faktoren $\alpha_i, \beta_i, A_P, A_L$ sind analog zu denen aus dem Romeo-Julia-Modell definiert. Dabei ist nun α_i der Vergessensfaktor. Damit ist gemeint, dass Liebe nach einiger Zeit ohne Auffrischung abebbt. Weiter stehen diesmal β_1 für die Reaktion auf Petrarcas Attraktivität und β_2 für die Reaktion sowohl auf die Attraktivität als auch auf die Inspiration und Liebe Lauras. Die dritte Gleichung beschreibt die Inspiration Petrarcas (Z) und geht linear aus den beiden Liebesfaktoren von Laura und Petrarcas hervor.

Die Reaktionsfunktion R_L in der ersten Gleichung, also die Reaktion von Laura auf Petrarcas Liebe, ist nicht-linear zu beschreiben. Man kann sie aus folgendem Zusammenhang herleiten: Sie fühlt sich zum

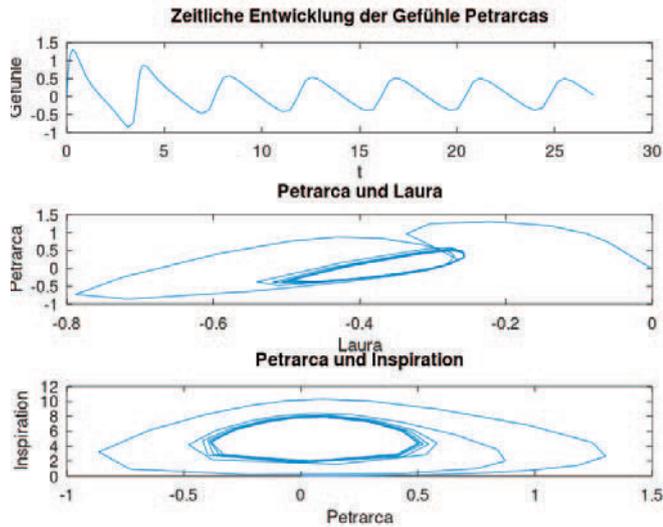


Abbildung 92: Zeitliche Entwicklung der Gefühle Petrarcas und Lauras.

einen von seine Liebeleien geschmeichelt. Aber wie es zu der Zeit und dem Stand üblich war, musste eine verheiratete Frau irgendwann diese Schmeicheleien negativ auffassen. Anders herum, wenn die Liebe von Petrarca in Verzweiflung umschlägt, empfindet sie irgendwann Mitleid. Dieses Mitleid führt dazu, dass sie positiv auf seine Schmeicheleien reagiert. Die daraus resultierende Reaktionsfunktion sieht wie folgt aus

$$R_L(P) = \beta_1 P \left(1 - \left(\frac{P}{\delta} \right)^2 \right)$$

Die zweite Gleichung wurde im Vergleich zum Grundmodell noch um die Inspiration (Z) erweitert. Um die Konstanten anzupassen, ist es wichtig die datierten Gedichte zu verstehen. Aus der Interpretation von Rinaldi ergibt sich folgender Zusammenhang für den Vergessensfaktor α_i . Er ist bei Laura dreimal so hoch wie bei Petrarca ($\alpha_1 = 3 \alpha_2$). Währenddessen die Inspiration zehnmal so lange anhält ($\alpha_2 = 10 \alpha_3$). Die Liebesreaktionsfaktoren β_i lässt sich auf ähnliche Weise beschreiben. Petrarca reagiert auf Lauras Liebe fünfmal so heftig wie sie ($\beta_2 = 5 \beta_1$). Die Inspiration ist sogar zehnmal so stark ($\beta_3 = 10 \beta_1$). Das resultiert aus der einseitigen Liebe. Daraus ergeben sich die Anfangswerte $L(0) = P(0) = Z(0) = 0$ und die folgenden Parameter

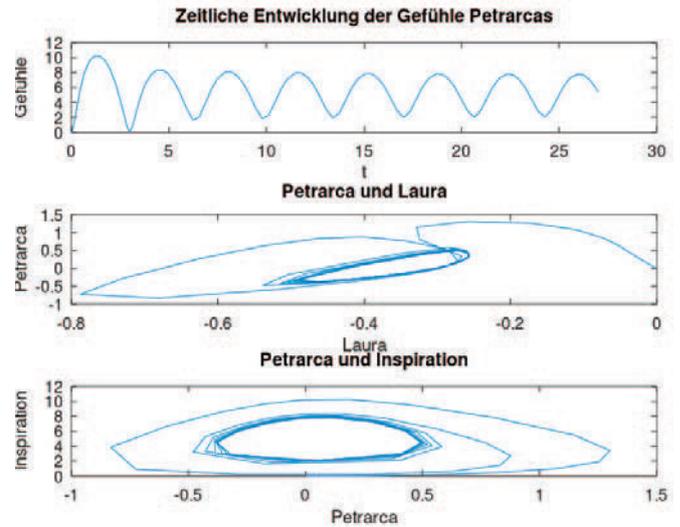


Abbildung 93: Gefühle Petrarcas und Lauras nach Fit an Gedichte.

$$\alpha_1 = 3, \quad \alpha_2 = 1, \quad \alpha_3 = 0.1, \quad \beta_1 = 1, \quad \beta_2 = 5, \quad \beta_3 = 10,$$

$$A_P = 1, \quad A_L = 2, \quad \delta = 1.$$

Die gesamte Laufzeit des Modells ist die Lebenszeit von Laura, also 27 Jahre. Die Werte wurden so gewählt, dass der Liebesverlauf von Petrarca analog wie die Analyse der datierten Gedichte verläuft. Es wird deutlich, dass das Modell die Beziehung aus der Sicht von Petrarca gut einfängt. Desweiteren sieht man, dass sich Petrarca und Laura im einem Wirbel der Gefühle befinden. Wie man in Abbildung 92 erkennen kann, ist die Liebe beider in einem Kreislauf (sog. LP -Kreis) gefangen.

Die Gefühlswelt von Petrarca schwankt von absoluter Liebe hin zu absoluter Verzweiflung und wieder zurück zu einem Liebeshoch. Dementsprechend entwickelt sich auch Petrarcas Inspiration. Man kann nach der Sortierung durch die Inhalte der Gedichte die Parameter des Modells noch anpassen (siehe Abbildung 93).

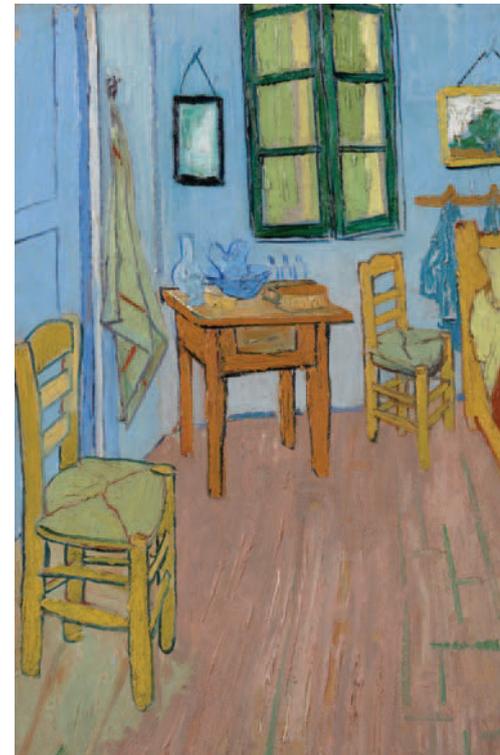
⁶⁰ S.H. Strogatz, Love affairs and differential equations, *Mathematics Magazine*, 61(1) (1988), 35-35.

⁶¹ S. Rinaldi, Laura and Petrarch: An intiguing case cyclical love dynamics, *SIAM J. Appl. Math.* 58(4) (1998), 1205-221

Der Zahn der Zeit

Die Gemälde von Vincent van Gogh haben sich im Laufe der Jahre langsam verfärbt. Digitale Rekonstruktionen zeigen, wie sie wahrscheinlich ursprünglich ausgesehen haben und wie sie verblasen werden.

Abbildung 94: In der Mitte 'Das Schlafzimmer', wie es heute aussieht. Links die Rekonstruktion dessen, wie die Leinwand ursprünglich aussah. Rechts die Prognose, wie die Leinwand nach Einwirkung einer bestimmten Lichtmenge in der Zukunft aussehen könnte. Die Rekonstruktionen wurden vom Van Gogh Museum (Vincent van Gogh Foundation), Amsterdam, in Zusammenarbeit mit RCE (Cultural Heritage Agency), Amsterdam, und Professor R.S. Berns, Munsell Color Science Laboratory, Rochester, NY, durchgeführt.



1888 beschrieb Vincent van Gogh das Gemälde, an dem er arbeitete, in einem Brief:

„Diesmal ist es nur mein Schlafzimmer. Aber hier muss die Farbe reichen. [...] Jedenfalls muss das Sehen dieses Gemäldes dem Kopf, oder besser gesagt, der Phantasie Ruhe geben. Die Wände sind blassviolett. Auf dem Boden liegen rote Fliesen. Das Holz des Bettes und der Stühle ist gelb wie frische Butter.“⁶²

Wenn Sie sich nun jedoch das Gemälde 'Das Schlafzimmer' im Amsterdamer Van Gogh-Museum ansehen, sehen Sie blaue Wände statt violetter. Ella Hendriks, leitende Konservatorin an diesem Museum und außerordentliche Professorin für Konservierung und Restaurierung an der Universität Amsterdam, erklärt, dass einige der Pigmente verschwunden sind. „Die Farbe, die Van Gogh kaufte, war nicht immer von guter Qualität. Zum Beispiel benutzte er Geranium Rot, einen sehr kurzlebigen synthetischen roten Farbstoff.“

Sammeln von Informationen

Hendriks war für die Restaurierung des Schlafzimmers verantwortlich und setzte ein Arsenal von Techniken ein, um die ursprünglichen Farben herauszufinden. „Wir hatten natürlich die Briefe Van Goghs, und auch alte Reproduktionen geben Auskunft. Wir wussten auch, dass das Klebeband unter dem Rahmen bei einer früheren Restaurierung im Jahr 1931 angebracht worden war. Die Farben darunter waren also so, wie das Gemälde zu der Zeit aussah.“

Die Erforschung der Farben der Vergangenheit war in hohem Maße multidisziplinär. Mit Hilfe von Mikroskopen wurde in den Tiefen der Farbschicht nach Pigmentkörnern gesucht. Scans zeigten, welche Elemente sich in den Farbschichten befanden. Hendriks: „Die Farbe des Geranium See enthielt das Element Brom, und wo wir diese Partikel gefunden haben, war früher diese rote Farbe.“ Im Labor wurde die Zusammensetzung der Farbe des 19. Jahrhunderts so gut wie möglich imitiert und künstlich gealtert.



Dadurch erhielten Hendriks und ihr Team eine Menge Informationen, aber noch kein vollständiges Bild davon, wie die Leinwand aussah. Und genau das war es, worum es in der Mathematik ging. Dieser Teil wurde in Zusammenarbeit mit dem amerikanischen Farbwissenschaftler Roy Berns durchgeführt. Er benutzte die Kubelka-Munk-Methode, benannt nach zwei deutschen Wissenschaftlern, die 1931 ein erstes theoretisches Modell für die Optik von Farbschichten erstellten – zufälligerweise ist ihr Werk so alt wie die frühere Restaurierung von 'Das Schlafzimmer'. Sie verwenden Differentialgleichungen, um zu beschreiben, wie Farbschichten einer bestimmten Dicke Licht streuen und absorbieren. Mit anderen Worten: wie ein aus verschiedenen Farben zusammengesetztes Gemälde aussieht.

Vergangenheit und Zukunft

So kam die Rekonstruktion des Gemäldes zustande, wie es ausgehen haben muss, als Van Gogh es malte. Und dann waren die Wände tatsächlich nicht blau, sondern färbten sich violett. Hendriks: „Van Gogh wird oft als Maler von Blau und Gelb bezeichnet, aber in Wirklichkeit verwendete er hauptsächlich Violett und Gelb. Deshalb ist diese Rekonstruktion für Kunsthistoriker so wichtig. Ich verstehe jetzt viel mehr darüber, wohin Van Gogh gehen wollte, sonst hätte ich es nicht gewusst.“

Übrigens wurde bei der Restaurierung der Leinwand nicht versucht, die Farben in ihren ursprünglichen Zustand zurückzusetzen. Hendriks: „Ich werde die Pinselstriche Van Goghs nicht mit meiner Farbe überdecken. Sie wollen die Authentizität seines Werkes so weit wie möglich bewahren. Darüber hinaus bleibt es zum Teil reine Vermutung: Wir wissen nie genau, wie es war, als er es gemalt hat.“ Das Museum entschied sich daher dafür, den Besuchern in einer digitalen Rekonstruktion zu zeigen, wie sich die Farben im Laufe der Jahre verändert haben.

Der Vorteil einer solchen digitalen Rekonstruktion ist, dass man sie auch in die Zukunft ausdehnen kann. Was passiert, wenn das Gemälde täglich mehrere Stunden lang bei einer bestimmten Lichtintensität zu sehen ist? Nach und nach wird der Boden immer kühler und die roten Linien im Bett verschwinden ganz. Hendriks machte es zu einem Quiz für die Museumsleitung: „Wie viel Schaden wollen wir akzeptieren? Und wie lange kann es dauern, bis die Leinwand so aussieht? Als Konservatoren wussten wir bereits, was vor sich ging, aber jetzt können wir die Schäden sichtbar machen. Diese Bilder machen viel mehr Eindruck als nur Worte. Das Museum zeigt

die Gemälde nun auf einem noch niedrigeren Lichtniveau, so dass die Verfärbung verlangsamt wird.“

Mit Mathematik kann man noch viel mehr, erklärt Prof. Dr. Ingrid Daubechies (Duke University) in ihrem Vortrag „The Master's Hand: Can Image Analysis Detect the Hand of the Master?“⁶³ Mittels sog. *Wavelets*, ein mathematisches Werkzeug, das für die Analyse und Komprimierung von Bildern (auch für das digitale Kino) verwendet wird, kann man entscheiden, ob ein Gemälde ein Original ist oder ob mehrere Teile eines Gemäldes von der gleichen Person gemalt wurden – oder nicht.

Darüberhinaus hilft Mathematik bei der Sichtbarmachung von übermalten Bildern. Diese werden zunächst mit Hochenergie-Röntgenstrahlen sichtbar gemacht, wobei allerdings großflächige Störstellen auf dem rekonstruierten Bild entstehen. Diese Störungen können mit der mathematischen Methode, der *Minimierung der totalen Variation*, entfernt werden.

⁶² Brief 705 an seinen Bruder Theo (Artes, 16. Oktober 1888) Vincent van Gogh: Die Briefe, Hrsg. Leo Jansen, Hans Luijten und Nienke Bakker. Für die Ausgabe der Website siehe: www.vangoghletters.org.

⁶³ <https://www.youtube.com/watch?v=-SyeN1hCYyA>



Wie man beim Beachvolleyball öfter gewinnt

Abbildung 95: Deutschland gegen Brasilien im Beach-Volleyball Finale der Olympischen Spiele, London 2012. Foto von Duncan Rawlinson (<https://duncan.co/>), ohne weitere Änderungen, Lizenz (CC BY-NC 2.0) <https://creativecommons.org/licenses/by-nc/2.0/deed.de>

Die Auswahl eines erfolgversprechenden Matchplans spielt im Spitzensport eine immer größere Rolle. Mathematische Optimierung und Spieltheorie kann diese Auswahl fundiert unterstützen.

Dr. Susanne Hoffmeister, Prof. Dr. Jörg Rambau und Ronan Richter (M.Sc.) vom Lehrstuhl Wirtschaftsmathematik der Universität Bayreuth berichten über Forschungen des Lehrstuhls zur Sport-Strategie-Optimierung im Beach-Volleyball. Dies war Gegenstand des Dissertationsprojekts von Susanne Hoffmeister (mittlerweile tätig als Operations Research Scientist bei der InVision AG am Standort Leipzig), in dem zum ersten Mal optimale Spielstrategien für ein Beach-Volleyball-Finale abhängig von der konkreten Final-Paarung ermittelt wurden.

Olympia-Finale im Beach-Volleyball. Wir treffen zum ersten Mal in diesem Turnier auf das gegnerische Team. Sollen wir bei allen Ballwechseln volles Risiko gehen, oder dürfen wir auch auf den einen oder anderen unerzwungenen Fehler der anderen hoffen? Sollten beide in unserem Team gleich aggressiv spielen?

Ob man bei einem Sportspiel gewinnt (Fußball, Tennis, Darts, . . .), hängt von den eigenen Fähigkeiten, den Fähigkeiten des gegnerischen Teams, dem Spiel-Glück, aber auch von der Qualität der strategischen und taktischen Entscheidungen ab. Laut den Olympiasiegern im Beach-Volleyball 2012 in London, Julius Brink und Jonas Reckermann, mit denen wir unmittelbar vor dem Verfassen dieses Beitrags gesprochen haben, ist ein auf das gegnerische Team zugeschnittener Matchplan eine zentrale Komponente, um die Gewinnchancen zu erhöhen. *Sportspiel-Strategieoptimierung* studiert Methoden, wie man Daten vergangener Spiele für das fundierte Aushecken eines Matchplans nutzbar machen kann. Solche Methoden kommen im Leistungssport in der letzten Dekade zunehmend als unterstützendes Instrument zum Einsatz.

Auf der Suche nach einer Erfolgsstrategie

Unser Ziel für diesen Beitrag ist es, zu erklären, was für eine Art von Information einem denn bei einer strategischen Entscheidung vor einem Match helfen könnte, wie man sie im Prinzip berechnen könnte und wie man daraus eine Erfolgsstrategie ablesen kann.

Eine nützliche Information kann z.B. so aussehen wie das Farbmuster in Tabelle 1. Wir nennen dieses Farbmuster in unserem Beitrag suggestiv unseren *Gewinnfahrplan*.

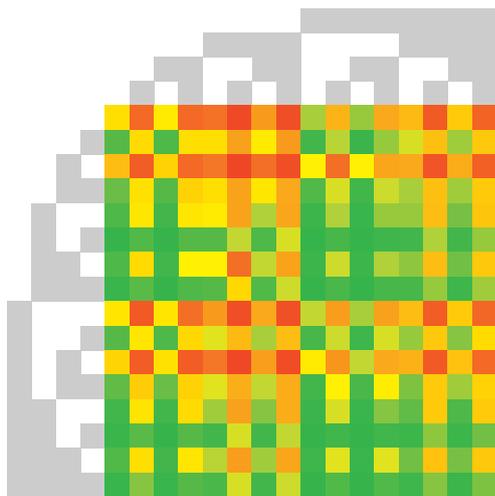


Tabelle 1: Gewinnwahrscheinlichkeiten in Abhängigkeit von Strategiekombinationen (Zeilen: unser Team; Spalten: gegnerisches Team). Von grün („wir gewinnen sicher“) über gelb („fifty-fifty“) bis rot („wir verlieren sicher“)

Wir können den Gewinnfahrplan ganz grob wie folgt interpretieren: Die Zeilen gehören zu unseren Strategien; die Spalten gehören zu den gegnerischen Strategien. Ein Farbfeld im Gewinnfahrplan gehört zu einer *Strategiekombination* aus einer Strategie für uns und einer gegnerischen Strategie. (Wir werden später noch auflösen, was die grau-weißen Muster zu bedeuten haben, die die Zeilen und die Spalten kennzeichnen.) Wenn wir vorher wissen, welche Strategie das gegnerische Team spielen wird, dann suchen wir uns die Strategie zu der Zeile aus, die in der Spalte der gegnerischen Strategie am grünsten ist. Genau dann haben wir die Erfolgsstrategie, das heißt, die Strategie mit maximal möglicher Gewinnwahrscheinlichkeit, gegen die bekannte gegnerische Strategie gefunden. Man nennt diese Erfolgsstrategie auch eine „beste Antwort“ auf die gegnerische Strategie. Wenn wir die gegnerische Strategie nicht vorher kennen, dann könnten wir uns zum Ziel setzen, im „schlimmsten“ Fall immer noch am besten abzuschneiden. Das heißt, wir wählen die Strategie zu der Zeile aus, in der das am wenigsten grüne Feld immer noch am grünsten ist. Man nennt so einen Gewinnfahrplan in der Mathematik (statt Farben könnten da natürlich auch Zahlen stehen) ein *Zwei-Personen-Nullsummenspiel in strategischer Form*. Hat man den Gewinnfahrplan, so hat man wertvolle Informationen für einen Matchplan.

Woher kommt aber nun genau der konkrete Gewinnfahrplan in Tabelle 1? In einem Dissertationsprojekt am Lehrstuhl für Wirtschaftsmathematik der Universität Bayreuth wurde untersucht, wie man ein solchen Gewinnfahrplan für ein Beach-Volleyball-Finale aus Beobachtungen in den Vor-Final-Spielen berechnen kann. Susanne Hoffmeister, selbst im Handball und Triathlon-Sport auf hohem Niveau aktiv, hat sich in ihrer Dissertation dieser Frage systematisch gewidmet. Als Beispiel kam Tabelle 1 heraus. Es handelt sich um den Gewinnfahrplan für das Team Brink/Reckermann (Deutschland) gegen Alison/Emanuel (Brasilien) vor dem Olympia-Finale 2012 in London.

Wie kann man denn nun ein Sport-Spiel mit mathematischen Methoden beschreiben? Nehmen wir das angesprochene Beispiel Beach-Volleyball. Beach-Volleyball gehört zu den *Rückschlag-Ballsportarten*. Das bedeutet, dass zwei Teams, die sich auf verschiedenen Seiten des Netzes gegenüberstehen, *abwechselnd* in Ballbesitz sind. (Das ist z.B. beim Tennis auch so, beim Fußball aber nicht.) Das Team in Ballbesitz darf gemäß den Regeln einen Angriff mit maximal drei abwechselnden Ballkontakten spielen. Am Ende eines erfolgreichen

Angriffs ist der Ball im Feld des gegnerischen Teams, das den Ball vor dem Bodenkontakt annehmen muss und dann seinerseits einen Angriff spielen kann. Ein nicht erfolgreicher Angriff (meistens Ball im Aus oder Ball im Netz) führt zu einem Punkt für das gegnerische Team. Wir nennen das einen „direkten Punktverlust“. Besonders erfolgreich ist ein Angriff, wenn das gegnerische Team die Ballannahme nicht mehr bewältigen kann. Wir nennen das einen „direkten Punktgewinn“. Wenn der Ballwechsel weitergeht, benennen wir das mit „Ball im Spiel“. Für eine Strategiekombination nennen wir die Wahrscheinlichkeiten für „direkte Punktgewinne“, „direkte Punktverluste“ und „Ball im Spiel“ die *Erfolgsverteilung* der Strategiekombination.

Nehmen wir für einen Moment an, dass die Erfolgsverteilungen für alle möglichen Strategiekombinationen bekannt sind und sich während des Spiels nicht ändern. Dann kann man über die Theorie und Algorithmen der *Markovketten* für jede Strategiekombination berechnen, wie groß die Gewinnwahrscheinlichkeit für uns ist. Mit der Theorie der *Markovschen Entscheidungsprobleme* kann man sogar ohne Simulation eine Entscheidungsregel bestimmen, die einem direkt eine optimale Strategie gegen jede feste gegnerische Strategie angibt.

Allerdings stellten sich an dieser Stelle dem Projekt zwei Hürden in den Weg: Erstens ist es anzunehmen, dass die Teams in unterschiedlichen Match-Phasen unterschiedlich spielen, und das ist in unserem Grundmodell nicht abgebildet. Zum Beispiel berichtete Jonas Reckermann, dass sich das Schlagverhalten bei Teams bei engen Spielständen am Satzende oft ändert: Es werden z.B. häufig die „Lieblingsschläge angebracht oder auch ganz bewusst Schläge, die man zuvor im Satz nicht gemacht und man sich sozusagen aufgespart hat“. Wenn die Spielweise abhängig vom Spielstand ist, dann ist es nicht ausgeschlossen (wenn auch nicht sicher), dass die Erfolgsverteilungen doch vom Spielstand abhängen. Zweitens sind die Erfolgsverteilungen nicht einfach so bekannt. Woher auch? Wenn man etwas gegen die erste Hürde unternimmt und die Erfolgsverteilungen spielstandabhängig macht, dann sind es ja noch mehr Wahrscheinlichkeiten, die man nicht kennt, und die zweite Hürde wird noch höher.

Das Projektteam entschied sich zunächst, einen Fortschritt bei der Überwindung der zweiten Hürde zu erzielen. (Wissenschaft geht nun einmal schrittweise vor, und man versucht eins nach dem ande-

ren zu verstehen.) Wie kann man also an diese Erfolgsverteilungen gelangen? Es war schnell klar, dass eine direkte Schätzung über die relativen Häufigkeiten von Erfolgen und Misserfolgen im Turnierverlauf bis vor dem Finale problematisch ist. Denn: Die Wahrscheinlichkeit für einen „direkten Punktgewinn“ hängt nicht nur von den Fähigkeiten des angreifenden Teams ab, sondern auch von denen des abwehrenden Teams. Aber was, wenn man gegen die Finalgegner bislang im Turnier noch gar nicht gespielt hat?

Hoffmeister hatte dazu folgende neue Idee: Die Erfolgsverteilung eines Angriffs wird zerlegt in die Skillverteilungen der beteiligten Einzelschläge (*Skill* ist neudeutsch für eine „Fertigkeit“). Eine Skillverteilung für einen Schlag gibt für alle Teammitglieder an, mit welchen Wahrscheinlichkeiten man den Schlag „genau wie geplant“, „mit Abweichungen“ oder „total verunglückt“ ausführt. Führt man einen Schlag aus, so hängt die Skillverteilung dabei nur davon ab, mit welchem vorigen Schlag man in welcher Qualität konfrontiert wurde (dazu gehören sowohl gegnerische Angriffsschläge als auch Zuspiele aus dem eigenen Team) und welche Fähigkeiten man selbst für den gewählten Schlag hat. Welcher Schlag gewählt wird, hängt dann von der eigenen Strategie ab (z.B. Smash oder Shot). Auf diese Weise sollen die Erfolgsverteilungen, die von *Teampaarungen* abhängen, berechnet werden aus Verteilungen, die nur noch vom *ausführenden Individuum* abhängen.

Aber wie geht nun das? Die zweite neue Idee von Hoffmeister war es, die Erfolgsverteilungen über ein detailliertes Spielmodell eines Beach-Volleyball-Ballwechsels zu ermitteln. Und zwar, indem „digitale Zwillinge“ (siehe Kapitel, Seite 136) mit genau den Skillverteilungen der echten beteiligten Personen in Computersimulationen eine riesige Anzahl von Beach-Volleyball-Ballwechselln bestreiten. Die Erfolgsverteilungen werden dann geschätzt aus der Anzahl der „direkten Punktgewinne“, der „direkten Punktverluste“ und der Vorkommen von „Ball im Spiel“ in diesen digitalen Ballwechselln.

Bleibt die Frage, wie man an die Skillverteilungen kommt. Auch das sind eine Menge Wahrscheinlichkeiten, die geschätzt werden müssen aus vorigen Spielen. Der Vorteil ist, dass wir nun jedes Match für die Schätzung heranziehen können, egal gegen wen. Für das eigene Team könnte man sogar entsprechende Testreihen im Training in großer Zahl auswerten. Da Letzteres für das Projektteam aber natürlich nicht möglich war, mussten Video-Analysen herhalten. Dies erforderte die langwierige Auswertung der Videoaufnahmen aller

Spiele von Brink/Reckermann und von Alison/Emanuel im olympischen Turnier bis zum Finale.

Olympia 2012

Wir haben ja versprochen, noch die merkwürdigen Strategiebezeichnungen aus dem Gewinnfahrplan in Tabelle 1 aufzulösen. Es handelt sich um die 16 Kombinationen aus folgenden Einzelstrategien (weiß = „nein“, grau = „ja“):

- Spieler 1 geht volles Risiko beim Aufschlag, im Wesentlichen charakterisiert durch *Jump Serves* (sehr harte Sprungaufschläge mit Anlauf), die nahe an den Spielfeldrand platziert werden, im Gegensatz zu (*Jump*) *Float Serves* (eher flatternde Aufschläge), bei denen eher die Spielfeldmitte anvisiert wird.
- Spieler 1 geht volles Risiko beim Feldangriff, im Wesentlichen charakterisiert durch *Smashes* (sehr harte Schmetterbälle, meist longline oder diagonal eng an die seitliche Außenlinie oder auf die „Sechs“ an der hinteren Außenlinie platziert) im Gegensatz zu *Shots* (im Bogen über den Block in eine verlassene Platzecke platziert).
- Spieler 2 geht volles Risiko beim Aufschlag (analog zu Spieler 1)
- Spieler 2 geht volles Risiko beim Feldangriff (analog zu Spieler 1).

Wir können uns nun in die Lage von vor dem Finale zurückversetzen, denn der Gewinnfahrplan entstand aus Daten, die allesamt vor dem Finale verfügbar waren.

Wenn beide Teams im Finale die gleichen Strategien gespielt hätten wie im Durchschnitt im Turnierverlauf bis vor dem Finale, so wäre nach unseren Auswertungen die Gewinnwahrscheinlichkeit nahe 50 % gewesen. Der spannende Verlauf des tatsächlichen Finales spricht jedenfalls nicht dagegen. Es gilt noch heute als eine Werbung für den Beach-Volleyball-Sport.

Was aber hätte unser Gewinnfahrplan dem Team Brink/Reckermann damals vorgeschlagen, um mit möglichst großer Wahrscheinlichkeit zu gewinnen? Hier nochmal das Muster:

Wir sind – grob gesprochen – auf der Suche nach einer besonders grünen Zeile. Es ist interessant, wie sensibel das menschliche Auge auf Farbunterschiede reagiert. Wir sehen auf einen Blick, dass die geraden Zeilen 2, 4, 6, . . . allesamt deutlich grüner sind als die ungeraden Zeilen 1, 3, 5, . . . Das deutet darauf hin, dass Spieler 2 im deutschen Team (Reckermann) auf jeden Fall volles Risiko im Feld spielen sollte. Und zwar unabhängig von der Strategie des brasilianischen Teams. Ganz ähnlich verhält es sich mit den Feldangriffen

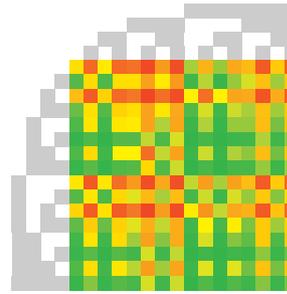


Tabelle 2: Wo ist die „grünste“ Zeile im Gewinnfahrplan?

von Julius Brink, auch wenn er nach der Aussage von Jonas Reckermann ein Top-Shot-Spieler ist. Auch ihm hätte der Gewinnfahrplan empfohlen, beim Angriffsschlag volles Risiko zu spielen, da die Zeilen 5, 6, 7, 8 und 13, 14, 15, 16 grüner als die anderen Zeilen sind.

Für die Aufschläge sieht es anders aus. Ob dieser Schlag mit mehr oder weniger Risiko ausgeführt wird, macht nach dem Gewinnfahrplan für das deutsche Team keinen sichtbaren Unterschied. Für Spieler 1 (Brink) sieht man das beispielsweise daran, dass die Zeilen 1 bis 8 den Zeilen 9 bis 16 sehr ähnlich sind.

Das ist nicht prinzipiell unplausibel: Jump-Float-Aufschläge, die bei uns als weniger riskant eingestuft werden als z.B. superharte Jump-Topspin-Aufschläge, sind bei der Annahme zuweilen auch unangenehm zu verarbeiten, vor allem z.B. bei Wind. Man muss mit dem Jump-Topspin schon viele direkte Punkte erzielen oder schlechte Angriffe provozieren, damit sich das Risiko in der Endabrechnung lohnt.

Wir sehen insgesamt, dass die vier Zeilen 6, 8, 14, 16 optisch schon sehr „grün“ aussehen. Nimmt man noch das Prinzip der mathematischen Spieltheorie mit ins Kalkül, nach dem man davon ausgeht, dass das gegnerische Team die beste Strategie gegen uns wählen könnte, dann sollten wir noch auf das am wenigsten grüne Feld in diesen Zeilen schauen. Und da fällt auf, dass bei der besten Antwort der Brasilianer (Spalte 6) die 8. Zeile nachteilig herausfällt, weil die entsprechende Gewinnwahrscheinlichkeit der Strategiekombination (GER 8, BRA 6) orange ist, also kleiner als 50 %. In den anderen Zeilen erreichen wir im Wesentlichen gelb (50 %) gegen die beste Antwort. Das heißt übersetzt, dass es etwas schlechter ist, wenn Spieler 1 (Brink) nicht volles Aufschlagrisiko geht, während Spieler 2 (Reckermann) das aber doch tut. Wir geraten bei den Farbunterschieden hier aber so langsam in den Bereich der Schätzfehler.

Auffallend ist noch, dass unsere Zahlen für Brasilien bessere Chancen prognostizieren, wenn Alison beim Aufschlag nicht volles Risiko (also eher Float Serves) spielt: Die Spalten 14 und 16 erreichen (im Gegensatz zu den Spalten 6 und 8) gegen die besten deutschen Strategien sichtbar keine 50 % Gewinnwahrscheinlichkeit (das am wenigsten grüne Feld ist immer noch grün).

Wenn man annimmt, dass Deutschland mit der Strategie spielt, deren am wenigsten grünes Feld noch am grünsten ist, und Brasilien mit der Strategie spielt, deren grünstes Feld am wenigsten grün ist, dann erhalten wir eine Gewinnwahrscheinlichkeit, die entsteht, wenn beide ihre Strategien so wählen, dass sie einseitig davon nicht abweichen würden. So eine Strategie-Kombination heißt *reines Nash-Gleichgewicht*, nach dem „Erfinder“ der nicht-kooperativen Spieltheorie John Nash (eine Hollywood-Fassung seiner Lebensgeschichte zeigt der Film „A Beautiful Mind“). Ein reines Nash-Gleichgewicht muss es nicht immer geben. Wenn man die Strategien „mischt“ (man spielt jede reine Strategie dann mit einer bestimmten Wahrscheinlichkeit), dann gibt es die schon. Man spricht dann von einem *gemischten Nash-Gleichgewicht*. In unserem Gewinnfahrplan sind alle in Frage kommenden Felder gelb (50 %). Das sind (GER 6, BRA 6), (GER 14, BRA 6) und (GER 16, BRA 6), und das verspricht ein spannendes Finale mit offenem Ausgang, was es am Ende ja auch geworden ist. Vielleicht war es ein kleiner Vorteil für das deutsche Team, mehr verschiedene Spielweisen mit ähnlichen Gewinnwahrscheinlichkeiten zur Verfügung haben.

Man sieht am Gewinnfahrplan sehr eindrücklich, dass auf dem hohen Niveau beide Teams vom Risiko im Feldangriff am meisten profitieren. Jedes Team hätte mit einer risikobereiten Strategie gegen eine risikoärmere Strategie des jeweils anderen Teams fast sichere Gewinnaussichten gehabt. Gegen die Top-Teams ist Zurückhaltung offenbar keine Option.

Aber Vorsicht: allgemeine, für *alle* Beach-Volleyball-Matches geltende Regeln kann man aus *unseren* Daten ohne Zusatzwissen *nicht* ableiten: Die Abhängigkeit der Schlussfolgerungen von den Fähigkeiten der Teams, die aufeinander treffen, ist ja gerade das Neue am Ansatz. Es hängt alles von den Details ab, insbesondere von den Skillverteilungen, also wie gut die einen auf die eine oder andere Art aufschlagen oder angreifen und wie gut die anderen die entsprechenden Schlagarten annehmen oder blocken. Für ein anderes Finale mit anderen Teams könnten die Schlussfolgerungen andere sein. Um

allgemeine Regeln aus Daten abzuleiten, würde man andere Methoden benutzen.

Detailliertere Untersuchungen zu der Abhängigkeit von den Skillverteilungen ergaben, dass man die aggressiven Schläge mit sehr hoher Konstanz ausführen muss, um etwas davon zu haben. Einen Smash z.B. sollte man schon mit um die 90 % Gelingwahrscheinlichkeit können, damit sich das Risiko lohnt. Wir konnten das bei den deutschen Hochschulmeisterschaften 2016 im Beach-Volleyball an der Universität Bayreuth aus dem Publikum heraus beobachten: In einigen Partien auf mittlerem Niveau konnten sich oft die Teams durchsetzen, die sehr konsequent mit viel Übersicht Shots in verlassene Platzecken gezielt haben, anstelle ständig Smashes aus eigentlich unzureichender Sprunghöhe zu versuchen (was aber natürlich viel Spaß macht).

Die Grenzen unseres Modells kann man deutlich sehen, wenn man die Skillauswertung des Finales zur Berechnung der Gewinnwahrscheinlichkeiten mitverwendet. Damit ergaben sich durch die Bank trotz des Finalsiegs deutlich niedrigere Gewinnwahrscheinlichkeiten für Brink/Reckermann, so als hätten sich die Brasilianer im Finale steigern können. Jonas Reckermann konnte einen anderen möglichen Grund für die Veränderung benennen: Alison/Emanuel hatten im Finale in der Gesamtzählung einen Punkt mehr erzielt, weil sie sich den zweiten Satz sehr deutlich holen konnten. Bei einem schwer einholbaren Vorsprung in einem Satz auf Top-Niveau legt das zurückliegende Team schonmal nicht alle Skills auf den Tisch und schenkt evtl. sogar mal einen Satz ab – und die in so einer Phase beobachteten Skills sind dann nicht mehr repräsentativ für andere Spielphasen. Bei uns zählen diese Beobachtungen aber genau wie alle anderen. Unter anderem liegt da also Verbesserungspotential.

Wir sehen, dass mathematische Modelle eigentlich nie fertig sind. Immer stellt man fest, dass man etwas bislang nicht auf der Rechnung hatte. Nur im Gespräch mit Fachleuten kann man den wichtigsten Auslassungen auf die Spur kommen. Der Austausch zwischen Mathematik und Anwendung ist daher einer der wichtigsten Erfolgsfaktoren in der anwendungsorientierte Mathematik-Forschung.

Wir danken sehr herzlich Julius Brink und Jonas Reckermann für Ihr Interesse an unserem Projekt und das überaus interessante Gespräch – und natürlich immer noch für das grandiose olympische Finale 2012!



Daten- und modellgestützte Talentförderung von Nachwuchsfußballern

Daten und deren Analyse spielen auch im Fußball eine immer größere Rolle. Zur Bestimmung und Optimierung der körperlichen und mentalen Leistungsfähigkeit werden an Nachwuchsleistungszentren regelmäßig leistungsdiagnostische Untersuchungen durchgeführt. Basierend auf den so erhobenen Daten lassen sich mit statistischen Ansätzen nicht nur Stärken-Schwächen-Profile ableiten, sondern mithilfe von mathematischen Modellen und Machine Learning auch Entwicklungsprognosen für den einzelnen Spieler entwickeln.

Prof. Jan Mayer und Dr. Sascha Härtel von der TSG Hoffenheim sowie Oliver Wohak von der d-fine GmbH berichten von einem gemeinsamen Projekt, bei dem mathematische Modelle und Machine Learning Methoden auf physiologische und psychologische Leistungsdaten von Nachwuchsfußballern der TSG Hoffenheim angewandt wurden. In dem innovativen Projekt konnten erfolgreich datengestützte Stärken-Schwächen-Profile und Kompensationsmechanismen aufgezeigt, sowie individuelle Entwicklungsprognosen für den einzelnen Spieler entwickelt werden.

Mathematische Modelle unterstützen die individuelle Förderung von Jugendspielern im Fußball

Im Fußball stecken viele Daten. Mittlerweile werden an einem Bundesligaspieltag 3,6 Millionen Datenpunkte pro Spiel erhoben. Bevor ein Spieler es jedoch bis in die höchste deutsche Spielklasse schafft,

durchläuft er in der Regel die Jugendakademien oder Nachwuchsleistungszentren. Und auch da – bereits ab den jüngsten Jahrgängen – gehört die Erfassung von Leistungsdaten dazu. Ein Vorreiter in der Nachwuchsleistungsförderung ist die TSG Hoffenheim. Neben Tracking Daten, die zu jedem Training und Spiel aktuelle Performance-Daten aufzeichnen, werden sehr umfangreich physiologische und psychologische Leistungstests durchgeführt. Mit Hilfe dieser interdisziplinären Daten soll ein möglichst ganzheitliches Bild über die einzelnen Spieler und Mannschaften entstehen. Die Bestimmung des Ist-Stands ermöglicht die Einschätzung der Grundvoraussetzungen (z.B. ist ein Spieler eher schnellkräftig oder ausdauerveranlagt; ist er im Kopf und/oder in den Beinen schnell) sowie einen Abgleich mit Soll-Werten (Benchmark-Werte Profis, positionsspezifische Anforderungsprofile).

Das aus den Testungen abgeleitete Stärken-Schwächen-Profil ermöglicht konkrete, individuelle Trainingsmaßnahmen und damit verbunden eine Optimierung der Leistungsfähigkeit. Mit Hilfe von Testwiederholungen kann die langfristige Entwicklung und damit die Wirksamkeit der gewählten Interventionen überprüft werden.

Die Verwertung von Leistungs- und Diagnostikdaten unterstützt die Talentförderung

Verglichen mit anderen Fußballvereinen hat die TSG schon sehr früh mit der strukturierten Datenerfassung begonnen. Somit liegen umfangreiche, mehrjährige Zeitreihen für verschiedene U-Mannschaften zu einem einheitlichen Satz von Kennzahlen vor, die sich sowohl zur explorativen Datenanalyse, als auch für die Entwicklung von Prognosemodellen nutzen lassen. Typische Fragestellungen in diesem Kontext sind: Wie verändern sich bestimmte Kennzahlen im Karriereverlauf? Wie entwickeln sich bestimmte Korrelationen im Zeitverlauf? Welche Prognosen lassen sich bzgl. des Übergangs in den U-Mannschaften treffen? Alles mit dem Ziel Muster in den Daten zu erkennen, welche die individuelle Förderung der Spieler unterstützen und das Leistungspotential heben.

Der Blick zurück – erste explorative Datenanalysen zeigen Zusammenhänge auf

Einfachere Muster lassen sich mittels explorativer Analysen in den Daten erkennen. Dies sind zum Beispiel die zeitliche Entwicklung einzelner Attribute (univariate Analysen) oder auch Korrelationen verschiedener Attribute (multivariate Analysen). Korrelationsanalysen sind insbesondere dann hilfreich, wenn die „Ähnlichkeit“ von histo-

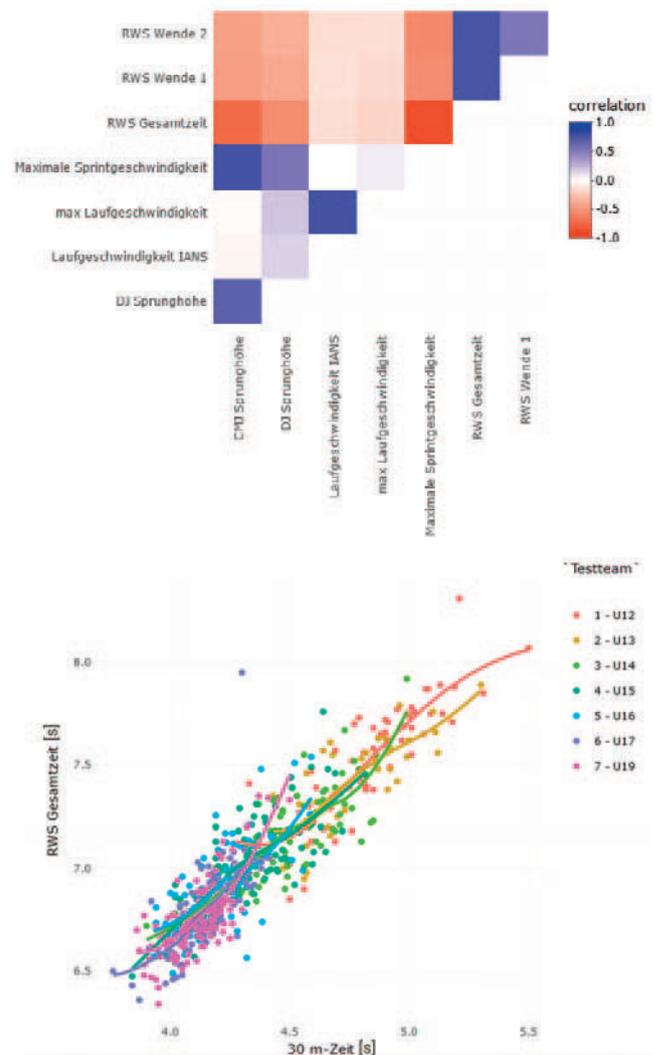


Abbildung 96: Korrelationsplot zu Schnelligkeits-, Ausdauer und Sprungkraftwerten (oben) und bivariate Analyse zu Richtungswechselsprint (RWS) und 30 m-Sprintzeit unten.

rischen Zeitreihen einzelner Attribute bewertet werden soll. Dabei gibt die Berechnung der Korrelation Erkenntnis über die Stärke des Zusammenhangs der Attribute sowie der Richtung: -1 ist maximal stark negativ korreliert und +1 maximal stark positiv korreliert. In Abbildung 96 sind die Korrelationswerte verschiedener Attribute aus den Leistungstests dargestellt.

Zum Beispiel ist dort zu erkennen, dass die Attribute der Sprungkraft (konkret hier die verschiedenen Angaben zur Sprunghöhe DJ Sprunghöhe und CMJ Sprunghöhe) sehr mit den Sprintwerten korrelieren -- stark negativ mit den Richtungswechselsprintzeiten (RWS) und stark positiv mit der maximalen Sprintgeschwindigkeit. Daraus lässt sich ableiten, dass Spieler mit guten Schnelligkeitswerten typischerweise auch eine starke Sprungkraft haben.

Neben den Korrelationsplots können bivariate Analysen helfen, um Zusammenhänge zwischen zwei bestimmten Attributen näher zu untersuchen. In dem Punktwolkendiagramm in Abbildung 96 sind exemplarisch die beiden Attribute RWS Gesamtzeit und 30 m-Zeit gegenübergestellt. Durch die inhaltliche Nähe der beiden Attribute lässt sich in der Analyse ein starker Zusammenhang erkennen. Spieler mit einer schnellen 30 m- Zeit absolvieren typischerweise auch den Richtungswechselsprint in einer kurzen Zeit. Allerdings ist ebenfalls zu erkennen, dass dieser Zusammenhang nicht perfekt ist. So kann im Einzelfall ein Spieler eine gute 30 m-Zeit aufweisen, während die Leistung beim Richtungswechseltest ausbaufähig ist. Zusammengefasst scheinen die Attribute 30 m-Zeit und RWS Gesamtzeit stark verwandt zu sein, jedoch noch spezifisch genug, um Stärken und Schwächen im jeweiligen Attribut getrennt identifizieren zu können. Darüber hinaus lässt sich aus den U-Mannschafts-spezifischen Trendlinien noch erkennen, dass die sich die Sprintzeiten insbesondere im Bereich der U12 bis U14 auf Grund der körperlichen Entwicklung von den anderen Mannschaften abheben und, dass die Unterschiede zwischen den Mannschaften unabhängig vom Alter von der U15 aufwärts sehr gering sind.

Der Blick in die Zukunft – Machine Learning Modelle ermöglichen Prognosen zur Spielerentwicklung

Neben den explorativen Datenanalysen ist es interessant zu untersuchen, inwiefern Prognosemodelle helfen können, Aussagen über zukünftige Entwicklungen der Spieler zu treffen. Während explorative Analysen die vorliegenden Daten lediglich beschreiben und nachträglich Zusammenhänge aufzeigen, ermöglichen Prognosemodelle eine Vorhersage der Spielerentwicklungen. In unserem Projekt erfolgte dies an Hand der Prognose der Übergangswahrscheinlichkeit in die nächst höhere U-Mannschaft, die für jeden Spieler ermittelt wurde. Dies ermöglicht somit das frühzeitige Eingreifen von Athletiktrainern oder Sportpsychologen, um die Weiterentwicklung der Spieler optimal zu unterstützen.

Hierfür bieten sich etablierte Machine Learning Methoden an -- also Mathematik -- um sowohl lineare Abhängigkeiten (z.B. mittels linearer oder logistischer Regression) als auch komplexere Zusammenhänge (z.B. mittels Baum-basierten Modellen oder neuronalen Netzen) in den Daten zu erkennen und zu verwerten. Diese Methoden haben gemeinsam, dass in einem Prozess -- der auch überwacht Lernen (engl. „supervised learning“) genannt wird -- in historischen Daten Gesetzmäßigkeiten erkannt werden, um basierend auf einem Satz an Eingabewerten (Modell Features) zugehörige Ergebnisse oder Ereignisse (Modell Label) möglichst gut und zuverlässig zu prognostizieren.

In diesem Projekt wurde ein so genanntes Random Forest Modell (eine Kombination aus mehreren Baum-basierten Modellen) trainiert, um ausgehend von den physiologischen und psychologischen Testergebnissen der Spieler eine U-Mannschaftsspezifische Übergangswahrscheinlichkeit zu prognostizieren. Die Features im Modell sind somit z.B. die körperliche Aktivität, die Laufgeschwindigkeit an der individuellen anaeroben Schwelle (Grundlagenausdauer) oder auch der Geburtsmonat (spät im Jahr geborene Kinder fallen häufig aus der Förderung, da sie meist gegenüber den früh im Jahr geborenen weniger weit entwickelt sind), während der Label der (nicht) erfolgreiche Übergang in die nächsthöhere U-Mannschaft am Ende der Saison ist.

Das Modell wird zuerst mit den historischen Daten trainiert und anschließend validiert. Dafür werden typischerweise die vorhandenen Daten in einen Trainingssatz (z.B. 80 % der Daten) und einen Validierungssatz (die restlichen 20 %) getrennt. Mithilfe des Validierungssatzes wird die Modellgüte getestet, zum Beispiel über die Anzahl der Richtig-Positiven und Falsch-Positiven Vorhersagen. Dabei spricht man dann von der ROC-Kurve (Receiver-Operating-Characteristics oder auf Deutsch Grenzwertoptimierungskurve) oder auch einem Gini-Koeffizient (der die Fläche unter der Kurve auf Werte zwischen 0 und 1 normalisiert). Es lässt sich somit mit einer gewissen Zuverlässigkeit vorhersagen, dass ein Mittelfeldspieler der U17 mit 16 Jahren, einer körperlichen Aktivität von 1 (Maximum), einer Laufgeschwindigkeit IANS (12,8 km/h) sowie einigen weiteren Attributen (Abbildung 97) mit einer Wahrscheinlichkeit von ca. 76 % den Sprung in die U19 schafft.

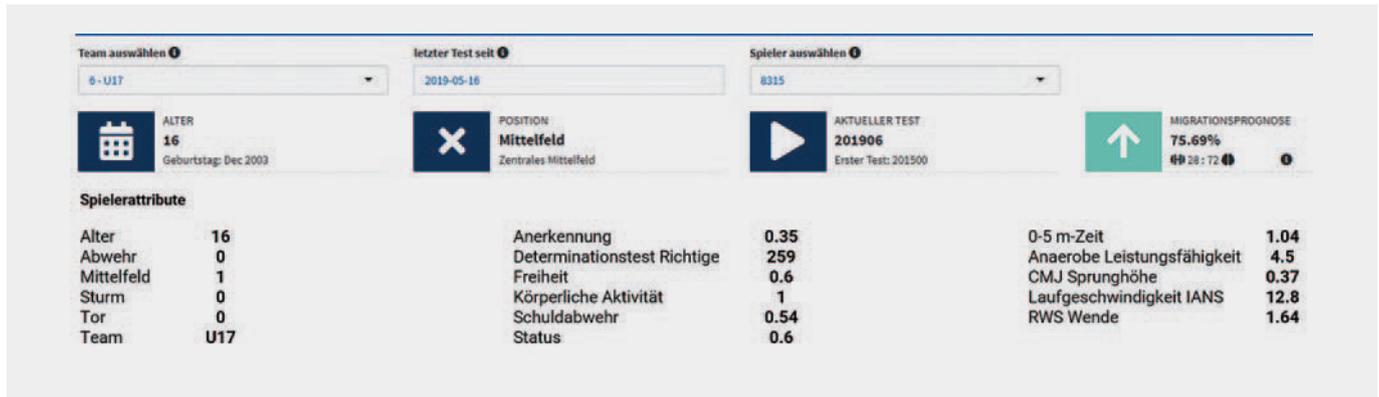


Abbildung 97: Darstellung der Attribute eines Spielers der U17 mit einer hohen Übergangswahrscheinlichkeit.

Modelle erklärbar zu machen liefert zusätzliche Erkenntnisse

Als Ergänzung zu den Expertenmeinungen der Trainer ist es hilfreich, einzelnen Spielern eine Übergangswahrscheinlichkeit zuzuordnen zu können und somit datengestützt Spieler mit einem hohen Potential bereits frühzeitig identifizieren zu können. Um darüber hinaus jedoch jeden Spieler individuell bestmöglich fördern zu können, müssen wir zusätzlich wissen, welche Eigenschaften bei den Spielern mit hohem Potential stark ausgeprägt sind und welche körperlichen oder mentalen Eigenschaften bei „schwächeren“ Spielern insbesondere noch trainiert werden sollten. Standardmäßig können Machine Learning Modelle dazu keine direkte Aussage treffen. Es gibt jedoch mathematische Verfahren, die es ermöglichen, den Anteil jedes Attributs an dem Prognoseergebnis zu quantifizieren. Dies wird dann auch „explainable AI“ (xAI) oder erklärbare künstliche Intelligenz genannt. Eines dieser Verfahren – das auch in diesem Projekt angewandt wurde – ist SHAP (Shapley Additive exPlanations). Es baut auf einer Idee auf, die ursprünglich aus der Spieltheorie kommt und versucht den marginalen Beitrag jedes Spielers in einem kooperativen Spiel mit messbarem Ausgang zu quantifizieren. Angewandt auf Machine Learning Methoden quantifiziert SHAP den marginalen Beitrag jedes

Modell Features auf die Modellprognose. Dies erreicht man, indem man die Ergebnisse vieler ähnlicher Modelle vergleicht. Diese Modelle unterscheiden sich nicht in ihrer Parametrisierung, decken jedoch alle möglichen Kombinationen an Feature-Zusammensetzungen ab. Durch den Vergleich der jeweiligen Modellprognosen kann zurückgerechnet werden, welchen Beitrag jedes einzelne Feature auf die Prognose des vollständigen Modells hat. Insbesondere ist es in unserem Modell somit möglich aufzuzeigen, welche Features die Übergangswahrscheinlichkeit bei einem bestimmten Spieler – im Vergleich zu dem Durchschnittsspieler – eher treiben bzw. drücken. Dies ist durch den „Force-Plot“ in Abbildung 98 dargestellt an dem erkennbar ist, dass der dargestellte Spieler insbesondere durch seine körperliche Aktivität und das gute Abschneiden im Determinationstest (Kognitionstest) heraussticht. Eine Eigenschaft, an der der Spieler noch arbeiten kann, um sich weiter zu verbessern, und wo er schlechter abschneidet als der Durchschnitt, ist seine Grundlagenausdauer – hier dargestellt als Laufgeschwindigkeit an der individuellen anaeroben Schwelle (IANS).



Abbildung 98: Force-Plot der anzeigt welche Attribute den dargestellten Spieler im Vergleich zu einem Durchschnittsspieler (Basiswert) auszeichnen.

- Körperliche Aktivität (201906): +16.5%
- Determinationstest-Richtige (201906): +3.2%
- 0-5 m-Zeit (201806): +2.2%

- Laufgeschwindigkeit IANS (201906): -1.6%
- CMJ Sprunghöhe (201906): -0.8%
- RWS Wende 1 (201806): -0.4%

28 : 72

Aus dem xAI-Ansatz und der Darstellung mittels „Force-Plots“ lassen sich somit daten- und modellgestützte Stärken-Schwächen Profile für die Spieler ableiten – auch nochmal dargestellt in [Abbildung 99](#). Darüber hinaus lassen sich weitere interessante Erkenntnisse ableiten, die von den Leistungsdiagnostikern und Sportpsychologen verwertet werden können. So wurden in dem Projekt zum Beispiel Kompensationseffekte analysiert. Diese geben an, ob die Stärken eines Spielers eher physiologischer oder psychologischer Art sind. Bei dem oben dargestellten Spieler der U17 beträgt das Verhältnis physiologischer zu psychologischer Stärken 28:72. Das bedeutet, dass der Spieler ausgeprägte mentale Fähigkeiten hat und seine athletischen Defizite darüber ausgleicht. Insbesondere lässt sich über die verschiedenen Jugendmannschaften hinweg erkennen, dass physiologische Eigenschaften in jüngeren Jahren für die Förderung wichtiger sind, im Alter für den Sprung in den Profi-Kader die psychologischen Kompetenzen jedoch umso wichtiger werden.

[Abbildung 99: Das Stärken-Schwächen-Profil und der Kompensationseffekt \(Hantel = physiologisch, Gehirn = psychologisch\).](#)

Daten und Modelle werden auch im Fußball immer wichtiger

Insgesamt hat das Projekt gezeigt, dass mathematische Methoden und Modelle helfen können vorhandene Daten im Fußball zu verwerten. Anhand historischer Daten zu Jugendspielern der TSG Hoffenheim wurde ein Machine Learning Modell implementiert, das die Übergangswahrscheinlichkeiten für einzelne Spieler ableitet und darüber hinaus die Definition von Stärken-Schwächen-Profilen sowie Kompensationseffekten ermöglicht. Weitergedacht können diese Art Daten im Fußball auch genutzt werden, um z.B. Verletzungsanfälligkeit oder Überbelastung zu identifizieren – zum Beispiel auch in Kombination mit den Trackingdaten aus Training und Spiel.

Der optimale Freiwurf im Basketball

Wenn ein Basketballer, wie etwa Shaquille O'Neal oder Dirk Nowitzki, zum Freiwurf antritt, überlegt er sich vermutlich nicht wie er den Abwurfwinkel und die Abwurfgeschwindigkeit des Balles wählt, sondern er wirft den Ball intuitiv Richtung Korb, um die Wahrscheinlichkeit eines Korberfolges zu maximieren.



In diesem Kapitel wollen wir anhand des Freiwurfs im Basketball den Prozess der Mathematischen Modellierung erklären und so die Frage beantworten „Wie kann man die Trefferquote beim Freiwurf verbessern?“ Nach einer Modellierung des betrachteten meist komplexen Vorganges muss häufig das Modell vereinfacht werden (z.B. um Lösungen in Echtzeit zu bekommen). Anschliessend muss man Verfahren für die numerische Berechnung von approximativen Lösungen diskutieren (hier: ableitungsfreie Mehr-Ziel-Optimierungsalgorithmen). Weiterhin wollen wir am Ende erklären, wie man mithilfe von sog. *Voronoi-Diagrammen* die Zonen-Abwehr im Basketball optimieren kann.

Modellierung der Flugbahn

Weitere wichtige Fragen wären: wie wirkt sich der Luftwiderstand auf die Flugbahn des Balles aus? Wohin sollte man eigentlich am besten zielen; an die Vorder- oder Rückseite des Ringes oder in die Mitte des Korbes oder ...? Welchen Fehler (d.h. Abweichung von den optimalen Parametern) darf man sich erlauben, um trotzdem einen Korberfolg zu erzielen? Wir betrachten also im folgenden einen Basketballspieler einer bestimmten Größe und wollen diese Fragen beantworten.

Bei einem Freiwurf verlässt der Basketball die Wurfhand des Spielers mit einem bestimmten Abwurfwinkel und einer Abwurfgeschwindigkeit (und evtl. einem Spin); die weitere Flugbahn wird durch Bewegungsgleichungen beschrieben. Der Korberfolg kann unterschiedlich passieren: mit/ohne Berührung des Rings, nach Abprallen vom sog. Backboard oder nach mehrmaligem Hin- und Herspringen zwischen Teilen des Rings und des Backboards.

Wir treffen machen einige Annahmen, um zu einem einfachen Modell zu kommen. Wir betrachten nur Würfe, bei denen der Basketball direkt ins Netz trifft oder zuerst die Hinterkante des Ringes trifft und dann direkt ins Netz geht. Wir vernachlässigen den Einfluss des Luftwiderstands und auch des Spins (weil wir Abpraller vom Backboard nicht betrachten). Weiterhin vernachlässigen wir seitliche Fehler in der Flugbahn, so dass uns Bewegungsgleichungen in zwei Dimensionen (Höhe, Entfernung) ausreichen.

Ein Max-Min-Problem

Mit diesen Annahmen kann man die Flugbahn des Basketballs bei vorgegebenen Abwurfwinkel θ^0 mithilfe von Differentialgleichungen beschreiben und bekommt nach einiger Rechnung⁶⁴ eine Aussage, welcher Abwurfwinkel θ^0 die größtmögliche Abweichung zulässt. Dazu bezeichnen wir mit θ_{\min}^0 und θ_{\max}^0 den kleinsten bzw. größten abweichenden Winkel θ , bei dem der Freiwurf gerade noch zum Treffer führt, d.h. bei einem erfolgreichen Wurf muss der Abwurfwinkel θ^0 im Intervall $[\theta_{\min}^0, \theta_{\max}^0]$ liegen.

Mathematisch kann man eine *Zielfunktion*

$$e(\theta^0) = \min \{(\theta^0 - \theta_{\min}^0), (\theta_{\max}^0 - \theta^0)\}$$

einführen und die Aufgabe als sog. *Max-Min-Problem* definieren: Bestimme denjenigen optimalen Abwurfwinkel θ^0 , für den die Ziel-



Abbildung 100: Ein Vortrag von Matthias Ehrhardt über den optimalen Freiwurf beim Basketball Zweitliga-Verein RSV Stahnsdorf

funktion $e(\theta^0)$ maximiert wird. ihr Maximum annimmt. Diese Zielfunktion hängt natürlich von vielen weiteren Parameter ab, z.B. Abwurfgeschwindigkeit v^0 , Größe des Spielers h_s , usw.. Übrigens, diese Zielfunktion $e(\theta^0)$ ist zwar stetig, aber nicht überall differenzierbar, denn sie hat offensichtlich einen „Knick“. Es stellt sich heraus, dass es besser ist Abwurfgeschwindigkeit und -winkel gleichzeitig zu optimieren, man spricht dann von *Mehrzieloptimierung*.

Optimierte Defense

Bei der Zonen-Abwehr im Basketball (oder auch Raumdeckung im Fußball) möchte man eher ein Min-Max-Problem lösen: der maximale Abstand von Abwehrspielern und gegnerischen Angriffsspielern soll minimiert werden. Dazu gibt es ein einfaches mathematisches Werkzeug, das *Voronoi-Diagramm*. Dabei wird das Spielfeld in Bereiche (sog. Zellen) zerlegt, die durch vorgegebene Punkte (sog. Zentren, also den Spielern) bestimmt werden. Jede solche Voronoi-Zelle hat genau ein Zentrum und umfasst alle Punkte des Spielfeldes, die (bzgl. der euklidischen Norm) näher an dem Zentrum der Zelle liegen als an jedem anderen Zentrum. Aus allen Punkten, die mehr als ein nächstgelegenes Zentrum besitzen und somit die Grenzen der Zellen bilden, entsteht das Voronoi-Diagramm. Man kann sich vorzustellen, dass ein Spieler den Punkt des Basketballfeldes, auf dem er steht, vollständig 'besitzt', dass er keine Kontrolle über einen sehr weit entfernten Punkt hat und nur eine geringe Kontrolle über einen Punkt

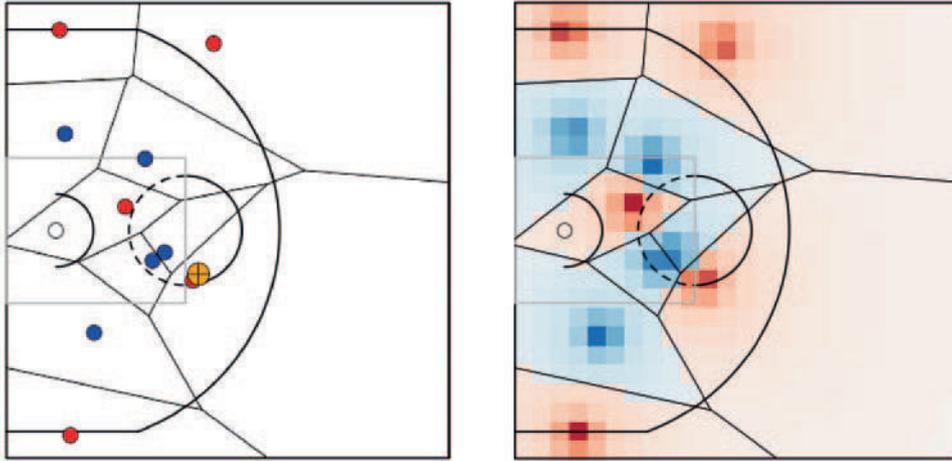


Abbildung 101: Beispiel einer Karte für den 'Besitz' von Spielfeldzellen. Die Kontrolle eines Spielers über jede Spielfeldzelle in seinem Voronoi-Segment (links) ist umgekehrt proportional zu seiner Entfernung von dieser Zelle (rechts).

in der Nähe hat (z.B. 1,5 m entfernt) – außer vielleicht, wenn ein anderer Spieler noch näher an diesem Punkt steht.

Investitionen in Immobilien

Auf dem Spielfeld sind die Regionen unterschiedlich wertvoll (wie bei Immobilien) und die Spieler handeln rational, d.h. sie tauschen bei ihren Bewegungen (und der Bewegung des Balls, Gegners) ihre Spielfeldbereiche derart, dass sie eine Gewinnstrategie für ihr Team darstellen.

In diesem Sinne führten Cervone, Bornn und Goldsberry ein ein „gewichtetes Voronoi“- Konzept ein, um den Platzbesitz eines jeden Spielers zu einem beliebigen Zeitpunkt zu definieren⁶⁵. Sie unterteilten das Halbfeld zunächst in $M = 576$ gleich große Zellen, jeweils etwa 60cm 60cm. Für Spieler i bezeichnet $X^i(t)$ einen M -dimensionalen Vektor, der seinen Besitzwert in jeder der M Spielfeldzellen zum Zeitpunkt t darstellt. Der m -te Eintrag von $X^i(t)$ ist nur dann umgekehrt proportional zur Entfernung ('dist') zwischen Spieler i und Spielfeldzelle m zum Zeitpunkt t , wenn kein anderer Spieler näher an der Spielfeldzelle m ist:

$$w_m^i(t) = \text{dist}(\text{Spieler } i, \text{ Spielfeldzelle } m) \quad \text{zur Zeit } t$$

$$X_m^i(t) = \begin{cases} \frac{1}{1+w_m^i(t)}, & i = \text{argmin}_j w_m^j(t) \\ 0, & \text{sonst} \end{cases}$$

So wird der Platz unter den Spielern entsprechend dem Voronoi-Diagramm der Spielerpositionen aufgeteilt, und innerhalb jedes von ihnen kontrollierten Segments ist die Investition der Spieler in den Platz umgekehrt proportional zu ihrer Entfernung von diesem Platz. Insbesondere kodiert der Platzbesitz eines Offensivspielers implizit Informationen über die Positionierung der Defensive. Wenn sich ein Verteidiger dem Spieler i nähert, verringert sich die Voronoi-Partition des Spielers i , was bedeutet, dass mehr Einträge im Spielfeldbesitzvektor $X^i(t)$ Null sind. Diese Voronoi-Diagramme werden übrigens auch in Computerspielen eingesetzt, um die Spielsituationen zu analysieren.⁶⁶

⁶⁴ J.M. Gablonsky, A.S. Lang, Modelling Basketball Free Throws, SIAM Review 47 (2005), 775-798.

⁶⁵ D. Cervone, L. Bornn, K. Goldsberry, NBA Court Realty, MIT Sloan Sports Analytics Conference 2016, http://www.lukebornn.com/papers/cervone_ssac_2016.pdf.

⁶⁶ S. Hu, S. Chang and J. Jiang, Voronoi State Management for Peer-to-Peer Massively Multiplayer Online Games, 5th IEEE Consumer Communications and Networking Conference, 2008, pp. 1134-1138.



Was ist das Geheimnis hinter einem erfolgreichen Endspurt im Radsport? Nicht nur ein Top-Sprinter, sondern auch eine Berechnung, die den optimalen Sprintzug bestimmt.

**Das Geheimnis
hinter einem
erfolgreichen
Endspurt
im Radsport**

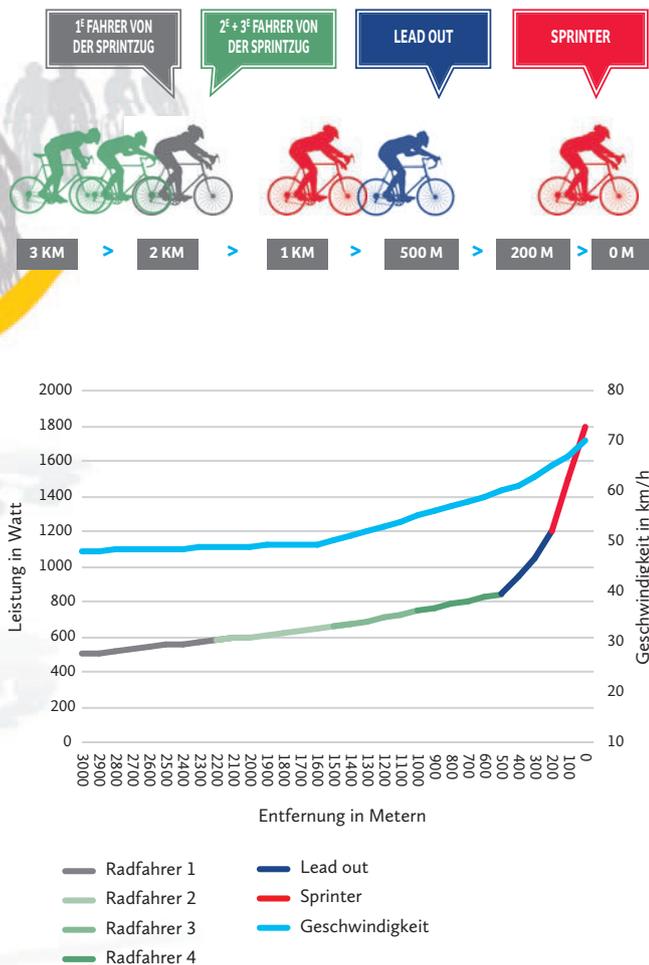


Abbildung 102: Graph von Leistung (untere Linie) und Geschwindigkeit (obere Linie) in Abhängigkeit von der Distanz bis zum Ziel (für fünf Fahrer aus dem Sprint-Team).

Zwei der besten Sprinter des aktuellen Radsportpelotons fahren für das niederländische Team Giant-Shimano: die Deutschen Marcel Kittel und John Degenkolb. Kittel zeichnet sich durch flache Ankünfte in flachen Etappen aus. Degenkolb hat Sprints in schwereren Rennen mit mehr Steigungen zu bewältigen. Sowohl bei der Tour de France 2013 als auch 2014 gewann Kittel nicht weniger als vier Etappen, darunter zweimal die prestigeträchtige Schlussetappe in Paris. Degenkolb gewann unter anderem den Klassiker Paris-Tours 2013 sowie eine Etappe der Italien-Rundfahrt und 2014 den Klassiker Gent-Wevelgem. Der Bewegungswissenschaftler Teun van Erp ist bei Giant-Shimano angestellt und analysiert unter anderem die Sprints von Kittel und Degenkolb.

Van Erp erhält von seinen Fahrern alle Messdaten, sowohl aus den Trainings als auch aus den Rennen: „Diese sogenannten SRM-Daten sagen genau aus, welche Leistung ein Fahrer in welchem Moment abgibt. Daraus kann ich ersehen, wie lange ein Reiter eine bestimmte Kraft aufrechterhalten kann und wie gut er in Form ist. Ich analysiere die Daten mit spezieller Software und wende alle möglichen mathematischen Tricks an.“

Er verwendet diese Daten auch zur Bestimmung des optimalen Sprintzuges. Van Erp: „Theoretisch wird der optimale Sprintzug dadurch bestimmt, wie lange ein Fahrer im Finale ein bestimmtes Leistungsniveau halten kann. In einem typischen flachen Ziel liegt der erste Mann auf unserem Sprintzug ein oder drei Kilometer vor dem Ziel in Führung. Er erhöht langsam die Geschwindigkeit des Pelotons über eineinhalb Kilometer, damit niemand entkommen kann. Er wird für eine bis eineinhalb Minuten gegen die fünfzig Stundenkilometer ausgewählt. Danach liegen die Nummern zwei und drei des Sprintzuges vorne, beide etwa fünfhundert Meter hintereinander. Sie erhöhen die Geschwindigkeit noch weiter, auf bis zu fünfundfünfzig Kilometer pro Stunde. Etwa fünfhundert Meter vor dem Ziel liegt der 'Lead-out' in Führung, mit dem Sprinter in seinem Rad. Der Führende kann auch gut sprinten und bringt unseren Spitzensprinter auf den Weg. Das sprintet mit einer Geschwindigkeit von mehr als siebzig pro Stunde und mit einer Spitzenleistung von etwa zweitausend Watt dem Ziel entgegen.“

Prof. Dr. Andreas Schuppert



Eine kurze Zusammenfassung Ihrer Karriere

Nach dem Studium der Physik und der Promotion in Mathematik an der Universität Stuttgart ging ich in die Industrie zur Zentralforschung der damaligen Hoechst AG, als Wissenschaftler in der Scientific-Computing Gruppe. Nach fast 10 Jahre bei Hoechst und einem berufsbegleitenden Studium der Wirtschaftswissenschaften wechselte als Leiter der Gruppe Verfahrenstechnische Software zur Zentralen Technik der Bayer AG in Leverkusen und übernahm dort nach zwei Jahren im Jahr 2000 die Leitung der Mathematik-Gruppe, die wir dann „Computational Solutions“ nannten. Das Umfeld dort war

damals einzigartig: Modellierung und Simulation waren eingeführt und hatten ihren Wert für das Unternehmen bewiesen, das aktuelle Arbeitsfeld, das wir erschließen konnten, war Hybride Modellierung als Kombination Maschinellen Lernens und mechanistischen, gleichungsbasierten Modellen. Hier entwickelten wir heute noch aktuelle mathematische Lösungen und Software und konnten sie in wichtigen Anwendungen, insbesondere in der Materialforschung und -Entwicklung erfolgreich anwenden. Fast parallel dazu griffen wir damals schon das Thema Data Mining auf und entwickelten es ebenfalls zu einem reifen Anwendungsfeld, insbesondere für Einsätze bei der Fehlersuche in komplexen chemischen und biotechnologischen Produktionsprozessen. Durch diese Aktivitäten, besonders das Thema Data Mining, kam ich mit den Life Sciences in Pharma Forschung und Entwicklung sowie in der Pflanzenschutzforschung in Kontakt. Bald darauf wurde ich aufgefordert, mich um das damals neue Gebiet der Systembiologie zu kümmern und dies für Bayer aufzubauen, wobei sich hier neue Möglichkeiten für die Entwicklung und Adaption auch komplexerer Mathematik für Fragestellungen in der Pharmakologie, klinischen Entwicklung und Diagnose ergaben.

All diese Aktivitäten verfolgten wir in Kooperation mit externen Partnern auch aus dem akademischen Bereich. Hierüber kam ich in sehr engen Kontakt mit der Angewandten Mathematik, besonders der Industriemathematik auch auf internationaler Ebene und engagierte mich in akademischen Gremien. 2007 wurde ich als Professor in Teilzeit für das Fach „Datengetriebene Modellierung in Computational Engineering Sciences“ an die RWTH Aachen, Fakultät für Maschinenbau, berufen. Das akademische Engagement wurde im Folgenden stückweise ausgeweitet, so dass ich 2010 bei Bayer meine Management-Position gegen die Stelle eines Key Expert für Industrial Mathematics eintauschte, die höchste Position der damals neu geschaffenen Expertenlaufbahn.

2013 wurde dann das „Joint Research Center for Computational Biomedicine“ als erstes Private-Public Partnership in Deutschland zwischen Bayer, RWTH Aachen und dem Universitätsklinikum Aachen unter meiner Leitung gegründet. 2017 wechselte ich als Professor für Computational Biomedicine in die Medizinische Fakultät, wo ich seit 2019 nach meinem Verlassen der Bayer AG zu 100% tätig bin. Wir bilden dabei die Brücke zwischen Computational Sciences, insbesondere Mathematik, und den Anforderungen der Medizin. Arbeitsfelder sind u.a. Hybride Modellierung, Digitale Patientenmodelle und auch Einsatzmöglichkeiten von Quantencomputing in biomedizinischen Fragestellungen.

Was hat Mathematik für Ihre Karriere bedeutet?

Mathematik war für meine Karriere extrem wichtig. Weniger die Details, als vielmehr das Verständnis der den mathematischen Methoden zu Grunde liegenden Konzepte und ihrer Stärken und Schwächen für die jeweiligen Anwendungen. Erst ein tiefgehendes Verständnis dieser Konzepte ermöglicht es, schnell problemadäquate Workflows für die Lösung komplexer Anwendungsprobleme zu entwickeln. Dabei kommt es weniger darauf an, „methodenrein“ zu arbeiten oder die Existenz einer Lösung zu beweisen, sondern es kommt auf die geschickte Kombination unterschiedlichster Methoden, deren Implementierung und Zuverlässigkeit an. Auch wenn ich die Fehlerschranken genau schätzen kann, hilft es nichts, wenn diese viel zu groß sind. Das geht nicht ohne ein tiefes Verständnis der jeweiligen mathematischen Konzepte, dazu gehören dann auch Beweise. Insofern schließt sich der Kreis zur „akademischen“ Mathematik.

Welche Rolle spielt Mathematik in Ihrem Institut/Ihrer Firma? Und allgemeiner in der Branche, in der Sie tätig sind?

Mathematik ist in der chemisch-pharmazeutischen Industrie eine typische „Enabling Technology“ und spielt als „Mathematik“ solche keine große eigenständige Rolle. Ihre Bedeutung gewinnt sie, indem sie zuverlässige Lösungen für sehr komplexe, oft multidisziplinäre Fragestellungen mit hohem Wert für die Firma ermöglicht. Idealerweise geht es weniger darum, noch ein wenig das Etablierte zu optimieren, sondern neue Optionen zu erschließen. Hierzu ist ein sehr gutes Verständnis der Fragestellung einerseits und der in Frage kommenden Mathematischen Methoden nötig. Oft scheitern solche Projekte zwischen Mathematik und Anwendung auch an kulturell bedingten Verständnisproblemen, das zu überbrücken war mein Studium der Physik und später der Betriebswirtschaftslehre sehr hilfreich.

Wie sieht ein typischer Arbeitstag aus?

Das ist sehr heterogen. Zu Anfang war ich in konkrete Forschungsprojekte eingebunden, mit konkreten Fragestellungen und Aufgaben. Diese galt es zu strukturieren, geeignete Workflows zu entwickeln und Implementieren sowie die Anwendungsfrage selbst zu lösen und die Lösung zu validieren. Hierzu ist, neben Arbeit mit Papier und Computer, ein nicht zu unterschätzender Kommunikationsaufwand und ein Einlassen auf die Arbeitsweisen und Probleme der Nicht-Mathematiker nötig, dazu das Lesen der entsprechenden Papers. Als Gruppenleiter hatte ich dann typische Aufgaben des mittleren Managements zu erfüllen, der Mathematikanteil sinkt dabei zwangsläufig und fokussiert sich auf konzeptionelle Arbeiten. Als Key Experte kommen dann wiederum mehr strategische Fragestellungen, insbesondere die Frage, welche mathematischen Themenfelder für zukünftige Herausforderungen des Unternehmens wichtig werden könnten, und deren Anarbeitung ins Spiel. Hier hat sich der mathematische Tätigkeitskreis wieder deutlich erweitert. Dazu gehört dann auch die Vertretung des Unternehmens und des Fachs Industriemathematik in der akademischen Community.

Sind, in Ihrer Hinsicht, Mathematiker aktiv genug, um Brücken zwischen Universität und Wirtschaft/Gesellschaft zu schlagen?

Es gibt seit langem eine Vielzahl von Aktivitäten von Seiten der Mathematik aus der Universität auf nationaler und internationaler Ebene, daran mangelt es sicher nicht.

Die größte Hürde stellen meiner Meinung nach die kulturellen Unterschiede sowie eine gewisse Furcht, die Komfortzone zu verlassen, dar. Das gilt natürlich für alle Wissenschaften, Mathematiker sind aber oft doch eher introvertiert, das hilft nicht unbedingt. Es braucht eine Neugier, sich auf Neues einzulassen, und auch eine gewisse Zähigkeit und Sturheit, auch dann, wenn die mathematische Fragestellung zum „Knacken“ der Schlüsselstelle des Problems nicht auf dem Präsentierteller gereicht wird, sondern erst mühsam freigelegt werden muss.

Was raten Sie jungen Leuten hinsichtlich Mathematik?

Machen sie Mathematik, machen sie das, was sie dabei ausfüllt, und machen sie das richtig. Ganz besonders: bleiben sie neugierig und offen. Sie werden erstaunt sein, wie wichtig die Mathematik für die Bewältigung der komplexen Herausforderungen unserer Welt ist und wieviel man als Mathematiker dazu beitragen kann.



Lineare Algebra liefert die Grundlage für Beleuchtungsmodelle und die dreidimensionale Bewegung in der Computergrafik.

Cyberpunk 2077: Die Mathematik hinter der Computergrafik

Wir wollen in diesem Kapitel untersuchen, inwieweit Mathematik in Computerspielen eine Rolle spielt. Dabei werden wir kennenlernen, wie uns die Vektorrechnung (lineare Algebra) hilft, Objekte (wie z.B. Wände in einem Labyrinth) im dreidimensionalen Raum zu charakterisieren und hilft diese Objekte zu bewegen, zu rotieren. Mathematik ist auch nötig, um zu berechnen, welche Teile eines Objektes überhaupt sichtbar sind und wie sich die Beleuchtung auf die Darstellung der Oberfläche auswirkt.

Die Grundidee der 3D-Grafik besteht darin, eine mathematische Beschreibung einer Welt in ein Bild davon zu verwandeln, wie diese Welt für jemanden im Inneren der Welt aussehen würde. Die mathematische Beschreibung könnte so sein: es gibt ein Objekt Box mit dem Mittelpunkt $(1,2,3)$, Seitenlänge 2 und Farbe Gelb. Der Betrachter (genaugenommen seine Augen) befindet sich bei $(5,14,9)$ und schaut direkt auf den Mittelpunkt der Box. Damit kann man berechnen, wie die Welt für diesen Betrachter aussehen würde.

Nehmen wir nun einen Betrachter mit Augen im Punkt A ; zwischen ihm und der Box ist eine Glasplatte (Ebene), auf die er die gesehene Box malen will. Eine der Boxecken befindet sich am Punkt B , und um herauszufinden, wo diese Ecke auf der Glasplatte sein sollte, zieht man eine Linie L von seinen Augen A zur Boxecke B und sieht dann, wo diese Linie durch die Glasplatte G , verläuft. So einen Schnittpunkt Gerade mit Ebene kann man leicht bestimmen. Verfolgt man diese Prozedur für jede Stelle der Box, erhält man ein Abbild der Box auf der Glasplatte, also dem Bildschirm. Und genau das passiert sehr grob im Computer, wenn man z.B. im Spiel wie Cyberpunk 2077 herumläuft, obwohl die Details doch etwas komplizierter sind.

Das Raytracing-Verfahren

Was gerade oben beschrieben wurde, ähnelt dem, was der Computer jedes Mal tut, wenn Sie im Spiel wie Cyberpunk 2077 herumlaufen, obwohl die Details etwas anders sind. Die neuesten Computerspiele verwenden kompliziertere Beschreibungen der Welt, indem sie gekrümmte Oberflächen, NURBS und andere seltsam klingende Dinge verwenden, aber am Ende reduziert sich das Ganze immer auf Dreiecke.

Abbildung 103: Titelbild Cyberpunk 2077 (Quelle: CD Projekt Red, Press Center)

Nehmen wir einmal an, dass wir unser Objekt in Dreiecke trianguliert haben. Der nächste Schritt ist das Beleuchtungsmodell: die Menge des Lichts oder Schattens, die auf jede Dreiecksfläche fällt, hängt von der Ausrichtung dieses Dreiecks in Bezug auf die Lichtquelle ab. Wenn ein Dreieck direkt der Lichtquelle zugewandt ist, wird es hell beleuchtet, aber wenn es von der Lichtquelle abgewandt ist, liegt es im starken Schatten. Die Ausrichtung jeder einzelnen Dreiecksfläche wird durch einen Normalenvektor beschrieben, der senkrecht aus der jeweiligen ebenen Fläche heraus zeigt. In ähnlicher Weise können die Strahlen der Lichtquelle durch einen Vektor beschrieben werden, und das Skalarprodukt dieser beiden Vektoren ergibt den Cosinus des Winkels zwischen diesen beiden Vektoren.

An dieser Stelle wollen wir kurz das *Skalarprodukt* vorstellen. Um das Skalarprodukt zweier Vektoren $u = (x_1, y_1, z_1)$ und $v = (x_2, y_2, z_2)$ zu berechnen, multipliziert man deren jeweilige Koordinaten und addiert sie:

$$(u, v) = x_1x_2 + y_1y_2 + z_1z_2.$$

Das Skalarprodukt kann auch als Produkt aus der Länge der beiden Vektoren und dem Winkel θ zwischen ihnen geschrieben werden:

$$(u, v) = (\text{Länge von } u) \cdot (\text{Länge von } v) \cdot \cos \theta.$$

Wenn ein Dreieck direkt der Lichtquelle zugewandt ist, beträgt der Winkel zwischen der Normalen und der Richtung der Lichtstrahlen 0° , und das Skalarprodukt dieser beiden Vektoren ergibt den Wert 1, d.h. es fällt die maximale Lichtmenge auf diese Fläche und sie sollte daher weiß schattiert werden. Wenn eine dreieckige Fläche fast seitlich zu den Lichtstrahlen liegt, so dass der Winkel zwischen der Flächennormalen und der Lichtrichtung nahe bei 90° liegt, ist der Wert des Skalarprodukts nahe bei 0, also sollte das Dreieck sehr dunkel schattiert werden, weil es im Schatten liegt. Ist ein Dreieck vom Licht abgewandt, also der Winkel größer als 90° ist, dann wird das Skalarprodukt negativ und dieser Wert wird normalerweise vom Computer auf Null gesetzt. Nun kommt hinzu, dass sich alle 3D-Objekte bewegen. Somit muss der Computer diese (relativ einfachen) Berechnungen ständig für alle Dreiecke wiederholen (das sog. *Rendering*). Hierzu gibt es inzwischen spezialisierte Hardware, die diese Berechnungen erledigt. Häufig wird auch die Triangulierung vergrößert, damit die Berechnungen für schnell bewegte Objekte effektiver wird; das menschliche Auge kann ohnehin nicht so fein auflösen.



Abbildung 104: Screenshot 2077 (Quelle: CD Projekt Red, Press Center)

Als nächsten Schritt möchte man die Oberflächen realistischer erscheinen lassen, ohne die Anzahl der Dreiecke weiter zu erhöhen. Das geschieht mit dem Konzept der Textur. Für Anwendungen, die eine schnellere Rendering-Zeit erfordern, wie z.B. das Echtzeit-Rendering in Computerspielen, verwendet man das *Normal Mapping*. Es entspricht dem zuvor beschriebenen Verfahren, bei dem die Textur mit ihren Unebenheiten durch lokal hochaufgelöste Modelle erzielt wird.

Für Anwendungen, wie z.B. in Filmen, die lange Zeiten für das Rendern einzelner Bilder zulassen, gibt es eine andere Methode. Man versucht mit dem *Displacement Mapping* mit positiven und negativen Werten genau zu kodieren, wie weit sich die Punkte auf der Oberfläche des vereinfachten Gitters von den Punkten auf dem Original-Gitters unterscheiden.

Mathematik hilft selbst
schwierigste Sudokus zu lösen.

Sudoku – Spaß, aber auch ernsthafte Mathematik

				9			5	
	1						3	
		2	3				7	
		4	5				7	
8							2	
					6	4		
	9			1				
	8			6				
		5	4					7

Das Lösen von Sudoku ist seit 2004 ein sehr beliebter Zeitvertreib und kann mit Recht als das Rätsel des 21. Jahrhunderts bezeichnet werden. In Wartezimmern, in Zügen und in Flugzeugen sind überall Menschen mit Sudoku-Rätselbüchern zu sehen. Was Sie tun müssen, um ein Sudoku-Rätsel zu lösen, kann in einer Minute erklärt werden. Aber wie man das macht, das ist eine andere Geschichte. Es gibt einfache Sudokus, und es gibt sehr schwierige. In den Rätselbüchern, die man kaufen kann, wird angegeben, wie schwierig die Rätsel sind, und die schwierigsten Rätsel werden mit 10, 11 oder sogar 15 Sternen bewertet.

Auch Mathematiker haben großes Interesse an Sudoku. Sie versuchen, sehr schwierige Sudokus zu erstellen, die für einen Menschen fast unmöglich zu lösen sind. Das Rätsel hier auf der Titelseite ist derzeit das schwierigste der Welt!

Dieses Rätsel kann mit den im Buch „Löse jedes Sudoku“ beschriebenen Techniken gelöst werden, erfordert aber viel Erfahrung und wird Sie trotzdem einen ganzen Tag Arbeit kosten.

Für Mathematiker gibt es aber noch viele andere interessante Fragen. Zum Beispiel: Wie viele Ziffern sollte man mindestens vorschreiben, um sicherzustellen, dass ein Sudoku eindeutig lösbar ist? Die Antwort ist 17 und wurde 2012 von einem irischen Mathematiker, Gary McGuire, bewiesen, aber er brauchte dafür 7 Millionen Stunden auf einem leistungsstarken Computersystem!

Eine weitere interessante Frage ist: Wie viele verschiedene Sudokus gibt es? „Unterscheidbar“ bedeutet, dass mindestens eine Zahl im Rätsel anders ist; Symmetriebeziehungen (wie Drehungen) werden nicht berücksichtigt, sie zählen als unterschiedlich. In einer Studie aus dem Jahr 2005 berechneten Bertram Felgenhauer und Frazer Jarvis die Anzahl der unterscheidbaren Sudokus mit mathematischen Mitteln (Gruppentheorie) und kamen auf die Zahl 6.670.903.752.021.072.936.960 und bestätigten damit den Wert, den Günther Stertenbrink bereits 2003 postuliert hatte. Bertram Felgenhauer ist ein sehr begabter Mathematiker, der 1995 bzw. 1996 die Silber- und Goldmedaille der Internationalen Mathematik-Olympiade gewonnen hat.

Mathematiker verwenden Sudokus oft, um ihre mathematischen Algorithmen zu testen, zum Beispiel bei Graphenfärbeproblemen. Interessieren Sie sich für weitere mathematische Fakten von Sudoku? Siehe die Webseite <https://www.loesejedessudoku.de>.



Die geheime Strategie der Bank

In der Fernsehshow „Deal Or No Deal“ bzw. „Millionendeal“ wählt der Kandidat einen von 26 (oder 20) Koffern aus, in denen sich jeweils eine Geldsumme zwischen 0,01 € und 2.000.000 € befindet. Es werden alle Geldbeträge auf einer Anzeigetafel im Fernsehstudio angezeigt, aber man weiß nicht, welcher Betrag in welchem Koffer ist. Das Spiel beginnt, indem sich der Kandidat einen Koffer aussucht, der zunächst verschlossen bleibt. Danach wählt der Kandidat nacheinander weitere Koffer aus, die geöffnet werden und deren Betrag von der Anzeigetafel verschwindet, da sich deren Betrag nicht im verschlossenen Koffer befinden kann.

Im Verlauf des Spiels, nach dem eine gewisse Anzahl von Koffern geöffnet wurde, macht die „Bank“ dem Kandidaten ein Angebot (ein „Deal“), wenn er seinen ausgewählten Koffer an sie verkauft. Dieses Angebot orientiert sich an den Geldbeträgen, die noch auf der Anzeigetafel sind. Je mehr höhere Beträge noch im Spiel sind, desto höher ist das Angebot der Bank. Es scheint, als ob das Angebot der Bank (irgendwie) auf dem Durchschnitt der Geldbeträge, die noch im Spiel sind (der Erwartungswert des Spiels zu diesem Zeitpunkt), basiert.

Nimmt der Kandidat dieses Angebot an, ist das Spiel vorbei und er erhält diesen Geldbetrag von der Bank. Lehnt er den Deal ab, setzt sich das Spiel fort und es werden weitere Koffer geöffnet und es kommt zu weiteren Angeboten der Bank. Lehnt der Kandidat alle Angebote der Bank ab, erhält er den Geldbetrag aus dem ursprünglich ausgewählten Koffer. Es gibt hierbei auch Varianten, bei denen (wenn der Kandidat vor Runde 7 seinen Deal macht) das im Hypothetischen zu Ende gespielt wird, d.h. es werden die verbleibenden Runden mit den Angeboten so ausgespielt, als wäre das Spiel noch nicht vorbei.

Die Frage hier ist, ob es sich um einen fairen Wettbewerb handelt oder ob die Show ihre Teilnehmer „betrügt“. Man kann mit einfachen mathematischen Begriffen, wie z.B. Erwartungswert und Ausgleichsrechnung (siehe auch Kapitel über robuste Inzidenzzahlen, Seite 14) untersuchen, ob Unregelmäßigkeiten oder spezielle Muster vorliegen.

Der Mathematik-Student Daniel R. Shifflet von der Bowling Green State University ist in seinem Artikel⁶⁷ genau dieser Frage nachgegangen und hat folgendes festgestellt. In den noch laufenden Spielrunden zeigten erste Berechnungen, dass der Banker den Teilnehmer Geldbeträge weit unter ihrem Erwartungswert anbot. Andererseits, in den oben beschriebenen hypothetischen Runden lagen diese Angebote der Bank nahe oder sogar über diesem Erwartungswert des Spiels. Hatte also der Banker fair gespielt oder „mit den Zahlen herumgespielt“?

Daniel R. Shifflet teilte die Daten in die beiden zu testenden Situationen auf:

- 1 offizielle Angebote an Teilnehmer, die noch im Spiel sind
- 2 hypothetische Angebote an Teilnehmer, die einen Deal angenommen haben.

Da Teilnehmer mit großen Geldbeträgen im Spiel offensichtlich größere Angebote bekommen würden als Teilnehmer mit nur kleinen Geldbeträgen, die noch übrig sind (offiziell oder hypothetisch), ist es besser, diese Angebote als einen Prozentsatz des Erwartungswerts des Spielergebnisses zu betrachten. Dies würde uns einen Wert geben, der unabhängig davon ist, ob der Kandidat einen „Glückstag“ hatte oder nicht. Wenn die Bank „fair spielt“, sollten die Mittelwerte dieser Prozentsätze Runde für Runde für die offiziellen und hypothetischen Situationen ungefähr gleich sein.

In einem Beispiel-Datensatz stellte Daniel R. Shifflet fest, dass die Teilnehmer, die noch im Spiel sind, Angebote von der Bank bekommen, die im Durchschnitt etwa 76 % ihres Erwartungswertes ausmachen, während den Teilnehmern, die nur hypothetisch spielen, Angebote von etwa 96 % desselben Erwartungswertes gemacht werden. Dieser Unterschied von 20 Prozentpunkten scheint darauf hinzudeuten, dass die Bank nicht die gleichen Regeln für die Abgabe von Angeboten befolgt, wenn das Spiel noch aktiv ist und wenn es hypothetisch ist.

Aufgrund der vielen Datensätze vermutete Daniel R. Shifflet einen linearen Zusammenhang zwischen dem Erwartungswert des Spiels und der Angebote der Bank. Mithilfe einer Ausgleichsrechnung fand er so eine Gerade, die am besten zu diesen Daten passte. Somit ergab sich ein Modell dafür, wie die Bank zu ihren Angeboten kommt. Sein Modell lieferte die folgenden Ergebnisse. Für Teilnehmer, die in Runde 6 noch im Spiel sind, würde die Bank Angebote in Höhe von 65,11 % des Erwartungswertes machen. Auf der anderen Seite, für Teilnehmer, die nur hypothetisch das Spiel in Runde 6 spielen, liefert sein Modell (nach Entfernen eines sog. Ausreißerwertes) ein Angebot von 94 % des Erwartungswertes, also eine drastische Steigerung gegenüber der früheren Situation.

In Anbetracht all dieser Informationen, gibt es starke Hinweise darauf, dass der Banker von seiner Formel zur Berechnung der Deals (was auch immer das ist) während der hypothetischen Runden des Spiels abweicht. Er scheint mehr Geld zu bieten als er es getan hätte, wenn die Teilnehmer noch tatsächlich spielen und in der Lage wären das Geld anzunehmen.

⁶⁷ D.R. Shifflet, Is Deal or No Deal Cheating Its Contestants?, *Ohio Journal of School Mathematics* (63) (2011), 5–10.
https://kb.osu.edu/bitstream/handle/1811/78122/OJSM_63_Spring2011_05.pdf

Danksagungen

Wir möchten uns an dieser Stelle bei Allen bedanken, die zu diesem Buch beigetragen haben

José Acosta Fernandez,
Wuppertal Institut für Klima, Umwelt, Energie

Dr. Maria Barbarossa,
Frankfurt Institute for Advanced Studies (FIAS)

Dr. Petra Bauer,
Siemens AG, Technology, München

Robin Befeldt,
Polizei Berlin

Dr. Andreas Binder,
MathConsult GmbH, Linz

Prof. Dr. Ralf Borndörfer,
FU Berlin und Zuse-Institute Berlin

Prof. Dr. med. Helmut Brunner,
Universität Düsseldorf, Bergische Universität Wuppertal

Dr. Yao-Li Chuang,
California State University, Los Angeles

Dr. Ágnes Cseh,
Hasso-Plattner-Institut, Potsdam

Sören Dittmer,
Universität Bremen

Prof. Dr. Maria Rita D'Orsogna,
California State University, Los Angeles

Dr. Martin Friesen,
Dublin City University

Daniela Gaul,
Bergische Universität Wuppertal

Dr. Sascha Härtel,
TSG Hoffenheim

Dr. Thomas Hahn,
Chief Expert Software, Siemens AG, Nürnberg

Dr. Dirk Hartmann,
Senior Principal Scientist, Simulation and Digital Twin, Siemens AG,
München

Dr. Markus Hinterkausen,
Zentralbereich Forschung und Voraentwicklung, Robert Bosch
GmbH

Leon Hoffe,
Bergische Universität Wuppertal

Dr. Susanne Hoffmeister,
InVision AG, Düsseldorf

Dr. Markus Hoschek,
HEAG Holding AG, Darmstadt

Prof. Dr. Wilhelm Huisinga,
Universität Potsdam

Dr. Uwe Iben,
Chief Expert Applied Mathematics, Robert Bosch GmbH, Renningen

Dr. Markus Kantner,
Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin

PD Dr. Jörg Kienitz,
Bergische Universität Wuppertal, Quaternion, AcadiaSoft

Prof. Dr. Kathrin Klamroth,
Bergische Universität Wuppertal

Dr. Margrit Klitz,
Deutsches Zentrum für Luftund Raumfahrt (DLR), Köln

Tobias Kluth,
Universität Bremen

Dr. Thomas Koprucki,
Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin

Dr. Igor Kossaczky,
Aptiv, Wuppertal

Prof. Dr. Ekaterina Kostina,
Ruprecht-Karls-Universität Heidelberg

Christine Krüger,
Wuppertal Institut für Klima, Umwelt, Energie

Mira Küpper,
Bergische Universität Wuppertal

Prof. Dr. Stefan Kurz,
TU Darmstadt und Robert Bosch GmbH

Benjamin Leonhardt,
Bergische Universität Wuppertal

Johannes Leuschner,
Universität Bremen

Prof. Dr. Peter Maass,
Universität Bremen

Corinna Maier,
AbbVie Deutschland GmbH & Co. KG

Prof. Jan Mayer,
TSG Hoffenheim

Prof. Dr. Nicole Megow,
Universität Bremen

Prof. Dr. Volker Mehrmann,
TU Berlin, Präsident European Mathematical Society

Dr. Michael Meister,
Parlamentarischer Staatssekretär, Bundesministerium für Bildung und
Forschung

Dr. Georg Mogk,
Principal Expert Applied Mathematics, Bayer AG, Leverkusen

Dr. Thomas Mrziglod,
Head of Applied Mathematics, Bayer AG, Leverkusen

Niels Neveling,
Senior Consultant, B&W Deloitte GmbH,

Prof. Dr. Jörg Rambau,
Universität Bayreuth

Ronan Richter,
Universität Bayreuth

Dr. Kim Rossmo,
School of Criminal Justice & Criminology, Texas State University

Dr. Matthias Rottmann,
Bergische Universität Wuppertal

Prof. Dr. Helga Rübsamen-Schaeff,
ehem. CEO von AiCuris GmbH, Wuppertal, Nationale Akademie der
Wissenschaften Leopoldina

Dr. Guillaume Sagnol,
TU Berlin

Dr. Hanna Schilar,
Bergische Universität Wuppertal

Maximilian Schmidt,
Universität Bremen

Prof. Dr. Anita Schöbel,
Institutsleiterin des Fraunhofer-Instituts für Techno- und
Wirtschaftsmathematik (ITWM), Kaiserslautern

Prof. Dr. Ingo Scholtes,
Bergische Universität Wuppertal

Dr. Britta Schulze,
Bergische Universität Wuppertal

Prof. Dr. Andreas Schuppert,
RWTH Aachen und Bayer AG

Prof. Dr. Armin Seyfried,
Bergische Universität Wuppertal

PD Dr. Michael Stiglmayr,
Bergische Universität Wuppertal

Matthias Stursberg,
R&D Manager Vaillant GmbH, Remscheid

Tobias Suszka,
Bergische Universität Wuppertal

Alexander Tesch,
Zuse-Institute Berlin

Jun.-Prof. Dr. Antoine Tordeux,
Bergische Universität Wuppertal

Dr. Uwe Vollmer,
Zentralbereich Forschung und Voraentwicklung, Robert Bosch
GmbH

Dr. Utz Wever,
Senior Principal Scientist, Simulation and Digital Twin, Siemens AG,
München

Oliver Wohak,
d-fine GmbH, Frankfurt am Main

Impressum

Autoren

Matthias Ehrhardt, Michael Günther, Wil Schilders
Bergische Universität Wuppertal
Fakultät 4 - Mathematik und Naturwissenschaften
Lehrstuhl für Angewandte Mathematik und Numerische Analysis
Gaußstrasse 20
D-42119 Wuppertal

weitere Autoren, siehe
www.erfolgsformeln.uni-wuppertal.de

Lektorat & Übersetzung

Matthias Ehrhardt
weitere Mitwirkende: siehe Danksagungen

Verlag & Druck

Books4You,
Helftheuvelweg 11, 5222 AV, 's Hertogenbosch, Niederlande
www.prijnsprinter.nl

Gestaltung

WAT ontwerpers, Utrecht, Niederlande
www.watontwerpers.nl

Bilder

Wenn nicht anders angegeben www.iStock.com

© 2021

Das Werk, einschließlich seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung der Autoren unzulässig. Dies gilt insbesondere für die elektronische oder sonstige Vervielfältigung, Übersetzung, Verbreitung und öffentliche Zugänglichmachung.

Bibliografische Information der Deutschen Nationalbibliothek:
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in
der Deutschen Nationalbibliografie; detaillierte bibliografische
Daten sind im Internet über
<http://dnb.d-nb.de> abrufbar.

Erfolgsformeln

Anwendungen der Mathematik

Der bekannte Mathematiker Lex Schrijver sagte einmal: "Mathematik ist wie Sauerstoff. Wenn sie da ist, merkt man sie nicht. Wenn sie nicht da wäre, würde man merken, dass man ohne sie nicht leben kann". Dieses Buch soll ein wenig von der Mathematik sichtbar machen, die wir jeden Tag einatmen.

Dabei stellt sich heraus, dass Mathematik auch im Sport präsent ist, in der Medizin und in der Wettervorhersage, im öffentlichen Verkehr, bei der Lösung von Mordfällen und vieles mehr. Neun prominente Wissenschaftler sprechen auch über die Rolle der Mathematik in ihren Beruf.

Dieses Buch ist ein Ergebnis der „Dr. Jörg Mittelsten Scheid-Gastprofessur“ von Wil Schilders an der Bergischen Universität Wuppertal im Wintersemester 2020/2021.



Matthias Ehrhardt ist Professor für Angewandte Mathematik und Numerische Analysis am IMACM, dem Institut für Mathematische Modellierung, Analysis und Computermathematik an der Bergischen Universität Wuppertal.

Seine Forschungsschwerpunkte sind die numerische Behandlung von partiellen Differentialgleichungen, insbesondere auf unbeschränkten Gebieten und numerische Methoden in der Finanzmathematik. Er bietet auch in seinem Programm "MATEMA - MAThematik Erlebt iM Alltag" Vorträge an Schulen an.



Michael Günther ist Professor für Angewandte Mathematik und Numerische Analysis am IMACM, dem Institut für Mathematische Modellierung, Analysis und Computermathematik an der Bergischen Universität Wuppertal.

Sein Hauptforschungsgebiet ist die Modellierung, Analyse und numerische Simulation von zeitabhängigen gekoppelten heterogenen Problemen mit Anwendungen in den Bereichen Computational Engineering, Computational Finance und Computational Physics. Er ist Vorsitzender von KoMSO e.V., dem Komitee für Modellierung, Simulation und Optimierung, das bundesweit die Kräfte der Industriemathematik bündelt.



Wil Schilders ist Professor für „Scientific Computing in the Industry“ an der Technischen Universität in Eindhoven, Niederlande.

Er ist Experte auf dem Gebiet der Entwicklung mathematischer Methoden, die bei Simulationen von Produkten und Prozessen in der Industrie eingesetzt werden. Er ist der Geschäftsführer der niederländischen Plattform für Mathematik, die 2014 ein ähnliches Buch veröffentlicht hat.