Mathematische Modelle von Epidemien

am Beispiel von SARS-CoV-2

Masterarbeit

zur Erlangung des akademischen Grades Master of Education (MEd)

an der Karl-Franzens-Universität Graz

vorgelegt von

Ian GOBETZ, BEd

am Institut für Mathematik und Wissenschaftliches Rechnen

Begutachter: Ao.Univ.-Prof. Dr.phil. Georg Propst

Graz, 2021



Danksagung

An erster Stelle gebührt mein Dank Herrn Ao.Univ.-Prof. Dr. Georg Propst, der meine Masterarbeit mühevoll betreut und begutachtet hat. Für die hilfreichen Anregungen und die konstruktive Kritik bei der Erstellung dieser Arbeit möchte ich mich herzlich bedanken.

Darüber hinaus möchte ich mich bei all denjenigen bedanken, die mich während der Erstellung der Masterarbeit unterstützt und motiviert haben.

Ein besonderer Dank gilt außerdem Sandra Klemm, BEd MA, die mich bei der Korrektur des Abstract in Englisch unterstützte.

Ich bedanke mich bei meinen Freunden, besonders bei Lena Haindl für das Korrekturlesen meiner Masterarbeit.

Abschließend möchte ich mich bei meinen Eltern bedanken, die mir mein Studium durch ihre Unterstützung überhaupt erst ermöglicht haben, stets ein offenes Ohr für mich hatten und mir auch in schwierigen Phasen immer wieder mit Rat und Tat zur Seite standen.

Abstract

Saisonale Grippewellen verursachen jedes Jahr zahlreiche gesundheitliche und soziale Probleme, die von Fehlzeiten und Krankenständen über Spitalsaufenthalte bis hin zu Todesfällen reichen. Das alltägliche Leben wird dadurch aber nicht sonderlich beeinflusst. Doch mit Ende des Jahres 2019 kam alles anders, als ein neuartiges Virus erhebliche gesundheitliche Probleme auslöste.

Zentrum des Ausbruchs ist die chinesische Millionenstadt Wuhan, in der am 31. Dezember 2020 erste Fälle von Lungenentzündungen mit unbekannter Ursache gemeldet wurden. Etwa drei Monate später, am 12. März 2020, wurde der COVID-19-Ausbruch durch die WHO zu einer weltweiten Pandemie erklärt.

Aufgrund der Neuartigkeit des Virus und den fehlenden Medikamenten, muss das soziale Leben stark eingeschränkt werden damit die Ausbreitung der Krankheit eingedämmt werden kann. Die Anzahl der Infizierten soll mit mathematischen Modellen beschrieben werden, um beispielsweise zukünftige Wachstumsraten zu prognostizieren.

Das Ziel der vorliegenden Masterarbeit ist es die Ausbreitung des Coronavirus in Österreich durch Epidemiemodelle zu modellieren. Dafür wird folgende Forschungsfrage gestellt: „Kann das epidemiologische Verhalten der Corona-Pandemie für Österreich mit Epidemiemodellen repräsentiert werden“.

Um diese Forschungsfrage beantworten zu können, werden die Epidemiemodelle zunächst theoretisch aufgearbeitet und diskutiert. Danach werden diese Modelle gemeinsam mit echten Daten in Matlab implementiert. Die Parameter der Differentialgleichungen der zugrunde liegenden Modelle werden in Matlab mittels Nelder-Mead-Verfahren optimiert und die Lösungen im Anschluss interpretiert.

Allgemein lässt sich festhalten, dass, je komplexer ein Modell in seiner Struktur aufgebaut ist, desto genauer kann die Lösungskurve an die wahren Daten angepasst werden.

Every year seasonal waves of influenza causes numerous health and social problems ranging from sick leaves and hospitalizations to deaths. However, everyday life is not particularly affected by this but in the end of 2019 everything turned out differently. A new type of virus caused significant health problems.
In December 31, 2020 the virus was first recognized in a Chinese city named Wuhan. The city reported about cases of pneumonia with an unknown cause. About three months later, on March 12, 2020 the WHO declared the COVID-19-outbreak as a global pandemic.
Due to the novelty of the virus and the lack of drugs, social life has to be severely limited to avoid the spread of the disease. The number of infected people can be described within mathematical models in order to predict future growth rates.
The aim of this master thesis is to model the Austrian spread of the coronavirus with mathematical models. Therefore, the following research question will be discussed: „Is it possible to represent the epidemiological spread of the coronavirus in Austria by mathematical models?“.
To answer this question some theoretical fundamentals of these models have to be discussed. Then both models and real data are implemented in Matlab where calculations are done.
Matlab uses Nelder-Mead Simplex Algorithm to find optimal parameters in the differential equations. After optimization the results need to be interpreted and discussed. Although the models do all have the same structure, the more complex a model is, the more precisely a solution curve can be adapted to the real data.

Inhalt

Abstract **III**

1. Einführung **1**
1.1. Problemstellung und Zielsetzung . . . 1
1.2. Mathematische Modellierung . . . 3
1.2.1. Analyse des Anwendungsproblems . . . 5
1.2.2. Modellbildung . . . 5
1.2.3. Mathematische Analyse des Modells . . . 6
1.2.4. Berechnungen und Simulationen . . . 6
1.2.5. Interpretation und Validierung . . . 7
1.3. Epidemiologische Grundbegriffe . . . 7
1.4. Geschichte der Epidemiologie . . . 8
1.5. Mathematische Modellierung von Infektionskrankheiten . . . 9
1.6. Grundlegende Definitionen . . . 10
1.7. Methodik . . . 16
1.7.1. Datenerhebung . . . 17
1.7.2. Datenauswertung . . . 18

2. Mathematische Modelle zur Modellierung von Infektionskrankheiten **20**
2.1. SI-Modell . . . 20
2.1.1. Mathematische Modellierung . . . 21
2.1.2. Berechnung der Lösungen . . . 22
2.1.3. Langzeitverhalten der Klassen S und I . . . 24
2.2. SIS-Modell . . . 25
2.2.1. Fixpunkte und Langzeitverhalten . . . 26
2.2.2. Stabilitätsanalyse . . . 27
2.3. SIR-Modell . . . 28
2.3.1. Basisreproduktionszahl . . . 29
2.3.2. Trajektorien und Phasenraum . . . 30
2.3.3. Maximale Infiziertenzahl und Langzeitverhalten der Klasse I . . . 31
2.3.4. Langzeitverhalten der Klasse S . . . 32
2.3.5. Langzeitverhalten von R . . . 34
2.3.6. Stabilitätsanalyse . . . 36

2.4. SEIR-Modell 37
2.4.1. Trajektorien und Phasenraum 39
2.4.2. Langzeitverhalten der Klassen S, I und R 39
2.5. SEIR Modell mit Populationsdynamiken 40
2.5.1. Basisreproduktionszahl 42
2.5.2. Fixpunkte und Stabilitätsanalyse 43
2.6. SEIQR Modell mit Populationsdynamik 45
2.6.1. Fixpunkte und Stabilität 47

3. Ergebnisse und Diskussion **49**
3.1. Daten 49
3.1.1. Minimumsuche in Matlab 50
3.2. SIR-Modell 52
3.2.1. Lösungen des allgemeinen SIR-Modells 52
3.2.2. Anpassung an die realen Daten 53
3.2.3. Vorhersage und Prognose 56
3.3. SEIR-Modell mit Populationsdynamik 58
3.3.1. Lösungen des SEIR-Modells mit Populationsdynamik 59
3.3.2. Anpassung an die realen Daten 60
3.3.3. Vorhersage und Prognose 65
3.4. SEIQR-Modell mit Populationsdynamik 67
3.4.1. Lösungen des allgemeinen SEIQR-Modells 67
3.4.2. Anpassung an die realen Daten 69
3.4.3. Vorhersage und Prognose 70

4. Zusammenfassung und Fazit **72**

Abbildungsverzeichnis **78**

Tabellenverzeichnis **80**

Anhang **81**
Infektionszahlen der beiden Corona Wellen 81

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen inländischen oder ausländischen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Die vorliegende Fassung entspricht der eingereichten elektronischen Version.

Graz, am 21. Juli 2021 .....

(Ian Gobetz)

Kapitel 1 Einführung

1.1 Problemstellung und Zielsetzung

„Mathematik ist das Alphabet, mit dessen Hilfe Gott das Universum beschrieben hat.“ Ein Zitat von Galileo Galilei, das die Wichtigkeit der Mathematik betont obwohl sie meist im Hintergrund (z.B. in technischen Anwendungen) passiert und so für viele Menschen unwichtig und nicht greifbar erscheint. Ein modernes Teilgebiet der Mathematik stellt die mathematische Biologie dar. Mit den umfassenden Arbeiten zur Populationsdynamik mit Räuber-Beute Modellen von Lotka und Volterra oder mit den mathematischen Modellen in der Epidemiologie von Kermack und McKendrick wurden im frühen 20. Jahrhundert wesentliche wissenschaftliche Beiträge für die mathematische Biologie geleistet. Seither ist dieses Teilgebiet der Mathematik wohletabliert und gewinnt immer mehr an Bedeutung.
Seit Bestehen der Menschheit sind Epidemien und Pandemien wie Influenza, Pest, Cholera usw. verantwortlich für Millionen Todesfälle weltweit. Tritt eine neue Krankheit erstmals in der Bevölkerung auf, sind in der Regel wenig bis keine Medikamente vorhanden und die Krankheit breitet sich in der Bevölkerung ungehindert aus. Abgesehen von der Arzneimittelforschung ist man daran interessiert zu erforschen wodurch die Krankheit ausgelöst wurde, damit ein erneuter Ausbruch bzw. die weitere Verbreitung möglichst verhindert werden kann. Des Weiteren stellt sich die Frage, wie sich eine Krankheit in der Bevölkerung – möglicherweise über Staatsgrenzen oder sogar über Kontinente hinweg – ausbreitet. Hierbei kommt nun die mathematische Biologie ins Spiel, mit deren Hilfe man unter anderem versucht Modelle zu entwickeln mit denen mögliche Krankheitsverläufe beschrieben und prognostiziert werden können. Andererseits kann nach Einsatz geeigneter Modelle auch die Frage nach dem Einnisten bzw. Fortbestehen einer Krankheit innerhalb der Bevölkerung beantwortet werden. Der grundlegende Aufbau derartiger Modelle wird in den nachfolgenden

Kapiteln beschrieben.
Die jüngste Pandemie, die weltweit die Menschheit, die Medizin und auch die Wirtschaft in Atem hält, ist die sogenannte Corona(virus)-Pandemie (COVID19-Pandemie). Coronaviren sind umschlossene Ribonukleinsäure-Viren (vgl. Tesini 2021), die vor allem unter Vögeln weit verbreitet sind (vgl. RKI 2021). Bislang haben Coronaviren im 21. Jahrhundert drei Mal die Barriere zwischen Tier und Mensch durchbrochen (vgl. Wit et al. 2016, S. 523).
Das Enzym ACE2, das im Körper normalerweise für die Regulation des Blutdrucks und zum Schutz vor Herz-Kreislauf-Erkrankungen verantwortlich ist, wird von den Coronaviren als Rezeptor ausgenutzt, um an die Körperzellen des Menschen zu gelangen und sie zu infizieren (vgl. Medizinische Universität Wien 2021).
Die Übertragung zwischen Menschen erfolgt vor allem durch soziale Kontakte mit infizierten Personen, zumeist über die Aufnahme von virushaltigen Partikeln über die Atemwege. Bei vielen infizierten Menschen löst der Befall überwiegend milde Erkältungssymptome aus. Es können aber auch schwere Krankheitsverläufe wie eine Pneumonie (Lungenentzündung), Leberfunktionsstörungen, Nierenerkrankungen, usw. auftreten. Studien zu Langzeitfolgen sind in der Entwicklung und liefern bisher keine einheitlichen Ergebnisse (vgl. RKI 2021).

Ausgebrochen ist die Pandemie in Wuhan, einer chinesischen Millionenstadt, als der WHO Fälle von Lungenentzündungen unbekannter Ursache gemeldet wurden. Am 7. Januar 2020 wurde bestätigt, dass diese durch ein neuartiges Coronavirus ausgelöst wurden. Bereits am 30. Januar 2020 wurde das Virus in 18 Ländern nachgewiesen und durch die WHO eine gesundheitliche Notlage ausgerufen. Am 11. März – also etwa drei Monate nach dem ersten Nachweis des Virus – wurde die Krankheit offiziell zu einer weltweiten Pandemie erklärt. Mitte März bildete Europa das Epizentrum der Pandemie (vgl. WHO 2021).
Am 25. Februar wurden die ersten beiden Fälle des Virus in Österreich bestätigt. Seit jeher breitete sich das Virus rasant aus (vgl. SALZBURG24 2021). Die Globalisierung und die weltweite Vernetzung spielen dabei eine wesentliche Rolle (vgl. Razum et al. 2016, S. 109ff).
Der Urpsrung des Virus ist bis heute unbekannt. Bislang existieren zwei Theorien als Quelle der Pandemie. Eine davon ist die sogenannte „Wildtier-Theorie“ und die andere die „Laborunfall-Theorie“ (vgl. Wiesendanger 2021). Auf beide wird in dieser Arbeit nicht näher eingegangen.
Das Ausgangsproblem dieser Arbeit ist die Modellierung der Infektionsausbreitung des neuartigen Coronavirus SARS-CoV-2. Durch die Neuartigkeit und Unbekanntheit,

den zu anfangs fehlenden Impfstoffen und Medikamenten und durch die fehlende Immunität gegen die Coronaviren innerhalb der Bevölkerung konnte sich das Virus zunächst ungehindert ausbreiten und wurde sehr rasch zu einer weltweiten gesundheitlichen Bedrohung.
Mithilfe von mathematischen Modellen kann ein möglicher Verlauf der Infektionsausbreitung modelliert und daraus Gesetzmäßigkeiten abgeleitet werden, um damit weitere Maßnahmen gegen die Ausbreitung treffen zu können.
Unter diesem Aspekt soll das Ziel der Arbeit sein, das von Kermack und McKendrick entwickelte und weitere, komplexere Epidemiemodelle an die realen Daten der Corona-Pandemie für Österreich anzupassen.

In der Literatur gibt es bereits einige Autoren, die sich mit ähnlichen Überlegungen zur Modellierung der Corona-Pandemie befassten und mathematische Modelle für die Anpassung an wahre Daten verwendeten. Für Österreich sind bis dato keine ähnlichen Veröffentlichungen bekannt.
Es sei hierbei ausdrücklich erwähnt, dass die Modelle, die dieser Arbeit zugrunde liegen sehr vereinfachte Abbildungen der Realität darstellen. Die theoretischen Überlegungen werden von Modell zu Modell zwar stets komplexer, dennoch werden aufgrund technisch eingeschränkter Ressourcen zahlreiche Vereinfachungen dargelegt bzw. bestimmte natürliche Vorgänge ausgeblendet.
Bevor näher darauf eingegangen wird, sollen im folgenden Kapitel die Grundprinzipien der Mathematischen Modellierung konkretisiert werden.

1.2 Mathematische Modellierung

Würde man einen Passanten in der Innenstadt damit konfrontieren, dass er sich gerade mit Modellierung beschäftige, würde er wahrscheinlich zunächst widersprechen. Doch genau das tut derjenige. Unbewusst tun wir dies alle jeden Tag: Über die Sinnesorgane nehmen wir Reize der Umwelt auf - sei es das Wetter, das Verkehrsgeschehen oder die Interaktion mit anderen Menschen – die in unserem Gehirn zu einem Abbild bzw. einem Modell der Umwelt verarbeitet werden.
Seit jeher beschäftigten sich Menschen mit Modellierung. Höhlenmalereien, auf denen bildhaft Jagdszenen modelliert wurden, zeigten beispielsweise die Rollenverteilung der beteiligten Personen.
Ein Modell dient aber oft nicht nur dazu, bereits existierende Objekte abzubilden, sondern auch dazu, neue Objekte zu entwerfen und systematisch weiterzuentwickeln.

Als Beispiel lassen sich hierbei architektonische Modelle nennen, die zuerst verkleinert nachgebaut oder auf Papier gezeichnet werden. Mit diesen Modellen kann dann in einem iterativen Prozess nach den optimalen Lösungen gesucht werden. Kleinere Details, die wenig Einfluss auf die Realisierung der Aufgabe haben, werden oft vernachlässigt, um Kosten, Zeit und Mühe zu sparen. Das Objekt wird dabei auch meist in unterschiedlichen Sichtweisen und mit unterschiedlichen Maßstäben modelliert bis es zu einer passenden Lösung kommt (vgl. Hausser et al. 2011, S. 3f). Daran erkennen wir bereits, worauf es auch bei der mathematischen Modellierung ankommt. Hinter vielen Phänomenen des Alltags, besonders in technischen Anwendungen, stecken mathematische Modelle.
Wir wollen nun auf den Prozess der mathematischen Modellierung näher eingehen und versuchen einen roten Faden zu finden, bei dem es beim Modellieren verschiedenster Phänomene ankommt. Wie eingangs bereits kurz erwähnt, kann mathematische Modellierung als iterativer Prozess beschrieben werden, in dem verschiedene Aktivitäten zyklisch-wiederkehrend durchlaufen werden. Dieser Prozess wird oftmals als Modellierungszyklus beschrieben und besteht aus folgenden Aktivitäten (nach Hausser et al. 2011, S. 52):

- Verständnis und Analyse des Anwendungsproblems
- Entwicklung eines Modells (Modellbildung)
- Mathematische Analyse des Modells (Sensitivitätsanalyse und mögliche Vereinfachung, Gleichgewichtslagen, deren Stabilität, Langzeitverhalten, etc.)
- Berechnung und Simulation
- Interpretation und Validierung (Vergleich der Lösungen mit realen Daten)

In Abb. 1.1 ist der Modellierungszyklus bildhaft dargestellt. Der Start ist das Anwendungsproblem. In der Praxis sind manche Modellierungsschritte oft nicht strikt zu trennen. Wie oft dieser Zyklus durchlaufen wird, hängt letztendlich von den Ergebnissen ab, die durch die Simulationen und Berechnungen entstehen. Je klarer die zu anfangs gestellten Fragenstellungen formuliert wurden, desto eher sieht man, wie gut diese Ziele erreicht wurden. Sind die Ergebnisse zufriedenstellend, kann der Zyklus abgebrochen werden, ansonsten muss dieser von vorne neu durchlaufen werden.

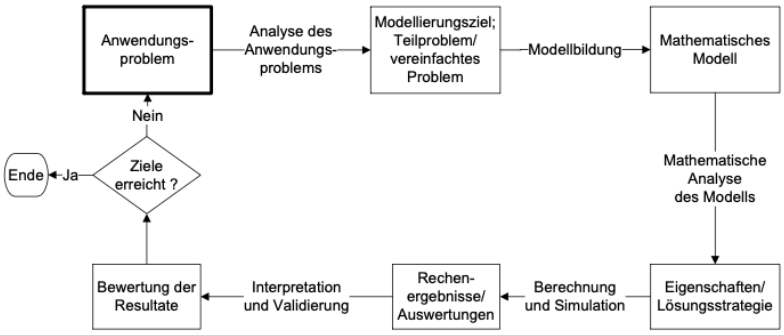


Abb. 1.1: Schematische Darstellung des Modellierungszyklus bestehend aus einer iterativen Abfolge von Aktivitäten, die zu bestimmten Ereignissen führen (vgl. Hausser et al. 2011, S. 53).

1.2.1 Analyse des Anwendungsproblems

Ein Anwendungsproblem ist in der Regel von komplexer Natur und besteht meist aus mehreren Teilproblemen. Deswegen ist es wichtig, das Problem mithilfe von explizit formulierten Fragestellungen passend einzuschränken. Außerdem lassen sich mit jedem mathematischen Modell nur Fragestellungen untersuchen, die im Modell repräsentiert sind. Es gibt keine Modelle, mit denen man „alles“ beantworten kann. Eine weitere wichtige Erkenntnis ist, dass aufgrund der Komplexität nicht das ganze System modelliert werden kann (vgl. Hausser et al. 2011, S. 53ff). Rückblickend auf die Fragestellungen sollen nur diejenigen Aspekte betrachtet werden, die für die Beantwortung zugänglich erscheinen. Deswegen muss das System von vornherein gut verstanden werden damit etwaige Nebeneffekte oder Einflüsse vernachlässigt werden können. Hierbei muss der Modellierungskreislauf wahrscheinlich mehrere Male durchlaufen werden damit man sich an die effizienteste Lösung herantasten kann.

1.2.2 Modellbildung

Der nächste Schritt nach der Analyse des Problems ist es, das Modell in die Sprache der Mathematik zu übersetzen. In der Regel besteht ein mathematisches Modell aus einer oder mehreren Gleichungen. Diese Gleichungen werden mit einer Liste von Parametern und Variablen gefüttert. Manche Parameter sind fest von außen vorgegeben, andere wiederum müssen mithilfe des Modells gefunden und festgelegt werden. Damit sich der Zustand des zu modellierenden Systems beschreiben lässt,

werden sogenannte Zustandsgrößen oder Zustandsvariablen benötigt. Für zeitabhängige Systeme sind dies meistens Funktionen von kontinuierlichen oder diskreten Zeitvariablen. Der Wertebereich der Zustandsvariablen wird Phasenraum genannt. Die Anzahl der benötigten Zustandsgrößen wird Dimension des Phasenraums genannt (vgl. Hausser et al. 2011, S. 58).

Letztendlich ist bei der Modellierung wiederum wichtig zu unterscheiden, welche Annahmen für das Modell unbedingt gebraucht werden und welche ohne größere Abweichungen vernachlässigt werden können. Als Beispiel kann hier der Wurf eines Basketballs genannt werden. Hier ist die Kraft des Luftwiderstands wesentlich kleiner als die Gravitationskraft (vgl. Hausser et al. 2011, S. 56) und kann daher vernachlässigt werden. Damit man aber überhaupt eine Ahnung von der Gesetzmäßigkeit hat, der das gegebene System unterliegt, ist es wichtig, sich mit dem theoretischen Wissen über entsprechenden Anwendungsgebiete auseinanderzusetzen. Beim oben erwähnten Modell des Ballwurfs wird man entsprechend andere Gesetzmäßigkeiten verwenden müssen als zum Beispiel bei der präzisen Modellierung der GPS-Koordinaten eines Objektes, wobei außerdem Vernachlässigungen und Vereinfachungen deutlich mehr ins Gewicht fallen werden, als bei der Modellierung eines Ballwurfs.
Am Schluss dieser Überlegungen soll die Problemstellung mithilfe einer präzisen mathematischen Aufgabenstellung formuliert werden, in der die zuvor definierten Parameter, Größen, Zustandsvariablen und Gesetzmäßigkeiten eingehen.

1.2.3 Mathematische Analyse des Modells

Die meisten Modellvariablen besitzen eine physikalische Dimension und werden mit einer Einheit angegeben. Für die mathematischen Untersuchungen ist es jedoch hilfreich, die Größen und Gleichungen unabhängig ihrer Einheiten zu verwenden. Diesen Vorgang bezeichnet man als Entdimensionalisierung (vgl. Hausser et al. 2011, S. 62). Ein Modell wird oft mit sehr vielen Parametern versehen. Wie stark Störungen der unterschiedlichen Parameter die Lösung des Modells beeinflussen ist ein weiterer wichtiger Aspekt, der als Sensitivitätsanalyse bezeichnet wird (vgl. Hausser et al. 2011, S. 63). Derartige Fragestellungen können ein weiteres Modellierungsziel darstellen.

1.2.4 Berechnungen und Simulationen

Sind die Gleichungen der Modelle erstmals aufgestellt, können die Modelle (meist mithilfe von geeigneten Computerprogrammen) durch Einsetzen von verschiedenen

Parameterwerten gelöst werden. Hier können zum Beispiel beim erneuten Durchlaufen einzelne Parameterwerte variiert und somit die Auswirkung dieser auf die Lösungen konkretisiert und optimiert werden. Wichtig ist, sich vorab geeignete Verfahren zur Lösung der verwendeten Gleichungen zu überlegen. In vielen Fällen können die Lösungen nicht exakt bestimmt werden. Hier muss man dann auf numerische Verfahren zurückgreifen, die eine Näherungslösung bestimmen.

1.2.5 Interpretation und Validierung

Nachdem die Berechnungen durchgeführt wurden, müssen sie noch interpretiert und auf ihre Gültigkeit überprüft (validiert) werden. Bei der Validierung wird lediglich überprüft, ob eine Übereinstimmung der Berechnungen mit den Daten vorliegt und nicht, ob das vorliegende Modell die Aufgabenstellung richtig beschreibt. Wie genau die Ergebnisse sein sollten, hängt auch von der Fragestellung ab. Beispielsweise können Rundungsfehler im Zehntelbereich bei der Modellierung eines Ballwurfs vernachlässigt werden, bei der Modellierung von GPS-Daten eher weniger.
Am Schluss müssen die Ergebnisse noch hinsichtlich der ursprünglichen Fragestellung interpretiert und rückübersetzt werden. Dabei ist entscheidend, ob man die Fragestellung anhand der Ergebnisse des Modells zufriedenstellend beantworten kann. Reicht die Genauigkeit nicht aus, so muss der Modellzyklus noch einmal durchlaufen und verfeinert werden. (vgl. Hausser et al. 2011, S. 68).
Lässt sich kein besseres Ergebnis zur zufriedenstellenden Beantwortung der Fragestellungen erreichen, sollten gegebenenfalls Erweiterungen des Modells angedacht werden.

1.3 Epidemiologische Grundbegriffe

Die Epidemiologie (altgriechisch: epi: „über“, demos: „das Volk“ und logos: „die Lehre“; im Sinnzusammenhang also „die Krankheit, die über das Volk kommt“) ist eine wissenschaftliche Disziplin, die sich vor allem mit der Entstehung, Verbreitung und Bekämpfung, sowie den sozialen Auswirkungen von Epidemien von Bevölkerungsgruppen bzw. Populationen beschäftigt. Die Epidemiologie forscht jedoch nicht nur in der Bekämpfung und Prävention von Krankheiten, sondern auch nach jenen Faktoren, die zur Verbesserung der Gesundheit der Grundgesamtheit beitragen (vgl. Kreienbrock et al. 2012, S. 1ff).
Grob gesagt kann die Epidemiologie in zwei Teilgebiete unterteilt werden: einerseits

stellt sie eine medizinische Wissenschaft dar, bei der nach Ursachen zur Krankheitsausbreitung und -eindämmung gesucht werden. Andererseits ist sie ein Teilgebiet der Statistik. Durch quantitative und qualitative Analysen mathematischer Modelle lassen sich etwaige Ursache-Wirkungsbeziehungen von Krankheitsausbreitungen erforschen (vgl. Bonita et al. 2008, S. 21f).
Die Ausbreitung einer Krankheit erfolgt in der Regel durch Krankheitserreger wie Viren, Bakterien oder Würmer, die durch ein infiziertes Mitglied der Population auf ein gesundes, infizierbares Individuum übertragen werden. Die Krankheit kann auch durch einen Zwischenträger (z.B. durch Insekten) übertragen werden. Nicht jeder Kontakt zwischen einer infizierbaren und einer infizierten Person muss zu einer Ansteckung führen. Es ist zunächst möglich, dass diese Person bereits dauerhaft (oder nur vorübergehend) immun gegen die Krankheit ist. Aber auch die Intensivität des Kontakts bzw. die Abwehrkräfte des natürlichen Organismus spielen hierbei eine entscheidende Rolle.
Kommt es aber zu einer Infektion der gesunden Person, folgt auf die Ansteckung die sogenannte Latenzzeit oder Latenzperiode. Diese bezeichnet eine Zeitdauer (Periode), in der eine infizierte Person weder Krankheitssymptome verspürt, noch die Krankheit an andere, gesunde Personen übertragen kann. Das Zeitintervall zwischen der Infektion und den ersten Krankheitssymptomen heißt Inkubationsperiode. Die Periode, in der die Krankheit übertragen werden kann, heißt infektiöse Periode; meist liegt der Beginn dieser vor dem Ablauf der Inkubationsperiode und endet auch oft vor dem Abklingen der Krankheitssymptome (vgl. Nöbauer et al. 1979, S. 122f). In der infektiösen Periode schränkt die Person (meist freiwillig oder auf Verordnung eines Mediziners), je nach Stärke der Symptome, ihre sozialen Kontakte ein, um eine weitere Verbreitung der Krankheit zu unterbinden.

1.4 Geschichte der Epidemiologie

Bereits vor mehr als 2000 Jahren hat Hippokrates den Begriff der Epidemie verwendet und erstaunliche Wirkungsfaktoren zwischen Krankheiten und Umweltfaktoren erläutert. Im 17. Jahrhundert wurde die erste Sterbestatistik „Bill of Mortality“ in London von John Graunt eingeführt. Darunter wurden erstmals die häufigsten Todesursachsen systematisch dokumentiert. Die Statistiken sollten als Warnsignal einer beginnenden Seuche dienen, die zeigten, dass die Sterblichkeit im Verlauf der Zeit nahezu konstant blieb, während sie zur Zeit der Pest Spitzenwerte erreichte. Dies war eine erste Andeutung, dass Umweltfaktoren einen wesentlichen Einfluss auf

die Gesamtsterblichkeit haben. William Farr baute darauf auf und sammelte und analysierte Sterbestatistiken von Großbritannien. Die damals verwendeten Praktiken dienen noch heute als Basis für die Analyse von Bevölkerungsstatistiken.
1854 stellte Jon Snow eine Hypothese zur Verbreitung von Cholera durch Trinkwasser auf, die er durch eine Reihe von Untersuchungen bestätigen konnte. Snow veröffentlichte im weiteren eine Karte von London, in der die Verbreitung der Krankheitsfälle visualisiert wurde (vgl. Zsifkovits 2016, S. 5). 1906 folgte mittels mathematischer Methoden die Erkenntnis von Hamer, Epidemien zu beenden, ohne der gänzlichen Erkrankung der Gesamtpopulation.
1927 wurde von William Ogilvy Kermack und Anderson Gray McKendrick erstmals ein mathematisches Modell für die Ausbreitung von Infektionskrankheiten entwickelt. Mit diesem Modell konnten sie die Pest-Epidemie in Bombay im Jahre 1905/1906 modellieren (vgl. Zsifkovits 2016, S. 6f). Deshalb bildeten diese Arbeiten die Basis der Entwicklung mathematischer Modelle zur Beschreibung der Ausbreitung von Infektionskrankheiten.

1.5 Mathematische Modellierung von Infektionskrankheiten

In der ersten Veröffentlichung von Kermack und McKendrick wurde ein sehr vereinfachtes Modell beschrieben: Personen, die sich mit einer Krankheit infizierten, waren am Krankheitsende vollständig immun oder tot. Während der gesamten Dynamik wurde die Bevölkerung als konstant angenommen – es gab keine Geburten oder Todesfälle, die in die Bevölkerung ein- oder austraten. Als Grundproblem wurde die Verbreitung der Krankheit in der Bevölkerung als eine Funktion der Zeit dargestellt. Die wohl interessanteste und wichtigste Frage dabei war, wie sich die Ausbreitung Krankheit im Laufe der Zeit entwickelte. Werden alle Individuen der Bevölkerung infiziert werden oder bringt ein bestimmtes Zusammenspiel von Infektions-, Genesungs- und Mortalitätsraten das Aussterben der Krankheit mit sich (vgl. Zacher et al. 2008, S. 21f)?
In ihrem Modell wurde die Gesamtpopulation (N) in verschiedene Klassen unterteilt, die den aktuellen gesundheitlichen Zustand eines Individuums beschrieben. Dieses Modell kann beliebig durch Hinzufügen weiterer Klassen erweitert werden. Durch etwaige Anpassungen der Modelle sollen dann Ausbreitungen unterschiedlicher Krankheiten modelliert werden.
Folgende Klassen bilden die Basis aller Epidemiemodelle (vgl. Mathea 2015, S. 2).

- Susceptibles (S): Gesunde bzw. infizierbare Personen, die mit der Krankheit noch nicht in Kontakt gekommen sind.
- Infectives (I): Infizierte bzw. erkrankte Personen, die die Krankheit an S weitergeben können.
- Gesamtpopulation (N): Diese ergibt sich aus der Summe aller Klassen und ist (meistens) konstant, sofern Geburten-, Sterberaten und Migration vernachlässigt werden.

Je nach Art des Modells können auch noch weitere Klassen hinzugefügt werden:

- Removed/Recovered (R): Personen, die die Krankheit aufgrund von Immunität oder Tot nicht mehr weitergeben können. Es gibt aber auch Krankheiten, die nach dem Tot noch ansteckend sein können. Dies wird in den Modellen dieser Arbeit hingegen gänzlich vernachlässigt.
- Exposed (E): Personen, die zwar infiziert sind, die Krankheit aber nicht weitergeben können. Sie befinden sich in einer Latenzphase.
- Quarantined (Q): Personen, die sich während einer Infektion in Quarantäne befinden.

Wie gut derartige Modelle die Krankheitsausbreitung beschreiben, hängt in erster Linie davon ab, wie viele Annahmen in das Modell hineingesteckt werden. Aufgrund unbekannter Daten oder der unbekannten Auswirkung bestimmter Systemgrößen und -parameter, werden viele Einflüsse zunächst vernachlässigt. Vor allem bei der Anpassung an verschiedene und neuartige Epidemien sind im Allgemeinen nur wenige verlässliche Daten bekannt. Auch mit einer guten Datenlage ist es in der Regel kein leichtes Unterfangen, jene Modelle und deren Parameter an die Daten anzupassen und somit einen Trend vorherzusagen.

1.6 Grundlegende Definitionen

Nachdem in den vorigen Kapiteln bereits einige wichtige Punkte in Bezug zur mathematischen Modellierung genannt wurden, werden in diesem Unterkapitel grundlegende Definitionen und Begrifflichkeiten eingeführt, die für den weiteren Verlauf und das Verständnis der Arbeit von zentraler Bedeutung sein werden. Um den Lesefluss während der Arbeit möglichst wenig zu stören, werden in diesem Unterkapitel auch etwaige Sätze oder Beweise angeführt, die für den weiteren Aufbau notwendig sein

werden.

Zunächst beginnen wir mit einem Parameter, der in exponentiellen Modellen eine wichtige Bedeutung hat.

Definition 1.1: Malthusianisches Wachstumsmodell und Malthus Parameter

Ein exponentielles Modell der Form $P(t) = P_0 \cdot e^{rt}$, das proportional zu der zum Zeitpunkt t vorhandenen Populationsgröße $P(t)$ wächst, nennt man *Malthusianisches Wachstumsmodell*. Der Parameter r ist der Proportionalitätsfaktor bzw. die Wachstumsrate im engeren Sinn und wird oft als *Malthus Parameter* bezeichnet.

Hinter vielen Modellgleichungen verbergen sich exponentielle Wachstumsraten. Die Krankheitsdauer kann in den vereinfachten SIR-Modellen als exponentialverteilt angenommen werden wie wir in den folgenden Kapiteln noch zeigen werden. Hierfür benötigt man vorab folgende Definitionen.

Definition 1.2: Exponentialverteilte Zufallsvariable

Eine Zufallsvariable X nennt man *exponentialverteilt* mit Parameter $\alpha \in \mathbb{R}$, wenn für die Dichte gilt

$$f_X(x) = \begin{cases} \alpha \cdot e^{-\alpha x} & \text{wenn } x \geq 0 \\ 0 & \text{wenn } x < 0 \end{cases} \tag{1.1}$$

Definition 1.3: Erwartungswert einer exponentialverteilten Zufallsvariable

Für den Erwartungswert einer exponentialverteilten Zufallsvariablen X mit Parameter α gilt

$$\mathbb{E}(X) = \frac{1}{\alpha} \tag{1.2}$$

(vgl. Teschl et al. 2014, S. 287)

Beweis

Der Erwartungswert stetiger Verteilungen ist wie folgt zu berechnen: $\int_{-\infty}^{\infty} x \cdot f(x)\,dx$. Also gilt für die exponentialverteilte Zufallsvariable

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot \alpha \cdot e^{-\alpha x}\,dx \overset{\text{Def. 1.2}}{=} \int_{0}^{\infty} x \cdot \alpha \cdot e^{-\alpha x}\,dx = \int_{0}^{\infty} x \cdot \Big(-e^{-\alpha x}\Big)'\,dx$$

(vgl. Teschl et al. 2014, S. 287).
Dieses Integral kann man durch partielle Integration lösen:

$$\int_{0}^{\infty} x \cdot \Big(-e^{-\alpha x}\Big)'\,dx = \underbrace{-x \cdot e^{-\alpha x}\Big|_{0}^{\infty}}_{=0} + \int_{0}^{\infty} e^{-\alpha x}\,dx = -\frac{1}{\alpha} \cdot e^{-\alpha x}\Big|_{0}^{\infty} = \frac{1}{\alpha}$$

□

Bei der Modellierung von Epidemiemodellen ist es wichtig zu wissen, wohin sich die Lösungskurven im Langzeitverhalten bewegen werden. Diesbezüglich untersucht man die Fixpunkte (Stabilitätspunkte/Gleichgewichtspunkte -lagen oder Equilibria) von Differentialgleichungssystemen.

Definition 1.4: Fixpunkt einer Differentialgleichung

Es sei eine Differentialgleichung von der Form

$$x'(t) = f(x(t))$$

gegeben. Ein Punkt x_0 heißt *Fixpunkt oder Gleichgewichtslage der Differentialgleichung*, wenn gilt, dass

$$f(x_0) = 0$$

Anmerkung: Es gibt verschiedene Eigenschaften von Fixpunkten. Für diese Arbeit ist vor allem die Eigenschaft „asymptotisch stabil“ von Bedeutung. (vgl. Teschl et al. 2014, S. 183).

Arten von Gleichgewichten

Versucht man einen Körper in seinem Schwerpunkt zu balancieren, ist dieser im Gleichgewicht. Wird dieser Gleichgewichtszustand von außen z.B. durch einen Windstoß gestört, ist es möglich, dass der Körper seine Gleichgewichtslage verlässt.
Genau so kann man sich das Gleichgewicht eines dynamischen Systems vorstellen. Ein System befindet sich im Gleichgewicht, wenn es sich zeitlich ohne Einwirkung von außen nicht verändert. Man kann im Allgemeinen nicht erwarten, dass Gleichgewichtslagen für immer eingenommen werden. Jedes System in der Natur ist Schwankungen oder Störungen ausgesetzt. Tritt eine Störung auf, kann ein System auf unterschiedliche Weisen auf diese Störung reagieren:

- Das System wird bei einer Störung aus seinem Gleichgewicht abgelenkt und bewegt sich danach wieder darauf zu. Diese Eigenschaft nennt man *„asymptotisch stabil“*.
  Gilt dies nur für „kleinere“ Störungen, nennt man den Gleichgewichtszustand des Systems *„lokal asymptotisch stabil“*.
  Strebt das System - egal von welcher Lage aus - immer dasselbe Gleichgewicht an, nennt man die Gleichgewichtslage *„global asymptotisch stabil“*.
- Wird das System vom Gleichgewicht ausgelenkt und strebt nicht weit entfernt vom alten Gleichgewicht in ein Neues, heißt das Gleichgewicht *„stabil“*.
- Wird es durch kleine Störungen weit ausgelenkt, heißt das Gleichgewicht *„instabil“*.

Definition 1.5: Stabilitätsbegriffe

Sei $x'(t) = f(x(t))$ mit $t \geq 0$ eine Differentialgleichung mit $x(t) \in \mathbb{R}^n$ und $f : \mathbb{R}^n \to \mathbb{R}^n$ eine Funktion. $x_0 \in \mathbb{R}^n$ sei eine Gleichgewichtslage, sodass $f(x_0) = 0$.
Das Gleichgewicht x_0 heißt

- global asymptotisch stabil, sodass für jede Lösung $\bar{x}(t)$ von $x(t)$ gilt, dass $\lim_{t\to\infty} \bar{x}(t) = x_0$.
- lokal asymptotisch stabil, wenn es ein $\delta > 0$ gibt, sodass jede Lösung $\bar{x}(t)$ mit Anfangswert $x(0)$ im Abstand δ von x_0 gegen x_0 strebt: $\mid x(0) - x_0 \mid \leq \delta \Rightarrow \lim_{t\to\infty} \bar{x}(t) = x_0$.
- stabil, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, sodass sich jede Lösung $\bar{x}(t)$

mit Anfangswert $x(0)$ im Abstand δ zu x_0 nie weiter als ϵ von x_0 entfernt:
$| x(0) - x_0 | \leq \delta \quad \Rightarrow \quad | \bar{x}(t) - x_0 | \leq \epsilon \quad \forall\, t.$

(vgl. Voßwinkel 2019, S. 9)

Satz 1.1: über die linearisierte Stabilität

Es sei $x'(t) = f(x(t))$ eine Differentialgleichung in n Komponenten mit Gleichgewichtslage x_0.

$$\begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix} = \begin{pmatrix} f_1(x_1(t), \ldots, x_n(t)) \\ \vdots \\ f_n(x_1(t), \ldots, x_n(t)) \end{pmatrix}$$

Die Jacobi-Matrix von f an der Stelle x_0 ist gegeben durch

$$\begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x_0), \ldots, \frac{\partial}{\partial x_n} f_1(x_0) \\ \vdots \\ \frac{\partial}{\partial x_1} f_n(x_0), \ldots, \frac{\partial}{\partial x_n} f_n(x_0) \end{pmatrix}. \tag{1.3}$$

Dann gilt:

i) Sind alle Realteile der Eigenwerte der Jacobi-Matrix kleiner als Null, ist x_0 lokal asymptotisch stabil.

ii) Ist mindestens ein Realteil eines Eigenwerts der Jacobi-Matrix größer als Null, ist x_0 instabil.

iii) Ist der größte Realteil aller Eigenwerte gleich Null, kann mit diesem Kriterium keine Aussage über die Stabilität geschlossen werden.

Die Visualisierung der Lösungen von Differentialgleichungen erfolgt durch sogenannte Trajektorien. Im Verlauf der Trajektorien lassen sich viele Eigenschaften der Differentialgleichungen einsehen.

Definition 1.6: Trajektorie

Unter einer Trajektorie versteht man in der Mathematik eine Lösungskurve einer Differentialgleichung (vgl. *DWDS – Digitales Wörterbuch der deutschen Sprache*).

Definition 1.7: Phasenraum

Eine gemeinsame Darstellung der Trajektorien eines Differentialgleichungssystems nennt man Phasenraum oder Phasenportrait (vgl. Kim 1987, S. 25).

Der nächste Satz ist ein grundlegender Satz in der Mathematik der wichtige Aussagen zur lokalen Existenz und Eindeutigkeit der Lösungen von Differentialgleichungen liefert.
1890 wurde dieser von Ernst Leonard Lindelöf veröffentlicht. Etwa zur gleichen Zeit arbeitete Émile Picard an der schrittweisen Approximation der Lösungen von gewöhnlichen Differentialgleichungen. Mithilfe der Picard-Interation wird der Satz von Lindelöf üblicherweise bewiesen. Deshalb wird dieser Satz meist als Satz von Picard-Lindelöf bezeichnet.
Bevor wir uns diesem Satz widmen, müssen wir noch eine zusätzliche Definition einführen, die auf dem Gebiet der Differentialgleichungen ebenfalls eine wichtige Rolle für eine Abschätzung des Änderungsverhaltens einer Funktion spielt.

Definition 1.8: Lipschitz-Bedingung

Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ eine stetige Funktion in n Komponenten, sowie das Anfangswertproblem (t_0, x_0) gegeben. Das Anfangswertproblem des Differentialgleichungssystems $x'(t) = f(x(t))$ genügt einer Lipschitzbedingung in G genau dann, wenn gilt

$$\exists L \geq 0 \; \forall (t, x), (t, \tilde{x}) \in G : \|f(t, x) - f(t, \tilde{x})\| \leq L\|x - \tilde{x}\| \tag{1.4}$$

L heißt eine *Lipschitz-Konstante*. (vgl. Hausser et al. 2011, S. 153)

Satz 1.2: Satz von Picard-Lindelöf

Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ und $f : G \to \mathbb{R}^n$ eine stetige Funktion in n Komponenten. Außerdem genüge f einer Lipschitzbedingung. Dann $\exists \epsilon > 0 \; \forall (t_0, x_0) \in G$, sodass $\psi : [t_0 - \epsilon; t_0 + \epsilon] \to \mathbb{R}^n$ eine eindeutige Lösung des Anfangswertproblems $x'(t) = f(x(t))$ mit $\psi(t_0) = x_0$ ist. (vgl. Rannacher 2017, S. 27)

Der Beweis dieses Satzes wird in dieser Arbeit nicht angeführt. Er kann aber zum Beispiel in Rannacher 2017, S. 27f nachgelesen werden.

1.7 Methodik

Manche Aspekte eines realen Vorgangs können mithilfe mathematischer Vorschriften, Gleichungen, etc. beschrieben werden - man spricht dabei von mathematischen Modellen bzw. von mathematischer Modellierung. Grundsätzlich unterscheidet man qualitative und quantitative Modelle in der mathematischen Modellierung.
Qualitative Modelle dienen vor allem zur Beschreibung der Struktur eines Prozesses und den Wechselwirkungen der Systemgrößen (z.B. welchen Einfluss bewirkt eine Veränderung der einen Größe auf die anderen Größen, usw.). Der Mehrwert jener Modelle liegt daher in der Beschreibung und Vorhersage gewisser Strukturen ohne explizite Verwendung von Werten für die eingesetzten Variablen (vgl. Roll 2003, S. 108ff).
Das Kapitel 2 der Arbeit beschäftigt sich mit der qualitativen Analyse mathematischer Epidemiemodelle. Von jedem Modell werden etwaige Grundannahmen und Vereinfachungen diskutiert und daraufhin schrittweise komplexere Modelle entwickelt. Gerade zu Beginn des 2. Kapitels sollen diese vereinfachten Modelle zum Verständnis der Grundprinzipien beitragen, die bei der mathematischen Modellierung von Epidemiemodellen auf analoge Weise in jedem Modell wirken.
Mit mathematischen Modellen berechnete Simulationsergebnisse können mit unabhängig davon vorhandenen Daten verglichen werden. Die Ergebnisse sind im Anschluss zu interpretieren und zu beurteilen, ehe diese zur Beschreibung von realen Vorgängen verwenden werden können (vgl. Hausser et al. 2011, S. 5f).
Dies wird vor allem im 3. Kapitel durchgeführt. Nach Einsatz der wahren Daten der Coronainfektionen von Österreich werden die Ergebnisse nach den Modellberechnungen überprüft und für die Beschreibung, den Vergleich bzw. die Vorhersage des Infektionsverlaufs verwendet.

Für den theoretischen Aufbau und die Beschreibung der Modelle wurde eine umfassende Literaturrecherche durchgeführt. Mithilfe von Bibliotheksdatenbanken wie zum Beispiel „uni-kat“ oder „Google Scholar“ wurde Fachliteratur zum gegebenen Thema recherchiert. Zahlreiche Bücher, Artikel und Internetquellen wurden zunächst nach Themenbereich sortiert und quergelesen. Im Anschluss erfolgte eine genaue Auseinandersetzung mit den Modellen. Es gab bereits sehr viele Veröffentlichungen zu Epidemiemodellen angewendet auf die Corona-Epidemie. In diesen Veröffentlichungen wurden – ähnlich zu dieser Arbeit – Modelle zur Beschreibung der Infiziertenzahlen entwickelt, mithilfe von Parametervariationen an gegebene Daten angepasst und zukünftige Entwicklungen vorhergesagt.

Die Aufarbeitung einzelner Modelle wurde mit verschiedener Literatur getätigt. So ist es möglich, dass ein Modell aus mehreren Quellen aufgebaut wurde um damit auch möglichst viele theoretische Zusammenhänge aufzuarbeiten.
Die Auswahl der Modelle entstand aus folgenden Überlegungen: Nicht jedes Modell kann auch für jede Krankheit sinnvoll eingesetzt werden. Die verwendeten Modelle zu Beginn des 2. Kapitels dienen rein dem Verständnis und für den weiteren Ausbau für darauffolgende Modelle. Auf Basis des realen Krankheitsverlaufs einer Coronainfektion wurden die Modelle schließlich schrittweise durch Hinzufügen weiterer realer Vorgänge komplexer wie z.B. durch die Einführung einer Latenzperiode. Selbst Quarantänemaßnahmen konnten mit den Modellen modelliert werden.
Theorie und Praxis liegen meist jedoch weit auseinander, dennoch muss die Praxis stets auf guter Theorie beruhen (sinngemäß nach Leonardo da Vinci).
Ein Hauptaugenmerkt der Arbeit liegt in der quantitativen Analyse der postulierten Modelle, die mit echten Daten gefüttert wurden.
Aus den obigen Überlegungen lässt sich damit folgende Forschungsfrage für die Arbeit formulieren: „Kann das epidemiologische Verhalten der Corona-Pandemie für Österreich mit Epidemiemodellen repräsentiert werden?“. Aus dieser Frage ergibt sich eine weitere Zusatzfrage, deren Beantwortung ebenfalls sehr interessant scheint: „Kann das epidemiologische Verhalten der Corona-Pandemie für Österreich mit Epidemiemodellen vorhersagt werden?“
Um diese Fragen beantworten zu können, müssen – wie oben erwähnt – die Modelle mit echten Daten verglichen und ausgewertet werden.

1.7.1 Datenerhebung

Die Daten, die für die Berechnungen in den Modellen verwendet wurden, stammen von der Homepage data.gv.at. Wie auf der Homepage beschrieben, handelt es sich um „Open Government Data (OGD)“, die unter einer offenen Creative Commons (CC) Lizenz frei zugänglich und weiter verwendbar sind.
Auf der Homepage selbst werden verschiedene Kategorien von Datensätzen aufgelistet, wie zum Beispiel die gemeldeten Covid19 Fälle je Altersgruppe. Für die Arbeit war jedoch der Datensatz mit der „zeitlichen Darstellung von Daten zu Covid19-Fällen je Bundesland“ relevant. Die Datensätze sind in einer csv-Datei verfügbar, die heruntergeladen und weiter bearbeitet wurden. Die Fallzahlen der jeweiligen Bundesländer sind für die Arbeit nicht relevant. Deshalb wurden die Zahlen für Österreich aus den Gesamtdaten extrahiert und in Matlab implementiert.
In der Arbeit sollen die Infektionszahlen vom Frühjahr 2020 und vom Herbst 2020

untersucht werden. Die Verläufe der Infiziertenzahlen werden hier auch immer wieder als „Wellen“ beschrieben; es handelt sich dabei um einen erheblichen Anstieg der Infektionsverläufe, die einer Welle ähneln. Wichtig ist zu wissen, dass die erste Welle vom Frühjahr 2020 anderen Gegebenheiten zugrunde liegt als jene vom Herbst 2020. Im Frühjahr war die medizinische Forschung mit Bezug auf das neue Coronavirus weniger fortgeschritten als dies im Herbst der Fall war; es gab kaum Testmöglichkeiten für die Bevölkerung. Durch den strengen Lockdown im Frühjahr 2020 flachte die Kurve auch relativ schnell wieder ab. Ihr Maximum lag in Österreich bei 1 050 Neuinfektionen pro Tag. Im Herbst hingegen waren die Fallzahlen deutlich höher, im Schnitt um das Zehnfache. Spitzenwerte lagen bei etwa 9 000 Neuinfektionen pro Tag. In der kühleren Jahreszeit kommt außerdem hinzu, dass sich Viren besser übertragen können (vgl. Taschwer 2021).
Aufgrund dieser Unterschiede ist es interessant zu wissen, wie gut sich beide Wellen mit den theoretischen Modellen berechnen lassen.
Der Datensatz zur ersten Corona-Welle enthält 61 Daten. Von 01.03.2020 bis einschließlich 30.04.2020 wird jedem Tag genau eine bestimmte Infiziertenzahl zugeordnet. Diese Daten sind als Koordinaten der Form (Tage nach dem 29.02.2020 | Infiziertenzahl) in den Abbildungen eingetragen. Der Punkt (5 | 9) gibt somit an, dass es am 05.03.2020 österreichweit neun registrierte Coronainfizierte gab.
Die „Grenzen“ der zweite Welle vom Herbst 2020 sind nicht so eindeutig wie jene der ersten Welle. In der Arbeit werden deshalb die Infektionszahlen zwischen 13.09.2020 und einschließlich 31.12.2020 verwendet wodurch sich 110 Datenpaare ergeben. Zur besseren Übersicht werden die Daten der Infektionszahlen der ersten und zweiten Welle in den Tabellen im Anhang zusammengefasst.

1.7.2 Datenauswertung

Nach der Implementierung der Rohdaten wurden diese mithilfe der kommerziellen Software Matlab ausgewertet. Matlab dient als Plattform zur Programmierung und numerischen Berechnungen von (mathematischen) Problemen und kann vor allem zur Datenerfassung, -analyse, -auswertung und numerischen Simulation eingesetzt werden (vgl. *MathWorks*).
Damit die Daten entsprechend ausgewertet werden konnten, musste ein Code in Matlab aufgerufen werden der zum jeweiligen Modell bzw. zur jeweiligen Aufgabenstellung (Datenanpassung oder Vorhersage) passte. Zur Anpassung der Lösungen an die realen Daten wurde in Matlab ein sogenannter Minimumsucher aufgerufen.

Dieser sucht optimale Parameterwerte für die Lösungskurven, sodass die Fehlerquadratsummen zu den realen Daten minimal werden. Die genaue Beschreibung dieser Funktion erfolgt im Kapitel 3.1.1.
Die Berechnungen der Differentialgleichungen mit den gefundenen optimalen Parameterwerten wurden anschließend als grafische Lösungskurven dargestellt. Mithilfe der Variation von Startbedingungen der Parameter konnten die Lösungen zusätzlich optimiert werden.
Jede wissenschaftliche Arbeit muss auf bestimmten Gütekriterien beruhen damit die Forschung überhaupt an Gültigkeit besitzt. Für diese Arbeit sind die Gütekriterien empirischer Forschung (vgl. Lienert et al. 2011, S. 7ff) erfüllt:

- Objektivität: Die Daten stammen aus einer vom Ministerium veröffentlichten Homepage. Sowohl deren Verarbeitung, deren Implementierung als auch deren Weiterverarbeitung erfolgte unabhängig des Autors.
  Die Auswertung und Interpretation der Ergebnisse erfolgte nach klaren Regeln der Modellierung und Durchlaufen des Modellierungszyklus.
  Das bedeutet letztendlich, dass die Ergebnisse nicht durch die auswertende Person variieren.

- Reliabilität: Reliabilität besagt, dass die Verwendung des Erhebungsinstruments unter gleichen Bedingungen immer wieder zum selben Ergebnis kommen muss. Die Auswertung der Daten erfolgte mittels Einsatz einer mathematischen Software. Die Berechnungen mit den gleichen Parameterwerten wurden außerdem mehrmals durchgeführt und führten in jedem Versuch zu den gleichen Ergebnissen. Unter gleichen Voraussetzungen wurden also immer dieselben Ergebnisse erreicht.

- Validität: Die Auswertung der Daten erfolgte unter präziser Berücksichtigung der Forschungsfrage. Letztendlich wurden genau jene Datensätze verwendet, ausgewertet und interpretiert, die für die Beantwortung der Forschungs- und Zusatzfrage relevant sind. Mit den Ergebnissen der Simulationen der jeweiligen Modelle kann somit genau entschieden werden, ob die Modelle für den praktischen Einsatz geeignet sind oder nicht.

KAPITEL 2

Mathematische Modelle zur Modellierung von Infektionskrankheiten

Die Basisreproduktionszahl $\mathcal{R}_0$ ist in vielen mathematischen Modellen eine besonders ausschlaggebende Maßzahl, die den Verlauf eines theoretisch formulierten Krankheitsmodells bestimmt. Oftmals ist fälschlicherweise der Begriff „Basisreproduktionsrate“ in Verwendung. Da $\mathcal{R}_0$ eine dimensionslose Zahl darstellt, ist der Wortlaut „-rate“ formal nicht korrekt. Die durchschnittliche Anzahl der angesteckten Personen durch eine ansteckende Person wird durch die Basisreproduktionszahl erfasst. Diese hängt von vielen Umweltbedingungen ab und ist daher keine biologische Konstante. In den Modellen der Arbeit wird die Reproduktionszahl gemäß den in den Modellen verwendeten Parametern angepasst. Für $\mathcal{R}_0 > 1$ wird sich die Krankheit in der Bevölkerung ausbreiten; je größer $\mathcal{R}_0$ desto schneller. Für $\mathcal{R}_0 < 1$ wird sich die Krankheit nicht weiter ausbreiten, es ist aber möglich, dass die Krankheit in der Bevölkerung bestehen bleibt, also endemisch wird. Die Anzahl der Individuen der jeweiligen Klassen streben im Langzeitverhalten somit einen (stabilen) Gleichgewichtszustand an (vgl. HealthKnowledge 2018).

2.1 SI-Modell

Das epidemische SI-Modell ist das einfachste aller Modelle, bietet aber eine gute Basis für die Modellierung infektiöser Krankheiten. In dem Modell wird die Gesamtpopulation N in zwei Klassen eingeteilt, nämlich in die Klasse S und in die Klasse I. Es wird davon ausgegangen, dass die Bevölkerung konstant bleibt; es gibt weder Geburten- noch Sterbefälle.
Es gilt also

$$S(t) + I(t) = N \quad \forall\, t \geq 0 \tag{2.1}$$

wobei $S(t)$ die gesunden, ansteckbaren und $I(t)$ die infektiösen, ansteckenden Individuen zum Zeitpunkt t sind. Zur besseren Übersicht schreiben wir von nun an nur S statt $S(t)$ bzw. I statt $I(t)$, usw; auch in den Folgekapiteln werden wir die kürzere Schreibweise verwenden.
Da in dem Modell nur zwei Klassen auftreten, entspricht der Abnahme der Gesunden pro Zeiteinheit der Zunahme der Infizierten pro Zeiteinheit. Es läuft letztendlich darauf hinaus, dass irgendwann alle Individuen der Klasse S in die Klasse I übergehen. Daraus ergibt sich folgenden Schema:

$$\boxed{\text{S}} \longrightarrow \boxed{\text{I}} \tag{2.2}$$

2.1.1 Mathematische Modellierung

Durch Mischung (vor allem durch soziale Kontakte) beginnt die Ausbreitung der Infektion. Die Klasse S nimmt dabei um denselben Faktor ab, um den die Klasse I zunimmt. Somit gilt $-S' = I'$. Man nimmt außerdem an, dass die Änderungsraten von den Anzahlen der Individuen in den jeweiligen Klassen abhängen. Somit gilt $S' = -f(S, I)$ und $I' = f(S, I)$ mit einer gegebenen Funktion f.
Welche Eigenschaften hat nun f? Um dies beantworten zu können und um f sinnvoll zu definieren, nimmt man an, dass eine Ansteckung proportional zur Anzahl der Individuen in S und proportional zur Anzahl der Individuen in I ist. Denn einerseits können doppelt so viele Infizierte in der Regel auch doppelt so viele Gesunde anstecken. Andererseits kommt es zu den doppelten Ansteckungen, wenn es doppelt so viele Gesunde gibt. Somit ist die Änderungsrate mit der die Individuen pro Zeiteinheit die Klassen wechseln gegeben durch βSI, wobei $\beta > 0$ die Stärke der Infektion beschreibt (vgl. Bürger 2004, S. 69).
Das vorher erwähnte Schema 2.2 kann folgendermaßen präzisiert werden:

$$\boxed{\text{S}} \xrightarrow{\beta SI} \boxed{\text{I}} \tag{2.3}$$

Die zeitlichen Änderungsraten können durch Differentialgleichungen modelliert werden. Es gilt also:

$$\frac{dS}{dt} = -\beta SI \tag{2.4}$$

$$\frac{dI}{dt} = \beta SI \tag{2.5}$$

mit gegebenen Anfangswerten $S(0) = S_0$, $I(0) = I_0 > 0$.
Ersetzt man in Gleichung 2.5 S durch Umformung der Beziehung 2.1 in $S = N - I$, ergibt sich:

$$\frac{dI}{dt} = \beta I(N - I) = \beta N I\left(1 - \frac{I}{N}\right) \tag{2.6}$$

Für das anfängliche exponentielle Wachstum gilt $I \ll N$, denn I kann zu Beginn als sehr viel kleiner als N angenommen werden. Somit ist $\frac{I}{N} \approx 0$ und $1 - \frac{I}{N} \approx 1$. Für $I \ll N$ gilt:

$$\frac{dI}{dt} = \beta N I$$

Wir setzen den Proportionalitätsfaktor $\beta N := r$.
r ist dabei der Malthus Parameter (vgl. Definition 1.1).

Nach Einsetzen von r in 2.6 ergibt sich:

$$\frac{dI}{dt} = rI\left(1 - \frac{I}{N}\right) \tag{2.7}$$

Obige Differentialgleichung wird durch Trennung der Variablen gelöst:

$$\begin{aligned} \frac{dI}{dt} &= rI\left(1 - \frac{I}{N}\right) \\ \frac{dI}{I(1 - \frac{I}{N})} &= r\ dt \\ \int \frac{1}{I(1 - \frac{I}{N})}\ dI &= \int r\ dt \end{aligned} \tag{2.8}$$

2.1.2 Berechnung der Lösungen

Bevor man das linke Integral von 2.8 lösen kann, wird es durch Partialbruchzerlegung wie folgt aufgeteilt:

$$\begin{aligned} & \frac{1}{I(1 - \frac{I}{N})} = \frac{A}{I} + \frac{B}{1 - \frac{I}{N}} \\ \overset{\text{gem. Nenner}}{\Leftrightarrow} \quad & \frac{1}{I(1 - \frac{I}{N})} = \frac{A(1 - \frac{I}{N})}{I(1 - \frac{I}{N})} + \frac{BI}{I(1 - \frac{I}{N})} \\ \Leftrightarrow\ & 1 = A\left(1 - \frac{I}{N}\right) + BI \\ \Leftrightarrow\ & A = 1 \text{ und } B = \frac{1}{N} \end{aligned} \tag{2.9}$$

Nun kann man die Zerlegung aus 2.9 verwenden um das Integral aus 2.8 zu berechnen:

$$\begin{aligned}
\int \frac{1}{I}\, dI + \int \frac{1}{N \cdot \left(1 - \frac{I}{N}\right)}\, dI &= \int r\, dt \\
\int \frac{1}{I}\, dI + \int \frac{1}{N - I}\, dI &= \int r\, dt \\
ln(I) - ln(N - I) &= rt + c \\
ln\left(\frac{I}{N - I}\right) &= rt + c \\
\frac{I}{N - I} &= e^{rt} \underbrace{e^{c}}_{:=c} \\
I &= e^{rt} cN - e^{rt} cI \\
I(1 + e^{rt} c) &= e^{rt} cN \\
I &= \frac{e^{rt} cN}{1 + e^{rt} c} \\
I &= \frac{e^{rt} cN}{e^{rt} c(1 + e^{-rt} \underbrace{c^{-1}}_{:=c})} \\
I &= \frac{N}{1 + e^{-rt} c} \qquad (2.10)
\end{aligned}$$

Für $t = 0$ gilt für 2.10:

$$\begin{aligned}
I(0) = I_0 &= \frac{N}{1 + e^{-r \cdot 0} c} \\
c &= \frac{N}{I_0} - 1
\end{aligned}$$

Setzt man c in 2.10 ein, erhält man:

$$I(t) = \frac{N}{1 + (\frac{N}{I_0} - 1) \cdot e^{-rt}} \qquad (2.11)$$

Analog gilt für S:

$$S(t) = \frac{N}{1 + (\frac{N}{I_0} - 1) \cdot e^{rt}} \qquad (2.12)$$

2.1.3 Langzeitverhalten der Klassen S und I

Die Infiziertenzahlen lassen sich für $t \to \infty$ abschätzen:
$\lim_{t\to\infty} I(t) = \lim_{t\to\infty} \frac{N}{1+(\frac{N}{I_0}-1)\cdot \underbrace{e^{-rt}}_{\to 0}} = N.$
Für S gilt:
$\lim_{t\to\infty} S(t) = \lim_{t\to\infty} \frac{N}{1+(\frac{N}{I_0}-1)\cdot \underbrace{e^{rt}}_{\to \infty}} = 0.$
Wir sehen also - wie vorher schon behauptet - dass alle Gesunden angesteckt werden. Interessant ist hierbei noch der Zeitpunkt der stärksten Zunahme der Infizierten bzw. anders formuliert, derjenige Zeitpunkt an dem die Wachstumsgeschwindigkeit von I am größten ist. Dazu betrachtet man den Wendepunkt, also die zweite Ableitung von I. Dafür leitet man 2.7 nochmal ab und erhält mit Anwendung der Produktregel:

$$\frac{d^2I}{dt^2} = rI' \cdot \left(1 - \frac{I}{N}\right) - rI \cdot \frac{I'}{N} = rI' \cdot \left(1 - \frac{2I}{N}\right) = 0$$

Da $I' > 0$ und $r > 0$ folgt daraus, dass $1 - \frac{2I}{N} = 0 \Leftrightarrow I = \frac{N}{2}$. Die Wachstumsgeschwindigkeit ist also am größten, wenn die Hälfte der Individuen infiziert sind. Um den Zeitpunkt dieses Maximums herauszufinden, setzt man I gleich $\frac{N}{2}$ in 2.11 und berechnet t.

$$\begin{aligned}
\frac{N}{2} &= \frac{N}{1 + (\frac{N}{I_0} - 1) \cdot e^{-rt}} \\
\Leftrightarrow 2 &= 1 + \left(\frac{N}{I_0} - 1\right) \cdot e^{-rt} \\
\Leftrightarrow \frac{1}{\frac{N}{I_0} - 1} &= e^{-rt} \\
\Leftrightarrow ln(1) - ln\left(\frac{N}{I_0} - 1\right) &= -rt \\
\Leftrightarrow t &= \frac{ln\left(\frac{N}{I_0} - 1\right)}{r}
\end{aligned}$$

Ab diesem Zeitpunkt nimmt die Wachstumgsgeschwindigkeit von I wieder stetig ab.

Aufgrund der Einfachheit des SI-Modells lässt sich dieses in der Regel auch kaum für reale Krankheitsausbreitungen einsetzen. Im nächsten Kapitel soll das Modell ausgebaut werden.

2.2 SIS-Modell

Dieses Modell ist eine erste Erweiterung des SI-Modells. Es modelliert Infektionsketten in denen Individuen nach deren Genesung wieder neu empfänglich für die Krankheit werden und beschreibt somit Infektionen ohne Immunisierung. Der Infektionsweg wird analog zu jenem im SI-Modell beschrieben. Somit liegt folgendes Schema vor:

$$\boxed{\text{S}} \leftrightarrows \boxed{\text{I}} \tag{2.13}$$

Im nächsten Schritt überlegen wir uns, dass eine Heilung mit einem festen Parameter γ erfolgt. γI ist also die Anzahl der Infizierten, die pro Zeiteinheit wieder in die Klasse S übergehen. 2.13 lässt sich wie folgt präzisieren:

$$\boxed{\text{S}} \xrightarrow{\beta SI} \boxed{\text{I}} \xrightarrow{\gamma I} \boxed{\text{S}} \tag{2.14}$$

Sowie

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI + \gamma I \\ \frac{dI}{dt} &= \beta SI - \gamma I \end{aligned} \tag{2.15}$$

mit Anfangswerten $S(0) = S_0$, $I(0) = I_0 > 0$. Nach dem Satz von Picard-Lindelöf 1.2 besitzt 2.15 genau eine Lösung. Außerdem gilt $\frac{dS}{dt} + \frac{dI}{dt} = -\beta SI + \gamma I + \beta SI - \gamma I = 0$, woraus folgt, dass $S + I = N$ ist. Die Grundpopulation bleibt also wieder konstant; es gibt weder Geburten- noch Sterbefälle in diesem Modell.

Zur Charakterisierung des Parameters γ betrachten wir zunächst isoliert den Heilungsvorgang und versuchen $\frac{dI}{dt} = -\gamma I$ zu lösen:

$$\begin{aligned} \frac{1}{I}\, dI &= -\gamma\, dt \\ \int \frac{1}{I}\, dI &= \int -\gamma\, dt \\ ln(I) + \gamma t &= \text{konstant} = ln(I_0) \\ I \cdot e^{\gamma t} &= I_0 \\ I &= I_0 \cdot e^{-\gamma t} \end{aligned} \tag{2.16}$$

Dividiert man die Gleichung 2.16 durch I_0 erhält man mit $\frac{I(t)}{I_0} = e^{-\gamma t}$ die relative Häufigkeit der Individuen gegenüber I_0, die zur Zeit t noch infektiös sind. Die Infektionsdauer ist demnach laut 1.2 exponentialverteilt. Der Erwartungswert der

exponentialverteilten Variablen gibt die mittlere Infektionsdauer an und ist nach 1.3 gegeben durch $\int_0^\infty t\gamma \cdot e^{-\gamma t}\, dt = \frac{1}{\gamma}$.

Als Reproduktionszahl definiert man $\mathcal{R}_0 := \frac{\beta N}{\gamma}$, wobei βN die Infektionszahl pro Individuum in N ist und $\frac{\beta N}{\gamma}$ somit angibt, wie viele Infektionen ein Individuum in der mittleren Infektionszeit $\frac{1}{\gamma}$ auslöst (vgl. Zacher et al. 2008, S. 23).
Darauf aufbauend entdimensionalisieren wir 2.15 zu relativen Häufigkeiten wie folgt: $u(\tau) := \frac{S}{N} \quad v(\tau) := \frac{I}{N} \quad$ und $\quad \frac{\tau}{\gamma} := t.$

$$\begin{aligned}
\frac{d(uN)}{d\frac{\tau}{\gamma}} &= -\beta uNvN + \gamma vN \Leftrightarrow \frac{du}{d\tau} = \frac{-\beta uvN}{\gamma} + v \overset{\mathcal{R}_0=\frac{\beta N}{\gamma}}{\Leftrightarrow} \frac{du}{d\tau} = -(\mathcal{R}_0 u - 1)v \\
\frac{d(vN)}{d\frac{\tau}{\gamma}} &= \beta uNvN - \gamma vN \Leftrightarrow \frac{dv}{d\tau} = \frac{\beta uvN}{\gamma} - v \overset{\mathcal{R}_0=\frac{\beta N}{\gamma}}{\Leftrightarrow} \frac{dv}{d\tau} = (\mathcal{R}_0 u - 1)v
\end{aligned} \tag{2.17}$$

mit den Anfangswerten u_0, v_0. Wiederum folgt aus dem Satz von Picard-Lindelöf 1.2, dass 2.17 eine eindeutige (sogar positive) Lösung besitzt (vgl. Zacher et al. 2008, S. 23). Aus 2.17 folgt $\frac{du}{d\tau} + \frac{dv}{d\tau} = 0$ und dass die Summe der relativen Anteile $u + v = \frac{S}{N} + \frac{I}{N} = 1$ ergibt.

2.2.1 Fixpunkte und Langzeitverhalten

Versuchen wir nun zu klären, wie sich die Differentialgleichungen für $t \to \infty$ verhalten. Dazu schätzen wir $\frac{dv}{d\tau}$ nach oben ab:

$$\frac{dv}{d\tau} = (\mathcal{R}_0 u - 1)v \underbrace{\leq}_{0<u<1} = (\mathcal{R}_0 - 1)v \tag{2.18}$$

$(\mathcal{R}_0 - 1)v$ ist für $\mathcal{R}_0 < 1$ negativ, es kommt also zu einem exponentiellen Abfall und somit zu einem Verschwinden der Krankheit. Was passiert hingegen, wenn $\mathcal{R}_0 > 1$ ist? Für $\mathcal{R}_0 > 1$ ist $\frac{dv}{d\tau} > 0$, das eine Ausbreitung der Krankheit zur Folge hat. Nun ist es noch interessant zu wissen, wohin sich die relativen Anteile u und v im Langzeitverhalten bewegen. Für $\mathcal{R}_0 > 1$ bestimmen wir das Gleichgewicht von $\frac{du}{d\tau}$:

$$\frac{dv}{d\tau} = 0 = (\mathcal{R}_0 u - 1)v \Leftrightarrow u_f = \frac{1}{\mathcal{R}_0} \tag{2.19}$$

Setzt man diesen Fixpunkt in $u + v = 1$ ein erhält man als Fixpunkt für $v_f = 1 - \frac{1}{\mathcal{R}_0}$. Man kann also festhalten, dass die Reproduktionszahl $\mathcal{R}_0$ in diesem Modell aus-

schlaggebend für den weiteren Verlauf der Krankheit ist. Die Krankheit kann nur dann ausbrechen, wenn ein Individuum im Mittel mehr als ein anderes Individuum der Bevölkerung ansteckt, also $\mathcal{R}_0 > 1$ ist (vgl. Zacher et al. 2008, S. 23f).

2.2.2 Stabilitätsanalyse

Zur Untersuchung der Stabilität dieser Fixpunkte betrachtet man zunächst

$$\begin{pmatrix} \frac{du}{d\tau} \\ \frac{dv}{d\tau} \end{pmatrix} = \begin{pmatrix} -(\mathcal{R}_0 u - 1)v \\ (\mathcal{R}_0 u - 1)v \end{pmatrix} \tag{2.20}$$

Zur Untersuchung der Stabilitätsbedingung wird die Jacobi-Matrix von 2.20 gebildet. Man differenziert die rechte Matrix von 2.20 zuerst nach u und dann nach v partiell:

$$J = \begin{pmatrix} -\mathcal{R}_0 v & -(\mathcal{R}_0 u - 1) \\ \mathcal{R}_0 v & \mathcal{R}_0 u - 1 \end{pmatrix}$$

Nun werden die Fixpunkte $u_f = \frac{1}{\mathcal{R}_0}$ und $v_f = 1 - \frac{1}{\mathcal{R}_0}$ in J eingesetzt und die Eigenwerte von J berechnet:

$$\det(J - \lambda E) = \det \begin{pmatrix} -(\mathcal{R}_0 - 1) & 0 \\ \mathcal{R}_0 - 1 & 0 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \det \begin{pmatrix} 1 - \mathcal{R}_0 - \lambda & 0 \\ \mathcal{R}_0 - 1 & -\lambda \end{pmatrix}$$

Für das charakteristische Polynom erhält man $-\lambda \cdot (1 - \mathcal{R}_0 - \lambda) = 0 \Leftrightarrow \lambda_1 = 0$ oder $\lambda_2 = 1 - \mathcal{R}_0$.
Für $\mathcal{R}_0 > 1$ ist $Re(\lambda_2) = \lambda_2 < 0$, die Krankheit strebt gegen das endemische Gleichgewicht.
Für $\mathcal{R}_0 \leq 1$ ist $Re(\lambda_2) \geq 0$. Das Gleichgewicht ist instabil und $\frac{dv}{d\tau} \to 0$.

Satz 2.1

Es sei die Reproduktionszahl $\mathcal{R}_0 = \frac{\beta N}{\gamma}$ gegeben. Im SIS-Modell 2.17 kommt es für $\mathcal{R}_0 \leq 1$ zum Aussterben der Krankheit. Für $\mathcal{R}_0 > 1$ bleibt die Krankheit in der Bevölkerung bestehen und konvergiert gegen das stabile Gleichgewicht $(u_f, v_f) = (\frac{1}{\mathcal{R}_0}, 1 - \frac{1}{\mathcal{R}_0})$.

2.3 SIR-Modell

Im vorigen Abschnitt betrachteten wir ein Modell ohne Immunisierung. Nun wollen wir uns einen Infektionsweg mit anschließender Immunisierung überlegen. Das Modell stammt ursprünglich von W.O. Kermack und A.G. McKendrick und wird deshalb auch manchmal Kermack-McKendrick-Modell genannt. Im Unterschied zum vorigen SIS-Modell gehen die Individuen nach Beendigung der Infektion in die Klasse R über. Die Klasse R bezeichnet dabei jene Individuen, die nicht mehr angesteckt werden können. Darunter fallen z.B. Personen, die dauerhaft immun, tot oder in Quarantäne sind. Hier liegt also folgendes Schema zugrunde:

$$\boxed{\text{S}} \rightarrow \boxed{\text{I}} \rightarrow \boxed{\text{R}} \tag{2.21}$$

Bevor wir das Differentialgleichungssystem aufstellen, wollen wir etwas genauer auf das Infektionsgeschehen eingehen:
Das Infektionsrisiko ist umso höher, je mehr infektiöse Individuen in der Gesamtpopulation sind. Wie auch schon in den vorhergehenden Modellen können wir dies durch βI angeben wobei β die Infektiösität angibt. Wir gehen auch hier wieder von keinen Geburten- oder Sterbefällen aus, die aus der Gesamtpopulation ein- oder ausscheiden; N bleibt also wieder konstant. Etwaige Individuen, die in Folge der Krankheit sterben, bleiben ebenfalls in der Klasse R.
$\frac{I}{N}$ ist der relative Anteil der Infektiösen an der Gesamtbevölkerung. Weiters nehmen wir an, dass eine infizierte Person in einem bestimmten Zeitintervall k Kontakte mit anderen Individuen hat. Die Kontaktrate zu Infizierten beträgt also $k \cdot \frac{I}{N}$. Bei Kontakt sei a die Ansteckungsrate - die Zahl der Kontakte, die zu einer Infektion führen ist also $a \cdot k \cdot \frac{I}{N}$ woraus $\beta \cdot I = a \cdot k \cdot \frac{I}{N}$ folgt. $\beta = \frac{ak}{N}$ wurde damit genauer spezifiziert (vgl. Murray 2002, S. 320). Je genauer ein Parameter zu Beginn der Modellierung charakterisiert wird desto genauer lässt er sich abschätzen, wenn sein Wert unbekannt ist.
Die Genesungsrate sei wieder γ, was 2.21 auf folgendes Schema präzisiert:

$$\boxed{\text{S}} \xrightarrow{\beta IS} \boxed{\text{I}} \xrightarrow{\gamma I} \boxed{\text{R}} \tag{2.22}$$

und

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{2.23}$$

Zusätzlich führen wir noch den Parameter $\rho := \frac{\gamma}{\beta}$ ein. 2.23 ergibt nun:

$$\begin{aligned} \frac{dS}{dt} &= -\beta IS \\ \frac{dI}{dt} &= \beta(S - \rho)I \\ \frac{dR}{dt} &= \beta\rho I \end{aligned} \tag{2.24}$$

mit gegebenen Anfangswerten $S(0) = S_0, I(0) = I_0$ und $R(0)$, wobei $I_0 \ll N$ sehr gering und $R(0) = 0$. Es gilt auch hier $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$ woraus, wie beschrieben, $S + I + R = N$ folgt.
Ist die Änderungsrate pro Zeiteinheit der Infektionen größer Null (also gilt $\frac{dI}{dt} > 0$) nimmt die Klasse I an Individuen zu. Dies gilt in 2.24 genau für $S > \rho$. Wegen $\beta > 0$ ist auch $\frac{dI}{dt} = \beta(\underbrace{S - \rho}_{>0})I > 0$, was zu einem Ausbruch der Epidemie führen kann, wenn die Anzahl der Gesunden, Ansteckbaren über dem Schwellenwert ρ liegt. Ansonsten (also für $\frac{dI}{dt} < 0$) flaut die Epidemie ab oder bricht gar nicht erst aus. Es ist also notwendig, eine Mindestanzahl von Infizierbaren zu haben. Wenn für $t > 0$ die Zahl der Infizierbaren unter diesen Schwellenwert sinkt, erlischt die Epidemie ebenfalls.

2.3.1 Basisreproduktionszahl

Den Kehrwert des vorher definierten Schwellenwerts ρ ($\frac{1}{\rho} = \frac{\beta}{\gamma}$) bezeichnet man auch als „Übertragungsrate“ einer Infektion (vgl. Jakob 2013, S. 11f). Aus Gleichung 2 in 2.23 lässt sich die Beziehung

$$\mathcal{R}_0 = \frac{\beta S}{\gamma} \tag{2.25}$$

bilden. $\mathcal{R}_0$ wird aus der Übertragungsrate $\frac{\beta}{\gamma}$ und deren Auswirkung zu S gebildet und wird Reproduktionszahl bzw. Basisreproduktionszahl genannt. Sie gibt den Anteil der Gesunden an, den ein infektiöses Individuum im Mittel ansteckt. Grundlegend

gilt wieder:

$$\frac{dI}{dt} = (\beta S - \gamma)I > 0 \text{ wenn } \beta S > \gamma \Leftrightarrow \frac{\beta}{\gamma} S > 1 \Leftrightarrow \mathcal{R}_0 > 1$$

Für $\mathcal{R}_0 > 1$ steigt die Zahl der Infizierten also an. Umgekehrt fällt sie für $\mathcal{R}_0 < 1$. Für $\mathcal{R}_0 = 1$ ist $\frac{dI}{dt} = 0$, denn wenn ein Infizierter im Mittel nur einen Gesunden ansteckt, bleibt die Anzahl der Individuen der Klasse I unverändert (vgl. Jakob 2013, S. 11f).

2.3.2 Trajektorien und Phasenraum

Im Folgenden betrachten wir die (I, S)-Trajektorie als Phasenraum zur Lösungsdarstellung (vgl. Zacher et al. 2008, S. 25f). Dazu dividiert man $\frac{\frac{dI}{dt}}{\frac{dS}{dt}}$ und erhält $\frac{dI}{dS} = \frac{\beta SI - \beta\rho I}{-\beta SI} = -1 + \frac{\rho}{S}$.

Wir suchen nun die Lösung dieser Differentialgleichung:

$$\begin{aligned}
\frac{dI}{dS} &= -1 + \frac{\rho}{S} \\
\int dI &= -\int 1 \, dS + \rho \int \frac{1}{S} \, dS \\
\int dI + \int 1 \, dS - \rho \int \frac{1}{S} \, dS &= 0 \\
I + S - \rho \cdot ln(S) &= \text{konstant} \\
I + S - \rho \cdot ln(S) &= \underbrace{I_0 + S_0}_{=N} - \rho \cdot ln(S_0) \\
I(S) &= \rho \cdot ln(S) - S + N - \rho \cdot ln(S_0) \qquad (2.26)
\end{aligned}$$

Die Lösungstrajektorien beginnen alle auf der Geraden $I + S = N$ und gehen für $t > 0$ immer gegen $I = 0$.

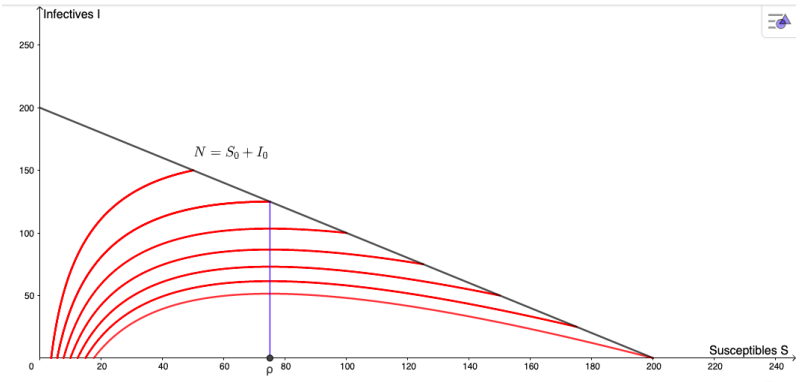


Abb. 2.1: Trajektorien für das SIR-Modell mit den Anfangsbedingungen S_0, I_0 und $N = 200$ (eigene Darstellung).

Aus dem obigen Phasendiagramm lässt sich deutlich erkennen, dass für $S_0 \leq \rho$ die Bahnkurven nicht mehr ansteigen. Es kommt also zu keiner weiteren Ausbreitung der Krankheit. Für $S_0 > \rho$ hingegen, steigen die Kurven vorerst an, wodurch sich eine mögliche Epidemie in der Bevölkerung ausbreitet (vgl. Jakob 2013, S. 10f).

2.3.3 Maximale Infiziertenzahl und Langzeitverhalten der Klasse I

Tritt nun der Fall ein, dass es für $S_0 > \rho$ zum Ausbruch einer Epidemie kommt, ist interessant zu wissen, welcher Höchstwert I_{max} von Erkrankten nach diesem Modell erreicht wird. Dazu bestimmt man jene Stelle von I an der die Wachstumsgeschwindigkeit gleich Null ist, also $\frac{dI}{dt} = 0 = \beta(S - \rho)I$.
Wir betrachten nur lokale Extrema, das heißt, $I > 0$ und sowieso $\beta > 0$. Für $0 = \beta(S - \rho)I$ ist dies genau dann der Fall, wenn $S = \rho$.
Eingesetzt in 2.26 erhält man:

$$\begin{aligned} I_{max} &= I(\rho) = \rho \cdot ln(\rho) - \rho + N - \rho \cdot ln(S_0) \\ \Leftrightarrow I_{max} &= N + \rho \cdot (ln(\rho) - ln(S_0) - 1) \\ \Leftrightarrow I_{max} &= N + \rho \cdot \left(ln\left(\frac{\rho}{S_0}\right) - 1 \right) \end{aligned} \tag{2.27}$$

Liegt der Startwert I_0 nahe an I_{max} fällt die Epidemie nicht sehr groß aus, denn die Differenz zwischen dem Startwert der Infizierten und der maximalen Anzahl der Infizierten ist relativ gering wie auch im Phasendiagramm zu erkennen ist. In jedem Fall fallen alle Trajektorien gegen $I = 0$ ab, denn zu einem bestimmten Zeitpunkt fällt S sicher unter den Schwellenwert ρ. Der Grund dafür ist, dass die Klasse S immer kleiner wird da die Individuen nach einer überstandenen bzw. tödlichen Infektion in

der Klasse R bleiben.
Am Schluss der Krankheitsausbreitung bleibt womöglich ein Restbestand an Gesunden übrig, was uns zu unserer nächsten Frage führt:

2.3.4 Langzeitverhalten der Klasse S

Verbreitet sich in der Bevölkerung die Infektion gilt für $0 < t_1 < t_2$, dass $0 \leq S(t_2) < S(t_1) < S(0) \leq N$. $S(t)$ ist also nach oben und nach unten beschränkt; folglich existiert ein Grenzwert für $t \to \infty$: $\lim_{t\to\infty} S(t) := S(\infty)$.
Damit man über diesen Grenzwert Aussagen treffen kann, betrachtet man zunächst den Quotienten $\frac{\frac{dS}{dt}}{\frac{dR}{dt}}$ (vgl. Brauer et al. 2012, S. 354ff). Daraus folgt: $\frac{dS}{dR} = \frac{-\beta SI}{\gamma I} = -\frac{S}{\rho}$.
Wieder löst man diese Differentialgleichung durch Trennung der Variablen auf:

$$\begin{aligned} \int \frac{dS}{S} &= -\frac{1}{\rho}\int dR \\ ln(S) + \frac{R}{\rho} &= \text{konstant} = ln(S_0) \\ S \cdot e^{\frac{R}{\rho}} &= S_0 \\ S(R) &= S_0 \cdot e^{-\frac{R}{\rho}} \end{aligned} \tag{2.28}$$

2.28 kann man nach unten durch $R < N$ abschätzen:

$$S(R) = S_0 \cdot e^{-\frac{R}{\rho}} > S_0 \cdot e^{-\frac{N}{\rho}} > 0 \tag{2.29}$$

Für $t \to \infty$ definieren wir $R(\infty) := N - S(\infty) - I(\infty)$, wobei $I(\infty) = 0$ und somit $R(\infty) := N - S(\infty)$. Aus 2.29 folgt

$$S(\infty) = S_0 \cdot e^{-\frac{R(\infty)}{\rho}} = S_0 \cdot e^{-\frac{N-S(\infty)}{\rho}} > 0 \tag{2.30}$$

Die Lösung dieser transzendenten Gleichung kann nicht explizit bestimmt werden, sondern muss z.B. mithilfe des Newton-Verfahrens numerisch gelöst werden (vgl. Jakob 2013, S. 15).
Es gilt somit: $0 < S(\infty) < N$.
Wir haben nun also gezeigt, dass nicht die gesamte Population infiziert wird; eine Restpopulation ohne Infektion bleibt bestehen.
Aus dem Phasendiagramm lässt sich erkennen, dass sogar $S(\infty) < \rho$ gilt.

Mit dem folgenden Satz könnte man (unter einigen vereinfachten Annahmen) die

Zahl der erkrankten Individuen am Ende der Epidemie abschätzen.

Satz 2.2: Schwellenwertsatz der Epidemiologie

Sei $S_0 := \rho + \epsilon$ für $\epsilon > 0$ und $\epsilon \ll \rho$ die Anzahl der gesunden, infizierbaren Individuen am Anfang der Epidemie. Die Anzahl der Infizierten sei sehr viel kleiner als die Anzahl der gesunden, infizierbaren Individuen, $I_0 \ll S_0$ und außerdem gelte $I_0 \ll \epsilon$, so werden schließlich etwa 2ϵ Individuen erkranken (vgl. Kohlmeier 2006, S. 76f).

Beweis

Es ist also zu zeigen, dass nach der Epidemie, wenn $\frac{dR}{dt} = 0$, ca. 2ϵ Individuen erkranken, also $R \approx 2\epsilon$ gilt. Aus 2.23 für $\frac{dR}{dt} = \gamma I$ und 2.28 folgt:

$$\frac{dR}{dt} = \gamma \cdot \underbrace{I}_{=N-R-S} \overset{2.28}{=} \gamma \cdot \left(N - R - S_0 \cdot e^{-\frac{R}{\rho}} \right) \tag{2.31}$$

2.31 ist nicht explizit lösbar, kann aber dennoch durch die Taylorreihenentwicklung für $T(R) = e^{-\frac{R}{\rho}}$ an der Stelle $R = 0$ angenähert werden (vgl. Kohlmeier 2006, S. 77).

$$\begin{aligned} T(R+0) &= T(0) + T'(0) \cdot R + \frac{1}{2} \cdot T''(0) \cdot R^2 \\ \Leftrightarrow T(R+0) &= 1 - \frac{R}{\rho} + \frac{R^2}{2\rho^2} \end{aligned}$$

Mit dieser Näherung erhält man:

$$\begin{aligned} \frac{dR}{dt} &= \gamma \cdot \left(N - R - S_0 \cdot \left(1 - \frac{R}{\rho} + \frac{R^2}{2\rho^2} \right) \right) \\ \Leftrightarrow \frac{dR}{dt} &= \gamma \cdot \left(\underbrace{N - S_0}_{=I_0} - R + \frac{S_0}{\rho} \cdot R - \frac{S_0}{2\rho^2} \cdot R^2 \right) \\ \Leftrightarrow \frac{dR}{dt} &= \gamma \cdot \left(I_0 + R \cdot \left(\frac{S_0}{\rho} - 1 \right) - \frac{S_0}{2\rho^2} \cdot R^2 \right) \end{aligned}$$

Laut Satz 2.2 gilt $S_0 = \rho + \epsilon$ mit $\epsilon \ll \rho$, sowie $I_0 \ll \epsilon$. Somit kann obige Gleichung wiederum approximiert werden:

$$\begin{aligned} \frac{dR}{dt} &= \gamma \cdot \left(\left(\frac{\rho + \epsilon}{\rho} - 1 \right) \cdot R - \frac{\rho + \epsilon}{2\rho^2} \cdot R^2 \right) \\ \Leftrightarrow \frac{dR}{dt} &= \gamma R \cdot \left(\frac{\epsilon}{\rho} - \frac{\rho + \epsilon}{2\rho^2} \cdot R \right) \end{aligned} \tag{2.32}$$

Zu Beginn und am Ende der Epidemie ist $\frac{dR}{dt} = 0 = \gamma R \cdot \left(\frac{\epsilon}{\rho} - \frac{\rho+\epsilon}{2\rho^2} \cdot R\right)$. Dies ist nach dem Produkt-Null-Satz genau dann der Fall, wenn $R_1 = 0$ (Beginn der Epidemie) und wenn $0 = \frac{\epsilon}{\rho} - \frac{\rho+\epsilon}{2\rho^2} \cdot R \Leftrightarrow R_2 = \frac{\epsilon}{\rho} \cdot \frac{2\rho^2}{\rho+\epsilon} = \frac{2\epsilon\rho}{\rho+\epsilon} \overset{\epsilon>0}{<} \frac{2\epsilon\rho}{\rho} = 2\epsilon$.

□

2.3.5 Langzeitverhalten von R

Oftmals ist es schwer abzuschätzen wie viele Infektionen pro Zeiteinheit stattfinden. Deshalb untersuchen wir nun die Klasse R da es manchmal einfacher ist, die Anzahl der geheilten (toten, in Quarantäne befindlichen) Individuen pro Zeiteinheit zu bestimmen (vgl. Jakob 2013, S. 16). Dazu betrachten wir 2.23:

$$\frac{dR}{dt} = \gamma I \overset{S+I+R=N}{=} \gamma \cdot (N - R - S) = \gamma \cdot \left(N - R - S_0 \cdot e^{-\frac{R}{\rho}}\right) \text{ und } R(0) = 0. \tag{2.33}$$

Obige Differentialgleichung hängt nur noch von R ab. R hängt wiederum von t ab. Wenn also die Parameter N, S_0, ρ und γ bekannt sind, kann mit wenig Aufwand eine Lösung für 2.33 approximiert werden.
Meistens ist es jedoch der Fall, dass nicht alle Parameter bekannt sind. Deshalb versuchen wir für 2.33 eine näherungsweise analytische Lösung herzuleiten. Dafür muss bereits zu anfangs eine wichtige Annahme getroffen werden: Folgende Annäherung gelte nur für Epidemien, die nicht sehr groß ausfallen (vgl. Jakob 2013, S. 16). Hierbei kann man annehmen, dass $\frac{R}{\rho}$ sehr klein ist, etwa $\frac{R}{\rho} < 1$, denn dann bleibt $S(R) = S_0 \cdot e^{-\frac{R}{\rho}}$ für $t \to \infty$ „nahe“ bei S_0.
Wie in Beweis zu Satz 2.2 schätzt man $e^{-\frac{R}{\rho}}$ mit der Taylor-Entwicklung um 0 ab und erhält:

$$\begin{aligned} \frac{dR}{dt} &= \gamma \cdot \Bigg(\underbrace{N - S_0}_{:=c} \underbrace{+\left(\frac{S_0}{\rho} - 1\right)}_{:=b} \cdot R \underbrace{- \frac{S_0}{2\rho^2} \cdot R^2}_{:=a}\Bigg) \\ \Leftrightarrow \frac{dR}{dt} &= \gamma \cdot \left(aR^2 + bR + c\right) \end{aligned} \tag{2.34}$$

Diese Differentialgleichung 2.34 wird wieder durch Trennung der Variablen und unter Zuhilfenahme eines Computeralgebrasystems, das speziell auf das Lösen von

Integralen ausgerichtet ist (vgl. David Scherfgen IT Services 2021), gelöst:

$$\begin{aligned}
\frac{dR}{dt} &= \gamma \cdot \Big(aR^2 + bR + c\Big) \\
\Leftrightarrow \int \frac{dR}{aR^2 + bR + c} &= \int \gamma \; dt \\
\Leftrightarrow \frac{2 \; \operatorname{artanh}\Big(\frac{2aR+b}{\sqrt{b^2-4ac}}\Big)}{\sqrt{b^2-4ac}} &= \gamma t + \psi \\
\operatorname{artanh}\Big(\frac{2aR+b}{\sqrt{b^2-4ac}}\Big) &= \frac{-(\gamma t + \psi) \cdot \sqrt{b^2-4ac}}{2} \\
\frac{2aR+b}{\sqrt{b^2-4ac}} &= \tanh\Big(\frac{-(\gamma t + \psi) \cdot \sqrt{b^2-4ac}}{2}\Big) \\
R(t) &= \frac{1}{2a}\Big[\tanh\Big(\frac{-(\gamma t + \psi) \cdot \sqrt{b^2-4ac}}{2}\Big) \cdot \sqrt{b^2-4ac} - b\Big]
\end{aligned} \tag{2.35}$$

mit $a = -\frac{S_0}{2\rho^2}, \quad b = \frac{S_0}{\rho} - 1, \quad c = N - S_0.$
Für die Konstante ψ ergibt sich durch Einsetzen der Anfangsbedingung $R(0) = 0$:

$$\begin{aligned}
\frac{2 \; \operatorname{artanh}\Big(\frac{2a\cdot 0+b}{\sqrt{b^2-4ac}}\Big)}{\sqrt{b^2-4ac}} &= \gamma \cdot 0 + \psi \\
\psi &= \frac{2 \; \operatorname{artanh}\Big(\frac{b}{\sqrt{b^2-4ac}}\Big)}{\sqrt{b^2-4ac}}
\end{aligned} \tag{2.36}$$

2.35 ergibt also die analytisch approximative Lösungsformel für die numerische Differentialgleichung 2.33. Diese Lösung ist - wie eingangs erwähnt - nur für $\frac{R}{\rho} < 1$ hinreichend gut approximiert. Für $\frac{R}{\rho} > 1$ kann auf die numerische Lösung von 2.33 zurückgegriffen werden. Die Idee der Näherungslösung stammt aus Murray 2002, S. 324.
Murray hat allerdings einen Fehler in seiner Lösung. Dies kann gezeigt werden, indem man die soeben hergeleitete Lösung mit der numerischen Lösung der exakten Differentialgleichung von 2.33 vergleicht. Daten, die mit Murrays Lösung geplottet werden, weichen jedoch weit von der richtigen numerischen Lösung ab (vgl. Jakob 2013, S. 18).

2.3.6 Stabilitätsanalyse

Wir haben uns bereits in den Kapiteln 2.3.3, 2.3.4 und 2.3.5 mit dem Langzeitverhalten der einzelnen Klassen beschäftigt. Nun wollen wir nochmals auf etwaige Gleichgewichte und deren Stabilität im SIR-Modell eingehen.
Zur Bestimmung der Gleichgewichtslage setzen wir die Mengenbilanzen 2.23 gleich Null, denn in einem Gleichgewicht bleiben die Bestände unverändert.

$$\begin{aligned} 0 &= -\beta SI \\ 0 &= \beta SI - \gamma I \\ 0 &= \gamma I \end{aligned} \tag{2.37}$$

Aus den vorigen Kapiteln ist bereits bekannt, dass die Klasse I gegen Null gehen wird. Aus 2.37 ergibt sich auch

$$\frac{dS}{dt} = 0 \Leftrightarrow I = 0 \text{ oder } S = 0$$
$$\frac{dI}{dt} = 0 \Leftrightarrow I = 0 \text{ oder } S = \frac{\gamma}{\beta} = \rho$$

Die Gleichgewichtspunkte von 2.23 sind demnach $(S_f,\, I_f) = (S,\, 0)$. Die gesamte S-Achse besteht also aus Fixpunkten, wobei dabei keine asymptotische Stabilität erwartet werden kann.
Zur Stabilitätsanalyse reicht es, wenn man die Größen S und I betrachtet, denn es gibt keine Abhängigkeiten bzw. Rückwirkungen von R auf S und I.

$$\begin{pmatrix} \frac{dS}{dt} \\ \frac{dI}{dt} \end{pmatrix} = \begin{pmatrix} -\beta SI \\ \beta SI - \gamma I \end{pmatrix} \tag{2.38}$$

Zur Untersuchung der Stabilitätsbedingung wird wieder die Jacobi-Matrix von 2.38 gebildet:

$$J = \begin{pmatrix} -\beta I & -\beta S \\ \beta I & \beta S - \gamma \end{pmatrix}$$

Nun setzt man für $I = 0$ ein und berechnet die Eigenwerte von J:

$$\det(J - \lambda E) = \det \begin{pmatrix} 0 & -\beta S \\ 0 & \beta S - \gamma \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \Big] = \det \Big[\begin{pmatrix} -\lambda & -\beta S \\ 0 & \beta S - \gamma - \lambda \end{pmatrix}$$

Das charakteristische Polynom ergibt $-\lambda(\beta S-\gamma-\lambda)=0 \Leftrightarrow \lambda_1=0$ oder $\lambda_2=\beta S-\gamma$. Für $S>\frac{\gamma}{\beta}=\rho$ ist der Realteil $Re(\lambda_2)=\lambda_2>0$ und somit ist der Fixpunkt $(S,\,0)$ instabil.

2.4 SEIR-Modell

Das SEIR-Modell stellt eine Erweiterung zum SIR-Modell dar, indem die Klasse E der Exposed (dt.: latent Infizierte) hinzugefügt wird. In diese Klasse werden jene Individuen gezählt, die sich mit der Krankheit angesteckt haben, aber noch nicht infektiös sind; sie befinden sich in der sogenannten Latenzzeit.
Der Krankheitsverlauf erfolgt in drei Schritten: Ein Infizierter trifft auf eine anfällige Person, die dadurch mit einer gewissen Wahrscheinlichkeit infiziert wird. Diese Person zeigt bis zum Ende der Inkubationszeit keinerlei Symptome. Nach Ende der Latenzzeit wechselt die bis dahin noch nicht infektiöse Person in die Klasse I und wird selbst ansteckend. Diese kann ihrerseits wiederum andere Personen infizieren und wird je nach Schwere der Krankheit sterben oder genesen (vgl. Schuster 2015, S. 7). Es lässt sich folgendes Schema aufstellen:

$$\boxed{\text{S}} \rightarrow \boxed{\text{E}} \rightarrow \boxed{\text{I}} \rightarrow \boxed{\text{R}} \tag{2.39}$$

Die Parameter für die Übergänge in die Klassen I und R bleiben gleich wie im SIR-Modell. Allerdings muss man hier näher auf die Klassen E eingehen. Anfällige Individuen wechseln nach einer Ansteckung in die Klasse E. Wir nehmen an, dass angesteckte Individuen $\frac{1}{\epsilon}$ Tage in der Klasse E verweilen bis sie ansteckend werden. $\frac{1}{\epsilon}$ gibt also die mittlere Latenzzeit an (vgl. Schuster 2015, S. 9). Diese Überlegung ist analog zu jener mit der mittleren Infektionszeit aus 2.16.
Schema 2.39 kann demnach wie folgt präzisiert werden:

$$\boxed{\text{S}} \xrightarrow{\beta IS} \boxed{\text{E}} \xrightarrow{\epsilon E} \boxed{\text{I}} \xrightarrow{\gamma I} \boxed{\text{R}} \tag{2.40}$$

Die Ausbreitungsdynamik der Krankheit wird durch folgendes nichtlineare Differentialgleichungssystem beschrieben:

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dE}{dt} &= \beta SI - \epsilon E \\ \frac{dI}{dt} &= \epsilon E - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned} \tag{2.41}$$

mit Anfangsbedingungen $S(0) = S_0$, $E(0) = E_0$, $I(0) = I_0$ und $R(0) = 0$.

$\frac{dS}{dt} + \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$ wodurch wieder $S + E + I + R = N$ gilt. Geburten- und Sterbefälle werden zunächst auch hier wieder vernachlässigt.

Lemma 2.1

Für die Lösungen von 2.41 für S_0, E_0, I_0, $R_0 \geq 0$ gilt S, E, I, $R \geq 0$.

Beweis

Die Differentialgleichungen von 2.41 sind stetig, erfüllen eine Lipschitzbedingung auf $\mathbb{R}^+$ und liefern daher für $t \geq 0$ eine eindeutige Lösung (vgl. Youssef et al. 2020, S. 3).

Schätzt man $\frac{dE}{dt}$ von 2.41 ab, erhält man: $\frac{dE}{dt} \geq -\epsilon E$ da $\beta SI \geq 0$. Löst man diese Differentialgleichung durch Trennung der Variablen, erhält man:

$$\begin{aligned} \int \frac{dE}{E} &= -\epsilon \int dt \\ \Leftrightarrow ln(E) + \epsilon t &= \text{konstant} \\ \Leftrightarrow E &= E_0 \cdot e^{-\epsilon t} \end{aligned}$$

Daher ist $E \geq E_0 \cdot e^{-\epsilon t} \geq 0$. Analog gilt für $\frac{dI}{dt} \geq -\gamma I$ da $\epsilon E \geq 0$, das nach Integration $I \geq I_0 \cdot e^{-\gamma t} \geq 0$ ergibt und für $\frac{dR}{dt} = \gamma I$ folgt die Behauptung ebenso. $\frac{dS}{dt}$ wird ebenfalls abgeschätzt mit $\frac{dS}{dt} \geq -\beta NS$. Durch Integration ergibt sich $S \geq S_0 \cdot e^{-\beta N t} \geq 0$.

□

2.4.1 Trajektorien und Phasenraum

Wir betrachten wieder die Lösungskurven in der S-$(E+I)$-Phasenebene. Dazu fassen wir zunächst alle Infizierten in die Klasse J zusammen und betrachten $\frac{\frac{dJ}{dt}}{\frac{dS}{dt}}$ aus dem obigen Differentialgleichungssystem 2.41.

$$\frac{dJ}{dS} = \frac{d(E+I)}{dS} = \frac{\beta IS - \gamma I}{-\beta IS} = -1 + \frac{\gamma}{\beta S} \tag{2.42}$$

Für $\rho = \frac{\gamma}{\beta}$ folgt $\frac{dJ}{dS} = -1 + \frac{\rho}{S}$ und durch Integration erhält man:

$$\begin{aligned} \int dJ &= \int -1 + \frac{\rho}{S}\, dS \\ \Leftrightarrow J + S - \rho \cdot ln(S) &= \text{konstant} = \underbrace{J_0}_{=E_0+I_0} + S_0 - \rho \cdot ln(S_0) \\ \Leftrightarrow J = E + I &= \rho \cdot ln(S) - S + E_0 + I_0 + S_0 - \rho \cdot ln(S_0) := \bar{f}(S) \end{aligned}$$

Für $\bar{f}'(S)$ gelten nun die gleichen Eigenschaften wie im SIR-Modell:

$$\bar{f}'(S) \begin{cases} > 0 & \text{falls } S < \rho \\ = 0 & \text{falls } S = \rho \\ < 0 & \text{falls } S > \rho \end{cases}$$

Liegt S unter dem vorher definierten Schwellenwert ρ, kommt es in Folge zu einem Anstieg der Klasse J, also der primär und latent Infizierten. Andererseits flaut die Epidemie ab.

2.4.2 Langzeitverhalten der Klassen S, I und R

Kommt es für $S > \rho$ zur Ausbreitung der Krankheit, ist auch hier wieder interessant zu wissen, welcher Höchstwert I_{max} an Infizierten erreicht wird. Dazu berechnet man das lokale Maximum von $J = \bar{f}(S)$. $\bar{f}(S)$ ist eine Funktion der Gestalt $f(x) = \rho \cdot ln(x) - x + c$.

$$\begin{aligned} f'(x) = \frac{\rho}{x} - 1 = 0 \Leftrightarrow x = \rho \\ f''(x) = -\frac{\rho}{x^2} < 0 \ \forall x \\ \Rightarrow x = \rho \Rightarrow \text{Maximum} \end{aligned}$$

Der maximal mögliche Wert der Infizierten hängt somit nur von ρ ab, wird für $S = \rho$ erreicht und ist

$$J_{max} = \rho \cdot ln(\rho) - \rho + E_0 + I_0 + S_0 - \rho \cdot ln(S_0)$$
$$\Leftrightarrow J_{max} = E_0 + I_0 + S_0 + \rho \cdot \Big(ln(\rho) - 1 - ln(S_0)\Big) = E_0 + I_0 + S_0 + \rho \cdot \left(ln\Big(\frac{\rho}{S_0}\Big) - 1\right)$$

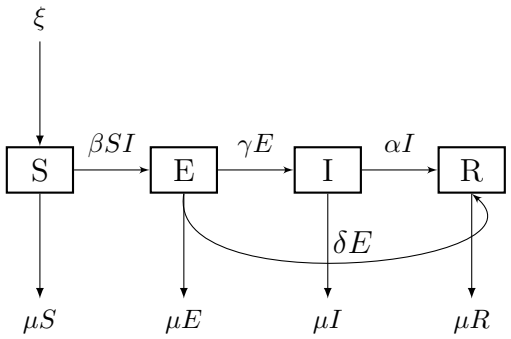
Tritt eine Krankheit, z.B. ausgelöst durch einen neuartigen Erreger auf, kann $S_0 = N$ und $E_0 = I_0 = 0$ angenommen werden. Somit wäre $J_{max} = N + \rho \cdot \Big(ln(\frac{\rho}{N}) - 1\Big)$.
Die obige Lösung und das Phasenportrait ist analog zu jenem im SIR-Modell. Deshalb wird nun in diesem Kapitel nicht mehr näher darauf eingegangen.

2.5 SEIR Modell mit Populationsdynamiken

Bislang betrachteten wir ausnahmslos Modelle ohne Populationsdynamiken. Das bedeutet, Geburten- und Sterbefälle wurden außer Acht gelassen. Werden Krankheiten nämlich nur über kurze Perioden hinweg modelliert, macht dies zur wesentlichen Vereinfachung des Modells durchaus Sinn.
Das SEIR-Modell, das wir bislang betrachteten, verhielt sich in dessen Eigenschaften gleich zum SIR-Modell. Die einzige Erweiterung war die Einführung der Klasse E der latent Infizierten. Mit der neuen Klasse J, die die Klassen E und I zusammenfasste, ließen sich schließlich die gleichen Eigenschaften z.B. im Langzeitverhalten nachweisen wie im SIR-Modell.
In diesem Kapitel wollen wir das SEIR-Modell erweitern, indem wir natürliche Todes- und Geburtenraten miteinbeziehen.
Wiederum teilt man die Gesamtpopulation in die vier Klassen S, E, I und R auf. Die Übergangsparameter bleiben gleich wie in den Kapiteln davor. Jedoch führt man zwei neue Parameter ein:

- ξ gibt die Anzahl der Geburten pro Zeiteinheit an.
- μ gibt die natürliche Todesrate (Mortalitätsrate) pro Zeiteinheit an.

Es liegt hier somit folgendes Schema zugrunde.



Die Gesamtpopulation sei zu jedem Zeitpunkt t gegeben durch $N = S + E + I + R$. Nach Youssef et al. 2020, S. 2) lässt sich folgendes nichtlineare Differentialgleichungssystem aufstellen:

$$\begin{aligned} \frac{dS}{dt} &= \xi - \beta SI - \mu S \\ \frac{dE}{dt} &= \beta SI - \epsilon_1 E \\ \frac{dI}{dt} &= \gamma E - \epsilon_2 I \\ \frac{dR}{dt} &= \delta E + \alpha I - \mu R \end{aligned} \tag{2.43}$$

wobei $S(0) = S_0$, $E(0) = E_0$, $I(0) = I_0$ und $R(0) = 0$ die Anfangsbedingungen angeben. Fasst man die Raten mit denen die Individuen die Klassen E und I verlassen zusammen, gilt: $\epsilon_1 = \gamma + \mu + \delta$ und $\epsilon_2 = \alpha + \mu$.
Bevor wir uns genauer mit dem Modell beschäftigen, prüfen wir wieder, ob die Lösungen von 2.43 allesamt positiv sind. Der Beweis ist ähnlich zum Beweis 2.4.

Lemma 2.2

Alle Lösungen mit den Anfangsbedingungen S_0, E_0, I_0, R_0 sind für $t \geq 0$ größer gleich Null.

Beweis

Da $\xi - \beta SI \geq 0$ schätzt man zunächst Gleichung 1 aus 2.43 ab und löst diese durch Integration:
$\frac{dS}{dt} \geq -\mu S \Leftrightarrow S \geq S_0 \cdot e^{-\mu t}$. S ist also eine nicht negative Funktion für $t \geq 0$.
Analog gilt für alle übrigen Gleichungen:
$\frac{dE}{dt} \overset{S \geq 0}{\geq} -\epsilon_1 E \Leftrightarrow E \geq E_0 \cdot e^{-\epsilon_1 t}$.

$\frac{dI}{dt} \overset{E\geq 0}{\geq} -\epsilon_2 I \Leftrightarrow I \geq I_0 \cdot e^{-\epsilon_2 t}$.
$\frac{dR}{dt} \overset{E,I\geq 0}{\geq} -\mu R \Leftrightarrow R \geq R_0 \cdot e^{-\mu t}$.

□

2.5.1 Basisreproduktionszahl

Um eine geeignete Formel für die Basisreproduktionszahl herzuleiten, setzt man die Mengenbilanzen aus 2.43 gleich Null und formt die Gleichungen 1, 2 und 3 um:
Aus 2.43 (3) erhält man:

$$E = \frac{\epsilon_2 I}{\gamma} \tag{2.44}$$

Aus 2.43 (2) folgt:

$$S = \frac{\epsilon_1 E}{\beta I} \overset{2.44}{\Rightarrow} \frac{\epsilon_1 \epsilon_2}{\beta \gamma} \tag{2.45}$$

Für 2.43 (1) ergibt sich:

$$I = \frac{\xi - \mu S}{\beta S} = \frac{\xi}{\beta S} - \frac{\mu}{\beta} \overset{2.45}{\Rightarrow} = \frac{\xi}{\beta \cdot \frac{\epsilon_1 \epsilon_2}{\beta \gamma}} - \frac{\mu}{\beta} = \frac{\xi \gamma}{\epsilon_1 \epsilon_2} - \frac{\mu}{\beta} = \frac{\mu}{\beta} \cdot \left(\frac{\beta \xi \gamma}{\epsilon_1 \epsilon_2 \mu} - 1 \right) \tag{2.46}$$

Daraus lässt sich nun die Basisreproduktionszahl (vgl. Youssef et al. 2020, S. 3) definieren mit $\mathcal{R}_0 = \frac{\beta \xi \gamma}{\epsilon_1 \epsilon_2 \mu} = \frac{\beta \xi \gamma}{\mu(\gamma+\mu+\delta)(\alpha+\mu)} > 0$.

$$\mathcal{R}_0 = \begin{cases} > 1 & \text{falls } \beta \xi \gamma > \mu(\gamma + \mu + \delta)(\alpha + \mu) \\ \leq 1 & \text{sonst} \end{cases}$$

Wenn die Übertragungsrate $\beta = 0$, ist ebenfalls $\mathcal{R}_0 = 0$.
Wir diskutieren nun wie sich einzelne Veränderungen der Parameter auf $\mathcal{R}_0$ auswirken und klären somit die Bewährtheit der Formel:
Wächst β oder ξ, wird $\mathcal{R}_0$ ebenfalls größer. Wächst nämlich die Übertragungsrate β, muss auch $\mathcal{R}_0$ wachsen. Wird die Anzahl der Geburten ξ erhöht, gibt es mehr Suscepibles, die angesteckt werden können und $\mathcal{R}_0$ wächst ebenfalls.
Werden hingegen μ, δ oder α größer, also jene Parameter, die dafür sorgen, dass Individuen nicht mehr für eine Infektion empfänglich sind, sinkt auch $\mathcal{R}_0$ durch Vergrößerung des Nenners.
γ kommt sowohl im Zähler als auch im Nenner von $\mathcal{R}_0$ vor und gibt die Übertragungsrate von der Klasse der latent Infizierten in die Klasse der Infizierten an. Eine Vergrößerung von γ sollte deshalb auch $\mathcal{R}_0$ vergrößern, was man mit einer leichten

Rechnung auch zeigen kann.
Die Formel für $\mathcal{R}_0$ scheint somit zum gegebenen Modell zu passen.

2.5.2 Fixpunkte und Stabilitätsanalyse

Um eventuelle Fixpunkte zu ermitteln, setzt man die Modellgleichungen wie im vorigen Kapitel wieder gleich 0 und löst das Gleichungssystem $(\frac{dS}{dt}, \frac{dE}{dt}, \frac{dI}{dt}, \frac{dR}{dt}) = (0, 0, 0, 0)$. Als mögliche Lösung ergibt sich: $S = \frac{\xi}{\mu}$, $E = I = R = 0$. Laut Modell erreicht die Population also ein krankheitsfreies Equilibrium bei $(\frac{\xi}{\mu}, 0, 0, 0)$.
Zur Stabilitätsuntersuchung dieser stationären Lösung berechnet man die Eigenwerte der Jacobi-Matrix des Differentialgleichungssystems 2.43:

$$J = \begin{pmatrix} -\beta I - \mu & 0 & -\beta S & 0 \\ \beta I & -\epsilon_1 & \beta S & 0 \\ 0 & \gamma & -\epsilon_2 & 0 \\ 0 & \delta & \alpha & -\mu \end{pmatrix}$$

Nach der Linearisierung um den Gleichgewichtspunkt (vgl. Youssef et al. 2020, S. 3f), erhält man folgende Jacobi-Matrix:

$$J = \begin{pmatrix} -\mu & 0 & \frac{-\beta\xi}{\mu} & 0 \\ 0 & -\epsilon_1 & \frac{\beta\xi}{\mu} & 0 \\ 0 & \gamma & -\epsilon_2 & 0 \\ 0 & \delta & \alpha & -\mu \end{pmatrix} \tag{2.47}$$

Die Stabilität des Equilibriums hängt von den Realteilen der Eigenwerte ab. Man berechnet dafür das charakteristische Polynom der Jacobi-Matrix, das gegeben ist durch $\det(J - \lambda E)$:

$$\det \begin{pmatrix} -\mu - \lambda & 0 & \frac{-\beta\xi}{\mu} & 0 \\ 0 & -\epsilon_1 - \lambda & \frac{\beta\xi}{\mu} & 0 \\ 0 & \gamma & -\epsilon_2 - \lambda & 0 \\ 0 & \delta & \alpha & -\mu - \lambda \end{pmatrix} \tag{2.48}$$

Da es sich hierbei um eine (4×4)-Matrix handelt, führen wir vorher den „Laplace'schen Entwicklungssatz“ ein, mit dem man allgemein die Determinante einer $(n \times n)$-Matrix berechnen kann.

Definition 2.1: Entwicklungssatz von Laplace

Sei $n \geq 2$ und $A \in \mathbb{R}^{n \times n}$. Dann gilt für alle $i \in \{1, ..., n\}$ (Entwicklung nach der i-ten Zeile):

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A'_{ij})$$

sowie für alle $j \in \{1, ..., n\}$ (Entwicklung nach der j-ten Spalte):

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A'_{ij})$$

A'_{ij} bezeichnet dabei die „Streichungsmatrix“, die man erhält, wenn von A die i-te Zeile und j-te Spalte gestrichten wird (vgl. Fischer 2008, S. 203).

Nach Anwendung dieses Satzes auf 2.48 erhält man $\det(J-\lambda E) = \frac{-(\lambda+\mu)^2(\beta\gamma\xi-(\epsilon_1+\lambda)(\epsilon_2+\lambda)\mu)}{\mu}$. Setzt man das charakteristische Polynom gleich 0 und löst es nach λ auf, ergeben sich folgende Lösungen:

$$\begin{aligned}
\lambda_1 &= -\mu \\
\lambda_2 &= -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 - \sqrt{4\beta\gamma\xi\mu + (\epsilon_1 - \epsilon_2)^2}\right) = -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 - \sqrt{4\epsilon_1\epsilon_2\mathcal{R}_0 + (\epsilon_1 - \epsilon_2)^2}\right) \\
\lambda_3 &= -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 + \sqrt{4\beta\gamma\xi\mu + (\epsilon_1 - \epsilon_2)^2}\right) = -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 + \sqrt{4\epsilon_1\epsilon_2\mathcal{R}_0 + (\epsilon_1 - \epsilon_2)^2}\right)
\end{aligned} \tag{2.49}$$

Für $\mu > 0$ ist immer $\lambda_1 < 0$. Deswegen hängt die Stabilität des Equilibriums rein von λ_2 und λ_3 ab.

Ist $\mathcal{R}_0 = 1$ gilt:
$\lambda_2 = -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 - \sqrt{4\epsilon_1\epsilon_2 + (\epsilon_1 - \epsilon_2)^2}\right) = -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 - \sqrt{(\epsilon_1 + \epsilon_2)^2}\right) = 0.$
$\lambda_3 = -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 + \sqrt{4\epsilon_1\epsilon_2 + (\epsilon_1 - \epsilon_2)^2}\right) = -\frac{1}{2} \cdot \left(\epsilon_1 + \epsilon_2 + \sqrt{(\epsilon_1 + \epsilon_2)^2}\right) = -(\epsilon_1 + \epsilon_2) < 0.$

Ist $\mathcal{R}_0 > 1$ gilt nach Abschätzung:
$\lambda_2 > \frac{1}{2} \cdot \left(-(\epsilon_1 + \epsilon_2) + \sqrt{4\epsilon_1\epsilon_2 + (\epsilon_1 + \epsilon_2)^2}\right) = \frac{1}{2} \cdot \left(-(\epsilon_1 + \epsilon_2) + \sqrt{(\epsilon_1 + \epsilon_2)^2}\right) = 0.$
Für λ_3 gilt, je größer $\mathcal{R}_0$ wird desto kleiner wird λ_3 und immer $\lambda_3 < 0$.
Mit den analogen Argumenten gilt für $\mathcal{R}_0 < 1$, dass $\lambda_2 < 0$ und $\lambda_3 < 0$.
Diese Ergebnisse können wir im folgenden Satz zusammenfassen:

Satz 2.3

Es sei die Reproduktionszahl $\mathcal{R}_0 = \frac{\beta\xi\gamma}{\epsilon_1\epsilon_2\mu}$ gegeben. Im populationsdynamischen SEIR-Modell 2.43 wird für $\mathcal{R}_0 < 1$ das krankheitsfreie Gleichgewicht $E_0 = (\frac{\xi}{\mu}, 0, 0, 0)$ angestrebt. Für $\mathcal{R}_0 \geq 1$ ist das Gleichgewicht instabil und die Krankheit bricht aus.

Lemma 2.3

Alle Lösungen von 2.50 sind für $t \to \infty$ positiv und bleiben im Gebiet $\psi := \left\{(S, E, I, R) : 0 \leq N \leq \frac{\xi}{\mu}\right\}$.

Beweis

Es gilt $N = S + E + I + R$. Differenziert man beide Seiten dieser Gleichung erhält man $N' = S' + E' + I' + R'$. Substituiert man hier die Gleichungen aus 2.43, erhält man

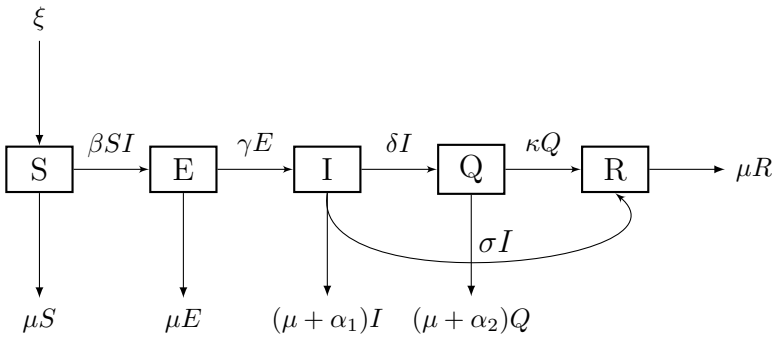
$$\begin{aligned} N' &= \xi - \beta IS - \mu S + \beta IS - \gamma E - \mu E - \delta E + \gamma E - \alpha I - \mu I + \delta E + \alpha I - \mu R \\ &= \xi - \mu \cdot (S + E + I + R) = \xi - \mu N \end{aligned}$$

Als Lösung der obigen Differentialgleichung ergibt sich $N = \frac{\xi}{\mu} + \left(N_0 - \frac{\xi}{\mu}\right)e^{-\mu t}$. Für $t \to \infty$ gilt $N \to \frac{\xi}{\mu}$.

□

2.6 SEIQR Modell mit Populationsdynamik

Die in den vorigen Kapiteln diskutierten Epidemiemodelle modellierten die Ausbreitung von Infektionskrankheiten auf verschiedene Weisen. Die jedem Modell zugrunde liegenden Annahmen wurden von Modell zu Modell komplexer. Quarantänemaßnahmen wurden bisher nur indirekt betrachtet, denn Individuen, in der Klasse R können auch als isoliert von der restlichen Bevölkerung angesehen werden. Zur besseren Differenzierung jener Individuen, die nach einer Infektion abgeschottet werden sollen, führt man eine neue Klasse Q - „quarantined“ ein. Das Schema des Modells sei nachfolgend dargestellt:



Aufgrund einiger grundlegender Unterschiede im Aufbau des Modells im Vergleich zu den anderen Modellen werden etwaige Parameter nun anders eingesetzt; einige wurden neu hinzugefügt. Deshalb ist deren Bedeutung nachfolgend erklärt:

- β, γ, ξ, μ sind ident zu den vorigen Modellen.
- α_1 gibt die natürliche Todesrate (Mortalitätsrate) der infizierten Individuen pro Zeiteinheit an.
- α_2 gibt die Mortalitätsrate der isolierten (in Quarantäne befindlichen) Individuen pro Zeiteinheit an.
- δ gibt die Übergangsrate der Individuen von Klasse I zu Klasse Q pro Zeiteinheit an.
- κ gibt die Übergangsrate der Individuen von Klasse Q zu Klasse R pro Zeiteinheit an.
- σ gibt die Übergangsrate der Individuen von Klasse I zu Klasse R pro Zeiteinheit an.

Es liegt diesem Modell folgendes Differentialgleichungssystem zugrunde:

$$\begin{aligned}
\frac{dS}{dt} &= \xi - \beta IS - \mu S \\
\frac{dE}{dt} &= \beta IS - \epsilon_1 E \\
\frac{dI}{dt} &= \gamma E - \epsilon_2 I \\
\frac{dQ}{dt} &= \delta I - \epsilon_3 Q \\
\frac{dR}{dt} &= \kappa Q + \sigma I - \mu R
\end{aligned} \tag{2.50}$$

wobei $\epsilon_1 = \gamma + \mu$, $\epsilon_2 = \mu + \alpha_1 + \delta + \sigma$, $\epsilon_3 = \mu + \alpha_2 + \kappa$ und die Anfangsbedingungen $S(0) = S_0$, $E(0) = E_0$, $I(0) = I_0$, $Q(0) = 0$ und $R(0) = 0$.

Lemma 2.4

Alle Lösungen von 2.50 sind positiv und bleiben im Gebiet $\psi := \left\{(S, E, I, Q, R) : 0 \leq N \leq \frac{\xi}{\mu}\right\}$.

Beweis

Der Beweis erfolgt analog zum Beweis 2.5.

□

2.6.1 Fixpunkte und Stabilität

Für die Basisreproduktionszahl gilt nach (vgl. Mishra et al. 2011, S. 4): $\mathcal{R}_0 = \frac{\beta\xi}{\mu\epsilon_1\epsilon_2}$. Um die Fixpunkte dieses Modells zu bestimmen, setzten wir die Mengenbilanzen aus 2.50 gleich Null und lösen dieses Gleichungssystem mittels Technologieeinsatz. Dafür wurden die jeweiligen Gleichungen mit einem Computeralgebrasystem (TI-Nspire CX CAS Premium Software) durch Nullsetzen gelöst.
Es ergeben sich zwei Lösungsmengen wobei sich für $\mathcal{R}_0 < 1$ das krankheitsfreie Equilibrium $G_0 = (\frac{\xi}{\mu}, 0, 0, 0, 0)$ analog zum vorigen Modell einstellt.
Ist die Basisreproduktionszahl hingegen $\mathcal{R}_0 > 1$ strebt das Modell gegen das endemische Gleichgewicht $G_e = (S_e, E_e, I_e, Q_e, R_e)$. Zur besseren Übersicht wurde in die nachfolgenden Lösungen $\mathcal{R}_0$ substituiert (vgl. Mishra et al. 2011, S. 4):

$$\begin{aligned} S_e &= \frac{\xi}{\mu\gamma\mathcal{R}_0} \\ E_e &= \frac{\xi(\gamma\mathcal{R}_0 - 1)}{\gamma\epsilon_1\mathcal{R}_0} \\ I_e &= \frac{\mu(\gamma\mathcal{R}_0 - 1)}{\beta} \\ Q_e &= \frac{\mu\delta(\gamma\mathcal{R}_0 - 1)}{\beta\epsilon_3} \\ R_e &= \frac{\kappa\delta(\gamma\mathcal{R}_0 - 1)}{\beta\epsilon_3} + \frac{\sigma(\gamma\mathcal{R}_0 - 1)}{\beta} \end{aligned} \tag{2.51}$$

Für die Stabilitätseigenschaften der Gleichgewichtspunkte, berechnet man die Nullstellen des charakteristischen Polynoms der Jacobi-Matrix analog zum vorigen Kapitel:

$\chi_A(J) = \det(J - \lambda E) = 0$. Für $\mathcal{R}_0 < 1$ ergeben sich folgende Lösungen:

$$\lambda_1 = -\mu; \quad \lambda_2 = -\epsilon_1; \quad \lambda_3 = -\epsilon_2; \quad \lambda_1 = -\epsilon_3$$

Alle Eigenwerte sind negativ und somit sind auch alle Realteile der Eigenwerte negativ. Das krankheitsfreie Equilibrium ist asymptotisch stabil für $\mathcal{R}_0 < 1$.
Analog berechnet man die Eigenwerte der Jacobi-Matrix für das endemische Gleichgewicht. Auch hier sind die Lösungen des charakteristischen Polynoms allesamt negativ. Das Gleichgewicht G_e ist asymptotisch stabil (vgl. Mishra et al. 2011, S. 4).

Satz 2.4

Für das populationsdynamische SEIQR-Modell sei die Reproduktionszahl $\mathcal{R}_0 = \frac{\beta\xi}{\mu\epsilon_1\epsilon_2}$ gegeben. Für $\mathcal{R}_0 < 1$ strebt das Modell ein krankheitsfreies Gleichgewicht $G_0 = (\frac{\xi}{\mu}, 0, 0, 0)$ an. Für $\mathcal{R}_0 >$ wird das endemische Gleichgewicht G_e von 2.51 erreicht.

KAPITEL 3

Ergebnisse und Diskussion

3.1 Daten

In diesem Kapitel wird der praktische Einsatz der theoretischen Modelle, die im vorherigen Kapitel eingeführt wurden, diskutiert. Der Aufbau wird analog zu jenem vom Kapitel 2 sein - beginnend von den einfachen zu den komplexeren Modellen werden deren Vor- und Nachteile für die Anpassung, sowie die Prognose der echten Daten zur Corona-Pandemie für Österreich diskutiert.
Für die Modellierung unterscheiden wir zwei Infektionswellen. Die erste Welle vom Frühjahr 2020 wird von 01.03.2020 bis einschließlich 30.04.2020 modelliert. Am 29.02.2020 waren $I_0 = 4$ Coronainfektionen für ganz Österreich gemeldet. Da es zu diesem Zeitpunkt noch keine Verstorbenen oder immunen Personen gab, ist $R_0 = 0$. Die zweite Welle wird von 13.09.2020 bis einschließlich 31.12.2020 modelliert. Dabei gilt für $I_0 = 413$ (Coronainfektionen am 12.09.2020). Laut einer Publikation der AGES, Medizinischen Universität Graz und der Universität Stanford ist eine Neuinfektion mit SARS-CoV-2 nach einer überstandenen Erkrankung um 91% geringer (vgl. Pilz et al. 2021, S. 2). Bis zum 12.09.2020 hatten $R_0 = 33\,204$ Personen bereits eine Covid-19-Infektion. Wir nehmen in dieser Arbeit eine hundertprozentige Immunität nach einer überstandenen Infektion an weswegen wir R_0 mit diesem Wert belegen. Die Daten stammen von der AGES (vgl. AGES 2021b) und sind als Anfangsbedingungen in jedem Modell gleich.
Damit für die Modelle optimale Parameterschätzungen durchgeführt werden können, wird in Matlab ein Code implementiert, der dieses Optimierungsproblem lösen soll. Es sollen Parameter gefunden werden, sodass die Modell-Kurven möglichst wenig von den echten Daten abweichen. Dieser Vorgang wird im folgenden Kapitel näher beschrieben.

3.1.1 Minimumsuche in Matlab

Viele natürliche, technische oder wirtschaftliche Aufgabenstellungen beschäftigen sich mit der Suche nach optimalen Bedingungen – also nach Minima oder Maxima – von Systemgrößen. Ein Optimierungsproblem richtet sich immer nach folgendem Schema: Gegeben ist ein Lösungsraum Ω mit Vektoren in $\mathbb{R}^n$, sowie eine Zielfunktion, die jedem Vektor von Ω eine entsprechende reelle Zahl zuordnet. Man ist nun an jenem Vektor interessiert, für welchen die Zielfunktion den kleinsten bzw. größten Wert besitzt (vgl. Jarre et al. 2019, S. IX).
Im Prinzip lassen sich Optimierungsaufgaben mit dem Nullsetzen der (partiellen) Ableitungen charakterisieren. Oft sind diese aber sehr aufwendig oder unmöglich zu lösen und man muss auf numerische Approximationen zurückgreifen. Für die numerische Optimierungssuche gibt es zahlreiche Algorithmen. Sie sind aber grundsätzlich nach dem gleichen Prinzip aufgebaut. Zunächst muss vom Benutzer ein Startwert bzw. ein Startvektor für die zu optimierenden Parameter eingegeben werden. Das Programm verändert nun die Parameter so lange bis eine optimale Lösung gefunden wird und sich die Zielfunktion kaum mehr ändert oder ein Abbruchkriterium unterschritten wird. Ein Problem ist dabei jedoch, wenn die Zielfunktion aus mehreren lokalen Extrema besteht. Das Programm findet zwar eine optimale Lösung, weiß aber in der Regel nicht, ob es sich dabei um die beste globale Lösung handelt. Ist im allgemeinen wenig über die Zielfunktion bekannt, sollte der Benutzer mehrere verschiedene Startwerte durchlaufen lassen um ein „besseres“ lokales Extremum zu finden (vgl. Benker 2010, S. 213ff).
Die Art der Suchrichtung unterscheidet sich je nach Optimierungsprogramm. In Matlab steht der Befehl bzw. die Funktion *fminsearch* für Optimierungsaufgaben zur Verfügung. Die Anwendung der Funktion gibt einen Vektor oder eine Matrix mit den gefundenen optimalen Lösungsskalaren zurück.
fminsearch folgt dem Nelder-Mead-Verfahren (vgl. The MathWorks, Inc 2005).

Nelder-Mead-Verfahren:
Das Nelder-Mead-Verfahren benutzt zur Optimumsuche keine partiellen Ableitungen der Parameter (vgl. Pedamkar 2021). Die Idee hierbei ist die optimale Lösung durch Minimierung eines Simplex zu ertasten. Besteht der Startvektor bzw. die Zielfunktion aus n Parametern, wird ein Simplex mit $n+1$ Ecken gebildet. Im zweidimensionalen Fall würde das Simplex ein Dreieck aufspannen, im dreidimensionalen Fall ein Tetraeder. Das heißt, dass die gewählten Punkte in $\mathbb{R}^n$ nicht allesamt in einer Hyperebene mit der Dimension $n-1$ liegen. Im dreidimensionalen Raum dürfen die Punkte

demnach nicht alle auf einer Geraden liegen (vgl. Scholz 2018, S. 120). Für jeden Eckpunkt des gebildeten $n+1$-Simplex werden die Zielfunktionswerte berechnet wobei ein Eckpunkt den größten Wert besitzt. Die berechneten Funktionswerte werden der Größe nach sortiert und der Schlechteste – also der Größte – durch einen neuen ersetzt. Dafür gibt es verschiedene Möglichkeiten (vgl. Sturm 2010, S. 19ff).: Die erste Möglichkeit besteht darin, den schlechtesten Eckpunkt zu spiegeln. Erfüllt der gespiegelte Punkt bessere Werte der Zielfunktion als alle übrigen Eckpunkte, wird dieser noch einmal in diese Richtung expandiert (z.B. mit den Faktor zwei). Ist auch dieser Wert am kleinsten, wird er beibehalten. Falls nicht, behält das Programm den ersten gespiegelten Eckpunkt.
Wenn der gespiegelte Eckpunkt hingegen größer als alle bis auf einen ist, führt die Spiegelung nicht hin zur optimalen Lösung. Das heißt, dass man mit dem Simplex bereits sehr nahe an der optimalen Lösung liegt.

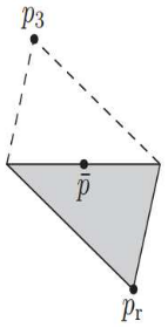


Abb. 3.1: Spiegelung des Punktes p_r im Mittelpunkt $\bar{p}$ der Strecke der beiden anderen Eckpunkte (vgl. Silva et al. 2018).

Die nächste Möglichkeit ist eine Kontraktion. Dabei unterscheidet man wiederum zwei Fälle: eine innere und eine äußere Kontraktion. Für den zweidimensionalen Fall würde dies bedeuten, dass der schlechteste Eckpunkt an dem Mittelpunkt der beiden anderen Punkte herangerückt wird.

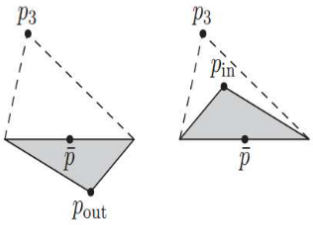


Abb. 3.2: Innere und äußere Kontraktion des Punktes p_3 (vgl. Silva et al. 2018).

Wird durch die Kontraktion auch kein besserer Zielfunktionswert erreicht, bleibt

noch das Verfahren der Komprimierung. Dazu werden jeweils die Mittelpunkte der Eckpunkte mit dem besten Eckpunkt berechnet. Die Seitenlängen des Simplex werden damit halbiert. Wenn die Seitenlängen kürzer als eine geforderte Toleranz sind, wird das Verfahren beendet und die Ecke mit dem kleinsten Zielfunktionswert berechnet.

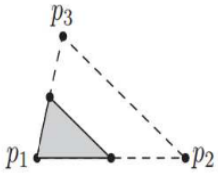


Abb. 3.3: Komprimierung des Dreiecks durch Verschiebung der Eckpunkte p_2 und p_3 in die Mittelpunkte der Strecken $\overline{p_1p_2}$ und $\overline{p_1p_3}$. (vgl. Silva et al. 2018).

3.2 SIR-Modell

Im SIR-Modell unterteilt man die Bevölkerung in drei Klassen: Susceptibles - Infected – Recovered/Removed. Angesteckte Individuen wechseln nach einer Infektion in die Klasse R, in der sie dann verweilen und nicht mehr ansteckbar sind. Daten der AGES zufolge (vgl. AGES 2021a) ist das Risiko, sich nach einer überstandenen Infektion erneut mit dem Virus anzustecken um 91% geringer als bei einer erstmaligen Infektion. Wir gehen aber davon aus, dass nach einer Infektion eine hundertprozentige Immunität erreicht wird.
Trotz dieser Annahme sollte kein wesentlicher Nachteil bei der Modellierung entstehen.

3.2.1 Lösungen des allgemeinen SIR-Modells

Bevor man genauer auf die Parameteranpassung eingeht, betrachtet man zunächst die Lösungskurven des SIR-Modells und plottet diese für $N = 1\ 000$, $S_0 = 999$, $I_0 = 1$, $R_0 = 0$, sowie $\beta = 0{,}3$ und $\gamma = 0{,}04$ und diskutiert, wie sich etwaige Änderungen der Parameter auf die Lösungskurven auswirken.
In den nachfolgenden Grafiken lässt sich erkennen, dass eine Vergrößerung von β einen steileren Anstieg und eine Linksverschiebung bewirkt. Wird γ größer, wird auch die Übergangsrate von I zu R größer und somit die Kurve flacher. Dies ist in der nächsten Abbildung dargestellt.

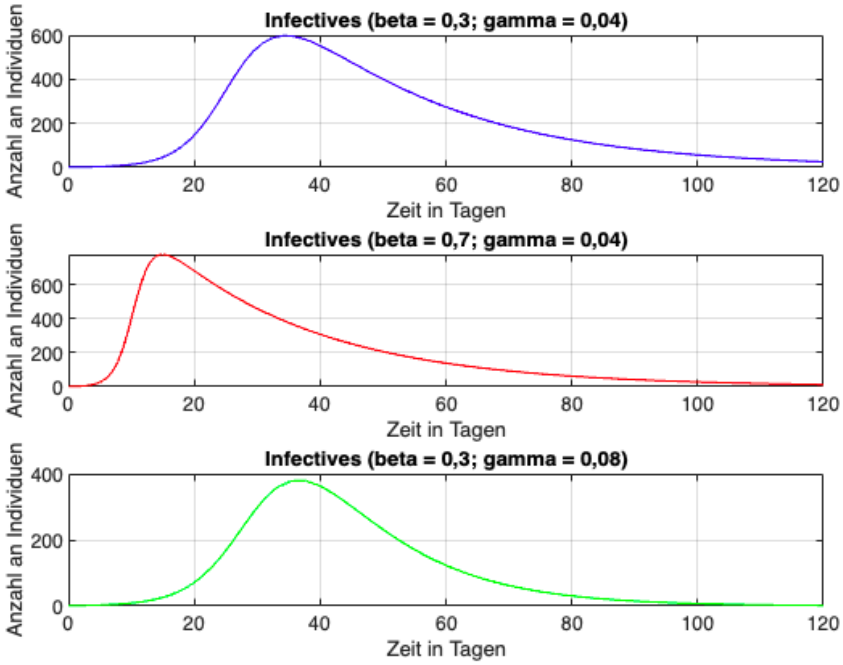


Abb. 3.4: Auswirkung der Parametervariation auf die Lösungskurven für β und γ (eigene Darstellung).

Mit diesen Überlegungen können wir nun versuchen, die Parameter so anzupassen, dass sich die Kurven an die wahren Daten schmiegen.

3.2.2 Anpassung an die realen Daten

Wie in 2.16 hergeleitet, ist die Heilungsdauer exponentialverteilt und die mittlere Infektionsdauer gegeben durch $\frac{1}{\gamma}$. Nachfolgend wird angenommen, dass die mittlere Krankheitsdauer mit SARS-CoV-2 sieben Tage dauert (vgl. RKI 2021), das heißt $\gamma = \frac{1}{7}$.
Nun überlegt man sich noch die Übertragungsrate β. Diese wählen wir derart, dass sie zu den anfänglichen Infektionszahlen für Österreich passt. Aus der durchschnittlichen Reproduktionsrate zwischen dem 01.09.2020 und dem 30.09.2020 lässt sich β durch Umformung der Beziehung $\mathcal{R}_0 = \frac{\beta}{\gamma}$ errechnen. Die durchschnittliche Reproduktionszahl betrug in diesem Zeitraum $\mathcal{R}_0 = 1{,}1978364$ wodurch sich $\beta = 0{,}1711195$ errechnen lässt.
Mit diesen Daten sollen nun die Differentialgleichungen des SIR-Modells in Matlab gelöst werden. Die y-Achse und somit die Anzahl der Individuen wird auf 1 skaliert.

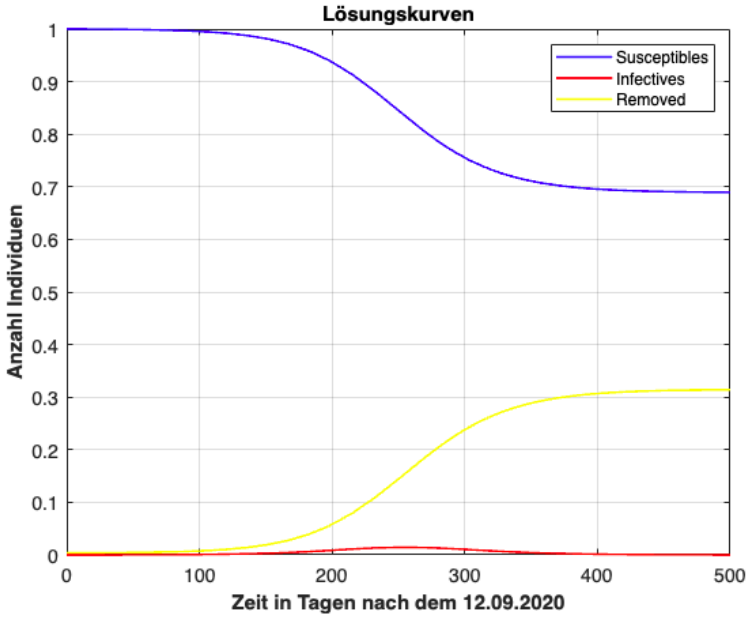


Abb. 3.5: Lösungskurven des SIR-Modells für $S_0 = 8\,901\,064$, $I_0 = 413$, $R_0 = 33\,204$, $\beta = 0{,}1711195$ und $\gamma = \frac{1}{7}$ (eigene Darstellung).

Man erkennt, dass die Infiziertenzahlen sehr langsam ansteigen. Erst nach etwa 270-280 Tagen ist der Peak mit etwa 1,4% gleichzeitig infizierter Bevölkerung erreicht. Das sind etwa 124 000 Personen. Aus den veröffentlichten Daten zur Pandemie weiß man, dass diese Zahl keineswegs richtig ist.
Die abgeschätzten Parameter führen also zu keinem repräsentativen Verlauf der Kurven.

Händisch, geeignete Parameter zu finden, stellt sich als ziemlich mühsam dar. Deshalb wollen wir nun versuchen, mithilfe von Matlab und mit der im Kapitel 3.1.1 beschriebenen Minimum-Suchfunktion, die Kurve an die gegebenen Daten anzupassen. Die verfügbaren Daten werden in Matlab implementiert, der Code für den Minimumsucher geschrieben und Berechnungen mit geeigneten Startbedingungen durchgeführt.
Die Startbedingungen werden händisch so lange angepasst bis sich eine scheinbar optimale Lösungskurve einstellt. Die Größen der Startbedingungen werden aufgrund kleinerer Zahlenwerte und zur anschaulicheren Darstellung normiert. Es wird jeder Wert der Anfangsbedingungen durch 890,1064 dividiert, sodass $S_0 = 10\,000$ und $I_0 = \frac{4}{890{,}1064}$.

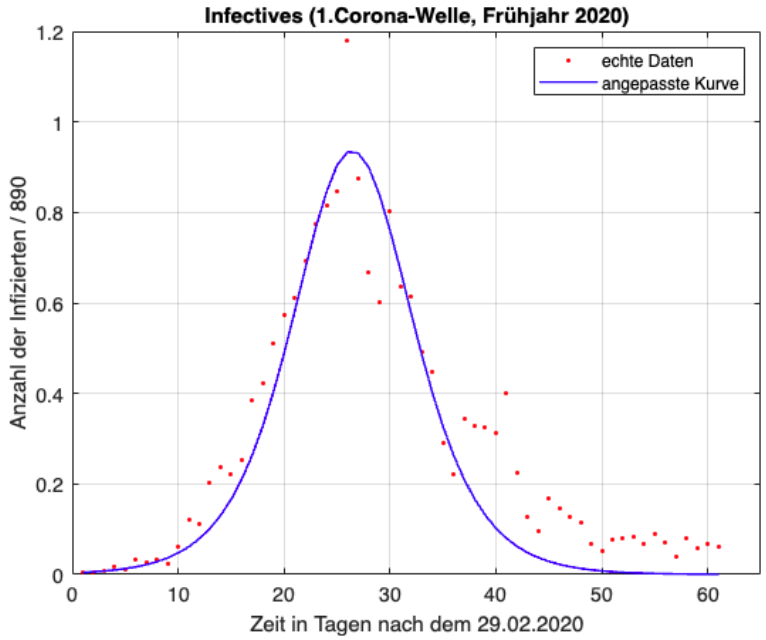


Abb. 3.6: Angepasste Lösungskurve des SIR-Modells für die erste Corona Welle mit Beginn 01.03.2020 (eigene Darstellung).

In der Grafik erkennt man, dass die Kurve die Daten - zumindest in der Anfangsphase - sehr gut beschreibt. Ab Tag Zehn liegen die echten Daten über der Kurve; die Kurve steigt also langsamer als die Neuinfektionen dies tun. Ab etwa dem 20. Tag liegen die Datenpunkte der wirklichen Infektionen bereits unter der Kurve. Der Peak der Kurve unterscheidet sich vom tatsächlichen Peak - also von den maximalen Infektionen der ersten Welle - um etwa 48 Personen. Danach fällt die Kurve schneller als die tatsächlichen Infektionen dies tun.
Im Großen und Ganzen entspricht die Lösungskurve einer relativ guten Näherung.

Das gleiche wollen wir nun für die zweite Corona Welle im Herbst versuchen. Wieder wird versucht, eine angepasste Lösungskurve für die wirklichen Daten mittels Minimumsucher in Matlab zu generieren. Für $\beta = 0{,}02$ und $\gamma = 0{,}15$ ist die Lösungskurve dargestellt. In der nachfolgenden Grafik sieht man bereits deutlich, dass die berechnete Lösungskurve die Daten weniger repräsentativ wiedergibt als beim ersten Mal. Etwa bis Tag 50 stimmen die wahren Daten mit der Kurve gut überein. Lediglich zwischen Tag 50 und Tag 70 übersteigen die echten Daten die geplottete Lösungskurve deutlich. Bis etwa zum 90. Tag passen die Daten wieder gut zur Kurve. Auch durch Variation der Startparameter können keine optimaleren Verläufe gefunden werden. Hier stößt das SIR-Modell in seiner Einfachheit anscheinend an seine Grenzen.

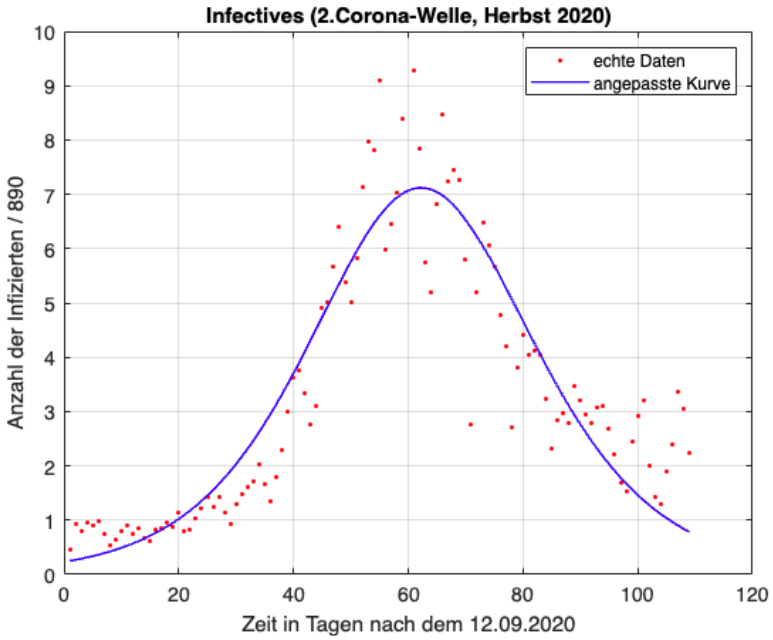


Abb. 3.7: Angepasste Lösungskurve des SIR-Modells für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung).

Bislang versuchten wir, Lösungskurven für gegebene Daten zu finden. Interessant ist am Beginn einer Pandemie aber auch, ob man mit dem gegebenen Modell und wenigen Daten den weiteren Krankheitsverlauf vorhersagen kann.

3.2.3 Vorhersage und Prognose

Wie dramatisch eine Pandemie ausfallen wird, ist wohl eine der wichtigsten Fragen, die man sich zu Beginn einer (unbekannten) Krankheit stellt. Ob sich mit dem gegebenen SIR-Modell mit nur einigen wenigen Anfangsdaten der Verlauf der Corona-Pandemie in Österreich vorhersagen lässt, wird in diesem Kapitel untersucht. Dazu verwenden wir wieder die Minimum-Suchfunktion in Matlab, indem mit den ersten Datenpunkten optimale Parameter gefunden werden sollen. Mit diesen Parametern wird danach die Lösungskurve geplottet und im Anschluss mit den restlichen Daten verglichen. Folgende Grafik zeigt ein Ergebnis für die ersten 15 Daten der ersten Welle vom März 2020.
Aus der Grafik lässt sich erkennen, dass die Lösungskurve zwar die ersten 15 Daten sehr gut wiedergibt, der weitere Verlauf jedoch überhaupt nicht repräsentativ ist.

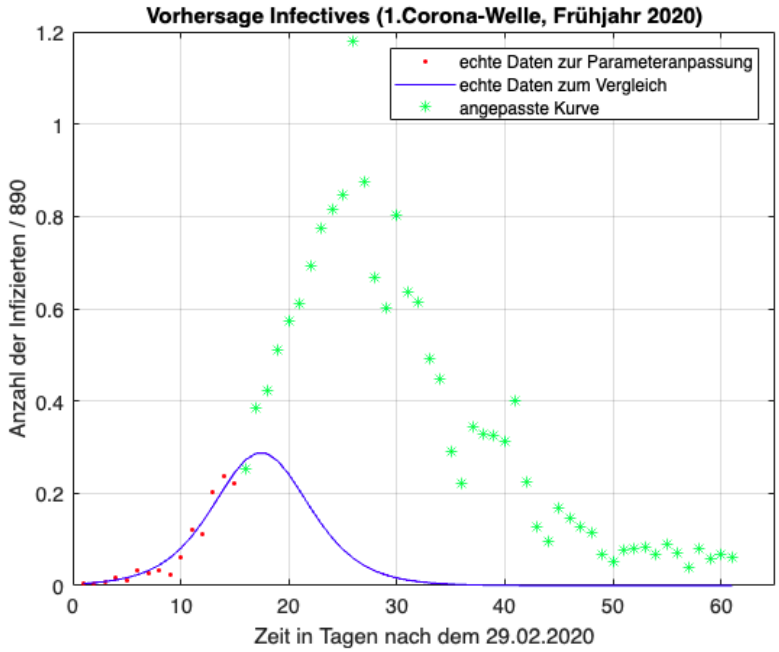


Abb. 3.8: Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung).

Wir fügen noch weitere zehn Daten hinzu und versuchen den Verlauf mithilfe der ersten 25 Daten der ersten Welle vorherzusagen.

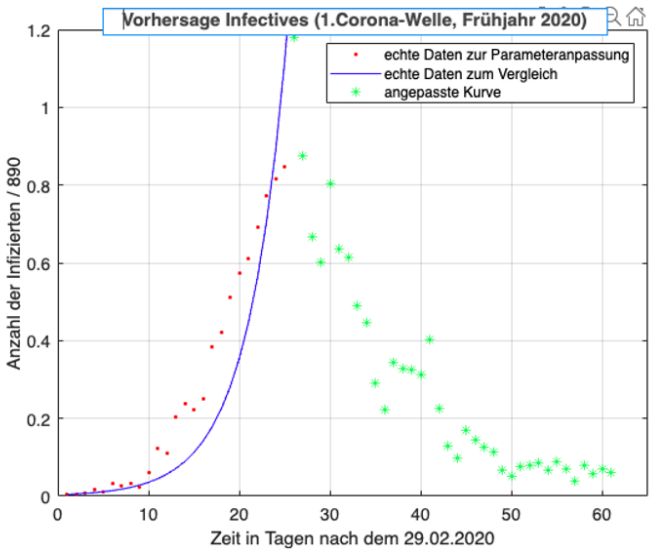


Abb. 3.9: Vorhergesagte Lösungskurve (blau) für 25 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung).

Auch hier sieht man, dass es zunächst zu keinem zufriedenstellenden Ergebnis kommt. Man erkennt, dass nicht einmal die bekannten Daten vom Modell erreicht werden. Nun bleibt nur noch die Möglichkeit, über die Änderung der Startparameter ein

besseres Ergebnis zu erzielen. Ändert man diese zum Beispiel von $\beta = 0,05$ auf $\beta = 0,01$ und von $\gamma = 0,1$ auf $\gamma = 0,2$ ab, stellt sich folgender Kurvenverlauf ein.

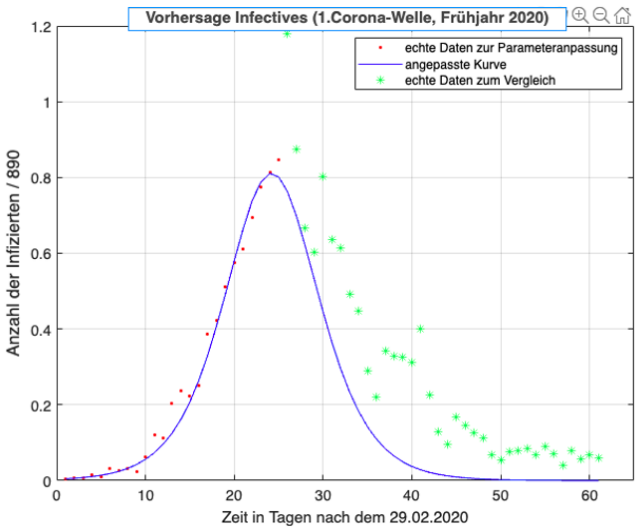


Abb. 3.10: Vorhergesagte Lösungskurve (blau) für 25 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung).

Hierbei wird nun ein deutlich besseres Ergebnis erzielt. Die Daten bis zum Maximum liegen sehr nahe an der Kurve. Diese sind jedoch für die Parameteranpassung gegeben. Jene Daten, die vorhergesagt werden sollen, weichen doch stärker von der geplotteten Lösungskurve ab.
Der Minimumsucher ist nicht im Stande für dieses Modell derartige Prognosen zu liefern. Die Prognose mit dem Modell ist sehr schwierig, wenn kaum Informationen über die Parameterwerte bekannt sind. Dazu müsste man möglicherweise ein (probabilistisches) Modell entwickeln, das unterschiedliche Prognosen aufgrund der vorhandenen Datensätze entwickeln und abschätzen kann. Ein Modell mit lediglich zwei freien Parametern ist zu grob.

3.3 SEIR-Modell mit Populationsdynamik

Im SEIR-Modell wird eine neue Klasse der latent infizierten Individuen eingeführt. Dies hatte den Grund, dass Individuen, die sich mit dem Virus infizierten, nicht sofort ansteckend sind.
In diesem Kapitel betrachten wir lediglich das SEIR-Modell mit Populationsdynamik da das SEIR-Modell ohne Populationsdynamik nur eine Vereinfachung darstellt und von der Grundstruktur sehr ähnlich zum SIR-Modell ist.
Die Überlegungen zum Modell scheinen gut zur tatsächlichen Ausbreitung des

Coronavirus zu passen. Laut Ma et al. 2020, S. 2 beträgt die obere Grenze der mittleren Latenzzeit zwischen 2,52 und 3,95 Tage. Der Parameter ξ beschreibt im Modell die Anzahl der Geburten pro Zeiteinheit, entspricht in Österreich in etwa 234 Geburten pro Tag und bleibt in den Modellen immer konstant; μ beschreibt die Sterberate. In Österreich verstarben im Jahr 2019 im Mittel 229 Menschen pro Tag (vgl. Statistik Austria 2021). Die Sterberate ist sehr schwer für das Modell abzuschätzen da jede Klasse den Parameter μ enthält. Die restlichen Parameter sind analog zum SIR-Modell. Die eingesetzten Parameter werden wieder mit dem Minimumsucher auf passende Werte abgeschätzt. Bevor wir uns aber mit der genauen Parameterabschätzung beschäftigen, gehen wir vorerst wieder auf die allgemeine Lösung des SEIR-Modells mit Populationsdynamik ein.

3.3.1 Lösungen des SEIR-Modells mit Populationsdynamik

Nachfolgend betrachten wir das Lösungsverhalten der Kurven durch verschiedene Parametervariationen. In der ersten Grafik links oben sind die Lösungskurven der Klassen E und I für $N = 1\,000$, $S_0 = 998$, $E_0 = 1$, $I_0 = 1$, $R_0 = 0$, sowie $\alpha = 0{,}03$, $\beta = 0{,}0018$, $\gamma = 0{,}1$, $\delta = 0{,}2$, $\xi = 4$ und $\mu = 0{,}00003$ dargestellt. Nach der Reihe werden einzelne Parameter vergrößert und die Änderungen beobachtet.

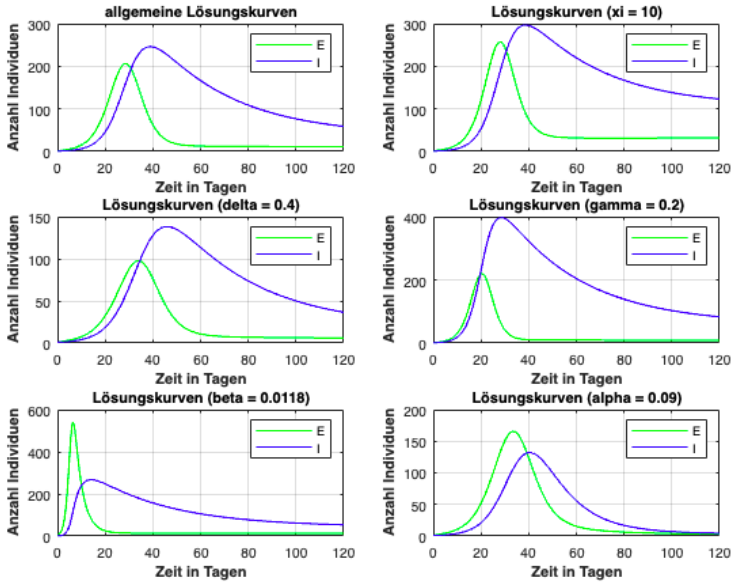


Abb. 3.11: Auswirkung der Parametervariationen auf die Lösungskurven für die Klassen E (grün) und I (blau) mit den Parametern α, β, γ, δ, ξ und μ (eigene Darstellung).

Es sei folgendes für die beiden Lösungskurven E und I festgehalten:

Wird μ größer, flachen die Kurven ab da mehr Individuen aus den Klassen ausscheiden.
Wird ξ größer, flacht zum einen die Kurve von I nach Erreichen des Maximums langsamer ab und zum anderen werden die Werte der Maxima beider Kurven größer da es durch Vergrößerung von ξ mehr Susceptibles gibt, wodurch die Rate βSI vergrößert wird. Dies bewirkt die Erhöhung der Maxima.
Eine Vergrößerung des Parameters δ lässt die Maxima von E und I sinken da eine Vergrößerung des Parameters einen schnelleren Übergang von der Klasse E in die Klasse R zur Folge hat.
Erhöht man γ, führt dies hingegen zu einem Anstieg des Maximums der Kurve von I.
Die Übertragungsrate β hat großen Einfluss auf die Klasse E. Durch Vergrößerung wird die Lösungskurve von E steiler und auch der Wert des Maximums wird größer. Dadurch wird auch der Maximalwert von I erhöht, jedoch mit einer deutlich geringeren Auswirkung als bei E.
Schließlich betrachten wir noch die Auswirkung einer Veränderung von α. Wird α erhöht, wird vor allem der Wert des Maximums von I geringer.
Diese Erkenntnisse können in der Folge wesentlich bei der Parameteranpassung eingesetzt werden.

3.3.2 Anpassung an die realen Daten

Wir wollen nun prüfen, ob sich mit dem SEIR-Modell der wahre Verlauf der zweiten Corona Welle entsprechend repräsentieren lässt. Dazu beachtet man die bekannten Parameter und Anfangsbedingungen und versucht, unbekannte Parameter anzupassen um einen optimalen Verlauf der Lösungskurven zu erhalten.
Wie im vorigen Kapitel erwähnt, beträgt die mittlere Latenzzeit - also diejenige Zeit, die die Individuen in der Klasse E verweilen - etwa 2,52 bis 3,95 Tage. Im Modell nehmen wir dafür als Durchschnitt 4 Tage; es gilt $\frac{1}{\gamma} = 4$. Des weiteren wählen wir für die Sterberate $\mu = 0{,}00002$. Die direkte Übergangsrate der Klasse E zur Klasse R ist wahrscheinlich auch sehr gering und setzen daher $\delta = 0{,}006$. Für die restlichen Parameter gelte anhand von Abschätzungen $\alpha = 0{,}1204$ und $\beta = 0{,}129$ woraus $\mathcal{R}_0 = 1{,}3750$ und sich folgende Lösungskurven erzeugen lassen. Die absoluten Zahlen von S_0, E_0, I_0 und R_0 sind auf relative Anteile zur Gesamtbevölkerung (8 901 064) normiert.

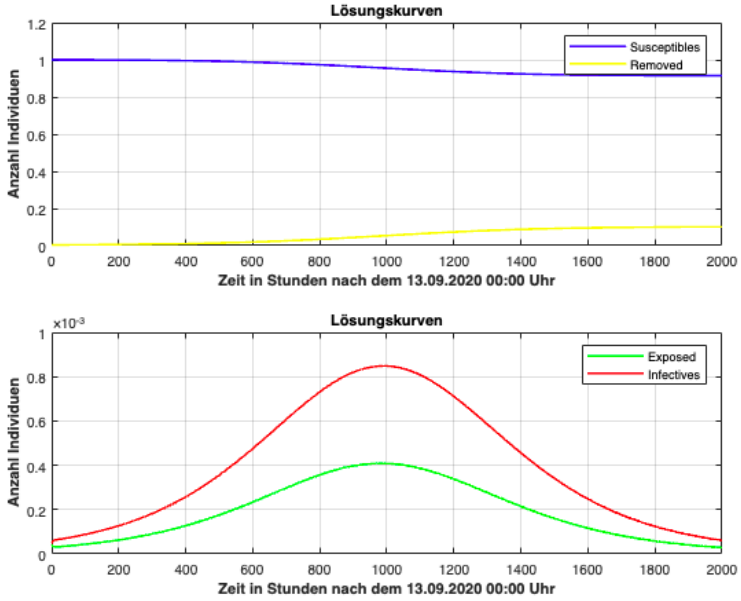


Abb. 3.12: Lösungskurven des SEIR-Modells für $S_0 = 1$, $E_0 = \frac{400}{8\ 901\ 064}$ $I_0 = \frac{413}{8\ 901\ 064}$, $R_0 = \frac{33\ 204}{8\ 901\ 064}$, $\alpha = 0{,}1204$, $\beta = 0{,}129$, $\gamma = \frac{1}{4}$, $\delta = 0{,}006$, $\xi = \frac{234}{8\ 901\ 064}$ und $\mu = 0{,}00002$ (eigene Darstellung).

Man erkennt, dass die Zeitwerte auf der x-Achse sehr groß sind. Eine Skalierung in Tagen macht daher wenig Sinn. Überlegt man sich aber die x-Werte in Stunden anzugeben, erkennt man, dass nach etwa 1 118 Stunden ($\approx$ 47 Tage) das Maximum der Infizierten mit ca. 7 540 Personen erreicht ist. Vergleicht man mit den realen Daten, wird auch dort das Maximum mit Anfang/Mitte November erreicht.
Bis etwa 10. Oktober 2020 streuen die Infektionszahlen nicht weit um die 1 000er-Marke. Erst ab 20. Oktober 2020 steigen die Infektionszahlen rasant an. Laut Modell müssten aber bereits Anfang Oktober zwischen 1 500 und 3 000 Personen infiziert sein. Man sieht, dass das Modell also zu früh ansteigt und das reale Maximum erreicht. Eine Skalierung der x-Achse in Stunden ist aber dennoch keine optimale Lösung. Pro Tag ist nämlich nur ein Datenpunkt bekannt.

Für die obigen Werte der Parameter war $\mathcal{R}_0 > 1$. Setzt man zum Beispiel $\beta = 0{,}1$ und $\mu = 0{,}000029$ ergibt sich für $\mathcal{R}_0 = 0{,}7350$. Wie in Kapitel 2.5 hergeleitet, strebt das Modell für $t \to \infty$ gegen das krankheitsfreie Equilibrium $E_0 = (\frac{\xi}{\mu}, 0, 0, 0)$. Das Langzeitverhalten der Klasse S und R ist für $t \in [0; 300\ 000]$ dargestellt.

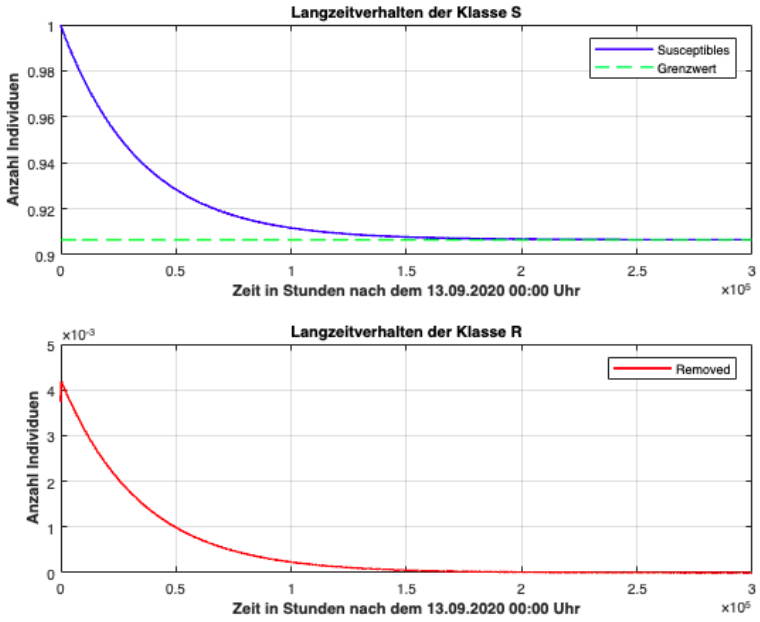


Abb. 3.13: Langzeitverhalten der Klassen S und R für eine Basisreproduktionsrate $\mathcal{R}_0 < 1$ (eigene Darstellung).

Man sieht, dass die Klasse S im Langzeitverhalten gegen den Grenzwert $\frac{\xi}{\mu} = 0{,}9065$ strebt. Für diese Werte von ξ und μ bleiben etwa 90% der Bevölkerung gesund. Die Klasse R strebt im Langzeitverhalten gegen 0 da für $\mathcal{R}_0 < 1$ die Klassen E und I „aushungern“ und in die Klasse R somit irgendwann keine Individuen mehr nachkommen.
Der Fall für $\mathcal{R}_0 < 1$ ist für dieses Kapitel eher uninteressant. Für einen groben Überblick sind die obigen Parameter und die Lösungskurven beider Fälle von $\mathcal{R}_0$ in Ordnung, für die Datenanpassung jedoch zu ungenau. Obige Parameteranpassungen führen zu keinem zufriedenstellenden Ergebnis. Deshalb versuchen wir wieder optimale Lösungen mithilfe des Minimumsuchers in Matlab zu finden.

Zunächst werden wieder die Daten der ersten Corona Welle vom Frühjahr 2020 untersucht. Für die Anfangsbedingungen gelten die normierten Werte für $S_0 = 10\,000$, $E_0 = \frac{1}{890{,}1064}$ und $I_0 = \frac{4}{890{,}1064}$. Der Parameter ξ, der die Anzahl der Geburten pro Tag angibt, bleibt mit $\xi = \frac{234}{890{,}1064}$ unverändert. Die Startbedingungen der übrigen Parameter α, β, γ, δ und μ werden so lange angepasst, bis sich die optimalsten Lösungskurven einstellen und die Summe der Fehlerquadrate am geringsten sind.

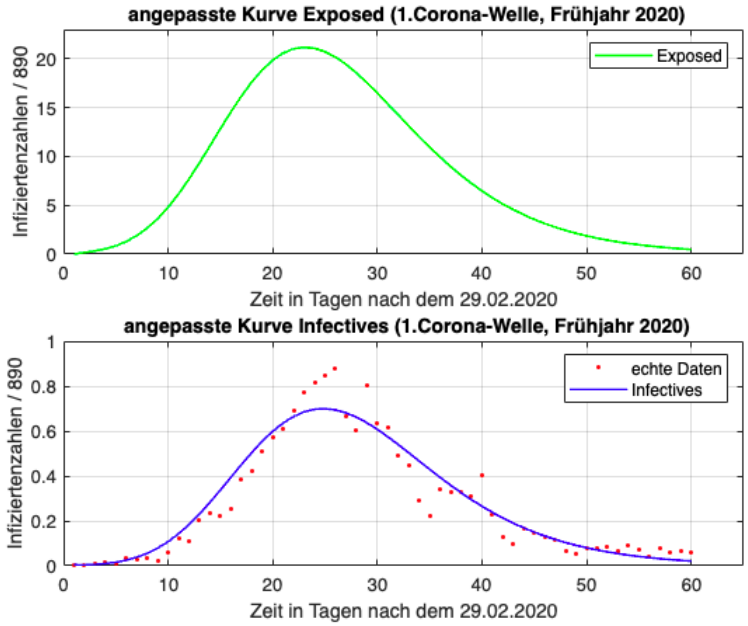


Abb. 3.14: Angepasste Lösungskurven des SEIR-Modells mit Populationsdynamik für die erste Corona Welle mit Beginn 01.03.2020 (eigene Darstellung).

In der zweiten Grafik lässt sich erkennen, dass die Kurve die Daten im Großen und Ganzen gut wiedergibt. Lediglich einige wenige Daten um das Maximum werden nicht erreicht, obwohl gerade dies für eine mögliche Anpassung von großer Bedeutung wäre. Hier ergeben sich Differenzen von etwa 131 Personen zwischen den Werten der Lösungskurve und den wahren Infektionszahlen. Die Grafik der Exposed zeigt die Anzahlen der latent infizierten Bevölkerung, die etwa um das 30-fache größer wäre als die Infiziertenzahlen.
Ein Grund für den Kurvenverlauf der Exposed könnte die Normierung der echten Zahlen sein, die jeweils durch den Faktor 890,1064 dividiert wurden. In der Differentialgleichung der Exposed kommt ein nichtlinearer Term SI vor, der eine derartige Normierung unzulässig machen könnte. Anstatt der Normierung der wahren Zahlen könnte eine Normierung der Parameter ein besseres Ergebnis liefern. Nelder-Mead arbeitet effizienter, wenn die zu optimierenden Parameter allesamt dieselbe Größenordnung haben.
Leider führt auch dieser Versuch zu keinem besseren Ergebnis.

Auch für die zweite Corona Welle im Herbst 2020 sollen wieder optimale Kurven gefunden werden. Dazu werden folgende Anfangsbedingungen verwendet wobei E_0 schätzungsweise in etwa gleich groß gewählt wird wie I_0:
$S_0 = 10\ 000$, $E_0 = \frac{400}{890{,}1064}$, $I_0 = \frac{413}{890{,}1064}$. α, β, γ, δ und μ sollen mittels Minimumsucher gefunden werden und der Parameter $\xi = \frac{234}{890{,}1064}$ bleibt wieder als konstante Geburtenrate gegeben. Es lassen sich folgende Kurven berechnen:

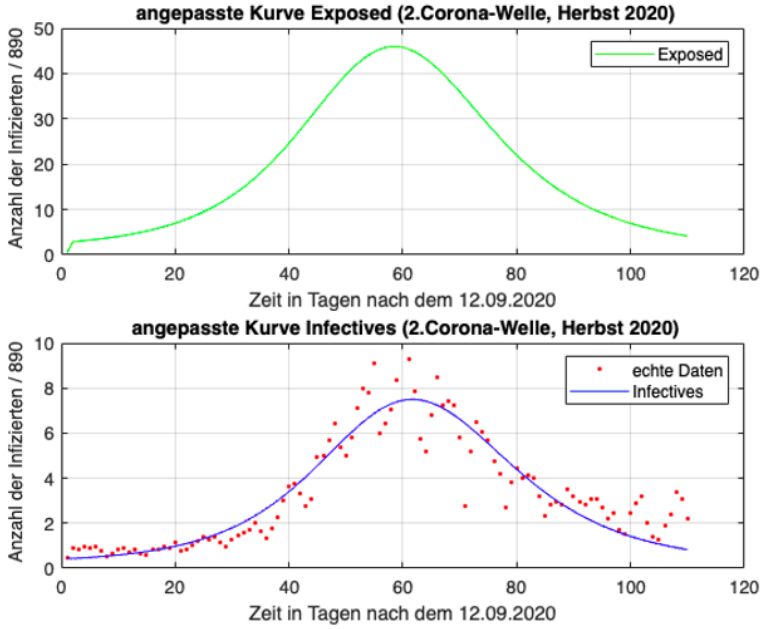


Abb. 3.15: Angepasste Lösungskurven des SEIR-Modells mit Populationsdynamik für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung).

Wie beim SIR-Modell ist es für die zweite Welle wesentlich schwieriger, eine passende Lösungskurve zu finden als für die Erste. Die Datenpunkte streuen nämlich weitaus stärker. Dies hat starken Einfluss auf die Parametersuche und auch die Summe der Fehlerquadrate ist hierbei etwas größer als jene vom Frühjahr 2020.
Das Maximum der Infektionszahlen wird auch von diesem Modell nicht erreicht. Zwischen dem 50. und dem 75. Tag liegen die Datenpunkte der tatsächlichen Infektionszahlen deutlich über der geplotteten Lösungskurve. Davor und danach schmiegt sie sich jedoch sehr passend an.
Die Anzahlen der latent infizierten Bevölkerung ist nun laut Modell etwa um das 5 bis 6-fache größer als jene der primär Infizierten.
Mit diesen Kurven lässt sich also kein optimales Ergebnis darstellen.
Verändert man die Anfangswerte β und δ auf $\beta = 0{,}0259$ und $\delta = 0{,}3$, stellen sich folgende Lösungskurven ein:

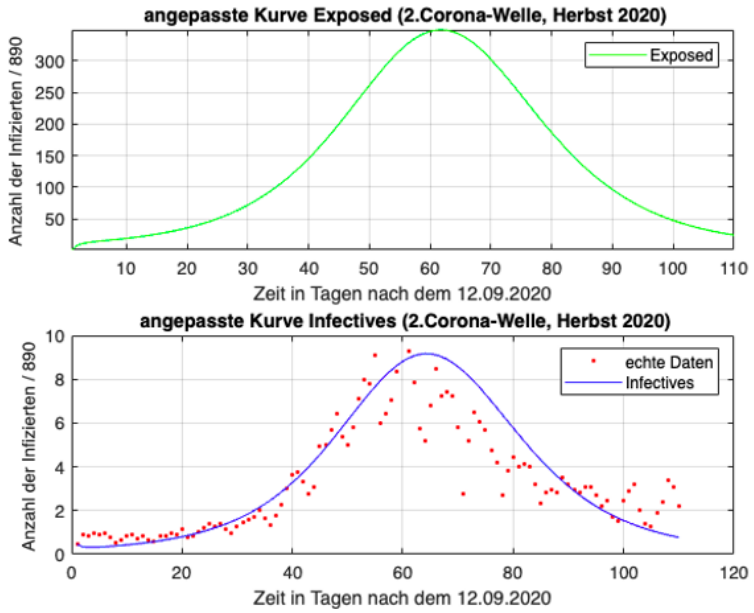


Abb. 3.16: Angepasste Lösungskurven mit minimalen Veränderungen der Startparameter des SEIR-Modells mit Populationsdynamik für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung).

Wie empfindlich die Lösungen auf minimale Parameterveränderungen reagieren, lässt sich mit obiger Abbildung verdeutlichen. Die Kurve der Infectives passt sich dem wahren Verlauf nun deutlich besser an als zuvor. Auch das Maximum wird nun fast optimal von der Kurve erreicht.
Betrachtet man hingegen die Kurve der Exposed, fällt auf, dass diese nun sehr hohe Zahlen wiedergibt. Das Maximum liegt bei etwa 311 000 latent infizierten Personen - eine nicht repräsentative Zahl. Die vom Programm gefundenen Parameterwerte durch den Minimumsucher passen zwar besser zum Kurvenverlauf der Infectives, jene Parameter führen aber zu einem völligen überspitzten Kurvenverlauf der Exposed. Man erhält also wieder nur ein teilweise zufriedenstellendes Ergebnis, das eventuell, wie zuvor erwähnt, durch die Normierung der Zahlen entstehen könnte.

3.3.3 Vorhersage und Prognose

Eine mögliche Vorhersage der Infektionszahlen soll nun ebenfalls mit dem SEIR-Modell modelliert werden. Wichtig bei einer Vorhersage ist, dass das Maximum der Infiziertenzahlen möglichst genau abgeschätzt werden kann. Um dies zu ermöglichen werden in der Regel verschiedene Szenarien dargestellt, denen unterschiedliche Parameterwerte zugrunde liegen. Diese Szenarien versucht man beispielsweise anhand

von Verhaltensmaßnahmen zu erklären (z.B. könnte der Peak der Kurve durch dreiwöchiges social distancing verkleinert werden).
Wir wollen testen, ob man mit wenigen Anfangsdaten den weiteren, wahren und für uns bekannten Verlauf der Infektionszahlen vorhersagen kann.
Dafür nehmen wir wiederum die ersten 15 Daten der Infiziertenzahlen der ersten Welle im Frühjahr 2020. Mithilfe der Minimum-Suchfunktion wird versucht, eine passende Kurve für die ersten 15 Daten zu finden und deren weiteren Verlauf mit den restlichen Daten zu vergleichen. Im ersten Versuch ergibt sich eine Lösungskurve, die die Daten bis zum 26. Tag sehr genau wiedergibt. Es werden also bereits elf Datensätze nach den ersten 15 Anfangsdaten sehr genau vorhergesagt. Anzumerken ist, dass der Datensatz vom 27. Tag mit 1 050 Neuinfektionen hierbei ausgelassen wird da sich dieser deutlich von den anderen Daten unterscheidet und als eventueller Ausreißer gewertet werden kann.
Die Lösungskurve steigt hingegen stärker an als die wahren Daten dies tun. So sieht man, dass sich das vorhergesagte Maximum deutlich vom tatsächlichen Maximum unterscheidet. Außerdem erkennt man, dass die Lösungskurve deutlich zu breit ist.

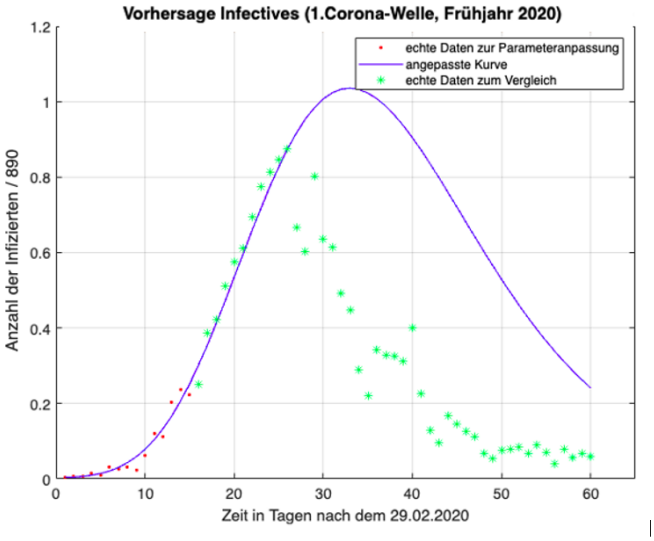


Abb. 3.17: Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung).

Verändert man wiederum die Startparameter, findet das Programm einen neuen Kurvenverlauf der wesentlich genauer ist als jener im Vergleich zur obigen Kurve.
Obwohl die Daten (besonders nach Erreichen des Maximums) von der Lösungskurve abweichen, konnten diese doch relativ genau prognostiziert werden. Man muss auf alle Fälle bedenken, dass zu anfangs lediglich 15 Daten bekannt waren.

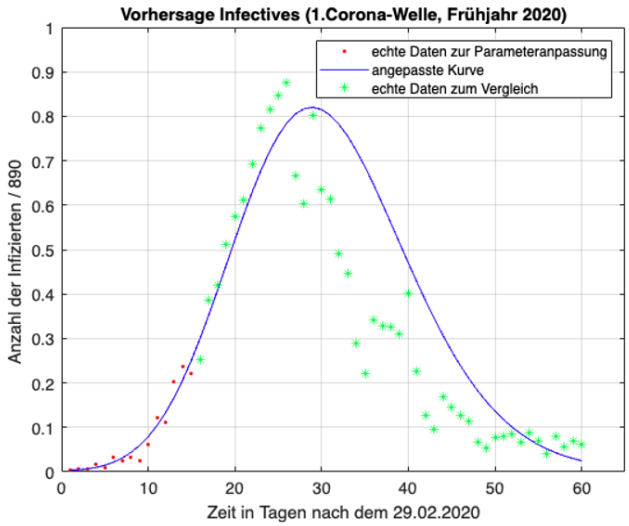


Abb. 3.18: Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung).

3.4 SEIQR-Modell mit Populationsdynamik

Personen, die sich mit dem Coronavirus infiziert haben, müssen einer behördlich angeordneten Quarantänepflicht nachkommen. Das bedeutet laut Epidemiegesetz, dass sich jene Personen über einen bestimmten Zeitraum sozial einzuschränken haben. All jene Personen, die mit Coronainfizierten Kontakt gehabt haben, müssen ebenfalls einer befristeten Absonderung nachkommen um durch einen möglichen Krankheitsausbruch keine weiteren Personen zu infizieren. Dieses Verfahren wird als Contact Tracing bezeichnet (vgl. Bioethikkommission 2020).
Das SEIQR-Modell ist das einzige Modell dieser Masterarbeit mit Quarantäne. Allerdings sei erwähnt, dass nur Personen, die tatsächlich infiziert wurden in die Gruppe Q der quarantined wechseln können. Dieses Modell schließt die Bedingung des Contact Tracings aus.

3.4.1 Lösungen des allgemeinen SEIQR-Modells

Zunächst sollen wieder die allgemeinen Lösungen des SEIQR-Modells diskutiert werden. Dafür betrachten wir eine Gesamtbevölkerung von $N = 1\,000$. Als Anfangsbedingungen werden $S_0 = 998$, $E_0 = 1$, $I_0 = 1$, $Q_0 = 0$, $R_0 = 0$ gesetzt. Für die verwendeten Parameter gilt $\alpha_1 = 0{,}12$, $\alpha_2 = 0{,}001$, $\beta = 0{,}002$, $\gamma = 0{,}08$, $\delta = 0{,}35$, $\kappa = 0{,}009$, $\sigma = 0{,}0005$, $\mu = 0{,}01$ und $\xi = 5$.
Aus der Grafik lässt sich erkennen, dass die Lösungskurven mehrere Wellen aufweisen.

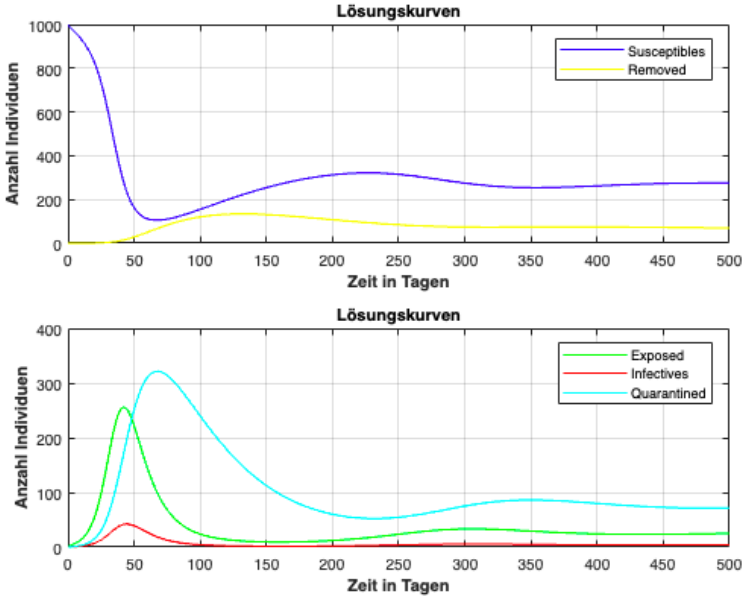


Abb. 3.19: Allgemeine Lösungskurven des SEIQR-Modells (eigene Darstellung).

Da in dieser Arbeit die Wellen vom Frühjahr und vom Herbst getrennt betrachtet werden, wird darauf aber nicht näher eingegangen.
Nun betrachten wir noch die Lösungskurven für $\mathcal{R}_0 > 1$ im Langzeitverhalten. Für die obigen Parameter streben die Lösungskurven gegen das endemische Gleichgewicht $G_e = (270{,}28;\ 25{,}52;\ 4{,}25;\ 74{,}37;\ 61{,}14)$. Nachfolgend sind die Lösungskurven im Intervall $[0; 800]$ dargestellt.

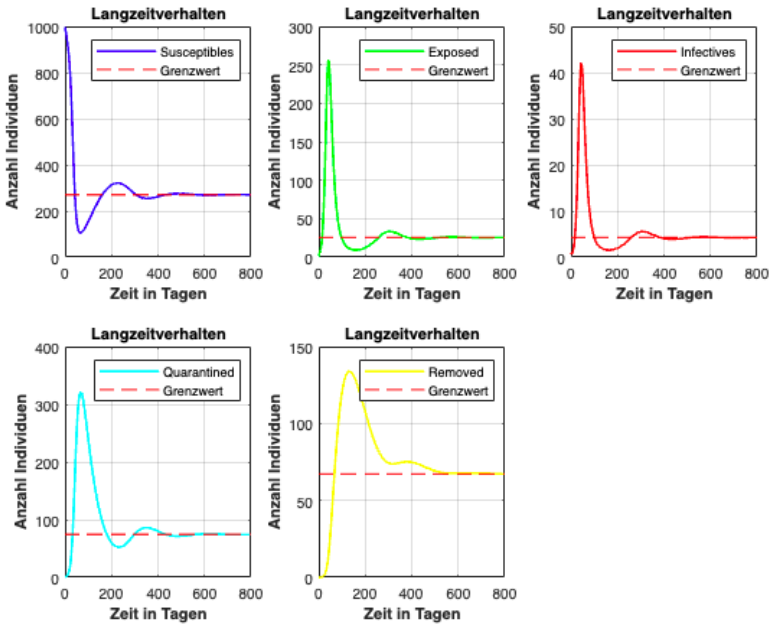


Abb. 3.20: Langzeitverhalten des SEIQR-Modells für $\mathcal{R}_0 > 1$ (eigene Darstellung).

3.4.2 Anpassung an die realen Daten

Die Parameter händisch an die wahren Daten anzupassen wird bei diesem Modell nicht mehr gezeigt. Es soll nun gleich versucht werden, mit dem Minimumsucher die Kurven an die Daten anzupassen. Getrennt betrachten wir wieder die erste Welle vom Frühjahr 2020 von der zweiten Welle vom Herbst 2020. Anstatt der latent infizierten Bevölkerung wird nun die Lösungskurve der in Quarantäne befindlichen Bevölkerung geplottet. Mit dem Minimumsucher sollen optimale Parameterwerte für α_1, α_2, β, γ, δ, κ, σ und μ gefunden werden. Der Parameter $\xi = \frac{234}{890{,}1064}$ bleibt konstant.

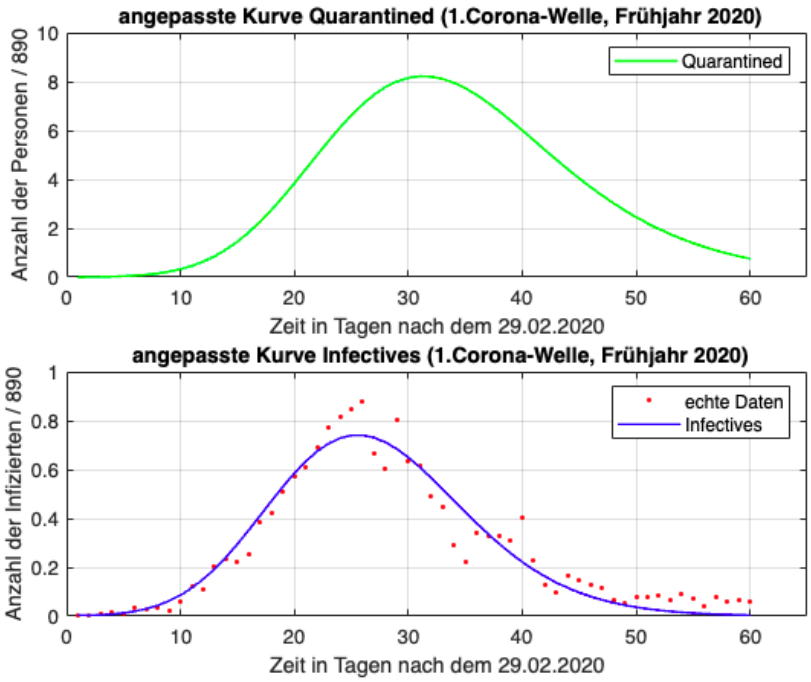


Abb. 3.21: Allgemeine Lösungskurven des SEIQR-Modells (eigene Darstellung).

Die Lösungskurve der infizierten Bevölkerung ist sehr ähnlich jener vom SEIR-Modell. Man beobachtet, dass sich die Kurve relativ genau an die Daten schmiegt. Lediglich die Spitzenwerte der Infektionen zwischen dem 20. und 30. Tag werden von der Kurve nicht erreicht - auch die Veränderungen der Startparameter führen zu keinem besseren Ergebnis.
Daten zu den Quarantänezahlen werden auf der Homepage der Ages nicht angeführt. Die Kurve der quarantined zeigt für dieses Modell, dass sich im Schnitt etwa zehn Mal so viele Individuen in Quarantäne befinden als infiziert sind. Außerdem ist die Kurve der quarantined im Vergleich zur Kurve der infectives zeitlich nach rechts verschoben. Dies ist durchaus berechtigt da Individuen erst nach einer Infektion in die Klasse Q wechseln.

Für die zweite Welle im Herbst ergeben sich mit nach der Parameteroptimierung die beiden folgenden Lösungskurven wobei für $Q_0 = \frac{400}{890{,}1064}$ angenommen wird.

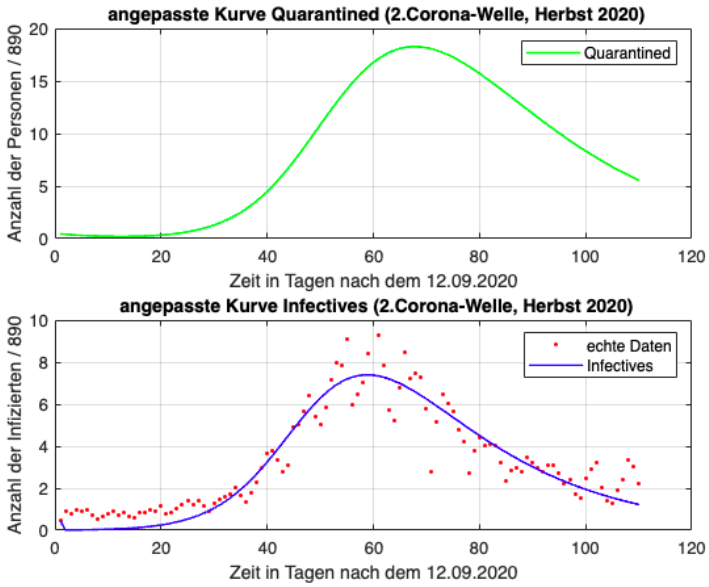


Abb. 3.22: Angepasste Lösungskurven des SEIQR-Modells für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung).

Wieder erkennt man, dass sich für die Lösungskurve der infizierten Bevölkerung kein optimaler Verlauf einstellt. Die Maximalwerte der echten Daten weichen weit von der Lösungskurve ab, sodass sich maximale Abweichungen von etwa 1 600 Personen ergeben. Vor und nach dem Maximum der Kurve passt sie sich jedoch gut an die Daten an. Für die Kurve der quarantined kann ähnliches zur Kurve im Frühjahr beobachtet werden. Die Anzahlen der Personen in Quarantäne sind nun aber nur mehr etwa das Zwei- bis Dreifache der Anzahlen der infizierten Bevölkerung.
Die Annahmen und die Belegung der im Modell vorkommenden Parameter reichen anscheinend nicht aus um optimale Lösungskurven für die Infektionszahlen der zweiten Corona Welle vom Herbst 2020 zu modellieren.

3.4.3 Vorhersage und Prognose

Auch wenn mit dem Modell kein optimaler Kurvenverlauf für die Infektionszahlen vom Frühjahr berechnet werden konnte, ist es dennoch interessant zu wissen, ob wenigstens ein ähnlich guter Verlauf vorhergesagt werden kann. In erster Linie ist erstrebenswert, dass die Kurve das Maximum möglichst genau erreichen kann.
Die Vorhersage basiert wieder auf den 15 Anfangsdaten der Infiziertenzahlen, die in

der nachfolgenden Grafik rot dargestellt sind. Zum Vergleich scheinen die restlichen Daten in der Grafik grün auf. Der Datensatz vom 27.03.2020 mit 1 050 Neuinfektionen hebt sich deutlich von den restlichen Infiziertenzahlen ab und wird daher für die Prognose wieder ausgelassen. Mithilfe des Minimumsuchers in Matlab werden mithilfe der wenigen Anfangsdaten optimale Parameterwerte gefunden mit denen die Differentialgleichungen des Modells gelöst und schließlich die Lösungskurve der Infizierten geplottet wird. Ähnlich zur Vorhersage des SEIR-Modells (3.18) erhält man folgende Kurve:

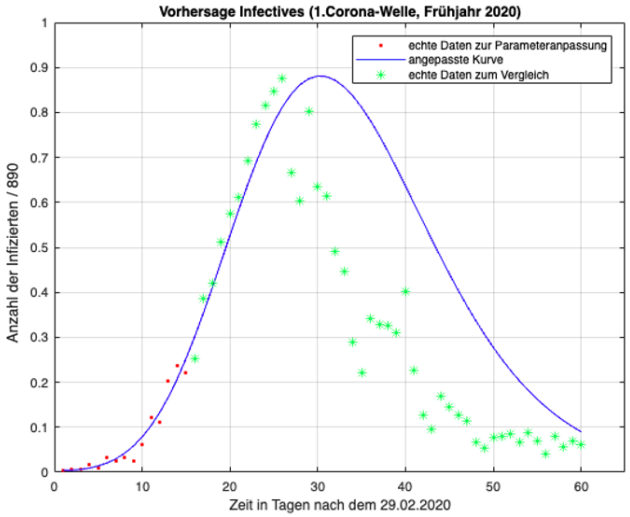


Abb. 3.23: Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung).

Man erkennt, dass die maximalen Infiziertenzahlen vom Modell relativ gut erreicht werden. Die Infiziertenzahlen bis einschließlich des 26. Tages liegen sehr nahe an der Lösungskurve. Danach fallen die wahren Infektionszahlen sehr schnell ab, die Kurve hingegen weniger schnell, sodass die Daten nach Erreichen des Maximums weit von der Kurve abweichen.
Ein Versuch, den „zweiten Teil“ der Kurve besser an die Daten anzupassen, hätte zur Folge, dass das Maximum nicht mehr erreicht wird. Ein optimalerer Kurvenverlauf als auf der dargestellten Grafik kann mit diesem Modell nicht erreicht werden.
Es ist außerdem zu erwähnen, dass die Lösungskurve sehr empfindlich gegenüber Veränderungen der Startparameter reagiert. Ändert man die Startparameter, könnte sich ein völlig anderer Kurvenverlauf einstellen. Würden also lediglich die ersten 15 Daten bekannt sein, müsste man - wie auch bei den anderen Modellen - mehrere Kurven erstellen um mögliche Verläufe prognostizieren zu können.

Kapitel 4

Zusammenfassung und Fazit

Ziel der vorliegenden Masterarbeit ist es, zu untersuchen, ob es möglich ist, die Infektionszahlen der Corona Pandemie für Österreich mit theoretischen Epidemiemodellen anzupassen und vorherzusagen. Aus diesem Grund werden die Modelle und deren Systemgrößen zunächst theoretisch aufgearbeitet. Dafür wurde eine systematische Literaturrecherche durchgeführt und die Aufarbeitung der jeweils gleichen Modelle aus unterschiedlichen Veröffentlichungen verglichen.
Im Wesentlichen sind die Modelle im Grundaufbau gleich. Beginnend mit dem SI-Modell, das das einfachste aller Modelle darstellt, werden schrittweise mehr und mehr Annahmen getroffen, die mit den Modellen modelliert werden können. Schließlich gelingt es mit dem SEIQR-Modell sogar latent infizierte und in Quarantäne befindliche Individuen zu modellieren.
Der Schwerpunkt der Arbeit liegt im praktischen Einsatz der Modelle. So soll versucht werden, mithilfe geeigneter Optimierungsverfahren von Matlab, den Verlauf der Infektionszahlen an die realen Daten anzupassen und im weiteren sogar zu prognostizieren.

Zusammenfassend lässt sich sagen, dass mit keinem der Modelle ein zufriedenstellendes Ergebnis erreicht werden konnte. Die Lösungskurven der Modelle spiegeln die Daten zwar überwiegend gut wieder, weichen aber oftmals rund um das Maximum zu stark von den wahren Daten ab. Es sei festgehalten, dass sich der Verlauf der ersten Welle vom Frühjahr 2020 wesentlich besser repräsentieren lässt als jener der zweiten Welle vom Herbst 2020. Dies hat den Grund, dass die Daten von der zweiten Corona Welle wesentlich mehr streuen als jene der ersten. Die Infektionszahlen konnten vor allem mit dem SEIR-Modell gut prognostiert werden. Das Problem bei derartigen Prognosen ist aber, dass diese Modelle nur einen einzigen Verlauf zeigen. Hätte man - wie in der Realität - keinerlei Daten der Infektionszahlen verfügbar, würde man nicht wissen, ob der Verlauf der Prognose tatsächlich zutrifft. Mit unterschiedlichen

Startwerten für die zu optimierenden Parameter ergeben sich nämlich unterschiedliche Verläufe. Für eine tatsächliche Prognose ist dieses Verfahren also ungeeignet.

Infektionskrankheiten sind ein hochkomplexes System, das unter anderem sehr vielen sozialen, geographischen und demographischen Aspekten unterliegt. In dieser Arbeit wurden diese Gesichtspunkte bei der Modellierung ausgeblendet. Weiterführende Forschung könnte diese Modelle unter Einbeziehung sozialer, geographischer und demographischer Gegebenheiten weiterentwickeln, um die Modellierung von Infektionskrankheiten aus einer neuen Perspektive zu betrachten.

Literaturverzeichnis

[1] AGES. *Coronavirus*. Eingesehen am 26.03.2021. 2021. URL: https://www.ages.at/themen/krankheitserreger/coronavirus/.

[2] AGES. *Epidemiologische Parameter des COVID19 Ausbruchs*. Eingesehen am 26.03.2021. 2021. URL: https://www.ages.at/en/wissen-aktuell/publikationen/epidemiologische-parameter-des-covid19-ausbruchs-oesterreich-20202021/.

[3] H. Benker. »Optimierung mit MATLAB«. In: *Ingenieurmathematik kompakt – Problemlösungen mit MATLAB: Einstieg und Nachschlagewerk für Ingenieure und Naturwissenschaftler*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, S. 213–226.

[4] Bioethikkommission. *Contact Tracing in der COVID-19 Pandemie*. Techn. Ber. Geschäftsstelle der Bioethikkommission, 2020.

[5] R. Bonita, R. Beaglehole und T. Kjellström. *Einführung in die Epidemiologie*. 2. Aufl. Verlag Hans Huber Bern, 2008.

[6] F. Brauer und C. Castillo-Chavez. *Mathematical Models in Population Biology and Epidemiology*. 2. Aufl. Texts in Applied Mathematics 40. Spektrum Akademischer Verlag, Heidelberg, 2012.

[7] R. Bürger. *Einführung in die Mathematische Biologie*. Institut für Angewandte Analysis und Numerische Simulation - Universität Stuttgart, 2004.

[8] David Scherfgen IT Services. *Integralrechner - Integrale online berechnen — mit Rechenweg und Graphen!, note = Eingesehen am 08.07.2021*. 2021. URL: https://www.integralrechner.de.

[9] *DWDS – Digitales Wörterbuch der deutschen Sprache*. Eingesehen am 19.02.2021. URL: https://www.dwds.de/.

[10] G. Fischer. *Lineare Algebra*. 16. Aufl. Vieweg+Teubner Verlag, Wiesbaden, 2008.

[11] F. Hausser und Y. Luchko. *Mathematische Modellierung mit MATLAB: Eine praxisorientierte Einführung*. Spektrum Akademischer Verlag, Heidelberg, 2011.

[12] HealthKnowledge. *Epidemic theory (effective & basic reproduction numbers, epidemic thresholds) & techniques for analysis of infectious disease data (construction & use of epidemic curves, generation numbers, exceptional reporting & identification of significant clusters)*. Eingesehen am 24.03.2021. 2018. URL: https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/epidemic-theory.

[13] A. Jakob. *Mathematische Modellierung von Epidemiemodellen*. Saarbrücken: Akademiker Verlag, 2013.

[14] F. Jarre und J. Stoer. *Optimierung*. 2. Aufl. Springer Spektrum, Berlin, 2019.

[15] Y. S. Kim. *The physics of phase space*. Berlin: Springer Verlag, 1987.

[16] C. Kohlmeier. »Einführung in die Mathematische Modellierung«. Vorlesungsskript. Carl von Ossietzky Universität Oldenburg, 2006.

[17] L. Kreienbrock, I. Pigeot und W. Ahrens. *Epidemiologische Methoden*. 5. Aufl. Springer-Verlag Berlin Heidelberg, 2012.

[18] G. A. Lienert und U. Raatz. *Testaufbau und Testanalyse*. 6. Aufl. Weinheim Basel: Beltz Verlag, 2011.

[19] S. Ma, J. Zhang, M. Zeng, Q. Yun, W. Guo, Y. Zheng, S. Zhao, M. H. Wang und Z. Yang. »Epidemiological parameters of coronavirus disease 2019: a pooled analysis of publicly reported individual data of 1155 cases from seven countries«. In: *medRxiv* (2020). DOI: 10.1101/2020.03.21.20040329.

[20] R. S. Mathea. »Angewandte Mathematik und Optimierung Schriftenreihe Applied Mathematics and Optimization Series«. Studienarbeit. Helmut-Schmidt-Universität / Universität der Bundeswehr Hamburg, 2015.

[21] *MathWorks*. Eingesehen am 17.04.2021. URL: https://de.mathworks.com/?s_tid=gn_logo.

[22] Medizinische Universität Wien. *Körpereigenes Enzym als Ansatzpunkt für COVID-19-Therapien*. Eingesehen am 26.04.2021. 2021. URL: https://www.meduniwien.ac.at/web/ueber-uns/news/news-im-maerz-2021/koerpereigenes-enzym-als-ansatzpunkt-fuer-covid-19-therapien/.

[23] B. Mishra und A. Singh. »Two Quarantine Models on the Attack of Malicious Objects in Computer Network«. In: *Mathematical Problems in Engineering* 2012 (Aug. 2011), S. 1–14.

[24] J. D. Murray. *Mathematical Biology*. 3. Aufl. Springer Verlag, Berlin Heidelberg, 2002.

[25] W. Nöbauer und W. Timischl. *Mathematische Modelle in der Biologie*. Heidelberg / Wiesbaden: Friedr. Vieweg & Sohn Verlag, 1979.

[26] P. Pedamkar. *fminsearch in Matlab*. Eingesehen am 02.05.2021. 2021. URL: https://www.educba.com/fminsearch-in-matlab/.

[27] S. Pilz, A. Chakeri, J. Ioannidis, V. Richter L. Theiler-Schwetz, C. Trummer, R. Krause und F. Allerberger. »SARS-CoV-2 re-infection risk in Austria«. In: *European Journal of Clinical investigation* 51 (Feb. 2021).

[28] R. Rannacher. *Numerik 1 - Numerik gewöhnlicher Differentialgleichungen*. Heidelberg University Publishing, Heidelberg, 2017.

[29] O. Razum und K. Bozorgmehr. »Globalisierung und Gesundheit«. In: *Soziologie von Gesundheit und Krankheit*. Hrsg. von M. Richter und K. Hurrelmann. Wiesbaden: Springer Fachmedien Wiesbaden, 2016, S. 107–119.

[30] RKI. *Epidemiologischer Steckbrief zu SARS-CoV-2 und COVID-19*. Eingesehen am 26.04.2021. 2021. URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html.

[31] O. Roll. »Das qualitative Forschungsmodell — Grundlagen und Forschungsdesign«. In: *Internetnutzung aus Konsumentensicht* (2003), S. 107–141.

[32] SALZBURG24. *Vom ersten Corona-Fall zum dritten Lockdown*. Eingesehen am 26.04.2021. 2021. URL: https://www.salzburg24.at/news/oesterreich/chronologie-vom-ersten-corona-fall-zum-dritten-lockdown-99952177.

[33] D. Scholz. *Optimierung interaktiv*. Berlin: Springer Spektrum, 2018.

[34] P. Schuster. »Ebola-challenge and revival of theoretical epidemiology: Why Extrapolations from early phases of epidemics are problematic«. In: *Complexity* 20(5) (2015). https://doi.org/10.1002/cplx.21694, S. 7–12.

[35] K. Gonçalves-e Silva, D. Aloise, S. Xavier-De-Souza und N. Munster Mladenović. »Less is more: simplified Nelder-Mead method for large unconstrained optimization«. In: *Yugoslav Journal of Operations Research* 28.2 (2018), S. 153–169.

[36] Statistik Austria. *Geborene*. Eingesehen am 30.03.2021. 2021. URL: https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bevoelkerung/geborene/index.html.

[37] L. Sturm. *Rechnerunterstützte Auswertung von Diodenkennlinien*. Projektarbeit. Friedrich-Alexander Universität Erlangen-Nürnberg, 2010.

[38] K. Taschwer. *Warum die zweite Welle anders ist*. Eingesehen am 06.05.2021. 2021. URL: https://www.derstandard.at/story/2000120952925/warum-die-zweite-welle-anders-ist.

[39] G. Teschl und S. Teschl. *Mathematik für Informatiker*. 3. Aufl. Texts in Applied Mathematics 40. Springer Verlag, Berlin Heidelberg, 2014.

[40] B. L. Tesini. *Coronaviren und akute respiratorische Syndrome (MERS und SARS)*. Techn. Ber. University of Rochester School of Medicine und Dentistry, 2021.

[41] The MathWorks, Inc. *fminsearch*. Eingesehen am 02.05.2021. 2005. URL: http://matlab.izmiran.ru/help/techdoc/ref/fminsearch.html.

[42] R. Voßwinkel. *Systematische Analyse und Entwurf von Regelungseinrichtungen auf Basis von Lyapunov's direkter Methode*. Springer Verlag, Wiesbaden, 2019.

[43] WHO. *Pandemie der Coronavirus-Krankheit (COVID-19)*. Eingesehen am 26.04.2021. 2021. URL: https://www.euro.who.int/de/health-topics/health-emergencies/coronavirus-covid-19/novel-coronavirus-2019-ncov.

[44] R. Wiesendanger. *Studie zum Ursprung der Coronavirus-Pandemie*. Studie. Universität Hamburg, 2021.

[45] E. de Wit, N. van Doremalen, D. Falzarano und V. J. Munster. »SARS and MERS: recent insights into emerging coronaviruses«. In: *Nature reviews. Microbiology* 14.8 (2016), S. 523–534.

[46] H. M. Youssef, N. A. Alghamdi, M. A. Ezzat, A. A. El-Bary und A. M. Shawky. »A modified SEIR model applied to the data of COVID-19 spread in Saudi Arabia«. In: *AIP Advances 10, 125210* 10.12 (2020). https://doi.org/10.1063/5.0029698.

[47] R. Zacher, R. Schnaubelt und J. Prüss. *Mathematische Modelle in der Biologie. Deterministische homogene Systeme*. Birkhäuser Verlag, 2008.

[48] I. Zsifkovits. »Aktuelle Epidemienmodelle Eine qualitative Evaluierung von Modellen zur Vorhersage und Berechnung des Übertragungsrisikos des Ebola-Virus«. Masterarbeit. Karl-Franzens-Universität Graz, 2016.

weitere Verzeichnisse

Abbildungsverzeichnis

1.1. Schematische Darstellung des Modellierungszyklus bestehend aus einer iterativen Abfolge von Aktivitäten, die zu bestimmten Ereignissen führen (vgl. Hausser et al. 2011, S. 53). 5

2.1. Trajektorien für das SIR-Modell mit den Anfangsbedingungen S_0, I_0 und $N = 200$ (eigene Darstellung). 31

3.1. Spiegelung des Punktes p_r im Mittelpunkt $\overline{p}$ der Strecke der beiden anderen Eckpunkte (vgl. Silva et al. 2018). 51
3.2. Innere und äußere Kontraktion des Punktes p_3 (vgl. Silva et al. 2018). 51
3.3. Komprimierung des Dreiecks durch Verschiebung der Eckpunkte p_2 und p_3 in die Mittelpunkte der Strecken $\overline{p_1p_2}$ und $\overline{p_1p_3}$. (vgl. Silva et al. 2018). 52
3.4. Auswirkung der Parametervariation auf die Lösungskurven für β und γ (eigene Darstellung). 53
3.5. Lösungskurven des SIR-Modells für $S_0 = 8\ 901\ 064$, $I_0 = 413$, $R_0 = 33\ 204$, $\beta = 0{,}1711195$ und $\gamma = \frac{1}{7}$ (eigene Darstellung). 54
3.6. Angepasste Lösungskurve des SIR-Modells für die erste Corona Welle mit Beginn 01.03.2020 (eigene Darstellung). 55
3.7. Angepasste Lösungskurve des SIR-Modells für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung). 56

3.8. Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung). 57
3.9. Vorhergesagte Lösungskurve (blau) für 25 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung). 57
3.10. Vorhergesagte Lösungskurve (blau) für 25 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung). 58
3.11. Auswirkung der Parametervariationen auf die Lösungskurven für die Klassen E (grün) und I (blau) mit den Parametern α, β, γ, δ, ξ und μ (eigene Darstellung). 59
3.12. Lösungskurven des SEIR-Modells für $S_0 = 1$, $E_0 = \frac{400}{8\ 901\ 064}$ $I_0 = \frac{413}{8\ 901\ 064}$, $R_0 = \frac{33\ 204}{8\ 901\ 064}$, $\alpha = 0{,}1204$, $\beta = 0{,}129$, $\gamma = \frac{1}{4}$, $\delta = 0{,}006$, $\xi = \frac{234}{8\ 901\ 064}$ und $\mu = 0{,}00002$ (eigene Darstellung). 61
3.13. Langzeitverhalten der Klassen S und R für eine Basisreproduktionsrate $\mathcal{R}_0 < 1$ (eigene Darstellung). 62
3.14. Angepasste Lösungskurven des SEIR-Modells mit Populationsdynamik für die erste Corona Welle mit Beginn 01.03.2020 (eigene Darstellung). 63
3.15. Angepasste Lösungskurven des SEIR-Modells mit Populationsdynamik für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung). 64
3.16. Angepasste Lösungskurven mit minimalen Veränderungen der Startparameter des SEIR-Modells mit Populationsdynamik für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung). 65
3.17. Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung). 66
3.18. Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung). 67
3.19. Allgemeine Lösungskurven des SEIQR-Modells (eigene Darstellung). 68
3.20. Langzeitverhalten des SEIQR-Modells für $\mathcal{R}_0 > 1$ (eigene Darstellung). 68
3.21. Allgemeine Lösungskurven des SEIQR-Modells (eigene Darstellung). 69
3.22. Angepasste Lösungskurven des SEIQR-Modells für die zweite Corona Welle mit Beginn 13.09.2020 (eigene Darstellung). 70
3.23. Vorhergesagte Lösungskurve (blau) für 15 Anfangsdaten (rot) und weiterer Verlauf (grün) (eigene Darstellung). 71

Tabellenverzeichnis

4.1. Infektionszahlen der ersten Corona Welle in Österreich von 01.03.2020 - 30.04.2020 81
4.2. Infektionszahlen der zweiten Corona Welle in Österreich von 13.09.2020 - 31.12.2020 82

Infektionszahlen der beiden Corona Wellen

Datum	Infektionszahlen
01.03.2020	4
02.03.2020	5
03.03.2020	6
04.03.2020	14
05.03.2020	9
06.03.2020	29
07.03.2020	23
08.03.2020	28
09.03.2020	22
10.03.2020	55
11.03.2020	108
12.03.2020	99
13.03.2020	181
14.03.2020	211
15.03.2020	198
16.03.2020	224
17.03.2020	343
18.03.2020	375
19.03.2020	455
20.03.2020	511
21.03.2020	543
22.03.2020	617
23.03.2020	689
24.03.2020	725
25.03.2020	754
26.03.2020	1050
27.03.2020	779
28.03.2020	594
29.03.2020	536
30.03.2020	714
31.03.2020	565

Datum	Infektionszahlen
01.04.2020	546
02.04.2020	437
03.04.2020	398
04.04.2020	258
05.04.2020	197
06.04.2020	305
07.04.2020	293
08.04.2020	290
09.04.2020	277
10.04.2020	357
11.04.2020	201
12.04.2020	114
13.04.2020	86
14.04.2020	150
15.04.2020	129
16.04.2020	112
17.04.2020	101
18.04.2020	60
19.04.2020	47
20.04.2020	68
21.04.2020	71
22.04.2020	75
23.04.2020	60
24.04.2020	79
25.04.2020	62
26.04.2020	35
27.04.2020	70
28.04.2020	51
29.04.2020	61
30.04.2020	54

Tab. 4.1: Infektionszahlen der ersten Corona Welle in Österreich von 01.03.2020 - 30.04.2020

Datum	Infektions-zahlen
13.09.2020	415
14.09.2020	819
15.09.2020	722
16.09.2020	855
17.09.2020	804
18.09.2020	875
19.09.2020	667
20.09.2020	469
21.09.2020	580
22.09.2020	719
23.09.2020	804
24.09.2020	658
25.09.2020	760
26.09.2020	587
27.09.2020	549
28.09.2020	737
29.09.2020	753
30.09.2020	851
01.10.2020	791
02.10.2020	1014
03.10.2020	717
04.10.2020	735
05.10.2020	928
06.10.2020	1081
07.10.2020	1271
08.10.2020	1110
09.10.2020	1259
10.10.2020	1009
11.10.2020	831
12.10.2020	1143
13.10.2020	1318
14.10.2020	1436
15.10.2020	1523
16.10.2020	1801
17.10.2020	1475
18.10.2020	1208
19.10.2020	1592
20.10.2020	2028
21.10.2020	2663
22.10.2020	3234
23.10.2020	3340
24.10.2020	2962
25.10.2020	2467
26.10.2020	2757
27.10.2020	4379
28.10.2020	4470
29.10.2020	5049
30.10.2020	5699
31.10.2020	4797
01.11.2020	4469
02.11.2020	5183
03.11.2020	6351
04.11.2020	7100
05.11.2020	6957
06.11.2020	8104
07.11.2020	5327
08.11.2020	5740
09.11.2020	6258
10.11.2020	7461
11.11.2020	9203
12.11.2020	8254
13.11.2020	6988
14.11.2020	5114
15.11.2020	4626
16.11.2020	6063
17.11.2020	7548
18.11.2020	6451
19.11.2020	6630
20.11.2020	6455
21.11.2020	5165
22.11.2020	2451
23.11.2020	4615
24.11.2020	5766
25.11.2020	5386
26.11.2020	5040
27.11.2020	4241
28.11.2020	3748
29.11.2020	2408
30.11.2020	3389
01.12.2020	3926
02.12.2020	3590
03.12.2020	3659
04.12.2020	3592
05.12.2020	2874
06.12.2020	2064
07.12.2020	2517
08.12.2020	2642
09.12.2020	2492
10.12.2020	3099
11.12.2020	2846
12.12.2020	2632
13.12.2020	2491
14.12.2020	2747
15.12.2020	2757
16.12.2020	2393
17.12.2020	1962
18.12.2020	2160
19.12.2020	1504
20.12.2020	1353
21.12.2020	2177
22.12.2020	2587
23.12.2020	2849
24.12.2020	1791
25.12.2020	1265
26.12.2020	1160
27.12.2020	1700
28.12.2020	2121
29.12.2020	2997
30.12.2020	2717
31.12.2020	1985

Tab. 4.2: Infektionszahlen der zweiten Corona Welle in Österreich von 13.09.2020 - 31.12.2020