
Der Linear Chain Trick in der epidemiologischen Modellierung als Kompromiss zwischen gewöhnlichen und Integro-Differentialgleichungen

Masterarbeit

vorgelegt von **Lena Plötzke**

am 13. Dezember 2023

am Mathematischen Institut der Universität zu Köln
in Kooperation mit dem Deutschen Zentrum für Luft- und Raumfahrt

Hauptbetreuung: **Dr. Martin Joachim Kühn**
Betreuung (UzK): **Prof. Dr.-Ing. Gregor Gassner**



Inhaltsverzeichnis

Abkürzungsverzeichnis	iii
Abbildungsverzeichnis	iv
Tabellenverzeichnis	v
1 Einleitung	1
2 SIR-Modelle basierend auf gewöhnlichen und Integro-Differentialgleichungen	5
2.1 Rahmenbedingungen und Problemabgrenzung	6
2.2 SIR-Modell auf der Basis von gewöhnlichen Differentialgleichungen	7
2.3 SIR-Modell auf der Basis von Integro-Differentialgleichungen	9
2.4 Verbindung der beiden Modelle und Eigenschaft der Gedächtnislosigkeit	12
3 SIR-Modell unter Ausnutzung des Linear Chain Tricks	15
3.1 Mathematische Grundlagen der Erlang-Verteilung	16
3.2 Herleitung des Linear Chain Trick-Modells aus dem IDE-SIR-Modell	19
3.3 Alternative Herleitungen und Literaturüberblick	24
3.4 Eigenschaften des Linear Chain Trick-Modells	26
3.4.1 Bezug zur Erlang-Verteilung	27
3.4.2 Einfluss der Parameterwahl	30
3.4.3 Untersuchung der Eigenschaft der Gedächtnislosigkeit	33
3.4.4 Überblick über die Beziehung der drei SIR-Modelle	36
3.5 Varianten und Erweiterungsmöglichkeiten des LCT-SIR-Modells	37
3.6 Vor- und Nachteile der Modelle	39
4 Umsetzung der Modellarten in die Praxis	45
4.1 Erweiterung der Modelle durch zusätzliche Kompartimente	45
4.1.1 Das ODE-SECIR-Modell	47
4.1.2 Das IDE-SECIR-Modell	49
4.1.3 Das LCT-SECIR-Modell	55
4.2 Numerische Lösungsverfahren	56
4.2.1 Lösungsverfahren für die Modelle basierend auf gewöhnlichen Differentialgleichungssystemen	57
4.2.2 Lösungsverfahren für das IDE-SECIR-Modell	57
4.3 Wahl der Anfangsbedingungen	62
4.4 Implementierung des LCT-SECIR-Modells	70
4.4.1 Vorhandene Voraussetzungen in der Software MEMilio	71
4.4.2 Einbindung des LCT-SECIR-Modells in die Software	74

5	Simulation und Ergebnisse am Beispiel von SARS-CoV-2 und COVID-19	79
5.1	Verwendete epidemiologische Parameter für SARS-CoV-2 und COVID-19	79
5.2	Simulation fiktiver Szenarien	84
5.2.1	Vergleich verschiedener Methoden zur Initialisierung des LCT-Modells	86
5.2.2	Vergleich der Modelle bei einer Reduzierung der Kontaktrate	91
5.2.3	Vergleich der Reaktion auf eine Erhöhung der Kontaktrate	98
5.2.4	Vergleich des prognostizierten Höhepunkts der Epidemie	100
5.3	Simulation realer Szenarien	104
5.3.1	Simulation über einen Zeitraum im Herbst	106
5.3.2	Simulation über einen Zeitraum im Sommer	109
6	Fazit und Ausblick	113
	Literatur	117
	Versicherung an Eides statt	127

Abkürzungsverzeichnis

Im Folgenden werden die wichtigsten in der Arbeit verwendeten Abkürzungen angeführt.

SIR-Modell	Modell basierend auf den drei Kompartimenten S , I und R
SEIR-Modell	Modell basierend auf den vier Kompartimenten S , E , I und R
SECIR-Modell	Modell basierend auf den acht Kompartimenten S , E , C , I , H , U , R und D
ODE	Engl. <i>ordinary differential equation</i> , gewöhnliche Differentialgleichung, häufig als Präfix für die zuvor genannten Modelle verwendet, beispielsweise ODE-SIR-Modell. Damit wird spezifiziert, dass das Modell auf gewöhnlichen Differentialgleichungen basiert.
IDE	Engl. <i>integro differential equation</i> , Integro-Differentialgleichung, ebenso häufig als Präfix für die zuvor genannten Modelle verwendet.
LCT	Linear Chain Trick, wird als alleinstehende Abkürzung sowie als Präfix für die zuvor genannten Modelle verwendet.
COVID-19	Coronavirus disease 2019
SARS-CoV-2	Severe acute respiratory syndrome coronavirus type 2
SARS	Severe acute respiratory syndrome
HIV	Human immunodeficiency virus
WHO	World Health Organization
RKI	Robert Koch-Institut
MEmlilio	Softwarebibliothek [61], a high performance Modular EpideMIcs simuLatioN software
UzK	Universität zu Köln
engl.	englisch

Abbildungsverzeichnis

3.1	Aufbau des LCT-SIR-Modells	21
3.2	Dichte und Überlebensfunktion der Erlang-Verteilung	31
4.1	Aufbau der SECIR-Modelle	46
4.2	Schema zur Initialisierung des ODE-Modells nach Koslow et al.	64
4.3	Schema einer Initialisierung des LCT-Modells	66
5.1	Tägliche Neuansteckungen für verschiedene Methoden zur Initialisierung des LCT-Modells	88
5.2	Kompartimente für verschiedene Methoden zur Initialisierung des LCT- Modells	89
5.3	Tägliche Neuansteckungen bei einer Verringerung der Kontaktrate	93
5.4	Kompartimente bei einer Verringerung der Kontaktrate in LCT-Modellen	95
5.5	Kompartimente bei einer Verringerung der Kontaktrate im Vergleich mit IDE-Modellen	96
5.6	Tägliche Neuansteckungen bei einer Erhöhung der Kontaktrate	99
5.7	Tägliche Neuansteckungen zum Vergleich der prognostizierten Höhepunkte der Epidemie	101
5.8	Kompartimente zum Vergleich der prognostizierten Höhepunkte der Epi- demie bei verschiedenen Modellarten	102
5.9	Tägliche Neuansteckungen mit Startdatum 01.10.2020	108
5.10	Infizierte und Todesfälle mit Startdatum 01.10.2020	109
5.11	Tägliche Neuansteckungen mit Startdatum 01.06.2020	111
5.12	Infizierte und Todesfälle mit Startdatum 01.06.2020	112

Tabellenverzeichnis

4.1	Butcher-Tableau des Runge-Kutta Cash-Karp 5(4) Verfahrens	57
5.1	Verwendete Werte epidemiologischer Parameter für SARS-CoV-2 und COVID-19	80
5.2	Verwendete Aufenthaltszeitverteilungen für das IDE-SECIR-Modell	83
5.3	Vergleich der benötigten Zeitschritte der Modelle bei einer Verringerung der Kontaktrate	92
5.4	Relative Abweichungen der Prognosen der Neuanssteckungen verschiedener Modelle vom ODE-Modell	94
5.5	Überblick über die implementierten nicht-pharmazeutischen Maßnahmen für die Simulation im Herbst	107
5.6	Überblick über die implementierten nicht-pharmazeutischen Maßnahmen für die Simulation im Sommer	110

1 Einleitung

Infektionskrankheiten stellen seit langer Zeit eine große Herausforderung für die Gesundheit und das Wohlergehen der Gesellschaft dar. Es wird angenommen, dass die Pockenkrankheit schon im Jahr 6000 vor Christus in den ersten Siedlungen auftrat [43, Kapitel 2.6]. In den Jahren 1918 bis 1919 forderte die spanische Grippe etwa 50 Millionen Todesopfer [11, S. 3]. Trotz verbesserter sanitärer Standards und medizinischer Erfolge wie Impfungen und Medikamente sind neu auftretende Infektionskrankheiten auch heute noch von Bedeutung [43, S. 599]. Ein aktuelles Beispiel dafür ist die Krankheit COVID-19, die durch das Coronavirus SARS-CoV-2 ausgelöst wird. Seit den ersten Ansteckungen im Jahr 2019 gab es weltweit bereits mehr als 760 Millionen bestätigte Fälle und 6.9 Millionen Todesfälle [96]. Die Krankheit hat in den vergangenen Jahren nahezu alle Aspekte des gesellschaftlichen Zusammenlebens, der öffentlichen Gesundheit und der globalen Wirtschaft beeinflusst, vergleiche [43, 44, 68]. Obgleich die *World Health Organization* (WHO) die Krankheit seit Mai 2023 nicht länger als *public health emergency of international concern* ansieht [98], ist die Analyse der Krankheit nach wie vor von Relevanz. Wenn wir die Dynamik der Ausbreitung von SARS-CoV-2 verstehen, können wir das Wissen zukünftig bei neu auftretenden oder wiederkehrenden Erregern anwenden. Andere übertragbare Krankheiten wie eine Infektion mit dem *Human immunodeficiency virus* (HIV) oder Tuberkulose sind vor allem in Entwicklungsländern eine der Hauptursachen für Todesfälle und hemmen die soziale und wirtschaftliche Entwicklung [43, 97].

Diese Herausforderungen veranschaulichen, wie bedeutend es ist, die Dynamik der Ausbreitung von Infektionskrankheiten zu verstehen und mögliche Gegenmaßnahmen zu analysieren. Mathematische Modelle haben sich in diesem Gebiet zu einem entscheidenden Werkzeug entwickelt [70, 89]. Die Verwendung von Modellen ermöglicht die Vorhersage von Trends in der Ausbreitung einer Krankheit. Ein wesentlicher Anwendungsbereich liegt in der Planung der Eindämmung und Kontrolle der Ausbreitung. Mathematische Modelle dienen der Einschätzung der Auswirkungen von Impfstrategien oder geplanten Interventionsmaßnahmen [70, S. 2069]. Mittels Simulationen soll die bestmögliche Kombination verschiedener Maßnahmen und der Einsatzzeitraum bestimmt werden, vergleiche [56, S. 22]. Weiterhin bieten Modelle die Möglichkeit, epidemiologische Parameter der betrachteten Infektionskrankheit zu schätzen [70, 93]. Auf diese Weise können politische Akteure¹ dabei unterstützt werden, angemessen auf übertragbare Krankheiten zu reagieren.

Es existiert eine Vielzahl methodischer Zugänge zur Prognose der Verbreitung von Infektionskrankheiten mittels mathematischer Modelle. Unter den Ansätzen gehören Modelle basierend auf gewöhnlichen Differentialgleichungen zu den weitverbreitetsten in der

¹In dieser Arbeit wird das generische Maskulinum zugunsten einer Verbesserung der Lesbarkeit verwendet. Personenbezogene Bezeichnungen gelten für alle Geschlechter gleichermaßen.

Wissenschaft [45, 46]. Diese Modelle zeichnen sich durch ihre einfache Formulierung sowie ihre vergleichsweise unkomplizierte mathematische Analyse und Umsetzung in Software aus [45, 46]. Allerdings gibt es bedeutende Kritik an dem Ansatz. Die Modelle beruhen auf der indirekten Annahme, dass die Verweildauer eines Individuums in einem Krankheitsstadium exponentialverteilt ist. Diese Annahme impliziert, dass die Wahrscheinlichkeit, in den nächsten Krankheitszustand überzugehen, unabhängig von der Zeit seit dem Eintritt in das aktuelle Stadium ist [22, 63]. Aus epidemiologischer Sicht ist die Modellannahme unrealistisch [22, 93]. Dies stellt die Anwendbarkeit von Modellen basierend auf gewöhnlichen Differentialgleichungen in der Epidemiologie infrage. Der Modellansatz wird in einer Arbeit von Kermack und McKendrick [53] aus dem Jahr 1927 eingeführt [28, S. 95ff.]. Für solche Modelle wird die Veröffentlichung häufig zitiert, obwohl der Ansatz in der Arbeit nur als Spezialfall eines allgemeineren Modells vorgestellt wird. Dieses verallgemeinerte Modell basiert auf Integro-Differentialgleichungen und erlaubt eine flexible Wahl der Verteilung der Aufenthaltszeit in den Krankheitszuständen [28, S. 96f.]. Somit beseitigt das Modell den primären Kritikpunkt des Modellansatzes basierend auf gewöhnlichen Differentialgleichungen. Allerdings sind Integro-Differentialgleichungen mathematisch anspruchsvoller zu analysieren, zu formulieren und in Software umzusetzen [45, S. 1833].

Um die Komplexität der Integro-Differentialgleichungen zu umgehen, richten einige Forschungen zu epidemiologischen Modellen mit nicht-exponentialverteilter Aufenthaltszeit den Fokus auf eine bestimmte Art von realistischen Verteilungen [58, Supplementary material]. Wenn die Aufenthaltszeit Erlang-verteilt gewählt wird, kann das Modell ebenso durch ein gewöhnliches Differentialgleichungssystem ausgedrückt werden. Dieses Konzept wird *Linear Chain Trick* genannt, da eine Kette von Zuständen mit linearen Übergangsraten verwendet wird, um eine Erlang-verteilte Verweildauer zu erzielen [45, 66]. Dabei wird ausgenutzt, dass die Erlang-Verteilung als Summe von Exponentialverteilungen darstellbar ist. Die Erlang-Verteilung repräsentiert einen Spezialfall der Gamma-Verteilung, wobei der Formparameter als natürliche Zahl gewählt wird [58, Supplementary material]. Die resultierenden gewöhnlichen Differentialgleichungssysteme weisen zwar eine höhere Dimension auf als die Systeme unter der Annahme von exponentialverteilten Aufenthaltszeiten, sind aber leichter zu handhaben als Integro-Differentialgleichungen [35, S. 17].

Das vorrangige Ziel dieser Arbeit besteht in der Herleitung der Modelle, die auf dem Linear Chain Trick basieren, sowie in der Analyse ihrer Eigenschaften. Dazu orientieren wir uns vor allem an den Ausarbeitungen zum Linear Chain Trick von Hurtado et al. [44, 45] und Feng et al. [31, 32]. Im Vordergrund steht zudem die Frage, ob die Modelle als Kompromiss zwischen den unkomplizierten, jedoch unrealistischen Modellen auf der Basis von gewöhnlichen Differentialgleichungen und den komplexen, jedoch flexiblen Modellen basierend auf Integro-Differentialgleichungen fungieren können. Zur Beantwortung dieser Frage stellen wir die Modelle basierend auf dem Linear Chain Trick den anderen beiden Modellklassen gegenüber und analysieren Vor- und Nachteile. Die theoretische Einleitung und Untersuchung erfolgt anhand von Modellen, welche die Ausbreitung einer Krankheit durch die Verwendung von lediglich drei Kompartimenten darstellen. Dieser Ansatz ermöglicht eine klare Darstellung der Konzepte anhand von einfachen Beispielen. In der praktischen Anwendung wird üblicherweise eine größere Anzahl von Kompartimenten

verwendet und eine feinere Unterteilung der Bevölkerung nach dem Krankheitszustand definiert, um die verschiedenen Aspekte der Erkrankung genauer zu erfassen. Zur Bewertung der Modelle in Übereinstimmung mit praxisnahen Bedingungen führen wir für die drei Klassen Modelle mit einer größeren Anzahl von Kompartimenten ein. Darüber hinaus prüfen wir das Verhalten in Simulationen, indem wir alle Modellarten in der Form mit mehr Kompartimenten implementieren. Zur Simulation verwenden wir Charakteristiken der Erkrankung COVID-19, um die Modelle unter realitätsnahen Rahmenbedingungen zu erproben.

Der folgende Teil der vorliegenden Arbeit gliedert sich wie folgt. In Kapitel 2 wird ein Modell auf der Basis gewöhnlicher Differentialgleichungen sowie eines basierend auf Integro-Differentialgleichungen eingeführt. Im Anschluss erfolgt in Kapitel 3 eine umfangreiche Herleitung und Analyse eines Modells, das den Linear Chain Trick verwendet. Ein besonderes Augenmerk liegt dabei auf dem Zusammenhang und dem Vergleich mit den in Kapitel 2 betrachteten Modellen. In Kapitel 4 wird eine praxisnahe Erweiterung der drei vorangestellten Modellarten auf acht Kompartimente vorgenommen. Des Weiteren untersuchen wir numerische Lösungsansätze für die Gleichungssysteme der Modelle und besprechen die Wahl der Anfangsbedingungen. Zudem wird eine Implementierung des Modells mit dem Linear Chain Trick präsentiert. Im anschließenden Kapitel 5 führen wir numerische Experimente für verschiedene fiktive und reale Szenarien durch und vergleichen die Modelle anhand der Simulationen. Zum Abschluss wird in Kapitel 6 ein Fazit gezogen und ein Ausblick auf mögliche weiterführende Arbeiten gegeben.

2 SIR-Modelle basierend auf gewöhnlichen und Integro-Differentialgleichungen

In diesem Kapitel stellen wir zwei Modellansätze zur Ausbreitung von Infektionskrankheiten vor. Hierbei beschränken wir uns auf Modelle mit lediglich drei Kompartimenten, um die Konzepte in einer simplen Form zu verdeutlichen. Zuerst führen wir die Modellierung auf der Grundlage von gewöhnlichen Differentialgleichungen ein. Diese Kategorie von Modellen ist mit Abstand die am häufigsten verwendete [38, S. 1]. Im Anschluss wenden wir uns einem Modell basierend auf Integro-Differentialgleichungen zu. Die Ausführungen zielen darauf ab, eine Basis für den Vergleich mit dem Modell, das den Linear Chain Trick nutzt, zu schaffen. Letzteres wird ausführlich im anschließenden Kapitel 3 hergeleitet und untersucht.

Das Modell auf der Basis gewöhnlicher Differentialgleichungen beruht auf der Annahme, dass die Austrittsrate aus einem Krankheitszustand unabhängig von der bereits im Zustand verbrachten Zeit ist. Diese Annahme ist zwar mathematisch vorteilhaft, impliziert jedoch, dass die Aufenthaltszeit im Krankheitszustand exponentialverteilt ist. Dies wird aus epidemiologischer Perspektive als unrealistisch angesehen [93, S. 0622]. Im Gegensatz dazu erlaubt das Modell auf der Grundlage von Integro-Differentialgleichungen den Einsatz vielfältiger Verteilungsfunktionen für die Aufenthaltszeit, wodurch es breite Möglichkeiten zur Modellierung realistischer Annahmen bietet [28, S. 96f.]. Beide Modelltypen werden in einer Arbeit von Kermack und McKendrick [53] aus dem Jahr 1927 vorgestellt. Dennoch zitieren die meisten Forscher die Publikation meist im Kontext des Modells auf der Grundlage gewöhnlicher Differentialgleichungen, obgleich dieses Modell dort lediglich als spezieller Fall des Modells basierend auf Integro-Differentialgleichungen aufgeführt wird [28, S. 96f.]. Obwohl einige Veröffentlichungen die weniger bekannten Modelle behandeln und zur Anwendung der Modelle basierend auf Integro-Differentialgleichungen raten, wie zum Beispiel [12, 28, 93], werden diese nach wie vor vergleichsweise selten für Simulationen genutzt.

Die beiden hier besprochenen Modelle werden in [72] ausführlich hergeleitet und verglichen. Siehe auch [11, Kapitel 2.1 & 2.4] für das Modell basierend auf gewöhnlichen Differentialgleichungen und [11, Kapitel 4.4 & 4.5] für das Modell basierend auf Integro-Differentialgleichungen sowie die anderen in [72, Kapitel 2] referenzierten Quellen. Vor diesem Hintergrund legen wir beide Modelle ohne erneute Herleitung dar. Wir untersuchen einige zusätzliche Eigenschaften der Modelle, die im Vergleich zu den Modellen, die den Linear Chain Trick ausnutzen, von Relevanz sind. Vorab ist es notwendig, einige zentrale Begriffe und Notationen zu klären sowie den Untersuchungsumfang abzugrenzen.

2.1 Rahmenbedingungen und Problemabgrenzung

Im Folgenden verfeinern wir unsere Problemstellung durch ihre Einbettung in einen Bezugsrahmen. In diesem Zusammenhang identifizieren wir Einschränkungen bezüglich der Anwendbarkeit der Modelle. Die Ausführungen finden auf sämtliche nachfolgende Modellarten Anwendung. Wir klären ausgewählte Begriffe und Notationen, wodurch eine Basis für die bevorstehenden Erörterungen geschaffen wird. Hierbei erfolgt die Präzisierung bereits zuvor verwendeter Terminologie.

Die Epidemiologie ist laut [54, S. 36] als „Lehre von der Untersuchung der Verteilung von Krankheiten, physiologischen Variablen und sozialen Krankheitsfolgen in menschlichen Bevölkerungsgruppen sowie der Faktoren, die diese Verteilung beeinflussen“ zu verstehen. Bei der Planung der Steuerung der Krankheitsdynamik spielen mathematische Modelle eine wichtige Rolle [11, S. 4]. Wir fokussieren uns auf Infektionskrankheiten, also jene Erkrankungen, die durch übertragbare Erreger beim Menschen hervorgerufen werden und bewegen uns somit auf dem Teilgebiet der mathematischen Infektionsepidemiologie [54, S. 67ff.].

Die hier betrachteten Modelle teilen die Bevölkerung in disjunkte Gruppen, sogenannte Kompartimente, entsprechend dem jeweiligen Krankheitszustand eines Individuums. In der einfachsten Form, die in den Kapiteln 2 und 3 behandelt wird, wird jede Person einem der drei Kompartimente S , I und R zugeordnet, vergleiche hierfür [11, S. 23]. Dabei sind im Kompartiment S (engl. *Susceptibles*) Individuen, die noch nicht infiziert wurden und daher anfällig für die betrachtete Krankheit sind. In der Gruppe I (engl. *Infectives*, *Infected* oder *Infectious*) befinden sich Personen, die infiziert sind und andere Individuen anstecken können. Genesene Individuen oder solche, die an der betrachteten Krankheit gestorben sind, werden dem Kompartiment R (engl. *Recovered* oder *Removed*) zugeordnet. Wir gehen davon aus, dass Personen nach der Genesung von der Krankheit immun sind und sich nicht erneut infizieren können. Die Anzahlen der Personen zu einem Zeitpunkt t im jeweiligen Kompartiment bezeichnen wir mit $S(t)$, $I(t)$ sowie $R(t)$. Als Zeiteinheit verwenden wir Tage, sodass $t \in \mathbb{R}$ die Anzahl der vergangenen Tage seit einem festgelegten Startzeitpunkt beschreibt. Im weiteren Verlauf beschränken wir uns auf Krankheiten, die von Mensch zu Mensch bei einem Kontakt übertragen werden können, siehe [11, S. 22ff.]. Die Übertragung der Krankheit kann daher durch das Kontaktgeschehen zwischen infektiösen Individuen und jenen im Kompartiment S modelliert werden. Ebenfalls beschäftigen wir uns mit Modellen für Epidemien, bei welchen das vermehrte Auftreten einer Krankheit ein zeitlich und räumlich begrenzter Prozess ist [54, S. 34]. Im Gegensatz dazu stehen Endemien, wo ein zeitlich unbegrenztes Vorkommen der Krankheit angenommen wird [54, S. 32]. Der Hauptunterschied in der Modellierung liegt darin, dass bei Endemien ein Wachstum der Gruppe S berücksichtigt wird, beispielsweise durch Geburten oder Immigration. Sterbeprozesse, die von der betrachteten Krankheit unabhängig sind, werden ebenso inkludiert. Bei Epidemien verläuft die Welle der Ansteckungen so rasch, dass solche Prozesse vernachlässigt werden können [11, S. 21f.]. Genau genommen wäre es angemessener, ein endemisches Modell zur Simulation der aktuellen Ausbreitungssituation von COVID-19 zu verwenden [49, S. 5379]. Da diese Krankheit lediglich als Beispiel in den Simulationen dient, beschränken

wir uns dort auf die epidemische Phase. Die Konzepte der epidemischen Modellierung können in zukünftigen Arbeiten zu endemischen Modellen erweitert werden.

Sofern ein Modell mit den drei Gruppen S , I und R betrachtet wird, kürzen wir es mit dem Namen „SIR-Modell“ ab. Der Bezeichnung kann ein Präfix für die verwendete Modellklasse vorangestellt werden. Die Gesamtbevölkerungszahl im betrachteten Gebiet zum Zeitpunkt t bezeichnen wir als $N(t)$. Da jede Person zu einem betrachteten Zeitpunkt genau einer Gruppe angehört, gilt die Beziehung $N(t) = S(t) + I(t) + R(t)$. Wie soeben konkretisiert, betrachten wir epidemische Modelle, sodass wir Geburten und Sterbefälle, die nicht durch die betrachtete Krankheit verursacht wurden, vernachlässigen. Hieraus folgt, dass die Bevölkerungszahl über den betrachteten Zeitraum konstant ist, das heißt $N(t) = N$. In der Bevölkerungszahl werden die durch die betrachtete Krankheit verursachten Todesopfer mitgezählt, da wir in der Gruppe R nicht zwischen den Genesenen und den Verstorbenen unterscheiden. Für die Personen, die an den Folgen der Krankheit gestorben sind, könnte eine eigene Gruppe eingeführt und von der Bevölkerungszahl abgezogen werden, sodass die Bevölkerungszahl von der Zeit abhängig wird.

2.2 SIR-Modell auf der Basis von gewöhnlichen Differentialgleichungen

Unter Berücksichtigung der Rahmenbedingungen aus dem vorigen Abschnitt stellen wir nun ein SIR-Modell vor, das auf gewöhnlichen Differentialgleichungen basiert. Dabei sei für die Herleitung der hier präsentierten Ergebnisse auf [72, Kapitel 2.2] verwiesen. Seit der Veröffentlichung von Kermack und McKendrick [53] hat sich das Modell zu einem etablierten Konzept in der Modellierung der Ausbreitung von Infektionskrankheiten entwickelt und ist auf verschiedene Arten erweitert worden, siehe [28, S. 95f.] oder [12, S. 103f.].

Das SIR-Modell basierend auf gewöhnlichen Differentialgleichungen ergibt sich als nicht-lineares gewöhnliches Differentialgleichungssystem

$$\begin{aligned} S'(t) &= -\rho(t) \phi(t) I(t) \frac{S(t)}{N} \\ I'(t) &= \rho(t) \phi(t) I(t) \frac{S(t)}{N} - \gamma I(t) \\ R'(t) &= \gamma I(t) \end{aligned} \tag{2.1}$$

für $t \geq 0$. Der Parameter $\rho(t) \in [0, 1]$ bezeichnet hier die Wahrscheinlichkeit der Übertragung der Krankheit bei einem Kontakt zwischen einem Infektiösen aus dem Zustand I und einer Person der Gruppe S zum Zeitpunkt t . Durch $\phi(t) \geq 0$ ist die durchschnittliche Anzahl an Kontakten eines Individuums pro Tag definiert. Beide Parameter sind von der Zeit t abhängig, da Virusvarianten oder nicht-pharmazeutische Maßnahmen, wie Schulschließungen, eine Maskenpflicht oder Quarantäne- und Isolationsregelungen, die Werte der Parameter verändern können. Im Modell wird angenommen, dass Infizierte die Gruppe I mit einer Rate γI verlassen, wobei $\gamma > 0$ ein konstanter Parameter ist. Wir definieren $T_I \geq 0$ als die erwartete Dauer der infektiösen Periode, das heißt die erwartete im Kompartiment I verbrachte Zeit bevor ein Individuum nach R übergeht. Die infektiöse

Periode bezeichnet den Zeitraum, in welchem ein Individuum andere Individuen anstecken kann [54, S. 75]. Für den Parameter γ gilt der Zusammenhang $\gamma = T_I^{-1}$, das heißt γ hat im epidemiologischen Kontext die Bedeutung der inversen infektiösen Periode. In der Modellgleichung stecken weitere Annahmen, beispielsweise bezüglich der Transmission der Krankheit. Siehe [11, S. 21ff.] und [72, Kapitel 2.1 & 2.2] für nähere Ausführungen. Wie bereits erwähnt, gehen wir davon aus, dass sich Individuen nach überstandener Krankheit nicht erneut infiziert können. Demnach sind durch die Modellgleichungen nur Übergänge von der Gruppe S nach I und von I nach R vorgesehen. Gelegentlich wird die Gleichung für das Kompartiment S durch

$$S'(t) = -\rho(t) \phi(t) I(t) \frac{S(t)}{N} = -\varphi(t) S(t)$$

umgeschrieben, wobei $\varphi(t) = \rho(t) \phi(t) I(t) N^{-1}$ den sogenannten Infektionsdruck (engl. *force of infection*) darstellt, vergleiche [44, S. 7]. Dieser beschreibt das Risiko einer Person aus der Gruppe S mit dem Erreger infiziert zu werden. Siehe dazu die Quelle [54, S. 68], aus welcher ebenfalls die deutsche Übersetzung übernommen wird.

Zusammen mit den Anfangswerten

$$S(0) = S_{t_0}, I(0) = I_{t_0} \text{ und } R(0) = R_{t_0} \tag{2.2}$$

ergibt das Differentialgleichungssystem (2.1) ein Anfangswertproblem. Dabei setzen wir im Folgenden stets den Anfangszeitpunkt t_0 als Zeitpunkt Null fest, sodass $t_0 = 0$ gilt. In Anlehnung an die englische Übersetzung von „gewöhnliche Differentialgleichung“ als *ordinary differential equation* (ODE) benennen wir dieses Modell im weiteren Verlauf als „ODE-SIR-Modell“. Sofern eindeutig ersichtlich ist, welche Kompartimente im Modell betrachtet werden, kürzen wir die Benennungen gelegentlich ab und schreiben nur „ODE-Modell“. Als Nächstes führen wir die für praktische Anwendungen relevanten Reproduktionszahlen ein und leiten eine Formel basierend auf den Parametern des ODE-Modells her.

Bemerkung 2.1 (Einführung der Reproduktionszahl). *Reproduktionszahlen sind ein bedeutsames Konzept auf dem Gebiet der mathematischen Infektionsepidemiologie. Die Basisreproduktionszahl \mathcal{R}_0 ist für den Anfang der Ausbreitung einer Infektionskrankheit definiert und beschreibt die Anzahl der Individuen, die ein einzelner Infizierter im Verlauf seiner Infektion im Durchschnitt ansteckt, falls die Bevölkerung ansonsten vollständig für die Krankheit anfällig ist [11, S. 6]. Dies entspricht der Annahme, dass alle anderen Personen in der Gruppe S sind. Die Basisreproduktionszahl wird häufig betrachtet, da sie für deterministische Modelle als Schwellenwert fungiert. Falls $\mathcal{R}_0 > 1$ gilt, steckt ein Infizierter im Durchschnitt mehr als ein weiteres Individuum an. Dementsprechend tritt eine epidemische Ausbreitung der Krankheit auf. Wenn $\mathcal{R}_0 < 1$ gilt, geht die Anzahl der Neuansteckungen zurück, siehe [20, S. 261] und [11, S. 7f.]. Wir verfolgen hier nur eine naive Vorgehensweise und leiten eine Formel für die Basisreproduktionszahl aus den Bedeutungen der Parameter des ODE-SIR-Modells her. Die Reproduktionszahl kann auch über andere Verfahren bestimmt werden. Beispielsweise wird in [93, S. 0623] ein Vorgehen über Eigenwerte und die Rate eines exponentiellen Wachstums verwendet. Besonders in detaillierteren Modellen ist die korrekte Berechnung der Basisreproduktionszahl herausfordernder [11, S. 28].*

Seien $\hat{\phi}$ und $\hat{\rho}$ die Kontaktrate beziehungsweise die Transmissionswahrscheinlichkeit zu Beginn der Ausbreitung der betrachteten Krankheit. Der Infizierte hat gemäß dem ODE-Modell durchschnittlich $\hat{\phi}\hat{\rho}$ Kontakte innerhalb eines Tages, die zu einer Ansteckung führen, falls sie mit einem Individuum der Gruppe S stattfinden. Die Basisreproduktionszahl ist unter der Annahme definiert, dass die Bevölkerung ansonsten vollständig anfällig ist, sodass jeder dieser Kontakte zu einer neuen Infektion führt. Innerhalb der erwarteten Dauer der infektiösen Periode T_I überträgt der Infizierte die Infektionskrankheit somit im Durchschnitt an

$$\mathcal{R}_0 = \hat{\phi} \hat{\rho} T_I = \hat{\phi} \hat{\rho} \frac{1}{\gamma}$$

Individuen. Vergleiche für die Erklärung [11, S. 28] mit [72, S. 9f.].

Im weiteren Verlauf der Verbreitung der betrachteten Krankheit lässt sich die Ausbreitungsdynamik besser anhand der effektiven Reproduktionszahl $\mathcal{R}_{\text{eff}}(t)$ zum Zeitpunkt t beschreiben. Diese inkludiert eine sinkende Anfälligkeit in der Bevölkerung und kann Veränderungen in der Kontaktrate und der Transmissionswahrscheinlichkeit berücksichtigen, vergleiche [20, S. 262] und [27, S. 320]. Die effektive Reproduktionszahl kann für einen Zeitpunkt t durch

$$\mathcal{R}_{\text{eff}}(t) = \rho(t) \phi(t) T_I \frac{S(t)}{N} = \frac{\rho(t) \phi(t)}{\gamma} \frac{S(t)}{N}$$

definiert werden. Zu einem fortgeschritteneren Zeitpunkt übernimmt $\mathcal{R}_{\text{eff}}(t)$ die Rolle als Schwellenwert für deterministische Modelle. Falls zu einem Zeitpunkt t die Beziehung $\mathcal{R}_{\text{eff}}(t) > 1$ gilt, wird sich die Krankheit weiter ausbreiten, wenn ρ und ϕ unverändert bleiben. Falls $\mathcal{R}_{\text{eff}}(t) < 1$ gilt, gehen die Fallzahlen bei gleichbleibenden Parametern ρ und ϕ zurück. Nicht-pharmazeutische Maßnahmen zur Eindämmung von Infektionskrankheiten sollten demnach darauf abzielen, die effektive Reproduktionszahl unter eins abzusenken. Siehe hierfür [20, S. 262] und [72, S. 9ff.].

2.3 SIR-Modell auf der Basis von Integro-Differentialgleichungen

Im folgenden Abschnitt präsentieren wir ein SIR-Modell, das auf Integro-Differentialgleichungen basiert. Für eine Herleitung verweisen wir erneut auf [72, Kapitel 2.3.1 & 2.3.2]. Im Vergleich zu seinem ODE-basierten Gegenstück aus Kapitel 2.2 erfährt das Modell eine wesentlich geringere Popularität [28, S. 96f.]. Das Modell wird im weiteren Verlauf als „IDE-SIR-Modell“ bezeichnet, wobei „IDE“ die Abkürzung für die englische Übersetzung von „Integro-Differentialgleichung“, *integro differential equation*, ist. Eine Integro-Differentialgleichung bezeichnet ein System, in dem die Ableitung einer gesuchten Funktion vorkommt, sowie ein Integral, bei welchem die gesuchte Funktion Teil des Integranden ist [90, S. 15f.].

Die Modellgleichungen des IDE-SIR-Modells ergeben sich durch

$$\begin{aligned} S'(t) &= \frac{S(t)}{N} \int_{\hat{T}}^t \rho(t, t-x) \phi(t, t-x) \gamma_I^R(t-x) S'(x) dx \\ I(t) &= - \int_{\hat{T}}^t \gamma_I^R(t-x) S'(x) dx \\ R(t) &= - \int_{\hat{T}}^t (1 - \gamma_I^R(t-x)) S'(x) dx \end{aligned} \quad (2.3)$$

für $t \geq 0$. Zur Lösung des Integro-Differentialgleichungssystems seien passende Anfangsbedingungen

$$S(t) = S_0(t), I(t) = I_0(t) \text{ und } R(t) = R_0(t) \text{ für alle } t \leq 0 \quad (2.4)$$

bekannt. Diese Anfangsfunktionen seien stetig differenzierbar. Mit $S'(t) = -\varphi(t)S(t)$ folgt, dass der Infektionsdruck in diesem Modell durch

$$\varphi(t) = -\frac{1}{N} \int_{\hat{T}}^t \rho(t, t-x) \phi(t, t-x) \gamma_I^R(t-x) S'(x) dx$$

gegeben ist. Zur Formulierung des IDE-SIR-Modells sind zusätzlich zu den bereits bekannten Bezeichnungen weitere Parameter notwendig. Die Größe \hat{T} mit $-\infty < \hat{T} < 0$ verweist auf einen Zeitpunkt, zu dem noch keine Individuen mit der betrachteten Krankheit infiziert sind. Dementsprechend gelten die Gleichheiten $S(\hat{T}) = N$ und $I(\hat{T}) = R(\hat{T}) = 0$. Ein solcher Zeitpunkt wird als bekannt angenommen. Analog zur Spezifikation im ODE-Modell bezeichnet $\rho(t, \tau)$ die Wahrscheinlichkeit der Krankheitsübertragung bei einem Kontakt, während $\phi(t, \tau)$ die durchschnittliche Anzahl der Kontakte eines Individuums pro Tag kennzeichnet. Jedoch sind die Parameter nun zusätzlich von der vergangenen Zeit seit der Infektion τ abhängig. Beispielsweise beschreibt $\rho(t, \tau)$ die durchschnittliche Übertragungswahrscheinlichkeit zum Zeitpunkt t von Individuen, die sich seit τ Tagen im Kompartiment I befinden. Die Infektiosität und das Kontaktverhalten von Individuen kann sich nun im Verlauf der infektiösen Periode verändern.

Eine wesentliche Parametrisierung des IDE-Modells ist die Funktion $\gamma_I^R : \mathbb{R}_0^+ \rightarrow [0, 1]$. Die Schreibweise \mathbb{R}^+ steht in dieser Arbeit für die positiven reellen Zahlen ohne die Null und \mathbb{R}_0^+ inkludiert die Null. Im IDE-Modell repräsentiert $\gamma_I^R(\tau)$ den erwarteten Anteil der Individuen, die τ Tage nach dem Eintritt in das Kompartiment I noch infektiös sind und nicht in das Kompartiment R übergegangen sind. Die Funktion wird als stetig differenzierbar und monoton fallend angenommen, das heißt

$$\gamma_I^R \in C^1(\mathbb{R}_0^+) \text{ und } \gamma_I^{R'}(\tau) \leq 0 \text{ für } \tau \geq 0.$$

Es wird vorausgesetzt, dass alle Individuen zum Zeitpunkt der Infektion im Kompartiment I sind, was bedeutet, dass $\gamma_I^R(0) = 1$ gilt [31, S. 1518]. Das Verständnis der Funktion γ_I^R erweist sich als fundamental für unsere anschließenden Ausführungen, da sich die von uns behandelten Modelle insbesondere in der Flexibilität der Auswahl dieser Funktion unterscheiden.

Zur Ermittlung der angemessensten Modellannahme für die Funktion γ_I^R eignet sich eine stochastische Interpretation, wofür wir auf Begriffe aus der Wahrscheinlichkeitstheorie

zurückgreifen. Es ist jedoch zu beachten, dass wir ausschließlich mit deterministischen Modellen arbeiten, das heißt mit nicht-stochastischen Modellen. Die Terminologie der Wahrscheinlichkeitstheorie verwenden wir lediglich zur Interpretation der Modelle, vergleiche [32, S. 91]. Gemäß [67, S. 8] sind die gesuchten Variablen in einem deterministischen Modell eindeutig durch die Modellparameter und die Anfangsbedingungen festgelegt. Im Gegensatz dazu zeichnen sich stochastische Modelle durch Zufälligkeit aus. In unseren Modellen spielt Zufall bei der Bestimmung der Größe der Modellvariablen keine Rolle. Hurtado et al. [44, 45] verwenden ein stochastisches Modell beziehungsweise stochastische Grundprinzipien, um aus diesen Annahmen das deterministische IDE-Modell als Approximation herzuleiten. Auch in [3] wird ein stochastisches Modell aufgestellt und ein deterministisches als Approximation abgeleitet, wobei direkt eine Verteilung der infektiösen Periode wie im Modell aus Kapitel 3 angenommen wird. Die deterministischen Modelle erfassen stets ein Durchschnittsverhalten beziehungsweise ein erwartetes Verhalten der stochastischen Modelle. Die Grundannahme hinter den deterministischen Approximationen ist, dass die Anzahl der Individuen, die am zu modellierenden Prozess beteiligt sind, groß genug ist [27, S. 27–29]. Die gesamte Herleitung des IDE-Modells aus stochastischen Annahmen ist in [44, Kapitel 3.5.1] zu finden und wird aufgrund der Zielrichtung der vorliegenden Arbeit nicht umfassend ausgeführt. Der Abschnitt [45, Anhang A] befasst sich mit einer analogen Herleitung in einem anderen biologischen Kontext.

Nun wenden wir uns der zu analysierenden Funktion γ_I^R zu und verleihen ihr eine Interpretation im stochastischen Kontext. In [72, Beweis zu Gleichung (2.13)] wird der Zusammenhang der Funktion $\gamma_I^R(\tau)$ zur Verteilung der infektiösen Periode beziehungsweise der Aufenthaltszeit im Kompartiment I aufgezeigt. Diesen Zusammenhang wollen wir hier in einem höheren Detaillierungsgrad ausführen. Sei dazu X eine Zufallsvariable, die einem Infizierten die Dauer der infektiösen Periode zuordnet, das heißt die Zeit vom Eintritt in das Kompartiment I bis zum Austritt. Nach Definition gibt $\gamma_I^R(\tau)$ den erwarteten Anteil der Personen an, die sich τ Tage nach dem Eintritt in I noch in dieser Gruppe befinden. Demnach ist die infektiöse Periode für diesen Anteil der Personen länger als τ Tage. Durch eine stochastische Interpretation erhalten wir das Ergebnis, dass die Wahrscheinlichkeit, eine infektiöse Periode länger als τ Tage zu haben, genau $\gamma_I^R(\tau)$ beträgt. Dies ergibt den Zusammenhang $\mathbb{P}(X > \tau) = \gamma_I^R(\tau)$ und für die Verteilungsfunktion F_X von X

$$F_X(\tau) = \mathbb{P}(X \leq \tau) = 1 - \mathbb{P}(X > \tau) = 1 - \gamma_I^R(\tau) \quad (2.5)$$

für $\tau \geq 0$. Dieser Zusammenhang wird in [72, Bemerkung 3] festgestellt. Hier sei nochmals darauf aufmerksam gemacht, dass wir die Sprache der Wahrscheinlichkeitstheorie nur zum besseren Verständnis der Modellgrößen verwenden. In den deterministischen IDE-Modellen ist kein Zufall inkludiert. Die erwähnten „Wahrscheinlichkeiten“ sind als erwarteter Anteil der Individuen zu verstehen, die noch im Kompartiment sind, siehe dazu [44, S. 62]. Aufgrund des Zusammenhangs $\gamma_I^R(\tau) = \mathbb{P}(X > \tau)$ wird die Funktion von Hurtado et al. in [44, 45] auch Überlebensfunktion (engl. *survival function*) genannt. Da die Dauer der infektiösen Periode stets nicht-negativ ist, ist X eine nicht-negative Zufallsvariable.

Somit folgt für die erwartete Dauer der infektiösen Periode

$$\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X \geq \tau) \, d\tau = \int_0^\infty \mathbb{P}(X > \tau) \, d\tau = \int_0^\infty \gamma_I^R(\tau) \, d\tau, \quad (2.6)$$

siehe [34, S. 125 & 390]. Demnach ist die zusätzliche Anforderung aus [31, S. 1518],

$$\int_0^\infty \gamma_I^R(\tau) \, d\tau < \infty, \quad (2.7)$$

an die Funktion γ_I^R sinnvoll, da wir erwarten, dass ein Individuum in endlicher Zeit wieder gesundet. Wenn die Parameter T_I und γ so definiert sind wie für das ODE-SIR-Modell in Kapitel 2.2, gilt außerdem die Beziehung

$$\gamma^{-1} = T_I = \int_0^\infty \gamma_I^R(\tau) \, d\tau.$$

Die wichtigsten Aspekte des IDE-SIR-Modells haben wir hiermit besprochen. Das Modell ist in der Literatur unter anderen Namen bekannt, was in der folgenden Bemerkung festgehalten wird.

Bemerkung 2.2 (Alternative Bezeichnungen des IDE-Modells). *In der Fachliteratur sind ähnliche Modelle wie das präsentierte IDE-SIR-Modell unter verschiedenen Bezeichnungen zu finden. Zur besseren Übersicht stellen wir einige dieser Bezeichnungen nachfolgend vor. Darüber hinaus können einige Benennungen zum Verständnis des Modells beitragen. In [11, Kapitel 4.5] wird ein Modell desselben Modelltyps mit dem Terminus „Age of Infection Model“ versehen. In [77, Challenge 4] findet sich der ähnliche Name „time-since-infection model“. Diese Begriffe beziehen sich auf die Abhängigkeit der Modellparameter vom „Infektionsalter“ τ , mit welchem die Zeit seit der Infektion gemeint ist. Eine allgemeinere Bezeichnung als „Stage-Age Structured Model“ wird in [84, Kapitel 20] verwendet, angelehnt an die bereits verbrachte Zeit in einer „Stage“ beziehungsweise in einer Krankheitsphase. In allgemeinen biologischen Kontexten wird auch von einem „distributed delay model“ gesprochen, siehe [8, S. 289] oder [49, S. 5379]. Durch die Schreibweise*

$$I(t) = - \int_{\hat{T}}^t \gamma_I^R(t-x) S'(x) \, dx = - \int_0^{t-\hat{T}} \gamma_I^R(\tau) S'(t-\tau) \, d\tau$$

ergibt sich, dass die Funktion γ_I^R als Gewichtung von Verzögerungstermen $S'(t-\tau)$ interpretiert werden kann, siehe dazu [80, S. 9].

2.4 Verbindung der beiden Modelle und Eigenschaft der Gedächtnislosigkeit

Wir betrachten nun den Zusammenhang zwischen dem ODE-SIR-Modell (2.1) und dem IDE-SIR-Modell (2.3). Der folgende Satz wird in [72, Kapitel 2.4] bewiesen.

Satz 2.3 (Zusammenhang von ODE- und IDE-Modell). *Das IDE-SIR-Modell reduziert sich unter der Verwendung von $\rho(t, \tau) = \rho(t)$ und $\phi(t, \tau) = \phi(t)$ sowie $\gamma_I^R(\tau) = e^{-\gamma\tau}$ für $\tau \geq 0$ auf das ODE-SIR-Modell. Diese Form für γ_I^R bedeutet, dass die Dauer der infektiösen Periode beziehungsweise die eingeführte Zufallsvariable X im ODE-SIR-Modell exponentialverteilt mit dem Parameter γ ist.*

Aus diesem Satz 2.3 folgt die Einschränkung des ODE-SIR-Modells, dass es nur exponentialverteilte infektiöse Perioden mit dem Parameter γ zulässt. Dies ist einer der Hauptkritikpunkte an ODE-basierten Modellen, da die Annahme einer exponentialverteilten infektiösen Periode für die meisten Infektionskrankheiten als unrealistisch angesehen wird, siehe [22, 31, 32, 47, 58, 63, 91, 93]. Dies könnte laut [32, S. 103] und [91, S. 2152] vor allem an der Eigenschaft der Gedächtnislosigkeit (engl. *memoryless*) der Exponentialverteilung beziehungsweise des ODE-Modells liegen. Die Bedeutung dieses Konzepts beleuchten wir im Folgenden näher. Wie in [48, Kapitel 4.8.1] erläutert wird, ist in diesem Zusammenhang die bedingte Wahrscheinlichkeitsverteilung der exponentialverteilten Zufallsvariable X von Bedeutung. Für die zuvor eingeführte Zufallsvariable X gilt im ODE-SIR-Modell $\mathbb{P}(X > \tau) = e^{-\gamma\tau}$. Für die Wahrscheinlichkeit, dass ein Individuum $s \geq 0$ weitere Tage im Kompartiment I verweilt, unter der Bedingung, dass die Aufenthaltszeit bereits $\tau \geq 0$ Tage beträgt, folgt demnach

$$\begin{aligned} \mathbb{P}(X > s + \tau | X > \tau) &= \frac{\mathbb{P}(X > s + \tau)}{\mathbb{P}(X > \tau)} = \frac{e^{-\gamma(s+\tau)}}{e^{-\gamma\tau}} \\ &= e^{-\gamma(s+\tau)+\gamma\tau} = e^{-\gamma s} = \mathbb{P}(X > s). \end{aligned} \quad (2.8)$$

Die Notation im ersten Ausdruck ist als bedingte Wahrscheinlichkeit zu verstehen. Entsprechend ist die Wahrscheinlichkeit, s weitere Tage im Kompartiment zu verbringen, unabhängig von der bereits in I verbrachten Zeit τ . Die Eigenschaft der Gedächtnislosigkeit ist durch den Zusammenhang

$$\mathbb{P}(X > s + \tau | X > \tau) = \mathbb{P}(X > s) \text{ für alle } s, \tau \geq 0 \quad (2.9)$$

passend zu der Interpretation definiert, dass die Vergangenheit „vergessen“ wird, vergleiche [48, S. 121 & Kapitel 4.8.1]. Für unser Modell bedeutet dies, dass die zukünftige Entwicklung der Modellgrößen nur von den derzeitigen Größen und nicht von den Entwicklungen in der Vergangenheit abhängt [44, S. 50].

Mit dieser Eigenschaft ist die Größe der erwarteten restlichen Aufenthaltszeit im Kompartiment I eng verbunden, siehe [31, S. 1521f.] oder [11, S. 99]. Die erwartete verbleibende Aufenthaltszeit $\mathcal{M}_I(\tau)$ im Kompartiment I eines Individuums mit dem Infektionsalter τ können wir für allgemeine nicht-negative Zufallsvariablen X durch

$$\mathcal{M}_I(\tau) = \int_0^\infty \mathbb{P}(X > s + \tau | X > \tau) ds = \int_0^\infty \frac{\mathbb{P}(X > s + \tau)}{\mathbb{P}(X > \tau)} ds \quad (2.10)$$

definieren. Dabei gleicht $\mathcal{M}_I(0)$ der erwarteten infektiösen Periode, da

$$\begin{aligned} \mathcal{M}_I(0) &= \int_0^\infty \mathbb{P}(X > s + 0 | X > 0) ds = \int_0^\infty \frac{\mathbb{P}(X > s)}{\mathbb{P}(X > 0)} ds \\ &= \int_0^\infty \frac{\mathbb{P}(X > s)}{1} ds = \int_0^\infty \mathbb{P}(X > s) ds \stackrel{(2.6)}{=} \mathbb{E}(X) \end{aligned}$$

gilt. Nehmen wir nun wieder an, dass die infektiöse Periode beziehungsweise die Zufallsvariable X exponentialverteilt ist, wie in dem Fall, in dem sich das IDE-Modell auf das ODE-Modell reduziert. Die Gedächtnislosigkeit der Exponentialverteilung impliziert für die erwartete restliche Aufenthaltszeit im Kompartiment I eines Individuums des Infektionsalters τ

$$\mathcal{M}_I(\tau) \stackrel{(2.10)}{=} \int_0^\infty \frac{\mathbb{P}(X > s + \tau)}{\mathbb{P}(X > \tau)} ds \stackrel{(2.8)}{=} \int_0^\infty \mathbb{P}(X > s) ds = \mathbb{E}(X).$$

Die erwartete verbleibende Aufenthaltszeit im Kompartiment I ist für unsere exponentialverteilt gewählte infektiöse Periode unabhängig von der bereits in I verbrachten Zeit.

Wir haben in diesem Kapitel sowohl das ODE-SIR-Modell als auch das IDE-SIR-Modell eingeführt. Dabei haben wir bemerkt, dass das ODE-Modell einen Spezialfall des IDE-Modells darstellt. Nutzen wir das IDE-Modell zur Simulation der Ausbreitung einer Infektionskrankheit, lässt es im Vergleich zur Nutzung des ODE-Modells somit stets eine größere Flexibilität zu. Wollen wir die Ergebnisse des ODE-Modells erzielen, können wir die Parameter für das IDE-Modell so wählen, dass sich das Modell auf das ODE-Modell reduziert. Dafür wird das IDE-Modell allerdings durch ein Integro-Differentialgleichungssystem formuliert, was es erschwert, das Modell zu verstehen. Wir haben hiermit die Basis dafür geschaffen, das Modell, welches wir im folgenden Kapitel 3 beleuchten werden, ausführlich mit dem ODE-SIR- und dem IDE-SIR-Modell zu vergleichen.

3 SIR-Modell unter Ausnutzung des Linear Chain Tricks

In diesem Kapitel leiten wir ein Modell her, das auf dem sogenannten Linear Chain Trick basiert. Im Abschnitt 2.4 wurden einige Quellen angegeben, welche die Aussage treffen, dass die Annahme des ODE-SIR-Modells einer exponentialverteilten infektiösen Periode im epidemiologischen Kontext unrealistisch ist. Eine für die meisten Krankheiten realistischere Annahme wäre die Berücksichtigung einer Gamma-verteilten Aufenthaltszeit [93, S. 0622]. Für das Modell basierend auf dem Linear Chain Trick wird sich herausstellen, dass die Aufenthaltszeit im Kompartiment I einer Erlang-Verteilung folgt. Erlang-Verteilungen sind Spezialfälle von Gamma-Verteilungen mit einer spezifischen Parameterwahl. Durch Einsetzen der Erlang-Verteilung in das IDE-Modell wird dieses auf ein gewöhnliches Differentialgleichungssystem reduziert, wobei die Dimension des Systems höher ist als die des ODE-Modells [35, S. 17]. Die Inklusion einer angemesseneren Annahme als im ODE-Modell ohne die Verwendung von komplexen Integro-Differentialgleichungssystemen motiviert die Untersuchung des Linear Chain Tricks. Wir erlangen in diesem Kapitel erste Erkenntnisse bezüglich der zentralen Frage, ob das Modell basierend auf dem Linear Chain Trick das Potential hat, die Vorteile der ODE- und IDE-Modelle zu vereinen.

Das Kapitel ist in folgender Weise strukturiert, um sich der Beantwortung dieser Fragestellung zu nähern. Die Ableitung des Linear Chain Trick-Modells in Kapitel 3.2 aus dem IDE-SIR-Modell erfordert eine fundierte mathematische Grundlage bezüglich der Erlang-Verteilung. Diese Grundlagen führen wir zunächst ein. Die Modellableitung basiert auf Arbeiten von Feng et al. [32] und Hurtado et al. [45]. Die mathematischen Voraussetzungen werden daher unter Berücksichtigung dieser Literatur ausgewählt. Für das Modell gelten ebenfalls die in Kapitel 2.1 aufgeführten Rahmenbedingungen. Erneut konzentrieren wir uns zu Anschauungszwecken auf ein Modell, das auf den drei Gruppen S , I und R basiert. Nach der Herleitung schließt sich in Kapitel 3.3 ein Literaturüberblick an, der alternative Beweisansätze präsentiert und einen umfassenden Einblick in die Quellen bietet, die sich mit dem Thema des Linear Chain Tricks beschäftigen. Um ein besseres Verständnis für das Verhalten des Modells zu erlangen, werden in Kapitel 3.4 verschiedene Eigenschaften des Modells untersucht. Als Ausblick für spätere Arbeiten stellen wir in Abschnitt 3.5 einige sinnvolle Erweiterungsmöglichkeiten zur Modellierung vor. Abschließend führen wir in Kapitel 3.6 die untersuchten Aspekte zur Bewertung der Modelle mit Argumenten aus der Literatur zusammen. Diese Erläuterung bildet die Grundlage für die Untersuchungsschwerpunkte der nachfolgenden Kapitel. Dort werden die Aussagen überprüft, indem die Modelle um praxisrelevante Überlegungen erweitert und mittels Simulationen verglichen werden.

3.1 Mathematische Grundlagen der Erlang-Verteilung

Für das Linear Chain Trick-Modell wird die Modellannahme einer Erlang-verteilten infektiösen Periode eine zentrale Rolle spielen. Zur Schaffung einer Basis für die Herleitung und Untersuchung der Eigenschaften des Modells sowie der Verbindung zu den zuvor behandelten Modellklassen, werden in diesem Kapitel diverse Definitionen und Ergebnisse in Bezug auf diese Verteilung erörtert. Die Erklärungen zur Erlang-Verteilung orientieren sich an den Darlegungen in [41], sofern nicht anders angegeben. Da die Erlang-Verteilung eine Unterklasse der Gamma-Verteilung darstellt, erfolgt zunächst eine Definition der letzteren Verteilung.

Definition 3.1 (Gamma-Verteilung). *Die Gamma-Verteilung mit den Parametern $\lambda \in \mathbb{R}^+$ und $\alpha \in \mathbb{R}^+$ ist durch die Dichtefunktion*

$$f_{\lambda,\alpha}(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{für } x > 0 \\ 0 & \text{für } x \leq 0 \end{cases}$$

beschrieben. Hierbei stellt $\Gamma(\alpha)$ die Gammafunktion

$$\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy$$

dar [78, Kapitel 5.6.1]. Der Parameter α bestimmt die Gestalt der Verteilung [42, S. 308], weshalb er Formparameter genannt wird, siehe [34, S. 47] oder [45, S. 1836]. Der Parameter λ bewirkt eine Skalenänderung [42, S. 308], weshalb er im Folgenden wie in [45, S. 1836] als Ratenparameter bezeichnet wird.

Wird der Formparameter einer Gamma-Verteilung als natürliche Zahl gewählt, das bedeutet $\alpha \in \mathbb{N}$, heißt die Verteilung Erlang-Verteilung [45, S. 1837]. In der vorliegenden Masterarbeit meint „ \mathbb{N} “ stets die natürlichen Zahlen exklusive der Null. Die Gammafunktion erfüllt für natürliche Zahlen $\alpha \in \mathbb{N}$ die Beziehung $\Gamma(\alpha) = (\alpha - 1)!$, wobei die Notation „!“ den Fakultätsoperator repräsentiert [42, S. 308]. An dieser Stelle sei daran erinnert, dass die Gleichheit $0! = 1$ besteht [33, S. 6]. Eine Erlang-Verteilung mit Ratenparameter $\lambda \in \mathbb{R}^+$ und einer natürlichen Zahl als Formparameter, $\alpha \in \mathbb{N}$, wird durch die Dichtefunktion

$$f_{\lambda,\alpha}(x) = \begin{cases} \frac{\lambda^\alpha}{(\alpha-1)!} x^{\alpha-1} e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

charakterisiert. Die zugehörige Verteilungsfunktion ergibt sich für $x \geq 0$ durch

$$F_{\lambda,\alpha}(x) = \int_{-\infty}^x f_{\lambda,\alpha}(t) dt = \int_0^x \frac{\lambda^\alpha}{(\alpha-1)!} t^{\alpha-1} e^{-\lambda t} dt \quad (3.1)$$

$$= \left[-e^{-\lambda t} \sum_{j=0}^{\alpha-1} \frac{(\lambda t)^j}{j!} \right]_0^x = 1 - e^{-\lambda x} \sum_{j=0}^{\alpha-1} \frac{(\lambda x)^j}{j!} \quad (3.2)$$

$$= 1 - e^{-\lambda x} \sum_{j=1}^{\alpha} \frac{(\lambda x)^{j-1}}{(j-1)!} = 1 - \sum_{j=1}^{\alpha} \frac{1}{\lambda} f_{\lambda,j}(x), \quad (3.3)$$

wobei wir die Gleichheit von (3.1) und (3.2) durch die Differentiation

$$\begin{aligned}
 \frac{d}{dt} \left(-e^{-\lambda t} \sum_{j=0}^{\alpha-1} \frac{(\lambda t)^j}{j!} \right) &= \lambda e^{-\lambda t} \sum_{j=0}^{\alpha-1} \frac{(\lambda t)^j}{j!} - e^{-\lambda t} \sum_{j=1}^{\alpha-1} j \lambda^j \frac{t^{j-1}}{j!} \\
 &= e^{-\lambda t} \sum_{j=0}^{\alpha-1} \lambda^{j+1} \frac{t^j}{j!} - e^{-\lambda t} \sum_{j=1}^{\alpha-1} \lambda^j \frac{t^{j-1}}{(j-1)!} \\
 &= e^{-\lambda t} \sum_{j=0}^{\alpha-1} \lambda^{j+1} \frac{t^j}{j!} - e^{-\lambda t} \sum_{j=0}^{\alpha-2} \lambda^{j+1} \frac{t^j}{j!} = e^{-\lambda t} \lambda^\alpha \frac{t^{\alpha-1}}{(\alpha-1)!} = f_{\lambda, \alpha}(t)
 \end{aligned} \tag{3.4}$$

für $t \geq 0$ einsehen können. Siehe dazu neben [34, S. 47] auch [45, S. 1837].

Anstatt als Spezialfall der Gamma-Verteilung kann die Erlang-Verteilung alternativ über Faltungen von Dichtefunktionen der Exponentialverteilung hergeleitet werden. Siehe für die folgende Behauptung [45, S. 1835]. Für den Beweis werden die Ideen aus [41, S. 12f.] konkretisiert.

Satz 3.2 (Erlang-Verteilung als Summe von Exponentialverteilungen). *Angenommen, die Zufallsvariable X stellt eine Summe von $n \in \mathbb{N}$ unabhängigen und identisch verteilten Zufallsvariablen X_j dar. Es gelte demnach*

$$X = \sum_{j=1}^n X_j.$$

Die Zufallsvariablen X_j seien exponentialverteilt mit dem Parameter $\lambda > 0$, das heißt, dass sie die Dichte

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases} \tag{3.5}$$

besitzen, vergleiche [78, Kapitel 5.5]. Dann ist X Erlang-verteilt mit dem Ratenparameter λ und dem Formparameter n .

Beweis. Wir beweisen den Satz per Induktion über n .

Induktionsanfang: Für $n = 1$ erhalten wir $X = X_1$, das heißt X hat genau die Dichte einer Exponentialverteilung. Für die Dichte der Erlang-Verteilung mit Ratenparameter λ und Formparameter $\alpha = n = 1$ ergibt sich für $x \geq 0$

$$f_{\lambda, 1}(x) = \frac{\lambda^1}{0!} x^0 e^{-\lambda x} = \lambda e^{-\lambda x}, \tag{3.6}$$

wobei die Gleichheit $0! = 1$ ausgenutzt wird. Diese Dichtefunktion entspricht der Dichte der Exponentialverteilung (3.5). Somit folgt die Behauptung für $n = 1$.

Induktionsschritt $n \mapsto n + 1$: Die Behauptung sei für ein $n \in \mathbb{N}$ erfüllt, das heißt, die Zufallsvariable $X = \sum_{j=1}^n X_j$ sei Erlang-verteilt mit den behaupteten Parametern. Zu zeigen ist nun, dass die Zufallsvariable

$$\tilde{X} := \sum_{j=1}^{n+1} X_j = X + X_{n+1}$$

Erlang-verteilt mit Ratenparameter λ und Formparameter $n + 1$ ist, wobei die Zufallsvariablen X_j mit $j \in \{1, \dots, n + 1\}$ die Voraussetzungen aus der Behauptung erfüllen. Nach [34, S. 81f.] entspricht die Dichte der Summe von zwei unabhängigen Zufallsvariablen der Faltung der Wahrscheinlichkeitsdichten der beiden Zufallsvariablen. Die Faltung von zwei Dichtefunktionen f_1 und f_2 ist definiert durch

$$(f_1 * f_2)(x) := \int_{\mathbb{R}} f_1(\tau) f_2(x - \tau) d\tau.$$

Die Variable \tilde{X} wird durch die Summe aus der Zufallsvariablen X , die nach Voraussetzung gemäß der Dichte $f_{\lambda,n}$ verteilt ist, und der Zufallsvariablen X_{n+1} gebildet. Dabei ist X_{n+1} exponentialverteilt gemäß der Dichte $f_{\lambda,1}$, siehe hierfür den Induktionsanfang. Wir erhalten durch Anwendung der Faltungsformel, dass \tilde{X} gemäß der Dichte

$$\begin{aligned} (f_{\lambda,n} * f_{\lambda,1})(x) &= \int_{\mathbb{R}} f_{\lambda,n}(\tau) f_{\lambda,1}(x - \tau) d\tau = \int_0^x f_{\lambda,n}(\tau) f_{\lambda,1}(x - \tau) d\tau \\ &= \int_0^x \frac{\lambda^n}{(n-1)!} \tau^{n-1} e^{-\lambda\tau} \lambda e^{-\lambda(x-\tau)} d\tau = e^{-\lambda x} \frac{\lambda^{n+1}}{(n-1)!} \int_0^x \tau^{n-1} d\tau \\ &= e^{-\lambda x} \frac{\lambda^{n+1}}{(n-1)!} \left(\frac{x^n}{n} - \frac{0^n}{n} \right) = \frac{\lambda^{n+1}}{n!} x^n e^{-\lambda x} = f_{\lambda,n+1}(x) \end{aligned}$$

verteilt ist. Damit ergibt sich die Behauptung. \square

Der Induktionsanfang des Beweises verdeutlicht, dass die Erlang-Verteilung nicht lediglich als Spezialfall der Gamma-Verteilung betrachtet werden kann, sondern ebenfalls eine Verallgemeinerung der Exponentialverteilung darstellt. Da diese Eigenschaft für den Vergleich der Modelle relevant ist, halten wir sie in einem Korollar fest. Der Beweis folgt aus Gleichung (3.6). Vergleiche für die Behauptung [78, S. 216].

Korollar 3.3 (Zusammenhang von Exponentialverteilung und Erlang-Verteilung). *Eine Erlang-Verteilung mit dem Formparameter $\alpha = 1$ und dem Ratenparameter $\lambda \in \mathbb{R}^+$ entspricht einer Exponentialverteilung mit dem Parameter λ .*

Für die Evaluierung einer geeigneten Kombination von Parametern für eine spezifische Infektionskrankheit sind die ersten beiden Momente der Erlang-Verteilung von Bedeutung.

Bemerkung 3.4 (Erwartungswert und Varianz der Erlang-Verteilung). *Ist X eine Erlang-verteilte Zufallsvariable mit den Parametern $\lambda \in \mathbb{R}^+$ und $\alpha \in \mathbb{N}$, ergeben sich der Erwartungswert und die Varianz von X als*

$$\mathbb{E}(X) = \frac{\alpha}{\lambda} \quad \text{und} \quad \text{Var}(X) = \frac{\alpha}{\lambda^2}.$$

Das Ergebnis ist mittels der in Satz 3.2 besprochenen Beziehung einer Erlang-verteilten Zufallsvariable zu einer Summe von unabhängigen exponentialverteilten Zufallsvariablen leicht einzusehen. Dafür müssen die Momente für die Exponentialverteilung bekannt sein und Rechenregeln für Erwartungswert und Varianz für die Summe von unabhängigen Zufallsvariablen ausgenutzt werden. Ein vollständiger Beweis findet sich in [48, Kapitel 4.9].

Wir beweisen nun einen Satz, der die Basis der Herleitung des Linear Chain Trick-Modells bilden wird. Vergleiche für die Behauptung [45, S. 1837]. Der Beweis wird eigenständig geführt. Das Resultat kann auch so interpretiert werden, dass die Dichtefunktion der Erlang-Verteilung das Anfangswertproblem, das durch die Gleichungen (3.7) und (3.8) definiert ist, löst [80, S. 123f.].

Satz 3.5 (Dichte der Erlang-Verteilung als Lösung eines Anfangswertproblems). *Die Dichtefunktion der Erlang-Verteilung $f_{\lambda,\alpha}$ mit den Parametern $\lambda \in \mathbb{R}^+$ und $\alpha \in \mathbb{N}$ erfüllt für $x \geq 0$ die Zusammenhänge*

$$f'_{\lambda,1}(x) = -\lambda f_{\lambda,1}(x) \quad \text{mit } f_{\lambda,1}(0) = \lambda \quad (3.7)$$

$$\text{und } f'_{\lambda,\alpha}(x) = \lambda (f_{\lambda,\alpha-1}(x) - f_{\lambda,\alpha}(x)) \quad \text{mit } f_{\lambda,\alpha}(0) = 0 \quad \text{für } \alpha \geq 2. \quad (3.8)$$

Beweis. Beweisen wir zunächst die Gleichheit (3.7), indem wir die Dichte $f_{\lambda,1}(x)$ für $x \geq 0$ nach x ableiten, das heißt

$$f'_{\lambda,1}(x) = \frac{d}{dx} (\lambda e^{-\lambda x}) = -\lambda^2 e^{-\lambda x} = -\lambda f_{\lambda,1}(x).$$

Dabei können wir für $f_{\lambda,1}(x)$ wegen der in Korollar 3.3 festgehaltenen Eigenschaft direkt die Dichte einer Exponentialverteilung einsetzen. Zusätzlich gilt

$$f_{\lambda,1}(0) = \lambda e^{-\lambda \cdot 0} = \lambda.$$

Beweisen wir nun den Zusammenhang in Gleichung (3.8). Für $\alpha \in \mathbb{N}$ mit $\alpha \geq 2$ gilt

$$\begin{aligned} f'_{\lambda,\alpha}(x) &= \frac{d}{dx} \left(\frac{\lambda^\alpha}{(\alpha-1)!} x^{\alpha-1} e^{-\lambda x} \right) = \frac{\lambda^\alpha}{(\alpha-1)!} \left((\alpha-1) x^{\alpha-2} e^{-\lambda x} - \lambda x^{\alpha-1} e^{-\lambda x} \right) \\ &= \lambda \left(\frac{\lambda^{\alpha-1}}{(\alpha-2)!} x^{\alpha-2} e^{-\lambda x} - \frac{\lambda^\alpha}{(\alpha-1)!} x^{\alpha-1} e^{-\lambda x} \right) \\ &= \lambda (f_{\lambda,\alpha-1}(x) - f_{\lambda,\alpha}(x)) \end{aligned}$$

und

$$f_{\lambda,\alpha}(0) = \frac{\lambda^\alpha}{(\alpha-1)!} 0^{\alpha-1} e^{-\lambda \cdot 0} = 0.$$

Dies schließt den Beweis ab. □

3.2 Herleitung des Linear Chain Trick-Modells aus dem IDE-SIR-Modell

Im folgenden Kapitel nutzen wir die mathematischen Eigenschaften aus dem vorigen Abschnitt, um ein neues Modell basierend auf dem Linear Chain Trick einzuführen. Wir orientieren uns bei der Herleitung an der mathematischen Vorgehensweise von Feng et al. in [32, Anhang B]. Zusätzlich ergänzen wir Aspekte von Hurtado et al. [45, Kapitel 3.2], da das dort beschriebene Modell besser zu dem hier vorgeschlagenen Modell passt. Die

Herleitung erfolgt durch den Beweis, dass das IDE-SIR-Modell mit einer Erlang-verteilten infektiösen Periode zu einem System gewöhnlicher Differentialgleichungen vereinfacht werden kann. Im folgenden Satz fassen wir die Entwicklung des Linear Chain Trick-Modells aus dem IDE-SIR-Modell zusammen.

Satz 3.6 (Zusammenhang von LCT- und IDE-Modell). *Das IDE-SIR-Modell (2.3) reduziert sich unter den Parameterwahlen $\rho(t, \tau) = \rho(t)$ sowie $\phi(t, \tau) = \phi(t)$ und der Überlebensfunktion*

$$\gamma_I^R(\tau) = \sum_{j=1}^n \frac{(n\gamma)^{j-1}}{(j-1)!} \tau^{j-1} e^{-n\gamma\tau} = \frac{1}{n\gamma} \sum_{j=1}^n f_{n\gamma,j}(\tau) \quad (3.9)$$

mit einem beliebigen $n \in \mathbb{N}$ und $\gamma \in \mathbb{R}^+$ auf das gewöhnliche Differentialgleichungssystem

$$\begin{aligned} S'(t) &= -\rho(t) \phi(t) S(t) \frac{I(t)}{N} \\ I_1'(t) &= \rho(t) \phi(t) S(t) \frac{I(t)}{N} - n\gamma I_1(t) \\ I_j'(t) &= n\gamma I_{j-1}(t) - n\gamma I_j(t) \quad \text{für } j \in \{2, \dots, n\} \\ R'(t) &= n\gamma I_n(t) \end{aligned} \quad (3.10)$$

für $t \geq 0$ mit
$$I(t) = \sum_{j=1}^n I_j(t).$$

In einem späteren Kapitel zeigen wir, dass der Parameter γ dieselbe Bedeutung wie im ODE-Modell hat. Dies wird gemeinsam damit demonstriert, dass die behauptete Überlebensfunktion (3.9) mit einer Erlang-verteilten infektiösen Periode einhergeht. Die Kernaussage des Satzes 3.6 ist, dass wir anstatt eine Erlang-verteilte Aufenthaltszeit im Kompartiment I in das IDE-Modell einzusetzen, ein Modell bestehend aus einem gewöhnlichen Differentialgleichungssystem mit einer „Kette“ von n hintereinander geschalteten Subkompartimenten I_j mit $j \in \{1, \dots, n\}$ von I verwenden können. Die Gruppe I wird somit in Untergruppen beziehungsweise Subkompartimente aufgeteilt. Die Analyse des Differentialgleichungssystems zeigt, dass zwei Kettenglieder I_j stets durch lineare Terme verbunden sind. Dieser „Trick“, eine Erlang-verteilte Aufenthaltszeit durch eine Kette von Zuständen mit linearen Übergängen zu ersetzen, wird „Linear Chain Trick“ genannt. Manchmal findet sich anstelle dieser Bezeichnung der Name „Gamma Chain Trick“, da die Erlang-Verteilung, wie in Kapitel 3.1 gesehen, einen Spezialfall der Gamma-Verteilung darstellt. Wir bezeichnen das SIR-Modell (3.10), das den **Linear Chain Trick** verwendet, als „LCT-SIR-Modell“. Gelegentlich kürzen wir auch den Terminus Linear Chain Trick durch „LCT“ ab. Diese Kurzform folgt ebenso [45]. Aus der Modellgleichung (3.10) ist ersichtlich, dass Individuen gemäß dem LCT-Modell vom Zustand S nach I_1 , von I_j nach I_{j+1} für $j \in \{1, \dots, n-1\}$ und von I_n nach R übergehen können. Dies ist schematisch in Abbildung 3.1 dargestellt. Um mögliche Verwechslungen auszuschließen, sei darauf hingewiesen, dass die Notation I_j in der Literatur manchmal auch verwendet wird, um das Kompartiment I in Altersgruppen zu unterteilen. Dies wird beispielsweise in [60] umgesetzt. In der vorliegenden Arbeit ist mit der Notation stets die Aufteilung eines

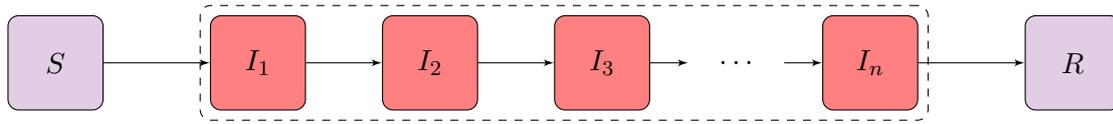


Abbildung 3.1: Aufbau des LCT-SIR-Modells. Schematische Darstellung der möglichen Übergänge zwischen Kompartimenten und Subkompartimenten gemäß dem LCT-SIR-Modell. Die zugehörigen Modellgleichungen sind in Gleichung (3.10) zu finden. Die Grafik ist an [99, Fig. 10.13] angelehnt. Rot hervorgehoben sind die $n \in \mathbb{N}$ Subkompartimente von I , in denen Individuen infektiös sind und Personen aus dem Kompartiment S anstecken können.

Krankheitszustands in eine Kette von Subkompartimenten gemeint. Die Idee des Linear Chain Tricks bewirkt, dass das SIR-Modell basierend auf Integro-Differentialgleichungen unter einer bestimmten Parametrisierung in ein gewöhnliches Differentialgleichungssystem einer höheren Dimension als die des ODE-Modells transformiert wird [35, S. 17]. Das LCT-Modell wird zwar ebenso wie das ODE-Modell (2.1) durch ein gewöhnliches Differentialgleichungssystem formuliert, jedoch ist mit „Modell basierend auf gewöhnlichen Differentialgleichungen“ oder einer ähnlichen Formulierung das ODE-Modell gemeint. Für das LCT-Modell wird entsprechend „Modell basierend auf dem Linear Chain Trick“ verwendet. Beweisen wir nun den Satz 3.6.

Beweis. Zuerst zeigen wir, dass die vorgeschlagene Parameterwahl für das IDE-SIR-Modell zulässig ist. Vergleiche Kapitel 2.3 für die Anforderungen an die Parameter. Die Wahlen $\rho(t, \tau) = \rho(t)$ und $\phi(t, \tau) = \phi(t)$ sind bei einer sinnvollen Wahl von $\rho(t)$ und $\phi(t)$ zulässig und bedürfen keiner weiteren Untersuchung. Der Parameter γ_I^R muss einige Eigenschaften erfüllen. Zunächst ist die Funktion (3.9) als Summe und Komposition stetig differenzierbarer Funktionen offensichtlich selbst stetig differenzierbar. Unter Verwendung von Satz 3.5 erhalten wir zudem

$$\gamma_I^R(0) = \sum_{j=1}^n \frac{1}{n\gamma} f_{n\gamma,j}(0) = \frac{1}{n\gamma} f_{n\gamma,1}(0) = \frac{n\gamma}{n\gamma} = 1.$$

Für die Wahl von γ_I^R aus Gleichung (3.9) für $\tau \geq 0$ kann die Gleichheit

$$\gamma_I^R(\tau) = 1 - F_{n\gamma,n}(\tau) \tag{3.11}$$

beobachtet werden, wobei $F_{n\gamma,n}$ die Verteilungsfunktion der Erlang-Verteilung aus den Gleichungen (3.1)–(3.3) mit $\lambda = n\gamma$ und $\alpha = n$ darstellt. Wegen dieses Zusammenhangs zu einer Verteilungsfunktion ist $\gamma_I^R : \mathbb{R}_0^+ \rightarrow [0, 1]$ erfüllt. Da Verteilungsfunktionen monoton steigend sind [34, S. 24], ist die Funktion für γ_I^R wie gewünscht monoton fallend. Auch die in Gleichung (2.7) spezifizierte Anforderung an γ_I^R ist erfüllt, da mit einer Erlangverteilten Zufallsvariable X mit Ratenparameter $n\gamma$ und Formparameter n folgt, dass

$$\int_0^\infty \gamma_I^R(\tau) d\tau = \int_0^\infty (1 - F_{n\gamma,n}(\tau)) d\tau = \int_0^\infty \mathbb{P}(X > \tau) d\tau \tag{3.12}$$

$$\stackrel{(*)}{=} \mathbb{E}(X) \stackrel{\text{Bemerkung 3.4}}{=} \frac{n}{n\gamma} = \frac{1}{\gamma} < \infty$$

gilt. Der Zusammenhang (*) gilt für Erwartungswerte von allgemeinen nicht-negativen Zufallsvariablen [34, S. 125 & 390] und die letzte Ungleichung folgt, da nach Definition $\gamma > 0$ erfüllt ist. Die Parametrisierung ist damit zulässig für das IDE-SIR-Modell.

Wir zeigen nun, dass sich das IDE-SIR-Modell unter der Parameterwahl auf das gewöhnliche Differentialgleichungssystem (3.10) reduziert. Betrachten wir dazu zunächst die Gleichung für das Kompartiment S des IDE-SIR-Modells (2.3). Mit $\rho(t, \tau) = \rho(t)$ und $\phi(t, \tau) = \phi(t)$ folgt unmittelbar, dass die Gleichung des IDE-Modells die Gleichung für S aus (3.10) ergibt:

$$\begin{aligned} S'(t) &= \frac{S(t)}{N} \int_{\hat{T}}^t \rho(t, t-x) \phi(t, t-x) \gamma_I^R(t-x) S'(x) dx \\ &= \frac{S(t)}{N} \rho(t) \phi(t) \int_{\hat{T}}^t \gamma_I^R(t-x) S'(x) dx = -\rho(t) \phi(t) S(t) \frac{I(t)}{N}, \end{aligned} \quad (3.13)$$

wobei bei der letzten Gleichheit die Formel für I des IDE-Modells (2.3) eingesetzt wird. Untersuchen wir als Nächstes die Gleichung für I des IDE-SIR-Modells. Wenn wir die Parameterwahl für γ_I^R aus Gleichung (3.9) in die Modellgleichung für I einsetzen, erhalten wir

$$\begin{aligned} I(t) &= - \int_{\hat{T}}^t \gamma_I^R(t-x) S'(x) dx = - \int_{\hat{T}}^t \frac{1}{n\gamma} \sum_{j=1}^n f_{n\gamma,j}(t-x) S'(x) dx \\ &= \sum_{j=1}^n \left(- \int_{\hat{T}}^t \frac{1}{n\gamma} f_{n\gamma,j}(t-x) S'(x) dx \right). \end{aligned}$$

Die letzte Gleichung gilt aufgrund der Linearität von Integralen. Wir benennen den j -ten Summanden der Summe durch $I_j(t)$, das heißt, es gelte

$$I_j(t) = - \int_{\hat{T}}^t \frac{1}{n\gamma} f_{n\gamma,j}(t-x) S'(x) dx \quad (3.14)$$

für $j \in \{1, \dots, n\}$ und

$$I(t) = \sum_{j=1}^n I_j(t).$$

Nun leiten wir die Formulierung für $I_j(t)$ ab, um zu prüfen, ob sich die Gleichungen auf die des Differentialgleichungssystems (3.10) reduzieren. Mittels Differentiation nach t der Gleichung (3.14), erhalten wir für $j = 1$ unter Verwendung der Leibniz-Regel für Parameterintegrale [57, Satz 4.29]

$$\begin{aligned} \frac{d}{dt} I_1(t) &= - \int_{\hat{T}}^t \frac{1}{n\gamma} f'_{n\gamma,1}(t-x) S'(x) dx - \frac{1}{n\gamma} f_{n\gamma,1}(0) S'(t) \\ &\stackrel{(3.7)}{=} \int_{\hat{T}}^t \frac{1}{n\gamma} n\gamma f_{n\gamma,1}(t-x) S'(x) dx - \frac{1}{n\gamma} n\gamma S'(t) \\ &= \int_{\hat{T}}^t f_{n\gamma,1}(t-x) S'(x) dx - S'(t) \\ &\stackrel{(3.14)}{=} -n\gamma I_1(t) - S'(t) \\ &\stackrel{(3.13)}{=} -n\gamma I_1(t) + \rho(t) \phi(t) S(t) \frac{I(t)}{N}. \end{aligned}$$

Dies ergibt die Gleichung für I_1 in (3.10). Hier wird der zentrale Satz 3.5 verwendet, der besagt, dass die Dichte der Erlang-Verteilung die Eigenschaft (3.7) erfüllt. Für $j \in \{2, \dots, n\}$ erhalten wir durch ein analoges Vorgehen die Gleichheit

$$\begin{aligned}
 \frac{d}{dt} I_j(t) &= - \int_{\hat{T}}^t \frac{1}{n\gamma} f'_{n\gamma,j}(t-x) S'(x) dx - \frac{1}{n\gamma} f_{n\gamma,j}(0) S'(t) \\
 &\stackrel{(3.8)}{=} - \int_{\hat{T}}^t \frac{1}{n\gamma} n\gamma (f_{n\gamma,j-1}(t-x) - f_{n\gamma,j}(t-x)) S'(x) dx \\
 &= - \int_{\hat{T}}^t f_{n\gamma,j-1}(t-x) S'(x) dx + \int_{\hat{T}}^t f_{n\gamma,j}(t-x) S'(x) dx \\
 &\stackrel{(3.14)}{=} n\gamma I_{j-1}(t) - n\gamma I_j(t),
 \end{aligned}$$

wobei abermals der Satz 3.5 mit der Eigenschaft (3.8) von Bedeutung ist. Die Berechnung liefert die behaupteten Gleichungen für I_j mit $j \in \{2, \dots, n\}$ aus (3.10). Beschäftigen wir uns zuletzt mit der Gleichung für das Kompartiment R . Es gilt laut den Modellgleichungen des IDE-Modells (2.3)

$$\begin{aligned}
 R(t) &= - \int_{\hat{T}}^t (1 - \gamma_I^R(t-x)) S'(x) dx \\
 &= - \int_{\hat{T}}^t S'(x) dx + \int_{\hat{T}}^t \gamma_I^R(t-x) S'(x) dx = -S(t) + S(\hat{T}) - I(t)
 \end{aligned}$$

und somit

$$R'(t) = -S'(t) - I'(t). \quad (3.15)$$

Die Modellgleichungen für S und I des IDE-Modells reduzieren sich, wie bereits gezeigt, auf die Formeln aus (3.10). Wenn wir $I'(t)$ mittels dieser Formeln berechnen, ergibt sich

$$\begin{aligned}
 I'(t) &= \sum_{j=1}^n I'_j(t) \stackrel{(3.10)}{=} \rho(t) \phi(t) S(t) \frac{I(t)}{N} - n\gamma I_1(t) + \sum_{j=2}^n (n\gamma I_{j-1}(t) - n\gamma I_j(t)) \\
 &= \rho(t) \phi(t) S(t) \frac{I(t)}{N} + \sum_{j=2}^n n\gamma I_{j-1}(t) - \sum_{j=1}^n n\gamma I_j(t) \\
 &= \rho(t) \phi(t) S(t) \frac{I(t)}{N} + \sum_{j=1}^{n-1} n\gamma I_j(t) - \sum_{j=1}^n n\gamma I_j(t) = \rho(t) \phi(t) S(t) \frac{I(t)}{N} - n\gamma I_n(t).
 \end{aligned}$$

Zusammen mit Gleichung (3.15) erhalten wir die folgende Formel für R im IDE-Modell unter der spezifizierten Parameterwahl

$$\begin{aligned}
 R'(t) &= -S'(t) - I'(t) = -S'(t) - \rho(t) \phi(t) S(t) \frac{I(t)}{N} + n\gamma I_n(t) \\
 &\stackrel{(3.13)}{=} -S'(t) + S'(t) + n\gamma I_n(t) = n\gamma I_n(t).
 \end{aligned}$$

Damit vereinfacht sich auch die Modellgleichung für das Kompartiment R vom IDE-SIR-Modell auf die aus Gleichung (3.10). Dies schließt den Beweis ab. \square

Zur vollständigen Spezifizierung des LCT-SIR-Modells benötigen wir Anfangswerte zum Zeitpunkt $t_0 = 0$. Diese sollten passend zu den Anfangsbedingungen des IDE-Modells gewählt werden, damit sich auch die Anfangsbedingungen des IDE-Modells unter den Annahmen aus Satz 3.6 auf die des LCT-Modells reduzieren. Unter Verwendung der für das IDE-Modell spezifizierten Anfangshistorien $S_0(t)$, $I_0(t)$ und $R_0(t)$ in (2.4) setzen wir deshalb für S und R

$$S_{t_0} = S_0(0) \quad \text{und} \quad R_{t_0} = R_0(0)$$

und für die Subkompartimente I_j mit $j \in \{1, \dots, n\}$ passend zu der Formel (3.14)

$$I_{j,t_0} = - \int_{\hat{T}}^0 S'_0(x) \frac{1}{n\gamma} f_{n\gamma,j}(-x) dx. \quad (3.16)$$

Hurtado et al. [45, S. 1836] wählen die Anfangswerte abweichend von der hier angegebenen Vorgehensweise. Für das IDE-Modell wird keine Anfangshistorie für $t \geq 0$ vorausgesetzt, sondern nur ein einzelner Anfangswert $I(0) = I_{t_0}$. Das IDE-Modell ist in der Publikation für diese Gegebenheit angepasst formuliert. Zudem wird angenommen, dass sich alle I_{t_0} Individuen im LCT-Modell zu Beginn im ersten Subkompartiment I_1 befinden. Demnach gilt dort $I_1(0) = I_{t_0}$ sowie $I_j(0) = 0$ für $j \in \{2, \dots, n\}$. In der Diskussion [45, S. 1874] wird angemerkt, dass dies eine vereinfachende Annahme ist. Einer der Vorteile des IDE-Modells und auch des LCT-Modells durch die Verwendung der Subkompartimente ist es, dass eine gewisse Zeit der Infektionsdynamik vor dem betrachteten Zeitraum in die Modellierung einfließen kann. Deshalb haben wir in der vorangegangenen Erläuterung direkt eine nicht-vereinfachte Form der Anfangsbedingungen für alle Modelle gewählt. Ein ähnliches Verfahren wird in [80, Kapitel 7.1.2] verwendet. Die Auswirkungen verschiedener Ansätze zur Bestimmung der Anfangswerte werden wir in Kapitel 5.2.1 anhand von Simulationen untersuchen.

3.3 Alternative Herleitungen und Literaturüberblick

Bei der Herleitung des LCT-Modells aus dem IDE-Modell gibt es Herangehensweisen, die von der in Kapitel 3.2 gewählten abweichen. Darüber geben wir im Folgenden einen Überblick und führen Quellen auf, in welchen Näheres zu den Ansätzen zu finden ist. Im Anschluss betrachten wir die Historie des Linear Chain Tricks und führen einige Autoren auf, die sich intensiv mit der Thematik beschäftigen. Bei den Erläuterungen wird kein Anspruch auf Vollständigkeit erhoben, da es eine umfangreiche Literatur zum LCT gibt, wenn auch in anderen biologischen Kontexten.

In unserer Herleitung sind wir vom IDE-Modell mit einer bestimmten Parameterwahl ausgegangen. Durch Benennung von Summanden und Differentiation haben wir daraus das gewöhnliche Differentialgleichungssystem (3.10) hergeleitet. Wie bereits erwähnt, ist diese Vorgehensweise in einer Publikation von Feng et al. [32, Anhang B] zu finden. Auch Greenhalgh et al. [37, Kapitel 2.3] wählen dieselbe Herangehensweise. Hurtado et al. [45, Kapitel 3.2] verfolgen einen vergleichbaren Ansatz, jedoch wird das LCT-Modell dort auf eine weniger anschauliche Weise entwickelt. In der Publikation wird zuerst eine

Integralform für I_j behauptet, die der Form in Gleichung (3.14) ähnelt. Für diese wird gezeigt, dass die Formel für I im IDE-Modell einer Summe der I_j entspricht und dass die Ableitung der Integralform das LCT-Modell (3.10) liefert. Die Integralform für die Subkompartimente I_j ergibt sich hier somit nicht im Verlauf des Beweises, sondern der Beweis wird von ihr ausgehend geführt. Eine ähnliche Vorgehensweise verwendet Smith in [80, Kapitel 7.1.2]. Hier unterscheidet sich das untersuchte Modell allerdings, sodass auch der Beweis verschieden ist. Ein anderer Beweisansatz wird von Brauer et al. [11, Kapitel 4.6] und von Champredon et al. [17, Kapitel 2.1–2.4] vorgestellt. Hier wird erst das LCT-Modell (3.10) behauptet und nachfolgend γ_I^R so bestimmt, dass das IDE-Modell mit dieser Wahl äquivalent zum LCT-Modell ist. Dafür wird ein Zwischenschritt über zusätzliche Variablen gegangen, die den Anteil der infizierten Individuen mit dem Infektionsalter τ im j -ten Subkompartiment angeben. Der in Kapitel 3.2 vorgestellte Beweis wurde ausgewählt, da sich das neue Modell auf natürliche Weise unter einer bestimmten Parameterwahl aus dem IDE-Modell ergibt. Zudem ist die passende Festlegung der Anfangsbedingungen klar. Auffällig ist die Tatsache, dass die sechs angeführten Quellen vergleichsweise aktuell sind. Obwohl der Linear Chain Trick schon Jahrzehnte bekannt ist [47, S. 253], stammt die älteste Beweisquelle von Smith aus dem Jahr 2011 [80]. Dies bemerken auch Feng et al. in [32, S. 90 & 103]. Eines ihrer Haupterkenntnisse sei der Beweis, dass sich ein IDE-Modell unter der Annahme einer Erlang-Verteilung auf ein gewöhnliches Differentialgleichungssystem reduziert, wobei ähnliche Argumente zuvor nur von Smith in [80] formuliert wurden.

Nachfolgend stellen wir den Ursprung und die Verwendung des Linear Chain Tricks in der Literatur dar. Die französischsprachigen Arbeiten von Vogel aus den Jahren 1961 [87] und 1965 [88] und die Publikation von Fargue [29] aus dem Jahr 1974 beschäftigen sich bereits mit Bedingungen, unter welchen eine allgemeine Integro-Differentialgleichung auf ein gewöhnliches Differentialgleichungssystem reduziert werden kann, siehe dazu [5, 7, 15, 45]. In der englischsprachigen Literatur hat MacDonald mit einer Veröffentlichung aus dem Jahr 1987 [66] die Technik, ein Integro-Differentialgleichungssystem mit der Annahme einer Erlang-Verteilung in ein gewöhnliches Differentialgleichungssystem zu überführen, bekannt gemacht, siehe [15, S. 5435]. Er führt den üblichen Begriff „Linear Chain Trick“ ein. MacDonald lässt die Thematik der Epidemiologie explizit aus und konzentriert sich stattdessen auf andere biologische Kontexte. Allgemein ist der Linear Chain Trick nicht auf die Epidemiologie beschränkt, sondern kann in einem breiten Anwendungsspektrum eingesetzt werden. Beispielsweise wird die Methodik in der Pharmakologie zur Modellierung der Wirkung von Arzneimitteln verwendet [15, S. 5436] oder zur Modellierung von anderen biologischen Problemstellungen, wie die Untersuchung von Populationsdynamiken, siehe dazu [8] oder S. III aus [66]. MacDonald [66, S. 3] beschreibt, dass Vogel [88] zwar die Technik hervorhebt, eine Integro-Differentialgleichung unter bestimmten Bedingungen auf ein gewöhnliches Differentialgleichungssystem zu reduzieren, jedoch habe die Vorgehensweise bereits eine lange Vorgeschichte im Kontext von Elastizitätsproblemen aus der Physik. Diese Quellen beweisen den Zusammenhang eines IDE-Modells und eines LCT-Modells nie explizit. Stattdessen beschäftigen sie sich mit der Reduzierbarkeit einer allgemeinen Integro-Differentialgleichung auf ein gewöhnliches Differentialgleichungssystem oder postulieren ein LCT-Modell ohne Herleitung oder

Beweis. Dennoch sind diese Arbeiten der Ursprung des Linear Chain Tricks.

In der Literatur sind weitere Autoren zu identifizieren, die sich mit LCT-Modellen auseinandersetzen. Im Folgenden werden einige dieser Autoren aufgeführt und alternative Benennungen des LCT-Modells vorgestellt. Diese Ausführungen können als Grundlage für eine über diese Masterarbeit hinausgehende Forschung dienen. Hurtado et al. beschäftigen sich neben den Grundlagen des Linear Chain Tricks in [44] vor allem mit einem sogenannten *Generalized Linear Chain Trick* [45, 46, 47], wovon wir einige Aspekte noch in Kapitel 3.5 betrachten werden. Feng et al. konzentrieren sich auf die Auswirkung der Verteilungsannahmen [31] und auf die Untersuchung von unterschiedlichen Annahmen bezüglich der Kompartimente unter Nutzung des LCTs [32]. Den Einfluss der Verwendung eines ODE-Modells anstatt eines LCT-Modells stellen Wearing et al. [93] heraus. Llyod ist ebenso ein bedeutsamer Autor im Kontext des LCTs, der Eigenschaften des LCT-Modells untersucht, siehe [62, 63, 64]. Der Linear Chain Trick wird hier als „method of stages“ bezeichnet [62, 64], siehe auch [5, S. 5060]. In [18, 35] werden LCT-Modelle „boxcar models“ genannt, was als „Güterwagen-Modell“ übersetzt werden kann. Dies könnte eine Verbildlichung der Darstellung in Abbildung 3.1 mit den Subkompartimenten als Güterwagen sein. Diesen Terminus führt Goudriaan in [36] für seine Methode ein, wobei ihm offenbar die früheren Arbeiten zu dieser Thematik nicht bekannt waren [50, S. 331]. Wie bereits in Kapitel 2.3 erwähnt, leiten Anderson und Watson [3] das deterministische LCT-Modell als Approximation eines stochastischen Modells her. Cassidy et al. [14, 15, 16] beschäftigen sich mit theoretischen Aspekten und Erweiterungen der LCT-Modelle. Als weitere Autoren seien noch Champredon [17, 18], Krylova und Earn [58] und im Kontext der Populationsbiologie Diekmann [25, 26] aufgeführt. Der Linear Chain Trick wird auch zur Modellierung verschiedener Aspekte der COVID-19 Pandemie verwendet. In [71, Kapitel 2.3] sowie in [49] werden Modelle betrachtet, bei welchen ein Immunitätsverlust nach einer Impfung oder nach der Gesundung möglich ist. Dieser Prozess wird mittels des Linear Chain Tricks modelliert. In [79] wird die Aufenthaltszeit im Kompartiment I mittels des LCTs als Erlang-verteilt gewählt, um nicht-pharmazeutische Interventionen zu beurteilen. Contento et al. [21] präsentieren ein Modell zur Modellierung von COVID-19, das zahlreiche Krankheitszustände umfasst und für viele davon Erlang-verteilte Aufenthaltszeiten berücksichtigt. Wie zuvor erwähnt, stellt die Zusammenstellung der Autoren, die sich mit dem Linear Chain Trick auseinandersetzen, keine vollständige Aufzählung dar. Die angeführten Quellen sowie deren Verweise bieten jedoch einen umfassenden Einblick in die behandelte Thematik.

3.4 Eigenschaften des Linear Chain Trick-Modells

Im folgenden Abschnitt untersuchen wir einige Besonderheiten und die Funktionsweise des LCT-SIR-Modells. Damit wird eine Grundlage zur Bewertung der Methode geschaffen und die Entscheidungsfindung bei der Auswahl eines geeigneten Modells für eine betrachtete Infektionskrankheit erleichtert. Anfangs stellen wir einen expliziten Bezug zwischen der Erlang-Verteilung und dem vorgestellten LCT-Modell her. Hierbei wird ein alternativer Blickwinkel auf den Linear Chain Trick hervorgehoben, der auf der Darstellung einer Erlang-Verteilung durch die Summe von Exponentialverteilungen basiert, siehe auch

Satz 3.2. In Abschnitt 3.4.2 analysieren wir den Einfluss der Parameterwahl auf die Funktionsweise des LCT-Modells. Hierbei konzentrieren wir uns vornehmlich auf die Bedeutung der Anzahl der Subkompartimente n . Zudem untersuchen wir für den Linear Chain Trick beziehungsweise für die Erlang-Verteilung die Eigenschaft der Gedächtnislosigkeit, die für das ODE-Modell in Kapitel 2.4 analysiert wurde, und betrachten die Größe der erwarteten restlichen Aufenthaltszeit. Durch Nutzung der Erkenntnisse dieses Kapitels können wir eine Zusammenfassung der Beziehung des LCT-Modells zu den beiden in Kapitel 2 eingeführten Modellen präsentieren. Obgleich das LCT-Modell bereits im Abschnitt 3.2 aus dem IDE-Modell abgeleitet wurde, ist ein gemeinsamer Vergleich aller drei Modelle gewinnbringend.

3.4.1 Bezug zur Erlang-Verteilung

Wir sehen zunächst ein, dass aus der Parametrisierung in Satz 3.6, wie bereits mehrfach behauptet, eine Erlang-verteilte infektiöse Periode folgt. Dazu sei X die Zufallsvariable aus Kapitel 2.3, welche die Dauer der infektiösen Periode von Individuen beschreibt. In der Gleichung (2.5) haben wir für das IDE-SIR-Modell festgestellt, dass für die Verteilungsfunktion F_X von X für $\tau \geq 0$

$$F_X(\tau) = \mathbb{P}(X \leq \tau) = 1 - \mathbb{P}(X > \tau) = 1 - \gamma_I^R(\tau)$$

gilt. Das LCT-SIR-Modell ist nach Satz 3.6 äquivalent zum IDE-SIR-Modell unter der Wahl $\rho(t, \tau) = \rho(t)$, $\phi(t, \tau) = \phi(t)$ und

$$\gamma_I^R(\tau) = 1 - F_{n\gamma, n}(\tau),$$

wie in Gleichung (3.11) festgestellt wurde. Wenden wir den Zusammenhang (2.5) auf das LCT-Modell beziehungsweise auf die spezifische Wahl von γ_I^R an, erhalten wir

$$F_X(\tau) \stackrel{(2.5)}{=} 1 - \gamma_I^R(\tau) \stackrel{(3.11)}{=} 1 - (1 - F_{n\gamma, n}(\tau)) = F_{n\gamma, n}(\tau).$$

Die Zufallsvariable, welche die Dauer der infektiösen Periode für das LCT-Modell beschreibt, ist damit Erlang-verteilt mit dem Ratenparameter $\lambda = n\gamma$ und dem Formparameter $\alpha = n$.

Nach [5, S. 5060] basiert der Linear Chain Trick darauf, dass die Erlang-Verteilung eine Summe von unabhängigen Exponentialverteilungen ist. Die Eigenschaft der Erlang-Verteilung haben wir in Satz 3.2 besprochen. Diese Behauptung betrachten wir ausführlich. Durch das Verständnis der Herangehensweise wird deutlich, wieso eine Kette von identischen linearen Subkompartimenten stets eine Erlang-verteilter Gesamtaufenthaltszeit ergibt. Wir können den Beweis der Erlang-verteilter Aufenthaltszeit folglich ohne einen expliziten Zusammenhang zu einem IDE-basierten Modell führen. Dies erleichtert die Erweiterung von allgemeinen ODE-Modellen, beispielsweise mit zusätzlichen Krankheitszuständen, um Erlang-verteilter Aufenthaltszeiten. Zunächst zeigen wir, dass die linearen Übergänge zwischen den Subkompartimenten im LCT-SIR-Modell jeweils exponentialverteilte Aufenthaltszeiten in den einzelnen Subkompartimenten I_j implizieren. Dieses Ziel formalisieren wir in Form eines Satzes. Der Beweis orientiert sich grob an [72, S. 31–38].

Satz 3.7 (Verteilung der Aufenthaltszeit in einem Subkompartiment). *Seien X_j für $j \in \{1, \dots, n\}$ die Zufallsvariablen, die jeweils die Aufenthaltszeit im Subkompartiment I_j beschreiben. Diese Zufallsvariablen X_j sind für alle $j \in \{1, \dots, n\}$ exponentialverteilt mit dem Parameter $n\gamma$.*

Beweis. Sei $\sigma_j(t)$ für $j \in \{1, \dots, n\}$ die Anzahl der Individuen, die zum Zeitpunkt t in das Subkompartiment I_j eintreten. Dann gilt nach den Formeln des LCT-Modells (3.10), dass

$$\sigma_1(t) = \rho(t) \phi(t) S(t) \frac{I(t)}{N} = -S'(t)$$

und $\sigma_j(t) = n\gamma I_{j-1}(t)$ für $j \in \{2, \dots, n\}$.

Setzen wir diese Definition in die Gleichungen für I_j aus (3.10) ein, ergibt sich für alle $j \in \{1, \dots, n\}$

$$I_j'(t) = \sigma_j(t) - n\gamma I_j(t). \tag{3.17}$$

Für die Subkompartimente I_j formulieren wir unter Verwendung der Größen $\sigma_j(t)$ eine Integralgleichung, um aus dieser die Exponentialverteilung der Aufenthaltszeiten abzuleiten. Interpretieren wir $\gamma_{I_j}^{I_j+1}(\tau) = \mathbb{P}(X_j > \tau)$ analog zu γ_I^R als Anteil der Individuen, die $\tau \geq 0$ Tage nach dem Eintritt in I_j noch im Subkompartiment verweilen und nicht nach I_{j+1} übergegangen sind. Zur einfacheren Notation setzen wir $I_{n+1} = R$. Aufgrund desselben Zusammenhangs zu einer Verteilung der Aufenthaltszeit erfülle $\gamma_{I_j}^{I_j+1}$ dieselben Anforderungen wie die an γ_I^R aus Kapitel 2.3. Dann entspricht $\gamma_{I_j}^{I_j+1}(\tau) \sigma_j(t - \tau)$ für $\tau \geq 0$ und $j \in \{1, \dots, n\}$ der Anzahl der Individuen, die sich zum Zeitpunkt t noch im Subkompartiment I_j befinden und genau vor τ Tagen in I_j eingetreten sind. Für die Gesamtzahl der Personen, die zum Zeitpunkt t in I_j sind, betrachten wir alle potentiellen Eintrittspunkte $t - \tau$, sodass sich

$$I_j(t) = \int_0^\infty \gamma_{I_j}^{I_j+1}(\tau) \sigma_j(t - \tau) d\tau = \int_{\hat{T}}^t \gamma_{I_j}^{I_j+1}(t - x) \sigma_j(x) dx \tag{3.18}$$

ergibt. Hier wird die Substitution $x = t - \tau$ durchgeführt und die Definition vom Zeitpunkt \hat{T} eingesetzt. Wir bestimmen $\gamma_{I_j}^{I_j+1}$ so, dass diese Formulierung zu Gleichung (3.17) äquivalent ist, um die Verteilungsfunktion der X_j daraus zu entwickeln. Wenn wir die Formulierung (3.18) ableiten, erhalten wir unter Verwendung der Leibniz-Regel für Parameterintegrale [57, Satz 4.29]

$$I_j'(t) = \int_{\hat{T}}^t \gamma_{I_j}^{I_j+1}'(t - x) \sigma_j(x) dx + \gamma_{I_j}^{I_j+1}(0) \sigma_j(t).$$

Aus der Bedingung $\gamma_{I_j}^{I_j+1}(0) = 1$ folgt, dass diese Gleichung der in (3.17) entspricht, falls die Funktion das Anfangswertproblem

$$\gamma_{I_j}^{I_j+1}'(\tau) = -n\gamma \gamma_{I_j}^{I_j+1}(\tau) \quad \text{für } \tau \geq 0$$

und $\gamma_{I_j}^{I_j+1}(0) = 1$

löst. Unter dieser Voraussetzung gilt

$$I_j'(t) = \int_{\hat{T}}^t \gamma_{I_j}^{I_j+1}'(t-x) \sigma_j(x) dx + \gamma_{I_j}^{I_j+1}(0) \sigma_j(t) = -n\gamma \int_{\hat{T}}^t \gamma_{I_j}^{I_j+1}(t-x) \sigma_j(x) dx + \sigma_j(t) \\ \stackrel{(3.18)}{=} \sigma_j(t) - n\gamma I_j(t),$$

sodass sich die gewünschte Modellgleichung (3.17) für alle $j \in \{1, \dots, n\}$ ergibt. Man kann zeigen, dass die eindeutige Lösung des Anfangswertproblems durch

$$\gamma_{I_j}^{I_j+1}(\tau) = e^{-\tau n\gamma}$$

für alle $j \in \{1, \dots, n\}$ gegeben ist, vergleiche dazu [72, S. 23f.]. Entsprechend gilt $\mathbb{P}(X_j > \tau) = e^{-\tau n\gamma}$ für das LCT-Modell mit den passend definierten Zufallsvariablen aus der Behauptung. Diese Form bedeutet, dass die X_j exponentialverteilt mit dem Parameter $n\gamma$ sind, was in ähnlicher Weise im Kontext des ODE-SIR-Modells in Satz 2.3 festgestellt wurde. \square

Der Satz 3.7 impliziert, dass die zuvor beschriebene Zufallsvariable X für die gesamte Aufenthaltszeit in der Menge der Subkompartimente von I Erlang-verteilt mit den Parametern $\lambda = n\gamma$ und $\alpha = n$ ist. Nach den Definitionen entspricht X der Summe der Zufallsvariablen X_j , das heißt $X = \sum_{j=1}^n X_j$. Wie wir im Satz 3.7 festgestellt haben, sind die X_j exponentialverteilt mit dem Parameter $n\gamma$. Die Zufallsvariablen X_j mit $j \in \{1, \dots, n\}$ sind außerdem unabhängig. Laut Satz 3.2 ist X folglich Erlang-verteilt mit Ratenparameter $n\gamma$ und Formparameter n . Wir haben damit die Verteilung der infektiösen Periode im LCT-Modell (3.10) hergeleitet, ohne explizit die IDE-SIR-Formulierung (2.3) zu benötigen. Dafür mussten wir einsehen, dass die Übergänge, die durch lineare Terme beschrieben werden, mit exponentialverteilten Aufenthaltszeiten in den Subkompartimenten einhergehen. Auf diese Art ist erkennbar, dass der Fakt aus Satz 3.2, dass die Erlang-Verteilung als Summe von exponentialverteilten Zufallsvariablen mit identischen Parametern dargestellt werden kann, eng mit dem Linear Chain Trick zusammenhängt.

In [64, S. 848] und [65, S. 684] findet sich die Behauptung, dass die Subkompartimente I_j mit $j \in \{1, \dots, n\}$ nur mathematische Mittel sind, um eine Erlang-verteilte infektiöse Periode zu erreichen. Den Subkompartimenten müsse keine biologische Bedeutung zugesprochen werden können. In komplexeren Modellen könnten einzelne Subkompartimente oder eine Gruppe von Subkompartimenten aber eine biologische Rolle übernehmen, wie zum Beispiel die Simulation einer Latenzzeit, in der Individuen infiziert, aber noch nicht infektiös sind, siehe [63, S. 60] mit [54, S. 85]. Die Beschreibung der Subkompartimente als rein mathematisches Mittel kann kritisch betrachtet werden, wie die folgende Ausführung verdeutlicht. Bestimmen wir die erwartete verbleibende Aufenthaltszeit im Krankheitszustand I für ein beliebiges Individuum aus dem Subkompartiment I_j mit $j \in \{1, \dots, n\}$. Nach Satz 3.7 sind die Aufenthaltszeiten in den Subkompartimenten jeweils exponentialverteilt mit dem Parameter $n\gamma$. Die erwartete Aufenthaltszeit in einem beliebigen Subkompartiment entspricht demnach $(n\gamma)^{-1}$, was sich aus dem Erwartungswert einer Exponentialverteilung mit dem passenden Parameter ergibt, vergleiche Bemerkung 3.4

für den Spezialfall $\alpha = 1$. Aufgrund der Eigenschaft der Gedächtnislosigkeit der Exponentialverteilung gleicht die restliche erwartete Aufenthaltszeit in I_j , unabhängig von der bereits in I_j verbrachten Zeit, dem Erwartungswert $(n\gamma)^{-1}$, siehe dazu Kapitel 2.4. Die erwartete restliche Aufenthaltszeit im Zustand I eines Individuums aus der Gruppe I_j entspricht der Summe aus der erwarteten restlichen Aufenthaltszeit in I_j und den erwarteten Aufenthaltszeiten in den übrigen Subkompartimenten I_{j+1} bis I_n . Demnach ergibt sich diese erwartete verbleibende Aufenthaltszeit durch

$$\frac{1}{n\gamma} + \sum_{m=j+1}^n \frac{1}{n\gamma} = (n-j+1) \frac{1}{n\gamma}.$$

Die verbleibende Dauer der infektiösen Periode ist für verschiedene Subkompartimente unterschiedlich, da das Ergebnis von j abhängt. Den Subkompartimenten kann in dem Sinne eine biologische Bedeutung zugesprochen werden, dass die erwartete restliche Aufenthaltszeit im Krankheitszustand I für steigendes j sinkt.

3.4.2 Einfluss der Parameterwahl

Im folgenden Abschnitt untersuchen wir, wie die Wahl der Parameter n und γ die Funktionsweise des LCT-Modells beziehungsweise die Annahme bezüglich der Verteilung der infektiösen Periode beeinflusst. Wir wollen die Bedeutung der Parameter besser verstehen, um eine fundierte Parameterwahl bei der Simulation der Ausbreitung einer Infektionskrankheit treffen zu können. Sei X die Erlang-verteilte Zufallsvariable mit Ratenparameter $n\gamma$ und Formparameter n für die infektiöse Periode im LCT-Modell. Mit der allgemeinen Formel für Erwartungswerte von Erlang-verteilten Zufallsvariablen aus Bemerkung 3.4 ergibt sich für die erwartete infektiöse Periode unter der Parameterwahl des LCT-Modells

$$\mathbb{E}(X) = \frac{\alpha}{\lambda} = \frac{n}{n\gamma} = \frac{1}{\gamma}. \quad (3.19)$$

Entsprechend stimmt die Bedeutung von γ im LCT-Modell mit der Definition von γ in Kapitel 2.2 für das ODE-SIR-Modell überein. Die erwartete infektiöse Periode ist von der Wahl der Anzahl n der Subkompartimente von I unabhängig. Die Varianz der infektiösen Periode ergibt sich abermals mittels Bemerkung 3.4 durch

$$\text{Var}(X) = \frac{\alpha}{\lambda^2} = \frac{n}{(n\gamma)^2} = \frac{1}{n\gamma^2}. \quad (3.20)$$

Die Anzahl der Subkompartimente n ist irrelevant für die erwartete infektiöse Periode, beeinflusst aber die Varianz der Verteilung. Die Varianz steigt zudem mit einer zunehmenden Aufenthaltszeit T_I in I . Verändern wir γ nicht, bleibt die erwartete infektiöse Periode konstant und wir können die Varianz durch den Parameter n steuern, siehe [92, S. 546]. Im Folgenden betrachten wir diesen Einfluss des Parameters n auf das Modell beziehungsweise auf die Verteilung der infektiösen Periode und halten den Erwartungswert und den Parameter γ konstant.

In Abbildung 3.2 ist die Dichte der Erlang-Verteilung $f_{n\gamma,n}(\tau)$ für verschiedene n dargestellt, um die Rolle der Wahl der Anzahl der Subkompartimente zu veranschaulichen.

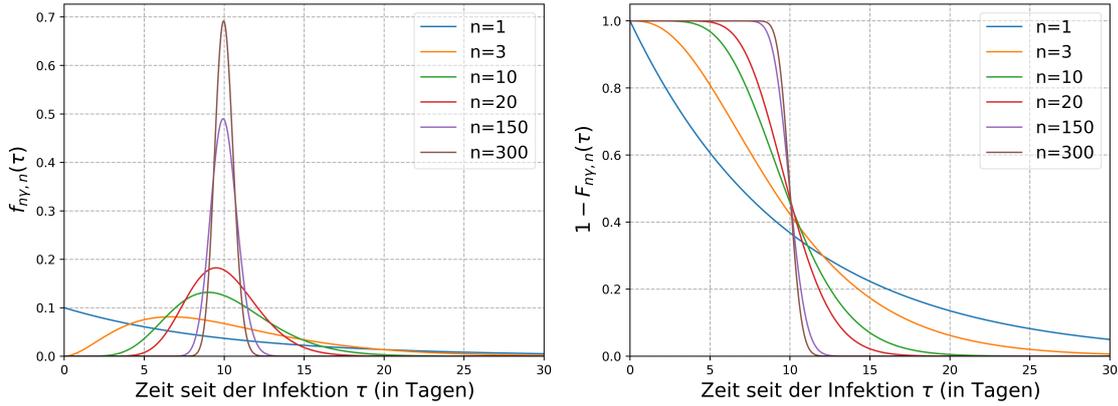


Abbildung 3.2: Dichte und Überlebensfunktion der Erlang-Verteilung. Darstellung der Dichtefunktion $f_{n\gamma,n}(\tau)$ der Erlang-Verteilung (links) und der zugehörigen Überlebensfunktion $1 - F_{n\gamma,n}(\tau)$ (rechts) für verschiedenen Wahlen des Parameters n . Für alle Funktionen ist $\gamma = \frac{1}{10}$ gesetzt. Ähnliche Abbildungen finden sich beispielsweise in [11, Fig. 3.6] oder in [31, Fig. 3] und [93, Figure 1 A].

Außerdem ist der zugehörige Verlauf der Überlebensfunktion $\gamma_I^R(\tau) = 1 - F_{n\gamma,n}(\tau)$ dargestellt. Über diese Funktion beeinflusst n im LCT-Modell als Spezialfall des IDE-Modells die Verteilung der infektiösen Periode. In den Abbildungen können wir beobachten, dass die Verteilung mit steigendem n immer näher um den Erwartungswert konzentriert ist, vergleiche [63, S. 60]. Dies bestätigt die Gleichung (3.20) für die Varianz. Die Varianz wird für größere Werte von n immer kleiner. In unserer Anwendung bedeutet dies, dass die Varianz der infektiösen Periode mit steigender Anzahl an Subkompartimenten im LCT-Modell sinkt. Somit konzentrieren sich die Aufenthaltszeiten der Individuen im Kompartiment I für steigende n mehr und mehr um den Mittelwert, siehe [5, S. 5060]. Die Form der Dichtefunktion wird zudem für wachsendes n immer symmetrischer [41, S. 24]. Für hinreichend große Werte von n nähert sich die gewählte Verteilung einer Normalverteilung, siehe [3, S. 191] oder [85, S. 11]. Dies ist eine relevante Eigenschaft, da eine normalverteilte infektiöse Periode beispielsweise für Masern realistisch ist [51, S. 65] und wir diese somit durch eine Erlang-Verteilung mit hohem n approximieren könnten.

Wir betrachten die beiden Extremfälle $n = 1$ und $n \rightarrow \infty$ genauer. In Korollar 3.3 haben wir festgestellt, dass die Erlang-Verteilung mit dem Formparameter $\alpha = 1$ die Exponentialverteilung ergibt. Mit $n = \alpha = 1$ erhalten wir folglich eine exponentialverteilte infektiöse Periode mit dem Parameter γ . Für die Wahl von γ_I^R im IDE-Modell, die äquivalent zu nur einem Subkompartiment im LCT-Modell ist, ergibt sich nach Satz 3.6

$$\gamma_I^R(\tau) = \sum_{j=1}^1 \frac{(1\gamma)^{j-1}}{(j-1)!} \tau^{j-1} e^{-1\gamma\tau} = e^{-\gamma\tau}.$$

In Satz 2.3 haben wir festgehalten, dass sich das IDE-SIR-Modell mit genau dieser Wahl von γ_I^R sowie $\rho(t, \tau) = \rho(t)$ und $\phi(t, \tau) = \phi(t)$ auf das ODE-SIR-Modell reduziert. Demnach ist das LCT-Modell mit $n = 1$ dasselbe wie das ODE-Modell. Dies ergibt sich

auch durch die Betrachtung der Modellgleichungen (3.10). Mit $n = 1$ erhalten wir

$$I(t) = I_1(t)$$

und

$$\begin{aligned} S'(t) &= -\rho(t) \phi(t) S(t) \frac{I(t)}{N} \\ I_1'(t) = I'(t) &= \rho(t) \phi(t) S(t) \frac{I(t)}{N} - 1 \gamma I_1(t) = \rho(t) \phi(t) S(t) \frac{I(t)}{N} - \gamma I(t) \\ R'(t) &= 1 \gamma I_1(t) = \gamma I(t), \end{aligned}$$

was den ODE-SIR-Modellgleichungen (2.1) entspricht. Für $n \rightarrow \infty$ gilt für die Varianz der Zufallsvariablen X , welche die Aufenthaltszeit im Kompartiment I bestimmt,

$$\lim_{n \rightarrow \infty} \text{Var}(X) = \lim_{n \rightarrow \infty} \frac{1}{n \gamma^2} = 0.$$

Nach wie vor wird der Erwartungswert konstant bei einem Wert von $\mathbb{E}(X) = \gamma^{-1} = T_I$ gehalten. Dabei sei T_I wie in Kapitel 2.2 durch $T_I = \gamma^{-1}$ als erwartete Dauer der infektiösen Periode definiert. Demnach entspricht die Erlang-Verteilung für $n \rightarrow \infty$ der Einpunkt-Verteilung beziehungsweise der Dirac-Verteilung (vergleiche dazu [34, S. 14]) im Punkt T_I . Die Konvergenz ist ebenfalls in Abbildung 3.2 zu beobachten. Die zugehörige Dichtefunktion ist

$$f_X(\tau) = \begin{cases} 1, & \text{für } \tau = T_I \\ 0, & \text{für } \tau \neq T_I \end{cases}$$

und die Verteilungsfunktion ergibt sich als

$$F_X(\tau) = \begin{cases} 0, & \text{für } \tau < T_I \\ 1, & \text{für } \tau \geq T_I. \end{cases}$$

Vergleiche für die vorangegangenen Erläuterungen [41, S. 25]. Wenn wir die Dirac-Verteilung beziehungsweise die zugehörige Überlebensfunktion $\gamma_I^R(\tau) = 1 - F_X(\tau)$ in das IDE-SIR-Modell (2.3) einsetzen, erhalten wir

$$\begin{aligned} I(t) &= - \int_{T_I}^t (1 - F_X(t-x)) S'(x) dx = - \int_{t-T_I}^t S'(x) dx \\ &= -(S(t) - S(t - T_I)). \end{aligned}$$

Individuen bleiben somit eine feste Zeit von T_I Tagen im Kompartiment I . Hier liegt eine „diskrete Verzögerung“ (engl. *discrete delay*) beschrieben durch eine *discrete delay differential equation* vor, siehe [44, Kapitel 3.5.4] oder [80, S. 120]. Im Gegensatz dazu steht der Terminus „distributed delay model“ aus Bemerkung 2.2 für das IDE-Modell. In der Gleichung für I wird im Fall $n \rightarrow \infty$ somit eine diskrete Verzögerung anstatt einer Gewichtung von Verzögerungstermen verwendet, vergleiche [80, S. 9].

Der zu der betrachteten Infektionskrankheit passende Wert für n sollte anhand von epidemiologischen Daten bestimmt werden. Dies gilt selbstverständlich auch für andere Parameter, wie beispielsweise T_I . Für einige Krankheiten ist eine Erlang-verteilte infektiöse Periode mit einem Parameter $n > 1$ eine gute Wahl und für andere ist eventuell eine Exponentialverteilung beziehungsweise das ODE-Modell eine ausreichende Approximation. Zur Beurteilung benötigen wir ein gutes Verständnis der Vor- und Nachteile der vorgestellten Modelle, wozu die vorliegende Masterarbeit einen Beitrag leisten soll. Zusammen mit ausreichendem Wissen über die betrachtete Krankheit können dann ein geeignetes Modell und die entsprechenden Parameter bestimmt werden. Siehe dazu [31, S. 1535]. Wenn der Erwartungswert und die Varianz der infektiösen Periode aus epidemiologischen Daten bekannt sind, können wir die passenden Parameter n und γ des LCT-SIR-Modells berechnen. Dafür formen wir die Formeln für den Erwartungswert (3.19) und die Varianz (3.20) um, was die Zusammenhänge

$$\gamma = \frac{1}{\mathbb{E}(X)} \quad \text{und} \quad n = \frac{(\mathbb{E}(X))^2}{\text{Var}(X)}$$

ergibt. Zusätzlich muss gesichert sein, dass $\alpha = n \in \mathbb{N}$ gilt, damit die Verteilung eine Erlang-Verteilung ist. Dafür kann der obige Ausdruck für n auf die nächste natürliche Zahl gerundet werden, sodass die aus epidemiologischen Daten bestimmte Varianz eventuell nur approximativ erreicht werden kann. Siehe dazu [45, S. 1837]. In [93, Table 1] findet sich eine Zusammenstellung von Schätzungen auf der Basis von epidemiologischen Daten des Parameters n für die infektiöse Periode von vier verschiedenen Infektionskrankheiten. Die dortigen Ausführungen berücksichtigen zusätzlich eine Latenzzeit (siehe zur Definition S. 29). Die Schätzungen der Anzahl der Subkompartimente für die Gruppe I reichen von 3 für SARS bis hin zu etwa 20 bei Masern.

3.4.3 Untersuchung der Eigenschaft der Gedächtnislosigkeit

Von Brauer et al. [11, S. 99] wird behauptet, dass sich Simulationen mit Modellen mit Exponentialverteilungen und mit Erlang-Verteilungen vor allem aufgrund der Eigenschaft der Gedächtnislosigkeit der Exponentialverteilung voneinander unterscheiden. Dieses Merkmal haben wir für das ODE-Modell in Kapitel 2.4 besprochen. Eine direkte Folgerung aus der Gedächtnislosigkeit ist, dass die erwartete restliche Aufenthaltszeit von der bereits in I verbrachten Zeit $\tau \geq 0$ unabhängig ist. Die erwartete restliche Aufenthaltszeit ist durch

$$\mathcal{M}_I(\tau) = \int_0^\infty \mathbb{P}(X > s + \tau \mid X > \tau) ds$$

beschrieben, siehe Gleichung (2.10). Hierbei ist X die in Kapitel 2.3 eingeführte Zufallsvariable. Wir verdeutlichen den Unterschied des ODE- und des LCT-Modells, indem wir die Größe der erwarteten restlichen Aufenthaltszeit in I für das LCT-Modell genauer untersuchen. Dazu zeigen wir, dass $\mathcal{M}_I(\tau)$ im LCT-SIR-Modell für eine Anzahl von $n \geq 2$ Subkompartimenten streng monoton fallend ist. Diese Behauptung halten wir in Form eines Satzes fest, wobei wir für den Beweis die Aussagen aus [31, S. 1529] und [11, S. 99] zeigen. Da die erwartete verbleibende Aufenthaltszeit $\mathcal{M}_I(\tau)$ damit nicht von τ

unabhängig ist, folgt damit insbesondere, dass die Erlang-Verteilung für $n \geq 2$ nicht gedächtnislos ist.

Satz 3.8 (Erwartete restliche Aufenthaltszeit in I im LCT-Modell). *Die Funktion $\mathcal{M}_I(\tau)$ ist unter der Annahme einer Erlang-verteilten infektiösen Periode mit $\lambda = n\gamma$ und $\alpha = n$ mit $n \geq 2$ streng monoton fallend, das heißt, es gilt $\mathcal{M}'_I(\tau) < 0$ für $\tau \geq 0$.*

Beweis. Sei X die Zufallsvariable, welche die Erlang-verteilte Dauer der infektiösen Periode im LCT-Modell mit den vorausgesetzten Parametern beschreibt. Zunächst setzen wir diese Annahme der Erlang-Verteilung in die Definition von $\mathcal{M}_I(\tau)$ ein und formen die resultierende Gleichung um:

$$\begin{aligned} \mathcal{M}_I(\tau) &= \int_0^\infty \frac{\mathbb{P}(X > s + \tau)}{\mathbb{P}(X > \tau)} ds = \frac{1}{\mathbb{P}(X > \tau)} \int_0^\infty \mathbb{P}(X > s + \tau) ds \\ &\stackrel{(**)}{=} \frac{1}{\mathbb{P}(X > \tau)} \int_\tau^\infty \mathbb{P}(X > x) dx \stackrel{(2.5)\&(3.9)}{=} \frac{1}{\mathbb{P}(X > \tau)} \int_\tau^\infty \frac{1}{n\gamma} \sum_{j=1}^n f_{n\gamma,j}(x) dx \\ &= \frac{1}{\mathbb{P}(X > \tau)} \frac{1}{n\gamma} \sum_{j=1}^n \int_\tau^\infty f_{n\gamma,j}(x) dx \\ &\stackrel{(3.4)}{=} \frac{1}{\mathbb{P}(X > \tau)} \frac{1}{n\gamma} \sum_{j=1}^n \left[-e^{-n\gamma x} \sum_{m=0}^{j-1} \frac{(n\gamma x)^m}{m!} \right]_\tau^\infty \\ &= \frac{1}{\mathbb{P}(X > \tau)} \frac{1}{n\gamma} \sum_{j=1}^n \left(\lim_{x \rightarrow \infty} \left(-e^{-n\gamma x} \sum_{m=0}^{j-1} \frac{(n\gamma x)^m}{m!} \right) + e^{-n\gamma \tau} \sum_{m=0}^{j-1} \frac{(n\gamma \tau)^m}{m!} \right), \end{aligned}$$

wobei bei Gleichung (**) die Substitution $x = s + \tau$ durchgeführt wird. Mit den Regeln von de l'Hospital [33, Satz 10 (2)] lässt sich folgern, dass der betrachtete Grenzwert gegen null konvergiert. Infolgedessen ergibt sich weiter

$$\begin{aligned} \mathcal{M}_I(\tau) &= \frac{1}{\mathbb{P}(X > \tau)} \frac{1}{n\gamma} \sum_{j=1}^n e^{-n\gamma \tau} \sum_{m=0}^{j-1} \frac{(n\gamma \tau)^m}{m!} \stackrel{(2.5)\&(3.9)}{=} \frac{1}{n\gamma} \frac{e^{-n\gamma \tau} \sum_{j=1}^n \sum_{m=0}^{j-1} \frac{(n\gamma \tau)^m}{m!}}{\sum_{j=1}^n \frac{(n\gamma)^{j-1}}{(j-1)!} \tau^{j-1} e^{-n\gamma \tau}} \\ &= \frac{1}{n\gamma} \frac{\sum_{j=0}^{n-1} \sum_{m=0}^j \frac{(n\gamma)^m}{m!} \tau^m}{\sum_{j=0}^{n-1} \frac{(n\gamma)^j}{j!} \tau^j}. \end{aligned}$$

Leiten wir diese Formel für die erwartete restliche Aufenthaltsdauer $\mathcal{M}_I(\tau)$ nach τ ab, ergibt sich

$$\begin{aligned} \mathcal{M}'_I(\tau) &= \frac{1}{n\gamma} \frac{d}{d\tau} \left(\frac{\sum_{j=0}^{n-1} \sum_{m=0}^j \frac{(n\gamma)^m}{m!} \tau^m}{\sum_{j=0}^{n-1} \frac{(n\gamma)^j}{j!} \tau^j} \right) \\ &= \frac{1}{n\gamma} \frac{1}{\left(\sum_{j=0}^{n-1} \frac{(n\gamma)^j}{j!} \tau^j \right)^2} \left(\left(\sum_{j=0}^{n-1} \sum_{m=1}^j \frac{(n\gamma)^m}{(m-1)!} \tau^{m-1} \right) \left(\sum_{\ell=0}^{n-1} \frac{(n\gamma)^\ell}{\ell!} \tau^\ell \right) \right. \\ &\quad \left. - \left(\sum_{j=0}^{n-1} \sum_{m=0}^j \frac{(n\gamma)^m}{m!} \tau^m \right) \left(\sum_{\ell=1}^{n-1} \frac{(n\gamma)^\ell}{(\ell-1)!} \tau^{\ell-1} \right) \right). \end{aligned}$$

Zur Differentiation wird die Quotientenregel [33, S. 170] angewendet. Wir definieren den Vorfaktor als

$$a := \frac{1}{n\gamma} \frac{1}{\left(\sum_{j=0}^{n-1} \frac{(n\gamma)^j}{j!} \tau^j\right)^2} = \frac{1}{n\gamma} \frac{1}{\left(1 + \sum_{j=1}^{n-1} \frac{(n\gamma)^j}{j!} \tau^j\right)^2}.$$

Für den Vorfaktor gilt für $\tau \geq 0$ die Ungleichung $a > 0$, da gemäß Definition n und γ jeweils größer als null sind. Wenn wir die obige Formel für die Ableitung weiter umformen, erhalten wir

$$\begin{aligned} \mathcal{M}'_I(\tau) &= a \left(\sum_{j=0}^{n-1} \sum_{m=1}^j \sum_{\ell=0}^{n-1} \frac{(n\gamma)^{m+\ell}}{(m-1)! \ell!} \tau^{m-1+\ell} - \sum_{j=0}^{n-1} \sum_{m=0}^j \sum_{\ell=1}^{n-1} \frac{(n\gamma)^{m+\ell}}{m! (\ell-1)!} \tau^{m+\ell-1} \right) \\ &= a \left(\sum_{j=0}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^{n-1} (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{(m-\ell)}{m! \ell!} \right. \\ &\quad \left. + \sum_{j=0}^{n-1} \sum_{m=1}^j \frac{(n\gamma)^m}{(m-1)!} \tau^{m-1} - \sum_{j=0}^{n-1} \sum_{\ell=1}^{n-1} \frac{(n\gamma)^\ell}{(\ell-1)!} \tau^{\ell-1} \right) \\ &= a \left(\sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^{n-1} (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{(m-\ell)}{m! \ell!} - \sum_{j=0}^{n-2} \sum_{\ell=j+1}^{n-1} \frac{(n\gamma)^\ell}{(\ell-1)!} \tau^{\ell-1} \right). \end{aligned}$$

Für $n \geq 2$ und $\tau \geq 0$ gilt

$$\sum_{j=0}^{n-2} \sum_{\ell=j+1}^{n-1} \frac{(n\gamma)^\ell}{(\ell-1)!} \tau^{\ell-1} = \sum_{j=1}^{n-2} \sum_{\ell=j+1}^{n-1} \frac{(n\gamma)^\ell}{(\ell-1)!} \tau^{\ell-1} + \sum_{\ell=2}^{n-1} \frac{(n\gamma)^\ell}{(\ell-1)!} \tau^{\ell-1} + n\gamma > 0,$$

da nach Definition $\gamma > 0$ gilt. Die Summen ergeben stets nicht-negative Werte, da sich auch die Fakultäten stets zu positiven Werten evaluieren. Gemeinsam mit der Feststellung $a > 0$ folgt

$$\begin{aligned} \mathcal{M}'_I(\tau) &< a \left(\sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^{n-1} (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{(m-\ell)}{m! \ell!} \right) \\ &= a \left(\sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^j (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{(m-\ell)}{m! \ell!} + \sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=j+1}^{n-1} (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{(m-\ell)}{m! \ell!} \right) \\ &\leq a \left(\sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^j (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{(m-\ell)}{m! \ell!} \right), \end{aligned}$$

wobei der zweite Summenausdruck nicht-positiv ist, da stets $m < \ell$ gilt und somit $m - \ell < 0$ folgt. Die übrigen Faktoren sind unter den Voraussetzungen des Satzes nicht-

negativ. Weiter erhalten wir

$$\begin{aligned}
 \mathcal{M}'_I(\tau) &< a \left(\sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^j (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{(m-\ell)}{m! \ell!} \right) \\
 &= a \left(\sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^j (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{m}{m! \ell!} - \sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^j (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{\ell}{m! \ell!} \right) \\
 &= a \left(\sum_{j=1}^{n-1} \sum_{m=1}^j \sum_{\ell=1}^j (n\gamma)^{m+\ell} \tau^{m-1+\ell} \frac{m}{m! \ell!} - \sum_{j=1}^{n-1} \sum_{i=1}^j \sum_{k=1}^j (n\gamma)^{i+k} \tau^{i-1+k} \frac{i}{i! k!} \right) \\
 &= 0.
 \end{aligned}$$

Damit folgt die Behauptung. \square

Aus dem bewiesenen Satz folgt, dass die erwartete restliche Aufenthaltszeit in I für das LCT-Modell mit $n \geq 2$ Subkompartimenten mit längerer bereits im Kompartiment I verbrachten Zeit τ stets kleiner wird. Für Infektionskrankheiten ist es plausibel, dass die erwartete verbleibende infektiöse Periode mit voranschreitender Infektion abnimmt. Für die Annahme einer exponentialverteilten Aufenthaltszeit beziehungsweise $n = 1$ gilt $\mathcal{M}_I(\tau) = T_I$, wie in Kapitel 2.4 gezeigt wurde. Wir erhalten für die Erlang-verteilte infektiöse Periode mit $n \geq 2$ folglich ein realistischeres Verhalten als für die exponentialverteilte Aufenthaltszeit in I , vergleiche [11, S. 99] und [31, S. 1529].

3.4.4 Überblick über die Beziehung der drei SIR-Modelle

An verschiedenen Stellen haben wir bereits den Zusammenhang der drei behandelten Modelle betrachtet, das heißt des ODE-, IDE- und des LCT-SIR-Modells. Da die Verbindungen essentiell für den Vergleich und die Auswahl eines Modells zur Simulation sind, halten wir die bisherigen Ergebnisse in Form einer zentralen Bemerkung fest.

Bemerkung 3.9 (Zusammenfassung der Beziehung der Modelle). *Das IDE-SIR-Modell (2.3) ermöglicht von den drei Vorgestellten die flexibelsten Möglichkeiten zur Anpassung. Die Parameter müssen wenige Eigenschaften erfüllen, die fast keine Einschränkungen für epidemiologisch sinnvolle Parameter bedeuten.*

Nach Satz 3.6 ist das LCT-SIR-Modell (3.10) ein Spezialfall des IDE-Modells mit den Parameterwahlen $\rho(t, \tau) = \rho(t)$ sowie $\phi(t, \tau) = \phi(t)$ und $\gamma_I^R(\tau) = 1 - F_{n\gamma, n}(\tau)$. Das LCT-Modell bietet insofern weniger Freiheiten in der Parametrisierung in dem Sinne, dass ρ und ϕ nicht von der Zeit seit der Infektion abhängen dürfen und die infektiöse Periode Erlang-verteilt angenommen wird.

In Satz 2.3 haben wir festgehalten, dass sich das ODE-SIR-Modell (2.1) ebenfalls aus dem IDE-Modell unter bestimmten Einschränkungen ergibt. Das IDE-Modell reduziert sich unter den Wahlen $\rho(t, \tau) = \rho(t)$ sowie $\phi(t, \tau) = \phi(t)$ und $\gamma_I^R(\tau) = e^{-\gamma\tau}$ auf das ODE-Modell. Für dieses Modell ist nur eine exponentialverteilte infektiöse Periode zulässig. In Abschnitt 3.4.2 wurde deutlich, dass das ODE-Modell auch ein besonderer Fall des LCT-Modells ist. Das ODE-Modell entsteht durch die Wahl $n = 1$.

Zusammenfassend haben wir festgestellt, dass das ODE-Modell einen Spezialfall des LCT-Modells darstellt, welches wiederum eine besondere Form des IDE-Modells ist. In dieser Reihenfolge betrachtet, bieten die Modelle einen aufsteigenden Grad an Flexibilität.

In diesem Kapitel haben wir verschiedene Eigenschaften des Linear Chain Trick-Modells analysiert. Wir haben zunächst untersucht, wie die Erlang-verteilte Aufenthaltszeit ohne Kenntnis des IDE-Modells abgeleitet werden kann. Daraufhin wurde die Bedeutung der Parameter für das Modell behandelt, wobei wir uns vor allem auf den Parameter n der Anzahl der Subkompartimente konzentriert haben. Anschließend haben wir gezeigt, dass die erwartete restliche Aufenthaltszeit in I für das LCT-Modell mit $n \geq 2$ für $\tau \geq 0$ streng monoton fallend ist. Der Zusammenhang der Modelle, den wir in diesem Abschnitt festgehalten haben, stellt die Ausgangslage für den Vergleich der Modellarten dar. Auf der Grundlage der untersuchten Eigenschaften und Zusammenhänge können erste Bewertungen der Modelle vorgenommen werden. Vorher besprechen wir einige mögliche Erweiterungen für das LCT-SIR-Modell. Hierdurch wird das Potential der LCT-Modelle für die Modellierung deutlich.

3.5 Varianten und Erweiterungsmöglichkeiten des LCT-SIR-Modells

Das vorgestellte SIR-Modell basierend auf dem Linear Chain Trick kann auf verschiedene Arten erweitert werden, um eine realistischere Abbildung der Infektionsdynamik zu erreichen. Einige Erweiterungen führen wir im Folgenden ein. Des Weiteren geben wir Ideen, wie das Modell darüber hinaus verbessert werden könnte. Dies ist als Ausblick auf Möglichkeiten für weiterführende Arbeiten gedacht.

Eine Möglichkeit zur Erweiterung besteht darin, weitere Kompartimente mit unterschiedlicher epidemiologischer Bedeutung einzuführen. Diese Zustände können jeweils eine Erlang-verteilte Aufenthaltszeit annehmen, sodass wir das resultierende Modell durch Subkompartimente und ein zugehöriges gewöhnliches Differentialgleichungssystem darstellen können. Diese Idee wird in Kapitel 4.1.3 realisiert. Eine Erweiterung, die in diesem Zuge in das LCT-Modell eingebaut werden könnte, um die Realität besser abzubilden, ist die Einführung eines Erlang-verteilten Immunitätsverlusts. Dies wurde bereits in Abschnitt 3.3 erwähnt und wird von [49, 71] angewendet. Sinnvoll wäre außerdem die Betrachtung von einer örtlichen Auflösung wie beim ODE-Modell in [60] sowie eine Teilung der Kompartimente nach dem Alter oder anderen Merkmalen der Bevölkerung, was für ein LCT-Modell zum Beispiel in [79] umgesetzt wird. Im LCT-SIR-Modell ist es leicht zu inkludieren, dass Infizierte in verschiedenen Subkompartimenten unterschiedlich infektiös sind oder verschiedene Anzahlen an durchschnittlichen Kontakten pro Tag haben. Dazu wird die Summe der Anzahlen der Individuen in den Subkompartimenten in der Gleichung für den Infektionsdruck des LCT-Modells, das heißt in

$$\varphi(t) = \frac{1}{N} \rho(t) \phi(t) I(t) = \frac{1}{N} \rho(t) \phi(t) \sum_{j=1}^n I_j(t),$$

durch eine gewichtete Summe mit den Gewichten $\rho_j(t)\phi_j(t)$ ersetzt. Vergleiche dazu [62, S. 986] in Verbindung mit [65, Kapitel 3–5]. Anstatt der Gleichung für $S'(t)$ aus (3.10) erhalten wir dann die Gleichung

$$S'(t) = -\frac{S(t)}{N} \sum_{j=1}^n \rho_j(t) \phi_j(t) I_j(t). \quad (3.21)$$

Auf diese Weise kann eine Latenzzeit simuliert werden, in der Individuen infiziert, aber noch nicht infektiös sind, siehe [54, S. 85]. Dies erreichen wir, indem wir $\rho_j(t)$ für die ersten $m \leq n$ Subkompartimente auf null setzen [64, S. 848]. Zudem können wir einen Spezialfall des ODE-SEIR-Modells

$$\begin{aligned} S'(t) &= -\rho(t) \phi(t) I(t) \frac{S(t)}{N} \\ E'(t) &= \rho(t) \phi(t) I(t) \frac{S(t)}{N} - \kappa E(t) \\ I'(t) &= \kappa E(t) - \gamma I(t) \\ R'(t) &= \gamma I(t) \end{aligned}$$

durch das LCT-SIR-Modell mit der Gleichung (3.21) für S ausdrücken. Siehe dazu [86, S. 3] und für das ODE-SEIR-Modell [72, Kapitel 3] oder [11, Kapitel 4.3.1]. Im SEIR-Modell fügen wir das zusätzliche Kompartiment E (engl. *Exposed*) hinzu, in dem sich die Individuen in der Latenzzeit befinden. Individuen verlassen die Gruppe E mit der Rate κE , wobei $\kappa > 0$ konstant ist. Dabei entspricht κ^{-1} der erwarteten Aufenthaltszeit in E beziehungsweise der erwarteten Latenzzeit. Betrachten wir nun den Spezialfall im ODE-SEIR-Modell, dass die erwarteten Aufenthaltszeiten in den Kompartimenten E und I identisch sind, das heißt $\kappa = \gamma$. Diesen Fall können wir durch Umbenennung der Variablen des LCT-SIR-Modells mit $n = 2$ Subkompartimenten und unterschiedlich wählbaren Infektiösitäten simulieren. Dafür ersetzen wir I_1 durch E , wählen $\rho_1(t) = 0$ und benennen $I_2 = I$, $\rho_2(t) = \rho(t)$ sowie $\phi_2(t) = \phi(t)$. Wir müssen zur Inklusion einer Latenzzeit somit nicht zwingend das zusätzliche Kompartiment E explizit einführen, sondern können das LCT-SIR-Modell mit der vorgestellten Modifikation nutzen.

Neben der Erlang-Verteilung existieren weitere Verteilungen, unter welchen sich das IDE-Modell auf ein gewöhnliches Differentialgleichungssystem reduziert. Wenn solch eine Verteilung die Erlang-Verteilung als Spezialfall inkludiert, kann das resultierende Modell als Erweiterung des LCT-Modells angesehen werden. Vergleiche für die folgende Erläuterung [25, S. 1442f.] und [66, S. 15]. Mit Bedingungen, unter welchen sich ein allgemeines Integro-Differentialgleichungssystem der Form

$$x'(t) = \int_{-\infty}^t K(t, \tau) x(\tau) d\tau$$

auf ein gewöhnliches Differentialgleichungssystem reduziert, beschäftigen sich [88] und [29], siehe auch Kapitel 3.3. Das Integro-Differentialgleichungssystem heißt dann „ODE-reduzierbar“. Dabei ist x die gesuchte Funktion und K eine Funktion, welche die Parametrisierung darstellt. Das obige System ist ODE-reduzierbar, falls die Funktion K eine

lineare gewöhnliche Differentialgleichung der Form

$$\frac{\partial^k K(t, \tau)}{\partial t^k} + \sum_{j=1}^{k-1} a_j(t) \frac{\partial^j K(t, \tau)}{\partial t^j} + a_0(t) K(t, \tau) = 0$$

für alle τ löst. Dabei kann $k \in \mathbb{N}$ beliebig sein und $a_j(t)$ für $j \in \{0, \dots, k-1\}$ seien von der Zeit abhängige Funktionen. Für $j \in \{1, \dots, k\}$ stellt $\frac{\partial^j K}{\partial t^j}$ die j -te partielle Ableitung von K nach t dar. Es könnte nun untersucht werden, wie dieses Ergebnis auf unser Modell anwendbar ist, um eine allgemeinere Klasse von Verteilungen zu definieren, unter der sich das IDE-Modell auf ein Modell basierend auf gewöhnlichen Differentialgleichungen reduziert. In [37] wird ein gewöhnliches Differentialgleichungssystem verwendet, um eine Pearson-verteilte Aufenthaltszeit zu erreichen. Hurtado et al. führen in [45] als Erweiterung des LCTs einen *Generalized Linear Chain Trick* ein. Hier wird die Idee umgesetzt, um allgemeinere Aufenthaltszeitverteilungen als die Erlang-Verteilung, sogenannte Phasenverteilungen, mittels eines gewöhnlichen Differentialgleichungssystems darzustellen. Eine Phasenverteilung wird durch eine Dichtefunktion

$$f(x) = \beta^T e^{x\mathbf{A}} (-\mathbf{A}\mathbb{1})$$

für $x \in \mathbb{R}^+$ beschrieben, wobei die beiden Parameter $\beta \in \mathbb{R}^n$ und $\mathbf{A} \in \mathbb{R}^{n \times n}$ verwendet werden. In der Funktion steht $\mathbb{1} \in \mathbb{R}^n$ für den Vektor bestehend aus Einsen, β^T für die Transposition des Vektors β und

$$e^{x\mathbf{A}} = \sum_{j=0}^{\infty} \frac{(x\mathbf{A})^j}{j!}$$

für das Matrixexponential. Vergleiche hierzu [46, S. 413] in Verbindung mit [6, S. 147] und [74, S. 2]. Die Klasse der Phasenverteilungen umfasst andere Verteilungen, wie die Exponentialverteilung, Hyper- und Hypoexponential-Verteilungen und die Coxian-Verteilung sowie auch die Erlang-Verteilung, siehe [47, S. 265]. Demnach stellt das Verfahren des Generalized Linear Chain Tricks eine Erweiterung des LCTs dar. Vorteilhaft könnte sein, dass sich Phasenverteilungen gut dazu eignen, andere Verteilungen zu approximieren oder sie an Daten anzupassen [47, S. 265f.]. Entsprechend könnte es sinnvoll sein, den Generalized Linear Chain Trick in weiteren Arbeiten näher zu betrachten.

3.6 Vor- und Nachteile der Modelle

An dieser Stelle haben wir die Prinzipien der ODE-, IDE- und LCT-Modelle umfassend eingeführt, sodass wir Vor- und Nachteile der Modellarten aus der Literatur verstehen können. Im Folgenden wird ein detaillierter Überblick über die Argumente aus verschiedenen Quellen sowie aus den bisherigen Kapiteln gegeben. Einige Aspekte wurden beispielsweise bereits in der Einleitung als Motivation angeführt. Nun haben wir die notwendige Basis, um die Punkte nachzuvollziehen und ausführlicher zu behandeln. Die Zusammenstellung dient als Grundlage für die Setzung der Untersuchungsschwerpunkte in den folgenden Kapiteln. Die angeführten Stärken und Schwächen aus der Literatur

werden wir untersuchen und validieren oder gegebenenfalls falsifizieren oder relativieren. In diesem Zusammenhang können wir erste Tendenzen bezüglich der Frage gewinnen, ob das LCT-Modell die Vorteile des IDE- und ODE-Modells vereinen kann und somit als eine Art Kompromiss dienen könnte.

Vorab ist zu hinterfragen, welche Merkmale sich dazu eignen, die Modelle zu bewerten und hinsichtlich ihrer Eignung zur Durchführung von Simulationen zu vergleichen. Laut [19, S. 119] und [91, S. 2150] stellt es eine Herausforderung bei der Auswahl mathematischer Modelle dar, ein angemessenes Maß an Modellkomplexität zu finden. Bei der Entwicklung und Anwendung von Modellen bedarf es stets einer Abwägung zwischen der Anpassungsfähigkeit und der Anwendbarkeit des Modells. Einerseits müssen Modelle hinreichend komplex sein, um die zugrundeliegenden Prozesse realistisch abzubilden. Andererseits sollten sie ausreichend einfach sein, um verständlich, analysierbar und leicht implementierbar zu bleiben. Führen wir die Vor- und Nachteile aus der Literatur und aus den letzten Kapiteln für die drei Modelle vor dem Hintergrund dieser beiden Aspekte zusammen.

Bewertung anhand der Flexibilität der Modelle

Stellen wir die drei Modellierungsansätze zunächst hinsichtlich der Flexibilität beziehungsweise Anpassungsfähigkeit gegenüber. In Bemerkung 3.9 haben wir festgehalten, dass das ODE-Modell einen Spezialfall des LCT-Modells darstellt und dieses wiederum eine spezielle Form des IDE-Modells ist. Entsprechend ist die gebotene Flexibilität in dieser Reihenfolge aufsteigend. Fraglich ist jedoch, welches Maß an Anpassungsfähigkeit für die Modellierung der Ausbreitung einer Infektionskrankheit notwendig ist.

In Kapitel 2.4 wurde festgestellt, dass das ODE-SIR-Modell die implizite Annahme einer exponentialverteilten Aufenthaltszeit im Kompartiment I nutzt. Eine Vielzahl an Quellen betiteln diese Annahme als für die meisten Infektionskrankheiten unrealistisch oder als Hauptkritikpunkt am Modell [22, 31, 32, 58, 63, 91, 93]. Nach [63, S. 60] und [22, S. 182] ist die Wahrscheinlichkeit nach R überzugehen in der Realität unmittelbar nach der Infektion klein und steigt erst im Verlauf der Infektion an. Eine realistische Verteilung der infektiösen Periode hat eine relativ kleine Varianz, sodass die Abweichung vom Mittelwert gering ist. Unter der Annahme einer Exponentialverteilung wird die Anzahl der Individuen überschätzt, deren infektiöse Periode bedeutend länger oder kürzer als der Mittelwert ist. Die Flexibilität des ODE-Modells scheint demnach nicht ausreichend zu sein, um realistische Annahmen in das Modell zu inkludieren.

Eine realistischere Wahl für die Verteilung der infektiösen Periode ist laut [31, 63, 91, 93, 99] die Erlang-Verteilung. Die beschriebene Form der Verteilung der infektiösen Periode kann mit einer Erlang-Verteilung besser getroffen werden, was auch durch die Form der Dichten in Abbildung 3.2 für $n = 1$ versus $n > 1$ ersichtlich ist. Somit lässt das LCT-Modell realistischere Verteilungen als das ODE-Modell zu. Die Eigenschaft der Gedächtnislosigkeit könnte maßgeblich dafür sein, dass die Annahme der exponentialverteilten infektiösen Periode unrealistisch ist, siehe Kapitel 2.4 sowie [32, S. 103] und [91, S. 2152]. Für die Erlang-Verteilung mit $n \geq 2$ ist die erwartete restliche Aufenthaltszeit streng monoton fallend, wie in Kapitel 3.4.3 gezeigt wurde. Die künftige Entwicklung des Modells kann

neben den aktuellen Modellgrößen auch von der Vergangenheit abhängen. Die Annahme der Erlang-Verteilung ist hinsichtlich dieser Merkmale folglich realistischer [31, S. 1529].

Aus den Ausführungen ergibt sich, dass die infektiöse Periode für die meisten Infektionskrankheiten nicht exponentialverteilt ist. Eine realistischere Wahl könnte eine Erlang-Verteilung sein. Unklar ist nun, inwiefern die Wahl die Modellierung beeinflusst und ob es vertretbar ist, eine weniger realistische Verteilung zugunsten anderer Vorteile zu verwenden. Laut [4, S. 1073] weisen die Modellvorhersagen eine hohe Sensitivität bezüglich der Annahme der Verteilung der infektiösen Periode von der betrachteten Krankheit auf. Die Annahme einer Exponentialverteilung kann nach [93, S. 0626] und [32, S. 90] zu fehlerbehafteten Beurteilungen der Effektivität von Maßnahmen führen, die zur Eindämmung der Ausbreitung der Infektionskrankheit in Betracht gezogen werden. Die Auswahl einer geeigneten Verteilung der Aufenthaltszeit in I könnte demnach von signifikanter Bedeutung sein. Demnach sollte der Vergleich der Modelle hinsichtlich der Flexibilität eines der Hauptargumente bei der Auswahl eines der Modelle bilden, falls verlässliche Ergebnisse und nicht nur Approximationen erzielt werden sollen.

Den Vorteil des LCT-Modells gegenüber dem ODE-Modell relativiert eine Ausführung von Cassidy et al. in [16]. Laut der Quelle gäbe es keinen Grund, warum die infektiöse Periode einer Erlang-Verteilung anstatt einer Gamma-Verteilung mit allgemeinem Formparameter $\alpha \in \mathbb{R}^+$ folgen sollte [16, S. 1046]. Wird eine Erlang-Verteilung verwendet, um eine Gamma-verteilte infektiöse Periode zu approximieren, kann es zu großen Approximationsfehlern kommen und das qualitative Verhalten kann sich unterscheiden [16, S. 1072]. Dies relativiert den Vorteil des LCT-Modells zwar, jedoch ist die Annahme einer Erlang-verteilten infektiösen Periode noch immer eine bessere Approximation für viele Verteilungen als die Annahme einer Exponentialverteilung [7, S. 17]. Das LCT-Modell ist nicht die vollständige Antwort auf die Probleme des ODE-Modells, da die Aufenthaltszeiten in jedem der Subkompartimente nach wie vor exponentialverteilt sind [31, S. 1534], siehe dazu auch Satz 3.7. Das LCT-Modell kann demnach keine allgemeinen Verteilungen simulieren, was zu potenziellen Fehlern führen kann. Aufgrund seiner höheren Anpassungsfähigkeit könnte es jedoch eine bessere Option sein als das ODE-Modell.

Wenn wir allgemeine Verteilungen zulassen wollen, ist das IDE-Modell die richtige Wahl. Hier kann die Verteilung der infektiösen Perioden mit nur wenigen Einschränkungen an die aus epidemiologischen Daten bestimmte Verteilung angepasst werden. Zur Simulation einer allgemeinen Gamma-Verteilung wäre gegebenenfalls ein IDE-Modell notwendig. Zudem lässt das IDE-Modell von der Zeit seit der Infektion τ abhängige Parameter ρ und ϕ zu, was ebenfalls einen Zuwachs der Anpassungsfähigkeit bedeutet. Laut [69, S. 239] sind IDE-basierte Modelle von steigender Beliebtheit, da die Infektionsdynamik der Vergangenheit in das Modell einfließen kann. Hiermit wird auf die Eigenschaft der Gedächtnislosigkeit verwiesen, die allerdings auch auf die LCT-Modelle nicht anzuwenden ist.

Abschließend ist festzuhalten, dass die begrenzte Flexibilität des ODE-Modells aus epidemiologischer Perspektive problematisch sein könnte. Die Annahme der infektiösen Periode hat eine große Auswirkung auf Simulationsergebnisse, sodass sich mit dem ODE-Modell lediglich Approximationen und Trends prognostizieren lassen. Insbesondere die Eigenschaft der Gedächtnislosigkeit macht das Modell unrealistisch. Soll die Vorhersage

als Grundlage für politische Entscheidungen dienen, werden möglichst genaue quantitative Ergebnisse benötigt [10, S. 117]. Das LCT-Modell lässt eine breitere Auswahl an Verteilungen zu. Die tatsächliche Verteilung der Aufenthaltszeit kann besser approximiert werden. Dennoch bleibt die Einschränkung bestehen, dass die reale Verteilung der infektiösen Periode im LCT-Modell nicht direkt verwendet werden kann. Hierfür ist ein IDE-Modell notwendig, das die höchste Anpassungsfähigkeit bietet. Den Unterschied der Modellvorhersagen unter verschiedenen Verteilungsannahmen werden wir in Kapitel 5 genauer untersuchen. Anhand eines Vergleichs mit realen Daten werden wir beurteilen, welches Maß an Flexibilität der Modellannahmen notwendig ist, um realitätsnahe Prognosen zu erstellen.

Bewertung anhand der Praktikabilität der Modelle

Nun vergleichen wir die Modelle im Kontext der Praktikabilität und Anwendbarkeit. Dazu betrachten wir zunächst das ODE-Modell. Aufgrund der vereinfachenden Annahmen sind Modelle basierend auf gewöhnlichen Differentialgleichungen vergleichsweise leicht zu formulieren. Zusätzlich sind mathematische Techniken zur Analyse unkompliziert anwendbar [44, S. 2]. Dies könnten Hauptgründe dafür sein, dass der ODE-Modellansatz in der Literatur verbreitet ist [46, S. 412]. Durch die Einfachheit sind Modelle basierend auf gewöhnlichen Differentialgleichungen für ein breites Spektrum von Wissenschaftlern zugänglich [45, S. 1873]. In mathematischer Software sind vorimplementierte numerische Methoden zur Lösung von gewöhnlichen Differentialgleichungssystemen Standard [44, S. 20f.]. ODE-Modelle sind daher leicht in Computersimulationen umsetzbar und die numerische Lösung ist effizient berechenbar [93, S. 0622].

Das LCT-Modell wird ebenfalls durch ein gewöhnliches Differentialgleichungssystem beschrieben. Daher verfügt das Modell über die genannten Vorteile einer vergleichsweise einfachen Analyse und Implementierung. Der Linear Chain Trick kann ohne einen Umweg über Integro-Differentialgleichungen oder Integralgleichungen angewendet und analysiert werden, wie in Kapitel 3.4.1 gezeigt wurde. Eine Erlang-verteilte Aufenthaltszeit kann stets durch eine Kette von Subkompartimenten mit linearen Übergängen erreicht werden. Somit lassen sich LCT-Modelle unkompliziert konstruieren, siehe [45, S. 1873]. Die vorherigen Ausführungen zeigen allerdings, dass die infektiöse Periode in der Realität nah um den Mittelwert verteilt ist und keine große Standardabweichung aufweist. Eine kleine Varianz wird im LCT-Modell bei festem Mittelwert durch die Erhöhung der Anzahl der Subkompartimente n erzeugt, vergleiche Abschnitt 3.4.2. Je kleiner die gewünschte Varianz ist, desto größer wird folglich die Dimension des resultierenden gewöhnlichen Differentialgleichungssystems. Dieser Anstieg der Systemgröße könnte zu Herausforderungen bei der Analyse und der Umsetzung in Computercode führen. Vergleiche dazu [45, S. 1874] mit [66, S. 15] und [37, S. 1089]. Gemäß [7, S. 17] ist die Annahme einer Erlang-verteilten infektiösen Periode selbst mit kleinem n eine bessere Approximation für viele realistische Verteilungen als die Annahme einer Exponentialverteilung. Allerdings bleibt das LCT-Modell auch für kleine $n > 1$ schwerer mathematisch zu analysieren als das ODE-Modell [91, S. 2152].

Rückblickend auf die Inhalte von Kapitel 2.3 lässt sich feststellen, dass die Formulierung

von Modellen basierend auf Integro-Differentialgleichungen vergleichsweise komplex ist, siehe [45, S. 1833f.]. Die Quellen [58, Supplementary material] und [22, S. 182] betonen, dass die mathematische Analyse von Modellen basierend auf Integro-Differentialgleichungen herausfordernd ist. Die Implementierung und Lösung dieses Differentialgleichungssystems gestaltet sich in der Regel als anspruchsvoller. Für Modelle basierend auf Integro-Differentialgleichungen sind im Gegensatz zu den beiden Modellen basierend auf gewöhnlichen Differentialgleichungen tiefere mathematische Kenntnisse erforderlich. Ein Schritt zur umfassenden Verwendung von IDE-Modellen wäre die Entwicklung von besseren Methoden zur Handhabung der Modelle, sowohl in theoretischer als auch in numerischer Hinsicht [77, S. 51].

Die Aspekte der Komplexität der Formulierung, der Implementierbarkeit und der Effizienz der Simulation werden wir in den Kapiteln 4 und 5 für die Modellklassen vergleichen. Dennoch lässt sich mittels der Überlegungen für die SIR-Modelle bereits festhalten, dass sich ODE-Modelle als anwendungsfreundlich erweisen. LCT-Modelle sind unwesentlich komplexer zu formulieren und zu analysieren. Wie ODE-Modelle sind LCT-Modelle für ein breites Spektrum von Wissenschaftlern zugänglich und im Vergleich zu den IDE-Modellen verständlicher, da die Modellgleichungen ein gewöhnliches Differentialgleichungssystem bilden. IDE-Modelle sind erheblich komplexer in ihrer Anwendung und der mathematischen Untersuchung.

Ein weiterer bedeutender Faktor in Bezug auf die Anwendbarkeit der Modelle ist die Verfügbarkeit von epidemiologischen Daten. Das IDE-Modell benötigt umfangreiche Informationen über Parameter wie die Verteilung der infektiösen Periode. Die Erhebung dieser Daten ist anspruchsvoller als beispielsweise die Erfassung der mittleren Aufenthaltsdauer, die für das ODE-Modell benötigt wird [30, S. 803]. Aufgrund der eingeschränkten Flexibilität erfordert auch das LCT-Modell im Vergleich zum IDE-Modell weniger Informationen. Für das LCT-SIR-Modell muss lediglich der Parameter n zusätzlich zu denen des ODE-SIR-Modells bestimmt werden. Krylova et al. [58, S. 7] bemerken, dass ein LCT-Modell mit Erlang-verteilter Latenzzeit und infektiöser Periode (mit einem Formparameter größer als eins) auch dann die Realität besser abbildet als ein ODE-SEIR-Modell, wenn lediglich die Erwartungswerte der Aufenthaltszeiten geschätzt werden können. Die Verfügbarkeit der epidemiologischen Parameter scheint unterschiedlich praktikabel zu sein. Dies werden wir in Kapitel 5.1 überprüfen. Zusätzlich könnte die realistische Setzung der passenden Anfangsbedingungen je nach verwendeter Methode einen unterschiedlichen Schwierigkeitsgrad aufweisen. Für das IDE-Modell sind Anfangsfunktionen gefordert, siehe Gleichung (2.4) und für das ODE-Modell einzelne Werte, siehe Gleichung (2.2). Den Anfangsbedingungen werden wir uns in Kapitel 4.3 zuwenden.

In dieser Argumentation haben wir verschiedene Aspekte in Bezug auf die Flexibilität und die Anwendbarkeit der Modelle aus der Literatur und den bisherigen Kapiteln angeführt. Es scheint, dass das IDE-Modell in Bereichen Vorteile bietet, in denen das ODE-Modell Nachteile aufweist und umgekehrt. Die Ausführungen legen nahe, dass das LCT-Modell immer zwischen dem ODE- und dem IDE-Modell steht. Insofern könnte das LCT-Modell die Vorteile beider Modellansätze vereinen. Allerdings wurden auch einige Bemerkungen angebracht, die diese Aussage einschränken. In der Analyse haben wir Argumente

aus verschiedenen Quellen vorgebracht, die in den folgenden Kapiteln näher untersucht und verifiziert werden. Mit dieser Besprechung haben wir eine Grundlage für unsere weiteren Untersuchungen gelegt und einen Überblick erhalten. Auf der Basis von den Ausführungen in den folgenden Kapiteln kann dann abschließend die Hauptfragestellung der Arbeit beantwortet werden.

4 Umsetzung der Modellarten in die Praxis

In diesem Kapitel stellen wir Überlegungen an, die für die Umsetzung der Modellarten in praktische Simulationen relevant sind. Ziel dieses Abschnitts ist es, sämtliche theoretische Grundlagen zur Verwendung der Modelle in praxisrelevanter Software zu klären und eine Implementierung zu beschreiben. Wir legen hiermit die Basis für Simulationen anhand einer beispielhaften Infektionskrankheit, für die im anschließenden Kapitel 5 verschiedene Szenarien definiert und deren Ergebnisse beschrieben werden. Daneben überprüfen wir einige Argumente aus dem Vergleich der Modellarten aus Abschnitt 3.6, wobei sich die Aspekte dieses Kapitels vornehmlich zur Beurteilung der Anwendbarkeit der Modelle eignen. Die Bewertung der Flexibilität führen wir später anhand der Simulationen durch. Die Grundlagen in diesem Kapitel sind im Allgemeinen für jede Infektionskrankheit anwendbar. Für Besonderheiten einer untersuchten Krankheit könnten geringfügige Änderungen notwendig sein. Mit der Software „MEmilio“ [61] von Kühn et al. existiert bereits eine Implementierung eines detaillierten ODE-Modells unter Berücksichtigung von mehr als drei Kompartimenten und mit einer Altersstruktur sowie einer örtlichen Auflösung. Ein IDE-Modell mit acht Kompartimenten ist dort ebenfalls implementiert. Zur Entwicklung wird die Programmiersprache C++ verwendet. Wir orientieren uns hier an bestimmten Aspekten der Software und der Publikation [60], um eine bessere Vergleichbarkeit und Wiederverwendbarkeit von bereits vorhandenen Implementierungen und Parametern zu erreichen. Vor diesem Hintergrund führen wir zunächst für die drei vorgestellten Modellansätze aus den Kapiteln 2 und 3 Erweiterungen durch mehr Kompartimente ein. Im Anschluss besprechen wir in Abschnitt 4.2 numerische Lösungsmethoden, die zur Umsetzung in eine praktische Implementierung verwendet werden können. In Kapitel 4.3 wenden wir uns der Festsetzung der Anfangsbedingungen zu. Abschließend stellen wir eine Implementierung des LCT-Modells eingebettet in die bestehende Software MEMilio vor. Daraufhin haben wir die Voraussetzungen dafür geschaffen, die Modelle zur Simulation der Ausbreitung von konkreten Erregern anzuwenden.

4.1 Erweiterung der Modelle durch zusätzliche Kompartimente

Um die Simulation der Ausbreitung einer betrachteten Infektionskrankheit besser an die Realität anpassen zu können, ist es gängige Praxis, in die Modelle mehr Kompartimente einzufügen, siehe zum Beispiel [9, 32]. Dadurch können zusätzliche Aspekte der Ausbreitung prognostiziert werden, beispielsweise eine Belegung der Intensivbetten aufgrund der betrachteten Krankheit oder die Anzahl der Todesfälle. Im Folgenden erweitern wir unsere drei vorgestellten SIR-Modelle aus den Kapiteln 2 und 3 um weitere Kompartimente. Wir orientieren uns bei der Anzahl und der Bedeutung der Gruppen sowie bei den möglichen Übergängen zwischen den Kompartimenten an dem ODE-basierten Modell

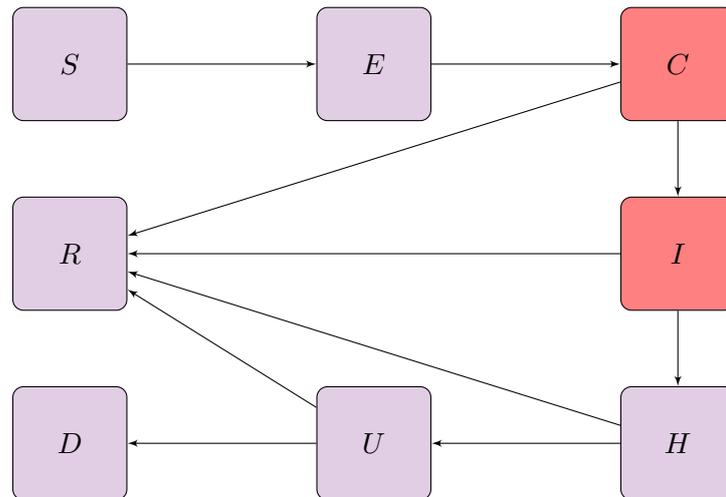


Abbildung 4.1: Aufbau der SECIR-Modelle. Schematische Darstellung der Kompartimente und der möglichen Übergänge zwischen den Kompartimenten in den SECIR-Modellen. Die Darstellung ist an [60, Fig. 1] angelehnt. Die Zustände, in denen Individuen infektiös sind und Personen aus dem Kompartiment S anstecken können, sind in Rot hervorgehoben.

von Kühn et al. [60] beziehungsweise der Implementierung in der Software MEmilio [61]. Auf andere dort betrachtete Erweiterungen, wie die Altersauflösung oder die örtliche Auflösung, verzichten wir hier. Auf die Herleitungen der Modellgleichungen legen wir keinen besonderen Fokus, da die Ideen aus den SIR-Modellen bekannt sind.

Wir beschreiben zunächst die acht Kompartimente, die in der Publikation [60] betrachtet werden. Die Bezeichnungen und Bedeutungen finden im Anschluss auf alle drei Modelle Anwendung. Siehe für die folgenden Erläuterungen entsprechend [60]. Im Kompartiment S (engl. *Susceptible*) befinden sich wie in den vorangegangenen Kapiteln die Individuen, die noch nicht mit der betrachteten Krankheit infiziert wurden und somit noch angesteckt werden können. Wurde eine Person angesteckt, wechselt sie nun zunächst in den Zustand E (*Exposed*), den wir bereits in Kapitel 3.5 eingeführt haben. In diesem Zustand ist ein Individuum zwar infiziert, aber selbst noch nicht ansteckend. Die Person befindet sich demnach in der Latenzzeit (siehe zur Definition S. 29). Ein Individuum geht nach der Latenzzeit in das Kompartiment C (*Carrier*) über, in dem Personen ansteckend sind, aber keine Symptome zeigen. Im Modell können Individuen entweder später Symptome entwickeln oder einen asymptomatischen Verlauf haben und von C direkt in die bereits bekannte Gruppe R wechseln. Der Buchstabe I (*Infected*) steht für die Personen, die ansteckend sind und Symptome zeigen. Die Definition unterscheidet sich von der Definition des Kompartiments I in den SIR-Modellen der vorangegangenen Kapitel und umfasst nun einen spezifischeren Teil der Infizierten. Individuen im Zustand H (*Hospitalized*) befinden sich im Krankenhaus und im Zustand U (*In Intensive Care Unit*) auf der Intensivstation. An der Krankheit verstorbene Personen werden im Kompartiment D (*Dead*) berücksichtigt. Individuen, die von der betrachteten Krankheit genesen und somit immun sind, werden dem Zustand R (*Recovered*) zugeordnet.

Die Menge der Krankheitszustände bezeichnen wir als $\mathcal{Z} = \{S, E, C, I, H, U, R, D\}$. Für ein Kompartiment $z \in \mathcal{Z}$ schreiben wir für die Anzahl der Individuen in diesem Kompartiment zum Zeitpunkt t wie zuvor $z(t)$, beispielsweise $S(t)$. Die möglichen Übergänge zwischen den Kompartimenten in einem Krankheitsverlauf sind in Abbildung 4.1 dargestellt. Die Übergänge implizieren einige Modellannahmen, wie zum Beispiel, dass Individuen nur an der Infektionskrankheit sterben können, wenn sie sich zuvor auf einer Intensivstation befunden haben. Diese Annahmen werden anhand der Abbildung und den Differentialgleichungssystemen deutlich. Die Modelle bezeichnen wir anstatt des naheliegenden Namens „SECIHURD-Modell“ abkürzend als „SECIR-Modell“. Wir wählen diese Buchstaben, da die zugehörigen Kompartimente für die Simulation der Infektionsdynamik am relevantesten sind. Die restlichen Gruppen liefern zusätzliche Informationen, wie die Anzahl der Todesfälle, die durch die Infektionskrankheit verursacht werden. Als Präfix fügen wir den Modelltyp hinzu, sodass wir in diesem Kapitel das ODE-SECIR-Modell, das IDE-SECIR-Modell sowie das LCT-SECIR-Modell einführen.

4.1.1 Das ODE-SECIR-Modell

Zunächst erweitern wir das ODE-SIR-Modell zu einem Modell mit den beschriebenen Kompartimenten. Das Modell basierend auf gewöhnlichen Differentialgleichungen wird in [60] eingeführt. Dort wird zusätzlich eine Altersauflösung sowie eine örtliche Auflösung betrachtet, die wir der Einfachheit halber weglassen. Bei der Wahl der Größen der Parameter für die Simulationen können wir auf die Publikation zurückgreifen. Die Notation passen wir an einigen Stellen an und nehmen wenige Änderungen an den Modellgleichungen und Parametern vor.

Für die Konkretisierung der Modellgleichungen sind zuvor Parameterdefinitionen erforderlich. Wir nutzen die Notation $\gamma_z > 0$ für die Parameter, welche die Rate des Austritts aus dem Kompartiment $z \in \mathcal{Z}$ beschreiben und T_z für die erwartete Aufenthaltszeit. Ein Individuum bleibt durchschnittlich $T_z = \gamma_z^{-1}$ Tage im Zustand z . Dabei seien γ_z und T_z nur für Kompartimente z definiert, für die eine erwartete Aufenthaltszeit sinnvoll ist. Für S wird der Übergang durch Kontakte zu Infektiosen beschrieben und von R und D kann nicht in einen anderen Zustand gewechselt werden. Von einigen Gruppen aus sind verschiedene Übergänge zu anderen Zuständen möglich, wie in Abbildung 4.1 ersichtlich ist. Beispielsweise können Individuen von C nach I oder direkt nach R übergehen. Zur Umsetzung der verschiedenen möglichen Krankheitsverläufe benötigen wir den Parameter $\mu_{z_1}^{z_2} \in [0, 1]$ für $z_1, z_2 \in \mathcal{Z}$ mit $z_1 \neq z_2$. Dieser gibt den erwarteten Anteil der Individuen an, die von Kompartiment z_1 aus im Verlauf ihrer Krankheit nach z_2 übergehen. Im Sinne der stochastischen Interpretation, die in Kapitel 2.3 erläutert wurde, kann der Parameter $\mu_{z_1}^{z_2}$ als die Wahrscheinlichkeit angesehen werden, dass ein einzelnes Individuum von z_1 nach z_2 wechselt. Wenn kein Zustandsübergang von z_1 nach z_2 möglich ist, setzen wir $\mu_{z_1}^{z_2} = 0$. Beispielsweise entspricht μ_C^I der Wahrscheinlichkeit, dass eine Person vom Kompartiment C aus nach I übergeht, und $\mu_C^R = 1 - \mu_C^I$ der Wahrscheinlichkeit, dass sie zum Zustand R wechselt. Ein Individuum muss nach dem Aufenthalt im Kompartiment $z_1 \in \mathcal{Z}$ in eine der anderen Gruppen wechseln. Die Summe der Wahrscheinlichkeiten von z_1 aus in eines der anderen Kompartimente überzugehen, muss demnach eins ergeben,

das bedeutet

$$1 = \sum_{z_2 \in \mathcal{Z} \setminus \{z_1\}} \mu_{z_1}^{z_2}.$$

In [60] werden abweichend von unserer Definition der Aufenthaltszeiten T_z die Parameter $T_{z_1}^{z_2}$ eingeführt. Die Größe $T_{z_1}^{z_2}$ beschreibt die erwartete Aufenthaltszeit in der Gruppe z_1 , bevor ein Individuum nach z_2 wechselt. Wir lassen den oberen Index in unserer Ausführung weg, da der Parameter $\mu_{z_1}^{z_2}$ sonst nicht als die definierte Wahrscheinlichkeit interpretierbar ist. Wenn vom Zustand $z_1 \in \mathcal{Z}$ aus die beiden Übergänge nach $z_2 \in \mathcal{Z}$ und $z_3 \in \mathcal{Z}$ möglich sind, setzen wir $T_{z_1}^{z_2} = T_{z_1}^{z_3} = T_{z_1}$. Zum Beispiel wird die erwartete Aufenthaltszeit in C als unabhängig davon angenommen, ob ein Individuum einen asymptomatischen Krankheitsverlauf hat und sofort nach R wechselt, oder ob die Krankheit symptomatisch verläuft und das Individuum nach I übergeht.

Nun betrachten wir die übrigen Parameter, die den Übergang von S nach E spezifizieren. Analog zu Kapitel 2.2 beschreibt $\rho(t)$ die Transmissionswahrscheinlichkeit bei einem Kontakt und $\phi(t)$ die durchschnittliche Anzahl an Kontakten eines Individuums pro Tag. In den SECIR-Modellen wird angenommen, dass Personen in den Kompartimenten C und I ansteckend sind. Wir spezifizieren den Anteil der Individuen im Kompartiment $z \in \{C, I\}$, die sich nicht isolieren oder in Quarantäne befinden, durch $\xi_z : \mathbb{R} \rightarrow [0, 1]$, wie es in [56] und [59] umgesetzt wird. Genauer beschreibt ξ_C den Anteil der Personen in C , die nicht detektiert wurden und sich nicht in Quarantäne befinden und ξ_I den Anteil der symptomatisch Infektiösen aus I in Isolation oder Quarantäne. In den Parametern ξ_C und ξ_I können politische Entscheidungen wie die Teststrategie berücksichtigt werden, beispielsweise wenn vornehmlich symptomatische Individuen getestet werden. Zusätzlich könnte eingeschlossen werden, dass symptomatische Personen im Verhältnis zu Individuen des Zustands C infektiöser sind. Eigentlich hängt ξ_z nicht nur von der Zeit ab, sondern ebenfalls von Größen wie $C(t)$ oder Ähnlichem. Je weniger neue Infektionen auftreten, desto mehr Kontakte können vom Gesundheitsamt detektiert werden, sodass sich potentiell ein größerer Anteil in Quarantäne befindet. Der Einfachheit halber ignorieren wir diese Abhängigkeit und simulieren nur die Abhängigkeit von der Zeit t . Unter Verwendung der Definitionen der Parameter lässt sich schließen, dass es

$$\rho(t) \phi(t) (\xi_C(t) C(t) + \xi_I(t) I(t))$$

kumulierte Kontakte von Infektiösen zum Zeitpunkt t gibt, die zu einer neuen Ansteckung führen, falls sie mit einem Individuum der Gruppe S stattfinden. Der Anteil dieser Kontakte mit Personen, die noch mit der Krankheit angesteckt werden können, ist wie im ODE-SIR-Modell durch $\frac{S(t)}{N(t)}$ beschrieben. Hieraus ergibt sich die Anzahl an Neuansteckungen und die Gleichung für S des Modells. Anders als im ODE-SIR-Modell, nehmen wir die Bevölkerungszahl nicht als konstant an, da wir die Todesfälle in einem eigenen Kompartiment zählen. Wir beschreiben die Anzahl der lebenden Individuen im betrachteten Gebiet durch $N(t) = N_0 - D(t)$. Dabei gibt N_0 die aktuelle Bevölkerungszahl zum Zeitpunkt $t_0 = 0$ an, zuzüglich der Individuen, die zum Startzeitpunkt bereits an der Krankheit gestorben sind. Somit ist gesichert, dass die Gleichheit $N(0) = N_0 - D(0)$

zur Definition von $N(t)$ passt. Es sollte für alle $t \in \mathbb{R}$

$$N_0 - D(t) = N(t) = \sum_{z \in \mathcal{Z} \setminus \{D\}} z(t) \quad (4.1)$$

gelten, da jede Person stets genau einem Krankheitszustand zugeordnet werden kann. In der Gleichung für S teilen wir durch $N(t)$, um den Anteil der anfälligen Individuen an der lebenden Bevölkerung zu erhalten und das Verhältnis nicht durch die Todesfälle zu verzerren. In [56] ist dies ebenfalls umgesetzt.

Wir verwenden alle vorgestellten Parameter zur Formulierung des gewöhnlichen Differentialgleichungssystems, welches das ODE-SECIR-Modell beschreibt, und erhalten

$$\begin{aligned} S'(t) &= -\frac{S(t)}{N(t)} \rho(t) \phi(t) (\xi_C(t) C(t) + \xi_I(t) I(t)) \\ E'(t) &= \frac{S(t)}{N(t)} \rho(t) \phi(t) (\xi_C(t) C(t) + \xi_I(t) I(t)) - \gamma_E E(t) \\ C'(t) &= \gamma_E E(t) - \gamma_C C(t) \\ I'(t) &= \mu_C^I \gamma_C C(t) - \gamma_I I(t) \\ H'(t) &= \mu_I^H \gamma_I I(t) - \gamma_H H(t) \\ U'(t) &= \mu_H^U \gamma_H H(t) - \gamma_U U(t) \\ R'(t) &= (1 - \mu_C^I) \gamma_C C(t) + (1 - \mu_I^H) \gamma_I I(t) \\ &\quad + (1 - \mu_H^U) \gamma_H H(t) + (1 - \mu_U^D) \gamma_U U(t) \\ D'(t) &= \mu_U^D \gamma_U U(t) \end{aligned} \quad (4.2)$$

für $t \geq 0$. Das Konzept der Übergänge zwischen den Kompartimenten folgt dem Schema in Abbildung 4.1. Zusammen mit Anfangswerten $z(0) = z_{t_0}$ für jedes der Kompartimente $z \in \mathcal{Z}$ haben wir ein Anfangswertproblem gegeben.

4.1.2 Das IDE-SECIR-Modell

Analog zum ODE-SECIR-Modell definieren wir ein IDE-SECIR-Modell als Erweiterung des IDE-SIR-Modells (2.3) auf die acht vorgestellten Kompartimente. Das Modell wird in [95] publiziert werden. Es gibt nur wenige Quellen, die eine allgemeine Verteilung der Aufenthaltszeit für mehr als zwei Kompartimente zulassen, vergleiche [83, S. 76]. Die Entwicklung eines solchen Modells ist vergleichsweise unerforscht.

Zur Formulierung des Modells benötigen wir Parameter, die über die Definitionen aus dem ODE-SECIR-Modell hinausgehen. Die Größe \hat{T} sei wie in Kapitel 2.3 für das IDE-SIR-Modell (2.3) definiert. Im Folgenden verwenden wir die Notationen z , z_1 und z_2 stets für Kompartimente aus der Menge \mathcal{Z} . Die Wahrscheinlichkeit, dass ein Individuum zum Zustand z_2 wechselt, wenn es sich aktuell in z_1 befindet, bezeichnen wir wie im ODE-SECIR-Modell als $\mu_{z_1}^{z_2}$. Die möglichen Übergänge zwischen Kompartimenten folgen abermals Abbildung 4.1. Für diese Paare von Zuständen kann $\mu_{z_1}^{z_2} \neq 0$ gelten. Wenn ein Wechsel von z_1 nach z_2 möglich ist, beschreiben wir die Anzahl der Individuen, die zum

Zeitpunkt $t \in \mathbb{R}$ von der Gruppe z_1 nach z_2 wechseln, mit $\sigma_{z_1}^{z_2}(t)$. Die Ableitung $z'(t)$ der Anzahl der Individuen im Kompartiment z zum Zeitpunkt t muss stets der Anzahl der Individuen entsprechen, die zum Zeitpunkt t nach z wechseln, abzüglich der Anzahl aller Individuen, die zum Zeitpunkt t aus dem Kompartiment hinaus wechseln. Zum Beispiel bedeutet dies gemäß Abbildung 4.1, dass für das Kompartiment C die Beziehung

$$C'(t) = \sigma_E^C(t) - \sigma_C^I(t) - \sigma_C^R(t) \quad (4.3)$$

gelten muss. Durch $\gamma_{z_1}^{z_2}(\tau)$ mit $\gamma_{z_1}^{z_2} : \mathbb{R}_0^+ \rightarrow [0, 1]$ sei der erwartete Anteil der Personen gegeben, die sich τ Tage nach dem Eintritt in das Kompartiment z_1 noch in diesem befinden und im späteren Krankheitsverlauf nach z_2 übergehen werden. In diesem Modell kann die Verteilung der Aufenthaltszeit in C folglich abhängig davon gewählt werden, ob Individuen nachfolgend direkt nach R wechseln oder Symptome entwickeln und nach I übergehen. Die erwartete Aufenthaltszeit ist von zwei Kompartimenten abhängig, sodass wir sie mit $T_{z_1}^{z_2}$ notieren. Für dieses Modell ist die Größe im Gegensatz zum ODE-SECIR-Modell unproblematisch, da die Modellgleichungen für die Kompartimente mit den Übergängen $\sigma_{z_1}^{z_2}(t)$ formuliert werden. Die Bedeutung der Parameter $\gamma_{z_1}^{z_2}$ entspricht der von γ_I^R im IDE-SIR-Modell (2.3). Die Funktionen müssen dieselben Eigenschaften erfüllen. Somit muss gelten, dass die Funktionen $\gamma_{z_1}^{z_2}$ stetig differenzierbar und monoton fallend sind sowie $\gamma_{z_1}^{z_2}(0) = 1$. Zudem sollte

$$T_{z_1}^{z_2} = \int_0^\infty \gamma_{z_1}^{z_2}(\tau) d\tau < \infty$$

erfüllt sein, vergleiche Gleichung (2.7).

Die Modellgleichungen erarbeiten wir anhand von ausgewählten Kompartimenten genauer, da sie im Vergleich zum ODE-Modell schwieriger zu formulieren sind. Wir werden die Gleichung des Kompartiments E besprechen, um das Prinzip des Modells deutlich zu machen. Zusätzlich untersuchen wir den Zustand C , von dem mehrere Übergänge möglich sind, sowie D und S . Wir beginnen mit dem Kompartiment E . Die Gleichung unterscheidet sich von der Form her nicht von der für die Gruppe I im IDE-SIR-Modell. Zum Zeitpunkt t sind $\gamma_E^C(t-x)\sigma_S^E(x)$ der Individuen, die zum Zeitpunkt $x \leq t$ in den Zustand E gewechselt sind, noch nicht nach C übergegangen. Wenn alle möglichen Eintrittszeitpunkte $x \leq t$ in E betrachtet werden, ergibt sich die Gesamtzahl der Personen in E als

$$E(t) = \int_{\hat{T}}^t \gamma_E^C(t-x)\sigma_S^E(x) dx.$$

Im IDE-SECIR-Modell müssen wir auch Formeln für die Übergänge $\sigma_{z_1}^{z_2}(t)$ finden, damit die Modellgleichungen vollständig beschrieben sind. In das Kompartiment E treten zum Zeitpunkt t alle Individuen ein, die aus der Gruppe S austreten, sodass wir

$$\sigma_S^E(t) = -S'(t)$$

erhalten. Zur Beschreibung von σ_E^C benötigen wir die Ableitung von der hergeleiteten Gleichung für E . Leiten wir die Formulierung ab, ergibt sich mittels der Leibniz-Regel

für Parameterintegrale [57, Satz 4.29]

$$\begin{aligned} E'(t) &= \int_{\hat{T}}^t \gamma_E^{C'}(t-x) \sigma_S^E(x) dx + \gamma_E^C(0) \sigma_S^E(t) \\ &= \int_{\hat{T}}^t \gamma_E^{C'}(t-x) \sigma_S^E(x) dx + \sigma_S^E(t). \end{aligned}$$

Wegen des Zusammenhangs der Übergänge $\sigma_{z_1}^{z_2}$ zur Ableitung der Kompartimente gilt für die Gruppe E

$$E'(t) = -\sigma_E^C(t) + \sigma_S^E(t).$$

Vergleichen wir die beiden Formeln, ergibt sich, dass die Anzahl der Personen, die zum Zeitpunkt t von E nach C übergehen, durch

$$\sigma_E^C(t) = - \int_{\hat{T}}^t \gamma_E^{C'}(t-x) \sigma_S^E(x) dx$$

beschrieben wird. Wir fahren nun mit der Gleichung für das Kompartiment C fort. Nach der Definition von μ_C^I , werden genau $\mu_C^I \sigma_E^C(x)$ der Individuen, die zum Zeitpunkt x in die Gruppe C gewechselt sind, in ihrem Krankheitsverlauf nach I übergehen und $(1 - \mu_C^I) \sigma_E^C(x)$ direkt nach R . Nach den Definitionen von γ_C^I und γ_C^R sind zum Zeitpunkt t

$$\gamma_C^I(t-x) \mu_C^I \sigma_E^C(x) + \gamma_C^R(t-x) (1 - \mu_C^I) \sigma_E^C(x)$$

Individuen von denen, die zum Zeitpunkt $x \leq t$ in C eingetreten sind, noch in diesem Kompartiment. Betrachten wir sämtliche potentielle Eintrittszeitpunkte, erhalten wir die Formel

$$\begin{aligned} C(t) &= \int_{\hat{T}}^t \gamma_C^I(t-x) \mu_C^I \sigma_E^C(x) + \gamma_C^R(t-x) (1 - \mu_C^I) \sigma_E^C(x) dx \\ &= \int_{\hat{T}}^t \left(\gamma_C^I(t-x) \mu_C^I + \gamma_C^R(t-x) (1 - \mu_C^I) \right) \sigma_E^C(x) dx. \end{aligned}$$

Für die Gleichungen der Übergänge von C aus, benötigen wir wieder die Ableitung der hergeleiteten Gleichung für C , das heißt

$$\begin{aligned} C'(t) &= \int_{\hat{T}}^t \left(\gamma_C^{I'}(t-x) \mu_C^I + \gamma_C^{R'}(t-x) (1 - \mu_C^I) \right) \sigma_E^C(x) dx \\ &\quad + \left(\gamma_C^I(0) \mu_C^I + \gamma_C^R(0) (1 - \mu_C^I) \right) \sigma_E^C(t) \\ &= \int_{\hat{T}}^t \gamma_C^{I'}(t-x) \mu_C^I \sigma_E^C(x) dx + \int_{\hat{T}}^t \gamma_C^{R'}(t-x) (1 - \mu_C^I) \sigma_E^C(x) dx + \sigma_E^C(t). \end{aligned}$$

In Anbetracht von Gleichung (4.3) und unter Berücksichtigung der Bedeutung der Parameter ist ersichtlich, dass

$$\sigma_C^I(t) = - \int_{\hat{T}}^t \gamma_C^{I'}(t-x) \mu_C^I \sigma_E^C(x) dx$$

den Übergang von C nach I beschreibt und

$$\sigma_C^R(t) = - \int_{\hat{T}}^t \gamma_C^R(t-x) (1 - \mu_C^I) \sigma_E^C(x) dx$$

den Übergang von C nach R . Die Gleichungen für die Kompartimente I , H und U sowie die Übergänge aus den Kompartimenten ergeben sich auf die gleiche Art wie für den Zustand C . Nun leiten wir die Gleichung für die Gruppe D der Verstorbenen her. Aus diesem Kompartiment ist kein Wechsel in eines der anderen Kompartimente möglich. Somit gleicht die Anzahl der Todesfälle der Summe aller Individuen, die je von U nach D übergegangen sind. Unter Berücksichtigung der Größe σ_U^D sowie aller vergangenen Zeitpunkte erhalten wir

$$D(t) = \int_{\hat{T}}^t \sigma_U^D(x) dx.$$

Die Modellgleichung für das Kompartiment R ist ähnlich zu der des Zustands D . Für die Gruppe S gehen wir von der Gleichung für S des ODE-SECIR-Modells (4.2) aus und erweitern sie um Aspekte des IDE-Modells. Wenn wir die hergeleitete Formel für C sowie die Gleichung

$$I(t) = \int_{\hat{T}}^t \left(\gamma_I^H(t-x) \mu_I^H + \gamma_I^R(t-x) (1 - \mu_I^H) \right) \sigma_C^I(x) dx$$

des IDE-SECIR-Modells in die Gleichung für S des ODE-Modells (4.2) einsetzen, erhalten wir

$$\begin{aligned} S'(t) &= - \frac{S(t)}{N(t)} \rho(t) \phi(t) \left(\xi_C(t) C(t) + \xi_I(t) I(t) \right) \\ &= - \frac{S(t)}{N(t)} \rho(t) \phi(t) \left(\xi_C(t) \int_{\hat{T}}^t \left(\gamma_C^I(t-x) \mu_C^I + \gamma_C^R(t-x) (1 - \mu_C^I) \right) \sigma_E^C(x) dx \right. \\ &\quad \left. + \xi_I(t) \int_{\hat{T}}^t \left(\gamma_I^H(t-x) \mu_I^H + \gamma_I^R(t-x) (1 - \mu_I^H) \right) \sigma_C^I(x) dx \right). \end{aligned}$$

Wir fügen für die Parameter ξ_C und ξ_I zusätzlich zur Abhängigkeit von der Zeit t eine Abhängigkeit von der Zeit seit dem Eintritt in das jeweilige Kompartiment τ ein. Hier betrachten wir $\phi(t)$ und $\rho(t)$ als maximale Anzahl der Kontakte beziehungsweise die maximale Transmissionswahrscheinlichkeit im Verlauf der Krankheit. Eine reduzierte Kontaktzahl sowie eine geringere Infektiösität in Abhängigkeit von τ kann individuell für die beiden Kompartimente in den Parametern $\xi_C(t, \tau) \in [0, 1]$ und $\xi_I(t, \tau) \in [0, 1]$ berücksichtigt werden, anstatt durch ϕ und ρ . Beim IDE-SIR-Modell haben wir für ϕ und ρ die Abhängigkeit von τ zugelassen, weil wir dadurch eine Reduktion der Kontakte durch Isolation oder Quarantäne oder eine sich ändernde Infektiösität im Verlauf der Infektion simulieren wollten. Im SECIR-Modell steuern wir dies nun durch die zusätzlichen Parameter ξ_C und ξ_I angepasst für die Kompartimente C und I . Inkludieren wir diese Abhängigkeit, erhalten wir die Formel

$$\begin{aligned} S'(t) &= - \frac{S(t)}{N(t)} \phi(t) \rho(t) \int_{\hat{T}}^t \left(\xi_C(t, t-x) \left(\gamma_C^I(t-x) \mu_C^I + \gamma_C^R(t-x) (1 - \mu_C^I) \right) \sigma_E^C(x) \right. \\ &\quad \left. + \xi_I(t, t-x) \left(\gamma_I^H(t-x) \mu_I^H + \gamma_I^R(t-x) (1 - \mu_I^H) \right) \sigma_C^I(x) \right) dx \end{aligned}$$

für S im IDE-SECIR-Modell. Der Infektionsdruck $\varphi(t)$ sei hier wie in Kapitel 2.3 so definiert, dass $S'(t) = -S(t)\varphi(t)$ gelte. Wir haben damit alle Parameter für das IDE-SECIR-Modell definiert und einige Formeln beispielhaft hergeleitet. Die Modellgleichungen des IDE-SECIR-Modells ergeben sich insgesamt durch

$$S'(t) = -S(t)\varphi(t) \quad (4.4)$$

$$E(t) = \int_{\hat{T}}^t \gamma_E^C(t-x) \sigma_S^E(x) dx \quad (4.5)$$

$$C(t) = \int_{\hat{T}}^t \left(\gamma_C^I(t-x) \mu_C^I + \gamma_C^R(t-x) (1 - \mu_C^I) \right) \sigma_E^C(x) dx \quad (4.6)$$

$$I(t) = \int_{\hat{T}}^t \left(\gamma_I^H(t-x) \mu_I^H + \gamma_I^R(t-x) (1 - \mu_I^H) \right) \sigma_C^I(x) dx \quad (4.7)$$

$$H(t) = \int_{\hat{T}}^t \left(\gamma_H^U(t-x) \mu_H^U + \gamma_H^R(t-x) (1 - \mu_H^U) \right) \sigma_I^H(x) dx \quad (4.8)$$

$$U(t) = \int_{\hat{T}}^t \left(\gamma_U^D(t-x) \mu_U^D + \gamma_U^R(t-x) (1 - \mu_U^D) \right) \sigma_H^U(x) dx \quad (4.9)$$

$$R(t) = \int_{\hat{T}}^t \left(\sigma_C^R(x) + \sigma_I^R(x) + \sigma_H^R(x) + \sigma_U^R(x) \right) dx \quad (4.10)$$

$$D(t) = \int_{\hat{T}}^t \sigma_U^D(x) dx \quad (4.11)$$

für $t \geq 0$ mit

$$\begin{aligned} \varphi(t) = & \frac{\phi(t)\rho(t)}{N(t)} \int_{\hat{T}}^t \left(\xi_C(t, t-x) \left(\gamma_C^I(t-x) \mu_C^I + \gamma_C^R(t-x) (1 - \mu_C^I) \right) \sigma_E^C(x) \right. \\ & \left. + \xi_I(t, t-x) \left(\gamma_I^H(t-x) \mu_I^H + \gamma_I^R(t-x) (1 - \mu_I^H) \right) \sigma_C^I(x) \right) dx. \end{aligned} \quad (4.12)$$

In den Gleichungen kommen die Übergänge $\sigma_{z_1}^{z_2}$ für einige $z_1, z_2 \in \mathcal{Z}$ vor. Zur vollständigen Modellformulierung benötigen wir Formeln für diese Übergänge für $t \geq 0$:

$$\sigma_S^E(t) = -S'(t) = S(t)\varphi(t) \quad (4.13)$$

$$\sigma_E^C(t) = - \int_{\hat{T}}^t \gamma_E^C(t-x) \sigma_S^E(x) dx \quad (4.14)$$

$$\sigma_C^I(t) = - \int_{\hat{T}}^t \gamma_C^I(t-x) \mu_C^I \sigma_E^C(x) dx \quad (4.15)$$

$$\sigma_C^R(t) = - \int_{\hat{T}}^t \gamma_C^R(t-x) (1 - \mu_C^I) \sigma_E^C(x) dx \quad (4.16)$$

$$\sigma_I^H(t) = - \int_{\hat{T}}^t \gamma_I^H(t-x) \mu_I^H \sigma_C^I(x) dx \quad (4.17)$$

$$\sigma_I^R(t) = - \int_{\hat{T}}^t \gamma_I^R(t-x) (1 - \mu_I^H) \sigma_C^I(x) dx \quad (4.18)$$

$$\sigma_H^U(t) = - \int_{\hat{T}}^t \gamma_H^U(t-x) \mu_H^U \sigma_I^H(x) dx \quad (4.19)$$

$$\sigma_H^R(t) = - \int_{\hat{T}}^t \gamma_H^R(t-x) (1 - \mu_H^U) \sigma_I^H(x) dx \quad (4.20)$$

$$\sigma_U^D(t) = - \int_{\hat{T}}^t \gamma_U^{D'}(t-x) \mu_U^D \sigma_H^U(x) dx \quad (4.21)$$

$$\sigma_U^R(t) = - \int_{\hat{T}}^t \gamma_U^{R'}(t-x) (1 - \mu_U^D) \sigma_H^U(x) dx. \quad (4.22)$$

Zusammen mit passenden Anfangsbedingungen formuliert dieses Integro-Differentialgleichungssystem (4.4)–(4.22) die Modellgleichungen des IDE-SECIR-Modells. Als Anfangsbedingungen benötigen wir für die in (4.13)–(4.22) spezifizierten Übergänge Werte für $\hat{T} \leq t \leq 0$, das heißt $\sigma_{z_1}^{z_2}(t)$ für alle $\hat{T} \leq t \leq 0$ für sinnvolle Paare $(z_1, z_2) \in \mathcal{Z} \times \mathcal{Z}$. Anfangswerte $z(0)$ für die Kompartimente $z \in \mathcal{Z} \setminus \{S\}$ ergeben sich aus den Anfangsbedingungen für die Übergänge mit den Gleichungen (4.5)–(4.11). Für S ist nur eine Gleichheit für die Ableitung gegeben. Der Wert $S(0)$ könnte mit der Gleichung (4.1) durch

$$S(0) = N_0 - \sum_{z \in \mathcal{Z} \setminus \{S\}} z(0)$$

berechnet werden. Wie bereits erwähnt, können wir ebenso Gleichungen formulieren, welche die Ableitungen der Kompartimente $z \in \mathcal{Z}$ mit den Übergängen verbinden, siehe zum Beispiel Gleichung (4.3). Dies ergibt das System

$$\begin{aligned} S'(t) &= -\sigma_S^E(t) \\ E'(t) &= \sigma_S^E(t) - \sigma_E^C(t) \\ C'(t) &= \sigma_E^C(t) - \sigma_C^I(t) - \sigma_C^R(t) \\ I'(t) &= \sigma_C^I(t) - \sigma_I^H(t) - \sigma_I^R(t) \\ H'(t) &= \sigma_I^H(t) - \sigma_H^U(t) - \sigma_H^R(t) \\ U'(t) &= \sigma_H^U(t) - \sigma_U^D(t) - \sigma_U^R(t) \\ R'(t) &= \sigma_C^R(t) + \sigma_I^R(t) + \sigma_H^R(t) + \sigma_U^R(t) \\ D'(t) &= \sigma_U^D(t). \end{aligned} \quad (4.23)$$

Man kann überprüfen, dass die Formeln (4.4)–(4.12) zu diesem Gleichungssystem (4.23) konsistent sind, wenn man die Gleichungen für die Übergänge (4.13)–(4.22) einsetzt. Tauschen wir die Formeln (4.4)–(4.11) durch das System (4.23) aus, erhalten wir eine ebenso valide Modellformulierung und Gleichungen für alle vorkommenden unbekannt Funktionen. Eventuell wären dann zusätzliche Anfangswerte im Zeitpunkt $t_0 = 0$ für $z \in \mathcal{Z}$ notwendig, wenn wir diese nicht mithilfe der Formeln (4.5)–(4.11) aus den Anfangsbedingungen für die Übergänge berechnen wollen.

Insgesamt ist das IDE-SECIR-Modell im Vergleich zum ODE-Modell wesentlich komplizierter zu formulieren und zu verstehen. Diese Feststellung wurde bereits in Kapitel 3.6 behauptet. Um die Komplexität zu reduzieren, könnte eine Idee von [32, S. 92] übernommen und einige Aufenthaltszeiten wie im ODE-Modell als exponentialverteilt angenommen werden. Dadurch würden sich die Gleichungen für diese Kompartimente auf ähnliche wie im ODE-Modell vereinfachen. Zum Beispiel könnte für die Kompartimente eine beliebige Verteilung der Aufenthaltszeit zugelassen werden, die für die Anzahl der Neuansteckungen bedeutsam sind, das heißt für E , C und I , nicht aber für H und U .

4.1.3 Das LCT-SECIR-Modell

Als letztes Modell erweitern wir das LCT-SIR-Modell um neue Kompartimente. Da wir die Funktionsweise des Linear Chain Tricks in Kapitel 3 gründlich untersucht haben, können wir das ODE-SECIR-Modell leicht auf Erlang-verteilte Aufenthaltszeiten verallgemeinern. Wir müssen zur Formulierung der Modellgleichungen nicht unbedingt das IDE-SECIR-Modell kennen. Aus Kapitel 3 wissen wir, dass eine Erlang-verteilte Aufenthaltszeit mit Formparameter $n \in \mathbb{N}$ und Ratenparameter $n\gamma$ mit $\gamma > 0$ durch eine Kette von n identischen Subkompartimenten erreicht werden kann. Die Subkompartimente sollten jeweils eine exponentialverteilte Aufenthaltszeit mit dem Parameter $n\gamma$ aufweisen. Dies lässt sich durch die Annahme erreichen, dass die Übergänge durch einen linearen Term mit dem Faktor $n\gamma$ beschrieben werden. Die Erkenntnis nutzen wir dazu, eine Erlang-verteilte Aufenthaltszeit für alle Zustände zu inkludieren, für die es sinnvoll ist. Sinnvoll könnte dies für die Menge aus Kompartimenten $\mathcal{A} = \{E, C, I, H, U\} \subset \mathcal{Z}$ sein. Für R und D ist keine Verteilung der Aufenthaltszeit nötig, da aus diesen Gruppen nicht in andere Zustände gewechselt werden kann. Beim Kompartiment S ist der Übergang durch einen nicht-linearen Term beschrieben.

Zur Definition des Modells benötigen wir zusätzliche Notation. Für ein Kompartiment $z \in \mathcal{A}$ wollen wir eine Erlang-verteilte Aufenthaltszeit mit dem Formparameter $n_z \in \mathbb{N}$ und Ratenparameter $n_z\gamma_z$ modellieren. Hierbei entspricht $\gamma_z = T_z^{-1} > 0$, analog zum ODE-SECIR-Modell, der inversen erwarteten Aufenthaltszeit im Kompartiment z . Wie in Kapitel 3.4.2 besprochen, bestimmt die Anzahl der Subkompartimente n_z die Varianz der Verteilung der Aufenthaltszeit. Die Bedeutung der restlichen Parameter bleibt dieselbe wie für das ODE-SECIR-Modell. Betrachten wir das ODE-SECIR-Modell (4.2) und wenden den Linear Chain Trick für alle Kompartimente in \mathcal{A} analog zur Vorgehensweise in Kapitel 3 für I an, erhalten wir das gewöhnliche Differentialgleichungssystem

$$S'(t) = -\frac{S(t)}{N(t)} \rho(t) \phi(t) \left(\xi_C(t) C(t) + \xi_I(t) I(t) \right) \quad (4.24)$$

$$E'_1(t) = \frac{S(t)}{N(t)} \rho(t) \phi(t) \left(\xi_C(t) C(t) + \xi_I(t) I(t) \right) - n_E \gamma_E E_1(t) \quad (4.25)$$

$$E'_j(t) = n_E \gamma_E E_{j-1}(t) - n_E \gamma_E E_j(t) \quad \text{für } j \in \{2, \dots, n_E\} \quad (4.26)$$

$$C'_1(t) = n_E \gamma_E E_{n_E}(t) - n_C \gamma_C C_1(t) \quad (4.27)$$

$$C'_j(t) = n_C \gamma_C C_{j-1}(t) - n_C \gamma_C C_j(t) \quad \text{für } j \in \{2, \dots, n_C\} \quad (4.28)$$

$$I'_1(t) = \mu_C^I n_C \gamma_C C_{n_C}(t) - n_I \gamma_I I_1(t) \quad (4.29)$$

$$I'_j(t) = n_I \gamma_I I_{j-1}(t) - n_I \gamma_I I_j(t) \quad \text{für } j \in \{2, \dots, n_I\} \quad (4.30)$$

$$H'_1(t) = \mu_I^H n_I \gamma_I I_{n_I}(t) - n_H \gamma_H H_1(t) \quad (4.31)$$

$$H'_j(t) = n_H \gamma_H H_{j-1}(t) - n_H \gamma_H H_j(t) \quad \text{für } j \in \{2, \dots, n_H\} \quad (4.32)$$

$$U'_1(t) = \mu_H^U n_H \gamma_H H_{n_H}(t) - n_U \gamma_U U_1(t) \quad (4.33)$$

$$U'_j(t) = n_U \gamma_U U_{j-1}(t) - n_U \gamma_U U_j(t) \quad \text{für } j \in \{2, \dots, n_U\} \quad (4.34)$$

$$R'(t) = \left(1 - \mu_C^I\right) n_C \gamma_C C_{n_C}(t) + \left(1 - \mu_I^H\right) n_I \gamma_I I_{n_I}(t) + \left(1 - \mu_H^U\right) n_H \gamma_H H_{n_H}(t) + \left(1 - \mu_U^D\right) n_U \gamma_U U_{n_U}(t) \quad (4.35)$$

$$D'(t) = \mu_U^D n_U \gamma_U U_{n_U}(t) \quad (4.36)$$

für $t \geq 0$, wobei $z(t) = \sum_{j=1}^{n_z} z_j(t)$ für alle $z \in \mathcal{A}$ gelte. Gemeinsam mit Anfangswerten für die Subkompartimente sowie für die übrigen Kompartimente, das heißt $z_j(0)$ für $z \in \mathcal{A}$ und $j \in \{1, \dots, n_z\}$ und $z(0)$ mit $z \in \{S, R, D\}$, ergibt das System (4.24)–(4.36) ein Anfangswertproblem.

Je nach Fokussierung der Untersuchung kann wie beim IDE-Modell diskutiert werden, ob es sinnvoll ist, für alle Kompartimente in \mathcal{A} gleichzeitig eine Erlang-verteilte Aufenthaltszeit einzuführen. In [32, S. 94] werden beispielsweise I und H in Subkompartimente unterteilt, während die Aufenthaltszeit in E als exponentialverteilt angenommen wird. Die beiden Kompartimente I und H werden im dortigen Modell als infektiös angenommen und der Schwerpunkt der Untersuchung liegt auf den Auswirkungen der Annahmen zu den Verteilungen der Aufenthaltszeiten in diesen infektiösen Gruppen. In [58, Supplementary material] werden in einem SEIR-Modell dagegen für beide Gruppen E und I Subkompartimente eingeführt. Wearing et al. [93, S. 0624ff.] betonen, dass die Annahme einer Exponentialverteilung anstatt einer Erlang-Verteilung für die Latenzzeit und die infektiöse Periode zu falschen Schätzungen führen kann. Um für alle Kompartimente in \mathcal{A} die Möglichkeit zu bieten, Erlang-verteilte Aufenthaltszeiten zu verwenden, haben wir für alle diese Kompartimente Subkompartimente eingeführt. Soll für ein Kompartiment $z \in \mathcal{A}$ eine exponentialverteilte Aufenthaltszeit betrachtet werden, kann $n_z = 1$ eingesetzt werden.

In diesem Kapitel haben wir Modellgleichungen für das ODE-SECIR-, das IDE-SECIR- sowie das LCT-SECIR-Modell vorgestellt. Hierdurch konnten wir die Hypothese aus Kapitel 3.6 bestätigen, dass die Formulierung und das Verständnis der IDE-Modelle deutlich herausfordernder ist als bei einem ODE-Modell. Dies lässt sich an den komplexeren SECIR-Modellen noch deutlicher als an den SIR-Modellen erkennen. Die Betrachtung des LCT-SECIR-Modells war zwar komplizierter als die des ODE-Modells, jedoch noch bedeutend einfacher als die Herleitung des IDE-Modells.

4.2 Numerische Lösungsverfahren

Im Folgenden stellen wir numerische Lösungsverfahren für die drei SECIR-Modelle vor. Die präsentierten Methoden verwenden wir in der Implementierung. Wir betrachten zunächst einen Löser für gewöhnliche Differentialgleichungssysteme. Da sowohl die Modellgleichungen des ODE-SECIR-Modells als auch die des LCT-SECIR-Modells ein gewöhnliches Differentialgleichungssystem bilden, nutzen wir die Methode für beide Modelle. Im Anschluss stellen wir ein Verfahren für das IDE-SECIR-Modell vor.

4.2.1 Lösungsverfahren für die Modelle basierend auf gewöhnlichen Differentialgleichungssystemen

Zur Lösung von gewöhnlichen Differentialgleichungssystemen existieren einige bereits implementierte Methoden aus Softwarepaketen. Die Nutzung dieser anstatt einer selbst implementierten Methode hat den Vorteil, dass die Funktionen häufig gut getestet und effizient implementiert sind. Auf diese Weise sind die Modelle leicht in Software umzusetzen [16, S. 1046]. Wir nutzen zur Lösung von gewöhnlichen Differentialgleichungssystemen die „Runge-Kutta Cash-Karp 5(4)“ Methode, die aktuell ebenso in der Software MEmilio [61] angewendet wird. Die Methode ist ein adaptives Verfahren, bei dem wir nicht zwangsläufig äquidistante Schrittweiten erhalten und uns a priori nicht damit beschäftigen müssen, welche Schrittweite für den Löser ausreichend klein ist. Wir benutzen die bereits implementierte Methode `runge_kutta_cash_karp54` [2] aus dem Modul „Odeint“ von der C++-Softwarebibliothek „boost“ [23]. Diese Lösungsmethode wenden wir sowohl bei dem ODE-SECIR- als auch bei dem LCT-SECIR-Modell an, weil beide Modellgleichungen gewöhnliche Differentialgleichungssysteme sind. Da wir eine Methode aus einer Softwarebibliothek verwenden, gehen wir auf die genaue Funktionsweise der Runge-Kutta Cash-Karp 5(4) Methode nicht näher ein. Der Vollständigkeit halber ist das Butcher-Tableau des Verfahrens jedoch in Tabelle 4.1 angegeben. Die explizite Runge-Kutta-Methode ist von fünfter Ordnung und der Fehler wird unter Verwendung einer Methode der Ordnung Vier geschätzt [1]. Für weitere Informationen zur Funktionsweise der Cash-Karp-Methode vergleiche [73, Kapitel 16.2] in Verbindung mit [13] und der Implementierung in boost.

0							
$\frac{1}{5}$	$\frac{1}{5}$						
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$					
$\frac{3}{5}$	$\frac{3}{10}$	$-\frac{9}{10}$	$\frac{6}{5}$				
1	$-\frac{11}{54}$	$\frac{5}{2}$	$-\frac{70}{27}$	$\frac{35}{27}$			
$\frac{7}{8}$	$\frac{1631}{55296}$	$\frac{175}{512}$	$\frac{575}{13824}$	$\frac{44275}{110592}$	$\frac{253}{4096}$		
	$\frac{37}{378}$	0	$\frac{250}{621}$	$\frac{125}{594}$	0	$\frac{512}{1771}$	Ordnung 5
	$\frac{2825}{27648}$	0	$\frac{18575}{48384}$	$\frac{13525}{55296}$	$\frac{277}{14336}$	$\frac{1}{4}$	Ordnung 4

Tabelle 4.1: Erweitertes Butcher-Tableau des Runge-Kutta Cash-Karp 5(4) Verfahrens. Das Tableau ist aus [13, S. 206] übernommen.

4.2.2 Lösungsverfahren für das IDE-SECIR-Modell

Softwarepakete für die Lösung von Integro-Differentialgleichungen beziehungsweise Systemen wie das des IDE-SECIR-Modells sind selten und häufig auf bestimmte Spezialfälle eingeschränkt [16, S. 1046]. Entsprechend ist in der Software MEmilio [61] ein Löser für die Modellgleichungen implementiert. Die Methode wird in [95] publiziert werden. Da das numerische Lösungsverfahren für das IDE-SECIR-Modell vergleichsweise komplex ist und die genaue Vorgehensweise noch unpubliziert ist, gehen wir auf die Methode genauer

ein. Wir leiten die Verfahrensvorschrift nicht explizit her, geben aber jede Formel an, die zur numerischen Lösung des IDE-SECIR-Modells verwendet wird. Die Methode orientiert sich an einer Arbeit von Messina et al. [69]. Hier wird allerdings nur die Gleichung für S und die Formel für den Infektionsdruck φ diskretisiert. Dabei wird ein SIR-Modell orientiert an [11, Kapitel 4.5] betrachtet. Messina et al. erhalten für ihr Verfahren eine Konvergenz- und Konsistenzordnung von eins [69, S. 242f.].

Beginnen wir mit der Betrachtung des numerischen Lösungsschemas für die Modellgleichungen (4.4)–(4.22) des IDE-SECIR-Modells. Dazu nehmen wir an, dass wir an der Lösung der Modellgleichungen auf einem Intervall $[0, T_{\max}]$ interessiert sind. Mit der Schrittweite $\Delta t > 0$ sei durch $t_m = m \Delta t$ mit $m \in \mathbb{Z}$ ein äquidistantes Gitter gegeben. Die Schrittweite und T_{\max} wählen wir so, dass $\frac{T_{\max}}{\Delta t} \in \mathbb{N}$ gilt. Somit sind Approximationen der Modellgrößen an den Gitterpunkten t_m mit $m \in \{0, 1, \dots, \frac{T_{\max}}{\Delta t}\}$ gesucht. Aus dem Grund der Unterscheidbarkeit zur Notation der Subkompartimente beim LCT-Modell verzichten wir darauf, wie sonst üblich, den approximierten Wert einer Funktion durch den Zeitindex zu kennzeichnen. Stattdessen seien $z(t_m)$ für $z \in \mathcal{Z}$ sowie $\varphi(t_m)$ und $\sigma_{z_1}^{z_2}(t_m)$ mit $z_1, z_2 \in \mathcal{Z}$ die numerischen Approximationen am Gitterpunkt t_m beziehungsweise ein vorausgesetzter Anfangswert. Messina et al. [69, S. 242] schlagen für das Kompartiment S die iterative Lösungsmethode

$$S(t_{m+1}) = \frac{S(t_m)}{1 + \Delta t \varphi(t_m)} \quad (4.37)$$

vor. Die Methode basiert auf einem Differenzenquotienten als Approximation der Ableitung von $S(t)$ und einer Approximation dieser Ableitung anhand von Gleichung (4.4):

$$\frac{S(t_{m+1}) - S(t_m)}{\Delta t} \approx S'(t_{m+1}) \approx -S(t_{m+1}) \varphi(t_m). \quad (4.38)$$

Wenn wir diesen Zusammenhang geeignet umformen, erhalten wir die Berechnungsvorschrift (4.37). Zur Berechnung einer approximativen Lösung von S im Punkt t_{m+1} müssen die Werte für φ und S im Gitterpunkt t_m zuvor bestimmt worden sein. Wenden wir das Vorgehen von [69, S. 242] auf die Gleichung (4.12) für den Infektionsdruck an, erhalten wir die Verfahrensvorschrift

$$\begin{aligned} \varphi(t_{m+1}) = & \frac{\phi(t_{m+1}) \rho(t_{m+1})}{N_0 - D(t_{m+1})} \Delta t \\ & \cdot \left(\sum_{i=a}^m \left(\xi_C(t_{m+1}, t_{m+1-i}) \left(\gamma_C^I(t_{m+1-i}) \mu_C^I + \gamma_C^R(t_{m+1-i}) (1 - \mu_C^I) \right) \sigma_E^C(t_{i+1}) \right. \right. \\ & \left. \left. + \xi_I(t_{m+1}, t_{m+1-i}) \left(\gamma_I^H(t_{m+1-i}) \mu_I^H + \gamma_I^R(t_{m+1-i}) (1 - \mu_I^H) \right) \sigma_C^I(t_{i+1}) \right) \right), \end{aligned}$$

wobei $a \in \mathbb{Z}$ mit $a < 0$ gelte und $t_a = a \Delta t \leq \hat{T}$ ein Zeitpunkt sei, zu dem es noch kein Infektionsgeschehen gab. Dieses numerische Lösungsschema ist ein nicht-standardmäßiges Finite-Differenzen-Verfahren (engl. *non-standard finite difference scheme*), da zur numerischen Quadratur für die Approximation der Funktion auf einem Teilintervall für die Übergänge $\sigma_{z_1}^{z_2}$ der rechte Randpunkt verwendet wird und für die Werte der Parameter

der linke. Nicht-standardmäßige Finite-Differenzen-Verfahren wurden ursprünglich für Differentialgleichungen eingeführt und später für die Anwendung auf Problemstellungen mit Integralen ausgeweitet [69, S. 242]. Eine ähnliche Vorgehensweise lässt sich bei dem Verfahren zur Lösung der Differentialgleichung für S in (4.37) beziehungsweise in (4.38) erkennen. Aus dem Sachzusammenhang folgt, dass die Funktion $S(t)$ nicht-negativ und monoton fallend sein sollte. Messina et al. [69, Theorem 3.4] zeigen für ihren Anwendungsfall, dass die mit der beschriebenen Diskretisierung berechnete approximative Lösung diese beiden Eigenschaften für alle $\Delta t \geq 0$ erfüllt. An einem Beispiel demonstrieren sie, dass die naheliegende Methode der Trapezregel (vergleiche hierfür auch [40, S. 317]) diese Eigenschaften nicht immer aufweist [69, S. 249]. Für unseren Anwendungsfall müsste das Verhalten validiert werden, jedoch spricht die Untersuchung für die Verwendung der nicht-standardmäßigen Methode.

Bei der Durchführung von Simulationen ersetzen wir die Variablen der Übergänge $\sigma_{z_1}^{z_2}$. Hierbei seien $z_1, z_2 \in \mathcal{Z}$ zwei sinnvolle und zueinander passende Zustände. Wir definieren $\Sigma_{z_1}^{z_2}(t_m)$ als die Anzahl der Individuen, die im Zeitintervall $(t_{m-1}, t_m]$ von z_1 nach z_2 wechseln. Es gilt folglich

$$\Sigma_{z_1}^{z_2}(t_m) = \int_{t_{m-1}}^{t_m} \sigma_{z_1}^{z_2}(t) dt.$$

Die Größe $\Sigma_{z_1}^{z_2}$ lässt sich aus realen Daten leichter bestimmen als die Anzahl der Personen $\sigma_{z_1}^{z_2}$, die genau zu einem Zeitpunkt das Kompartiment wechseln. Wenn wir die Annahme treffen, dass $\sigma_{z_1}^{z_2}$ auf den Intervallen $(t_{m-1}, t_m]$ jeweils konstant ist, folgt $\Sigma_{z_1}^{z_2}(t_m) = \Delta t \sigma_{z_1}^{z_2}(t_m)$. In der Praxis ist Δt in der Regel so klein, dass Änderungen im Intervall $(t_{m-1}, t_m]$ von $\sigma_{z_1}^{z_2}(t)$ nicht in den realen Daten registriert werden können, zum Beispiel $\Delta t \leq 1$ Tag. Ersetzen wir $\Delta t \sigma_{z_1}^{z_2}(t_m)$ in den Modellformulierungen durch $\Sigma_{z_1}^{z_2}(t_m)$, erhalten wir für den Infektionsdruck anstatt der zuvor formulierten Gleichung

$$\begin{aligned} \varphi(t_{m+1}) = & \frac{\phi(t_{m+1}) \rho(t_{m+1})}{N_0 - D(t_{m+1})} \sum_{i=a}^m \left(\xi_C(t_{m+1}, t_{m+1-i}) \left(\gamma_C^I(t_{m+1-i}) \mu_C^I \right. \right. \\ & + \gamma_C^R(t_{m+1-i}) (1 - \mu_C^I) \left. \right) \Sigma_E^C(t_{i+1}) + \xi_I(t_{m+1}, t_{m+1-i}) \\ & \cdot \left(\gamma_I^H(t_{m+1-i}) \mu_I^H + \gamma_I^R(t_{m+1-i}) (1 - \mu_I^H) \right) \Sigma_C^I(t_{i+1}) \left. \right). \end{aligned} \quad (4.39)$$

Für die Berechnung von $\varphi(t_{m+1})$ müssen einige andere Größen zuvor bestimmt worden sein, zum Beispiel $\Sigma_E^C(t_{i+1})$ für alle $a \leq i \leq m$. Für die Verfahrensvorschriften der Übergänge übernehmen wir das nicht-standardmäßige Verfahren. Unter Verwendung der eingeführten Variablen $\Sigma_{z_1}^{z_2}(t_m) = \Delta t \sigma_{z_1}^{z_2}(t_m)$ erhalten wir als Lösungsmethode für die Gleichungen (4.13)–(4.22) entsprechend

$$\Sigma_S^E(t_{m+1}) = \Delta t \varphi(t_m) S(t_{m+1}) \quad (4.40)$$

$$\Sigma_E^C(t_{m+1}) = -\Delta t \sum_{i=a}^m \gamma_E^C(t_{m+1-i}) \Sigma_S^E(t_{i+1}) \quad (4.41)$$

$$\Sigma_C^I(t_{m+1}) = -\Delta t \mu_C^I \sum_{i=a}^m \gamma_C^I(t_{m+1-i}) \Sigma_E^C(t_{i+1}) \quad (4.42)$$

$$\Sigma_C^R(t_{m+1}) = -\Delta t \left(1 - \mu_C^I\right) \sum_{i=a}^m \gamma_C^{R'}(t_{m+1-i}) \Sigma_E^C(t_{i+1}) \quad (4.43)$$

$$\Sigma_I^H(t_{m+1}) = -\Delta t \mu_I^H \sum_{i=a}^m \gamma_I^{H'}(t_{m+1-i}) \Sigma_C^I(t_{i+1}) \quad (4.44)$$

$$\Sigma_I^R(t_{m+1}) = -\Delta t \left(1 - \mu_I^H\right) \sum_{i=a}^m \gamma_I^{R'}(t_{m+1-i}) \Sigma_C^I(t_{i+1}) \quad (4.45)$$

$$\Sigma_H^U(t_{m+1}) = -\Delta t \mu_H^U \sum_{i=a}^m \gamma_H^{U'}(t_{m+1-i}) \Sigma_I^H(t_{i+1}) \quad (4.46)$$

$$\Sigma_H^R(t_{m+1}) = -\Delta t \left(1 - \mu_H^U\right) \sum_{i=a}^m \gamma_H^{R'}(t_{m+1-i}) \Sigma_I^H(t_{i+1}) \quad (4.47)$$

$$\Sigma_U^D(t_{m+1}) = -\Delta t \mu_U^D \sum_{i=a}^m \gamma_U^{D'}(t_{m+1-i}) \Sigma_H^U(t_{i+1}) \quad (4.48)$$

$$\Sigma_U^R(t_{m+1}) = -\Delta t \left(1 - \mu_U^D\right) \sum_{i=a}^m \gamma_U^{R'}(t_{m+1-i}) \Sigma_H^U(t_{i+1}). \quad (4.49)$$

Die Gleichung für $\Sigma_S^E(t_{m+1})$ basiert auf dem Zusammenhang (4.13) und der Herleitung des numerischen Verfahrens für S in Gleichung (4.38). Bei den übrigen Gleichungen wird zur numerischen Quadratur für die Approximation des Integranden auf einem Teilintervall für die Funktionen $\gamma_{z_1}^{z_2'}$ jeweils der linke Randpunkt verwendet und für $\Sigma_{z_1}^{z_2}$ der rechte, sodass wir ein nicht-standardmäßiges Verfahren erhalten. In den Formeln finden sich Auswertungen der Ableitungen der Funktionen $\gamma_{z_1}^{z_2}$. Diese Ableitungen sind nicht immer berechenbar, sodass wir sie numerisch approximieren. Dazu verwenden wir einen Differenzenquotienten, sodass wir die Approximation

$$\gamma_{z_1}^{z_2'}(t_i) = \frac{\gamma_{z_1}^{z_2}(t_i) - \gamma_{z_1}^{z_2}(t_{i-1})}{\Delta t}$$

erhalten. Zuletzt benötigen wir eine Diskretisierung für die übrigen Kompartimente. Wir wenden die nicht-standardmäßige Methode auf die Modellgleichungen (4.5)–(4.11) an und erhalten

$$E(t_{m+1}) = \sum_{i=a}^m \gamma_E^C(t_{m+1-i}) \Sigma_S^E(t_{i+1}) \quad (4.50)$$

$$C(t_{m+1}) = \sum_{i=a}^m \left(\gamma_C^I(t_{m+1-i}) \mu_C^I + \gamma_C^R(t_{m+1-i}) \left(1 - \mu_C^I\right) \right) \Sigma_E^C(t_{i+1}) \quad (4.51)$$

$$I(t_{m+1}) = \sum_{i=a}^m \left(\gamma_I^H(t_{m+1-i}) \mu_I^H + \gamma_I^R(t_{m+1-i}) \left(1 - \mu_I^H\right) \right) \Sigma_C^I(t_{i+1}) \quad (4.52)$$

$$H(t_{m+1}) = \sum_{i=a}^m \left(\gamma_H^U(t_{m+1-i}) \mu_H^U + \gamma_H^R(t_{m+1-i}) \left(1 - \mu_H^U\right) \right) \Sigma_I^H(t_{i+1}) \quad (4.53)$$

$$U(t_{m+1}) = \sum_{i=a}^m \left(\gamma_U^D(t_{m+1-i}) \mu_U^D + \gamma_U^R(t_{m+1-i}) \left(1 - \mu_U^D\right) \right) \Sigma_H^U(t_{i+1}) \quad (4.54)$$

$$R(t_{m+1}) = R(t_m) + \Sigma_C^R(t_{m+1}) + \Sigma_I^R(t_{m+1}) + \Sigma_H^R(t_{m+1}) + \Sigma_U^R(t_{m+1}) \quad (4.55)$$

$$D(t_{m+1}) = D(t_m) + \Sigma_U^D(t_{m+1}). \quad (4.56)$$

Die behaupteten Gleichungen der Größen R und D entsprechen nicht direkt den Approximationen der Integrale aus den Gleichungen (4.10) und (4.11). Dies würde für die Gruppe R die Vorschrift

$$R(t_{m+1}) = \sum_{i=a}^m \Sigma_C^R(t_{i+1}) + \Sigma_I^R(t_{i+1}) + \Sigma_H^R(t_{i+1}) + \Sigma_U^R(t_{i+1}) \quad (4.57)$$

ergeben. Für das obige Verfahren (4.55) setzen wir in diese Gleichung für $R(t_{m+1})$ die Formulierung für $R(t_m)$ ein, sodass die im vorigen Schritt berechnete Summe wiederverwendet wird. Für D wenden wir ein analoges Vorgehen an. Zur Bestimmung von $R(t_0)$ und $D(t_0)$ könnte einmalig die Summe in (4.57) und die entsprechende für D berechnet werden. Andere Berechnungen von Summen aus verschiedenen Zeitschritten oder von den Übergängen $\Sigma_{z_1}^{z_2}$ zu den Kompartimenten können nicht wiederverwendet werden, da sich die Summanden stets unterscheiden. Die Berechnungsvorschriften (4.55) und (4.56) entsprechen der Anwendung eines impliziten Euler-Verfahrens (siehe hierzu [40, Kapitel 75]) auf die Gleichungen für R und D aus dem Gleichungssystem (4.23). Die direkte Verwendung der Formeln (4.23), ausgenommen der Formel für S , als Grundlage zur Diskretisierung könnte Vorteile haben, da keine Summen approximiert werden müssen und weniger Berechnungen notwendig sind.

Damit haben wir für alle vorkommenden Modellgrößen numerische Lösungsverfahren besprochen. Einige Ergebnisse hängen von anderen Berechnungen aus demselben Zeitschritt ab. Deshalb ist die richtige Reihenfolge der Berechnung essentiell. Der Übersichtlichkeit halber halten wir die Vorgehensweise als Algorithmus mit Verweisen auf die jeweilige Berechnungsformel fest. Die Implementierung des IDE-Modells werden wir nicht näher beschreiben, sodass aus dem Algorithmus 1 die Funktionsweise und die richtige Reihenfolge der Berechnung deutlich wird. Die Abhängigkeiten der Modellgrößen innerhalb eines Zeitschritts könnten bei einer möglichen späteren Parallelisierung zu Problemen führen. Als Anfangsbedingungen benötigen wir wie in Kapitel 4.1.2, übersetzt in die Schreibweise dieses Kapitels, die Übergänge $\Sigma_{z_1}^{z_2}(t_m)$ für alle $a \leq m \leq 0$. Zusätzlich

Algorithmus 1: Vorgehensweise zur numerischen Lösung der Modellgleichungen des IDE-SECIR-Modells auf dem Intervall $[0, T_{\max}]$.

Input: Verlauf der Übergänge $\Sigma_{z_1}^{z_2}(t_m)$ für $a \leq m \leq 0$ und sinnvolle $z_1, z_2 \in \mathcal{Z}$; $z(t_0)$ für alle $z \in \mathcal{Z}$; alle benötigten Parameter; Schrittweite Δt ; T_{\max}

Berechne $\varphi(t_0)$ mittels (4.39)

for $m = 0$ **to** $\frac{T_{\max}}{\Delta t} - 1$ **do**

 Berechne $S(t_{m+1})$ mittels (4.37)

 Berechne $\Sigma_{z_1}^{z_2}(t_{m+1})$ mittels (4.40)–(4.49) in der dortigen Reihenfolge

 Berechne $z(t_{m+1})$ für $z \in \mathcal{Z} \setminus \{S\}$ mittels (4.50)–(4.56)

 Berechne $\varphi(t_{m+1})$ mittels (4.39)

end for

seien die Anfangswerte $z(t_0)$ für alle Kompartimente $z \in \mathcal{Z}$ vorausgesetzt. Wie wir diese bestimmen können, werden wir im nächsten Abschnitt untersuchen.

Aus diesem Abschnitt geht hervor, dass nicht nur die Formulierung der Modellgleichungen des IDE-Modells anspruchsvoller ist als die der übrigen Modelle. Auch die numerische Lösung stellt sich als erheblich aufwendiger heraus. Im Fall der beiden Modelle, deren Modellgleichungen als gewöhnliche Differentialgleichungssysteme formuliert sind, stehen praktikable Methoden in Softwarebibliotheken zur Verfügung und die Theorie der Löser ist vergleichsweise gut erforscht.

4.3 Wahl der Anfangsbedingungen

In diesem Kapitel verschaffen wir uns einen Überblick über verschiedene Möglichkeiten zur Initialisierung der drei SECIR-Modelle. Die Werte sind als Eingabe für die Simulationen mittels der numerischen Lösungsverfahren aus Kapitel 4.2 notwendig. Dabei wird auf die Verfügbarkeit von realen Daten zur Bestimmung der Anfangswerte geachtet. Für die Simulationen in Kapitel 5 am Beispiel von COVID-19 können die Methoden angewendet werden.

Initialisierung des ODE-SECIR-Modells nach Koslow et al.

Zunächst besprechen wir eine Möglichkeit, das ODE-SECIR-Modell mittels realer Daten zu initialisieren, die sich an der Umsetzung in der Software MEmilio [61] beziehungsweise an der Ausführung in [56, S2 Appendix] orientiert. Für SARS-CoV-2 veröffentlicht das „Robert Koch-Institut“ (RKI) tägliche Werte bezüglich der gesamten bestätigten Fälle sowie der Todesfälle [76]. Diese können mittels der Datei `pycode/memilio-epidata/memilio/epidata/getCaseData.py` aus der Software MEmilio [61] heruntergeladen und aufbereitet werden. Ähnliche Werte könnten für andere Krankheiten, für welche eine Ausbreitung simuliert werden soll, von öffentlichen Quellen bereitgestellt werden. Daher ist die Methode zur Initialisierung nicht auf den Erreger SARS-CoV-2 beschränkt. Im Folgenden besprechen wir die Initialisierung beispielhaft an der Krankheit COVID-19 und legen die verfügbaren gemeldeten Daten zugrunde. Wir berechnen für die bestätigten kumulierten Fälle und für die Todesfälle einen gleitenden Durchschnitt unter Berücksichtigung von sieben Tagen. Entsprechend übernehmen wir nicht direkt die gemeldeten Werte, sondern ermitteln für jeden Tag den Durchschnitt aus dem Tageswert sowie den vorigen und den folgenden drei Tagen. Auf diese Weise könnten Verzerrungseffekte, wie beispielsweise durch Wochenenden, geglättet werden. Wenn nicht alle Krankheitsfälle gemeldet werden, wird die Anzahl der bestätigten Fälle die tatsächliche Fallzahl unterschätzen. Möchten wir dies in unseren Simulationen berücksichtigen, können wir die bestätigten täglichen Fallzahlen mit einem Faktor $d > 1$ skalieren. Weitere Effekte, wie beispielsweise eine Verzögerung bei den Fallmeldungen [21, S. 2], vernachlässigen wir. Bezeichnen wir die extrapolierten Daten der kumulierten bestätigten Fälle am Tag t mit $F_{\text{MAR}}(t)$ und die Anzahl an Todesfällen mit $D_{\text{MAR}}(t)$. Hierbei steht „MAR“ für *moving average reported*, sodass gekennzeichnet wird, dass wir einen gleitenden Durchschnitt der gemeldeten Werte verwenden. Aus [56, S2 Appendix] übernehmen wir die Annahme,

dass die bestätigten Fälle die symptomatischen Individuen widerspiegeln. Das bedeutet, dass $F_{\text{MAR}}(t)$ der Gesamtzahl aller Individuen entspricht, die aktuell symptomatisch sind oder in der Vergangenheit symptomatisch waren. Der Zeitpunkt des Symptombeginns entspricht dem Zeitpunkt des positiven Tests und der Berücksichtigung in der Statistik des RKIs.

Wir schätzen aus den beiden Größen $F_{\text{MAR}}(t)$ und $D_{\text{MAR}}(t)$ Anfangswerte $z(t_0)$ für alle $z \in \mathcal{Z}$. Dafür gehen wir davon aus, dass jedes Individuum exakt die durchschnittliche Verweildauer von T_z Tagen in einem Zustand $z \in \mathcal{Z}$ verbleibt. Wir beginnen mit der Gruppe I , da diese nach der vorigen Annahme direkt mit F_{MAR} zusammenhängt. Eine Person, die vor dem Zeitpunkt $t_0 - T_I$ positiv getestet wurde, ist zum Zeitpunkt t_0 nach der durchschnittlichen Aufenthaltszeit bereits in das nächste Kompartiment gewechselt. Individuen, die zwischen den Zeitpunkten $t_0 - T_I$ und t_0 als bestätigte Fälle registriert wurden, befinden sich gemäß der durchschnittlichen Aufenthaltszeit T_I noch im Kompartiment I . Entsprechend gleicht $I(t_0)$ der Anzahl der Fälle, die in diesem Zeitraum bestätigt wurden. Den Wert können wir berechnen, indem wir die bestätigten kumulierten Fälle zum Zeitpunkt t_0 verwenden und die Fälle abziehen, die bereits zum Zeitpunkt $t_0 - T_I$ bestätigt waren, das heißt

$$I(t_0) = F_{\text{MAR}}(t_0) - F_{\text{MAR}}(t_0 - T_I). \quad (4.58)$$

Hospitalisierte Individuen wurden im Zeitraum zwischen $t_0 - T_I - T_H$ und $t_0 - T_I$ symptomatisch. Dabei wechselt ein Anteil von μ_I^H der symptomatischen Personen im Krankheitsverlauf nach H , sodass wir den Wert

$$H(t_0) = \mu_I^H \left(F_{\text{MAR}}(t_0 - T_I) - F_{\text{MAR}}(t_0 - T_I - T_H) \right)$$

als Anfangswert für H erhalten. Analog ergibt sich der Anfangswert für U durch

$$U(t_0) = \mu_H^U \mu_I^H \left(F_{\text{MAR}}(t_0 - T_I - T_H) - F_{\text{MAR}}(t_0 - T_I - T_H - T_U) \right).$$

In [56, S2 Appendix] wird eine alternative Initialisierung der Anzahl der Intensivpatienten über eine zusätzliche Datenquelle angegeben. Auf diese Möglichkeit wird hier verzichtet. Die bestätigten Todesfälle $D_{\text{MAR}}(t)$ werden nicht mit dem Tag des Todesfalls, sondern mit dem Symptombeginn des Verstorbenen veröffentlicht. Entsprechend verschieben wir die Todesfälle zeitlich und bekommen die Formel

$$D(t_0) = D_{\text{MAR}}(t_0 - T_I - T_H - T_U). \quad (4.59)$$

Die Anzahl $R(t_0)$ der genesenen Individuen entspricht den gesamten bestätigten Fällen $F_{\text{MAR}}(t_0)$ abzüglich der Todesfälle und der Fälle, die zum Zeitpunkt t_0 noch erkrankt sind. Hieraus folgt der Anfangswert

$$R(t_0) = F_{\text{MAR}}(t_0) - I(t_0) - H(t_0) - U(t_0) - D(t_0). \quad (4.60)$$

Vereinfachend berücksichtigen wir hier die Individuen mit einem asymptomatischen Verlauf nicht, das heißt die Personen, die direkt von C nach R gewechselt sind. Individuen, die sich gegenwärtig im Zustand C befinden und im Verlauf ihrer Erkrankung nach I

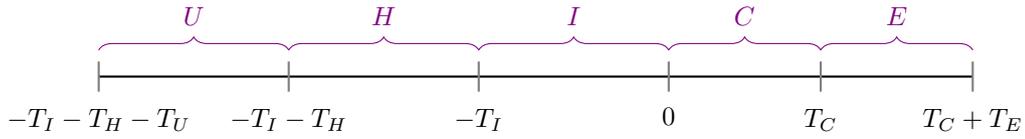


Abbildung 4.2: Schema zur Initialisierung des ODE-Modells nach Koslow et al. Schaubild zur Bestimmung der Anfangswerte für das ODE-SECIR-Modell für den Zeitpunkt $t_0 = 0$. Die Klammern kennzeichnen, dass die Anzahl der in dem Zeitintervall bestätigten Krankheitsfälle den Anfangswert in dem angegebenen Kompartiment bestimmt. Die Abstände zwischen den Zeiten sind willkürlich gesetzt und entsprechen nicht den später verwendeten Parametern. Diese Abbildung ist angelehnt an [56, S2 Appendix].

übergehen, zeigen zwischen den Zeitpunkten t_0 und $t_0 + T_C$ Symptome. Dies deckt nur einen Anteil von μ_C^I aller Personen in C ab, weshalb wir mit der passenden Skalierung den Wert

$$C(t_0) = \frac{1}{\mu_C^I} \left(F_{\text{MAR}}(t_0 + T_C) - F_{\text{MAR}}(t_0) \right)$$

erhalten. An dieser Stelle sind Daten $F_{\text{MAR}}(t)$ erforderlich, die zeitlich nach dem Startzeitpunkt $t_0 = 0$ liegen. In unserem Anwendungsfall stellt dies kein Problem dar, da wir Zeiträume in der Vergangenheit betrachten, um die Ergebnisse mit realen Daten vergleichen zu können. Die entsprechende Überlegung für die Anzahl der Individuen, die sich in der Latenzzeit befinden, führt zu dem Ausdruck

$$E(t_0) = \frac{1}{\mu_C^I} \left(F_{\text{MAR}}(t_0 + T_C + T_E) - F_{\text{MAR}}(t_0 + T_C) \right). \quad (4.61)$$

Die jeweils relevanten Zeitintervalle für die Bestimmung des Anfangswerts eines Zustands sind schematisch in Abbildung 4.2 dargestellt. Das Kompartiment S mit Personen, die sich noch mit der betrachteten Krankheit anstecken können, initialisieren wir durch

$$S(t_0) = N_0 - \sum_{z \in \mathcal{Z} \setminus \{S\}} z(t_0), \quad (4.62)$$

da die Gesamtzahl der Individuen aus allen Kompartimenten der Bevölkerungszahl N_0 entsprechen soll. Damit haben wir Formeln für alle Anfangswerte für das ODE-SECIR-Modell auf der Grundlage der Daten des RKIs kennengelernt.

Initialisierung des LCT-Modells nach Hurtado et al.

Nun widmen wir uns einer Methodik zur Ermittlung einer Verteilung der Anfangswerte auf die Subkompartimente des LCT-SECIR-Modells anhand der verfügbaren realen Daten. Wie bereits in Kapitel 3.2 erläutert, wird von Hurtado et al. [45, S. 1836] eine Initialisierung verwendet, bei welcher nur das erste Subkompartiment von I einen Wert zugewiesen bekommt. Die übrigen Subkompartimente werden mit null initialisiert. Für mehr Zustände mit Erlang-verteilten Aufenthaltszeiten ergibt sich aus [45, Corollary 1 mit

Theorem 1], dass diese Methode für jeden Zustand angewendet wird. Für alle Gruppen kann jeweils das erste Subkompartiment mit einem Anfangswert initialisiert werden und die übrigen erhalten den Wert Null. Wenden wir diese Initialisierung auf das LCT-SECIR-Modell an, erhalten wir für alle $z \in \mathcal{A}$, dass $z_1(t_0) = z(t_0)$ und $z_j(t_0) = 0$ für alle $j \in \{2, \dots, n_z\}$. Die Werte $z(t_0)$ könnten wir auf die für das ODE-Modell beschriebene Weise nach Koslow et al. [56, S2 Appendix] berechnen. Die übrigen Anfangswerte für die Gruppen $z \in \mathcal{Z} \setminus \mathcal{A}$ können mit derselben Methode bestimmt werden. Feng et al. verfolgen eine ähnliche Technik zur Initialisierung, siehe [32, S. 92f.]. In ihrem Modell mit sechs Krankheitszuständen werden nur die Kompartimente S , E und I mit Werten größer als null initialisiert. Der Einfachheit halber wird beispielsweise die Anzahl der Hospitalisierten zum Startzeitpunkt als null angenommen. Für den Zustand I wird in der Publikation durch den Linear Chain Trick eine Erlang-verteilte Aufenthaltszeit erreicht. Wie in [45] wird nur das erste Subkompartiment mit einem Anfangswert initialisiert und die übrigen mit null. In Kapitel 3.2 wurde erwähnt, dass die Methode von Hurtado et al. [45, S. 1874] als Vereinfachung bezeichnet wird. Entsprechend betrachten wir noch andere Möglichkeiten zur Initialisierung des LCT-Modells.

Erweiterung der Methode nach Koslow et al. zur Initialisierung des LCT-SECIR-Modells

Zu Beginn dieses Abschnitts haben wir eine Initialisierung des ODE-SECIR-Modells diskutiert, wie sie in [56, S2 Appendix] auf der Grundlage realer RKI-Daten erfolgt. Die Methode basiert auf erwarteten Aufenthaltszeiten in den Krankheitszuständen. Dieses Verfahren überführen wir unter Berücksichtigung der erwarteten Aufenthaltszeiten in den Subkompartimenten auf das LCT-SECIR-Modell. Grundsätzlich verfolgen wir denselben Ansatz wie zuvor. Wir teilen das für den jeweiligen Krankheitszustand relevante Zeitintervall jedoch zusätzlich in Teilintervalle für die Subkompartimente auf. Betrachten wir den Zustand I mit einer Unterteilung in $n_I \in \mathbb{N}$ Subkompartimente und der erwarteten gesamten Aufenthaltszeit T_I . Die erwartete Aufenthaltszeit in jedem der Subkompartimente entspricht $\frac{T_I}{n_I}$, vergleiche Kapitel 3.4.1. Individuen aus dem Subkompartiment I_1 wurden erst kürzlich symptomatisch, sodass alle bestätigten Fälle zwischen $t_0 - \frac{T_I}{n_I}$ und t_0 diesem Subkompartiment zugeordnet werden. Falls $n_I \geq 2$ gilt, befinden sich im Subkompartiment I_2 entsprechend alle Individuen, die zwischen den Zeitpunkten $t_0 - 2\frac{T_I}{n_I}$ und $t_0 - \frac{T_I}{n_I}$ symptomatisch wurden. Setzen wir dies für die übrigen Subkompartimente fort, erhalten wir für $j \in \{1, \dots, n_I\}$ die Anfangswerte

$$I_j(t_0) = F_{\text{MAR}}\left(t_0 - (j-1)\frac{T_I}{n_I}\right) - F_{\text{MAR}}\left(t_0 - j\frac{T_I}{n_I}\right).$$

Die Anzahl der kumulierten bestätigten Fälle wird vom RKI pro Tag angegeben. Das bedeutet, dass wir die extrapolierten Daten $F_{\text{MAR}}(t)$ für $t \in \mathbb{Z}$ kennen. Für die angegebene Formel für die Subkompartimente von I benötigen wir allerdings auch Werte $F_{\text{MAR}}(t)$ mit $t \notin \mathbb{Z}$. Diesen Fall haben wir in der Initialisierung des ODE-Modells ebenso, da die erwarteten Aufenthaltszeiten T_z nicht unbedingt ganzen Zahlen entsprechen müssen. In der Software MEmilio [61] wird das Problem gelöst, indem der benötigte Zeitwert auf

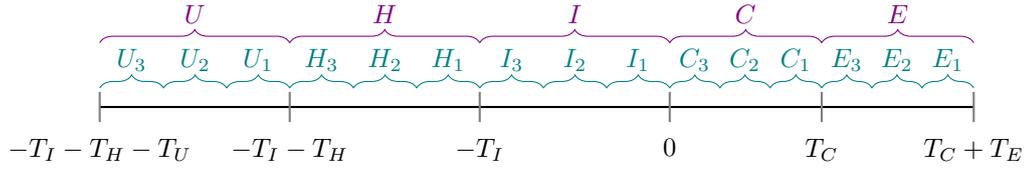


Abbildung 4.3: Schema einer Initialisierung des LCT-Modells. Visualisierung der Bestimmung der Anfangswerte für das LCT-SECIR-Modell nach der Methode von Koslow et al. für den Zeitpunkt $t_0 = 0$. Das Diagramm ist für ein Modell mit $n_z = 3$ Subkompartimenten für alle $z \in \mathcal{A}$ erstellt. Die Klammern geben an, dass die Anzahl der in dem Zeitintervall bestätigten Fälle den Anfangswert für das angegebene Kompartiment oder Subkompartiment bestimmt. Die Abstände zwischen den Zeiten sind willkürlich gesetzt und entsprechen nicht den später verwendeten Parametern. Diese Darstellung ist von [56, S2 Appendix] inspiriert.

die nächste ganze Zahl gerundet wird und damit der passende Wert in den Daten des RKIs zu finden ist. Für den Anwendungsfall des LCT-Modells ist dies keine Option, da beispielsweise $\frac{T_I}{n_I}$ klein sein könnte und dann viele Subkompartimente von I mit null initialisiert werden würden. Aus diesem Grund verwenden wir eine lineare Interpolation zwischen den Tageswerten, vergleiche hierfür [73, S. 113f.]. Für einen beliebigen Zeitpunkt $t \in \mathbb{R}$ nutzen wir somit die Formel

$$F_{\text{MAR}}(t) = F_{\text{MAR}}(\lfloor t \rfloor) + (t - \lfloor t \rfloor) \left(F_{\text{MAR}}(\lceil t \rceil) - F_{\text{MAR}}(\lfloor t \rfloor) \right).$$

Hier werden untere und obere Gauß-Klammern genutzt, wobei $\lfloor t \rfloor$ die größte ganze Zahl mit $\lfloor t \rfloor \leq t$ bezeichnet und $\lceil t \rceil$ die kleinste ganze Zahl mit $t \leq \lceil t \rceil$, vergleiche [33, S. 32]. Die Verwendung von linear interpolierten Werten ist auch für das ODE-Modell die genauere Option, sodass wir die Berechnung auch für das Modell anwenden. Teilen wir die Formeln für die Anfangswerte der Kompartimente H und U des ODE-Modells mittels derselben Überlegungen wie für I auf die Subkompartimente auf, erhalten wir für $j \in \{1, \dots, n_H\}$

$$H_j(t_0) = \mu_I^H \left(F_{\text{MAR}} \left(t_0 - T_I - (j-1) \frac{T_H}{n_H} \right) - F_{\text{MAR}} \left(t_0 - T_I - j \frac{T_H}{n_H} \right) \right)$$

und für $j \in \{1, \dots, n_U\}$

$$U_j(t_0) = \mu_H^U \mu_I^H \left(F_{\text{MAR}} \left(t_0 - T_I - T_H - (j-1) \frac{T_U}{n_U} \right) - F_{\text{MAR}} \left(t_0 - T_I - T_H - j \frac{T_U}{n_U} \right) \right).$$

Für die Zustände E und C sind wieder Datenpunkte $F_{\text{MAR}}(t)$ nach dem Startzeitpunkt t_0 notwendig. Personen in einem fortgeschrittenen Subkompartiment in C werden zuerst Symptome entwickeln. In C_{n_C} befinden sich zum Startzeitpunkt t_0 Individuen, die zwischen dem Zeitpunkt t_0 und $t_0 + \frac{T_C}{n_C}$ Symptome entwickeln oder gesunden, falls sie einen asymptomatischen Verlauf haben. Gilt $n_C \geq 2$, werden Personen in C_{n_C-1} mit

einem symptomatischen Verlauf zwischen $t_0 + \frac{T_C}{n_C}$ und $t_0 + 2\frac{T_C}{n_C}$ positiv getestet. Fortgesetzt auf die übrigen Subkompartimente erhalten wir für $j \in \{1, \dots, n_C\}$ die Initialisierung

$$C_j(t_0) = \frac{1}{\mu_C^I} \left(F_{\text{MAR}} \left(t_0 + (n_C - j + 1) \frac{T_C}{n_C} \right) - F_{\text{MAR}} \left(t_0 + (n_C - j) \frac{T_C}{n_C} \right) \right).$$

Ein analoges Vorgehen für den Krankheitszustand E ergibt die Anfangswerte

$$E_j(t_0) = \frac{1}{\mu_C^I} \left(F_{\text{MAR}} \left(t_0 + T_C + (n_E - j + 1) \frac{T_E}{n_E} \right) - F_{\text{MAR}} \left(t_0 + T_C + (n_E - j) \frac{T_E}{n_E} \right) \right)$$

für $j \in \{1, \dots, n_E\}$. In Abbildung 4.3 ist die Aufteilung der jeweils relevanten Zeitintervalle für die Krankheitszustände auf die Subkompartimente schematisch dargestellt. Für die Gruppen, die sich nicht für die Einführung von Subkompartimenten eignen, können wir die Initialisierung des ODE-Modells übernehmen. Entsprechend berechnen wir $D(t_0)$, $R(t_0)$ und $S(t_0)$ nach den Formeln (4.59), (4.60) und (4.62). Damit haben wir das Verfahren zur Initialisierung nach Koslow et al. [56, S2 Appendix] für das ODE-SECIR-Modell auf das LCT-SECIR-Modell übertragen. Wenn wir in die Formeln die Wahl $n_z = 1$ für alle $z \in \mathcal{A}$ einsetzen, unter der sich das LCT-Modell auf das ODE-Modell reduziert, erhalten wir die Methode nach Koslow et al. [56, S2 Appendix] für das ODE-Modell.

Initialisierung der Modelle mit gegebenen Werten für die Übergänge $\Sigma_{z_1}^{z_2}$

Als Nächstes besprechen wir eine Initialisierung für das IDE-SECIR-Modell. Wir gehen dazu davon aus, dass Werte für die Übergänge $\Sigma_{z_1}^{z_2}(t_m)$ für alle $a \leq m \leq 0$ und für alle sinnvollen Paare $(z_1, z_2) \in \mathcal{Z} \times \mathcal{Z}$ bekannt sind, siehe für die Notation auch Kapitel 4.2.2. Hieraus leiten wir die benötigten Anfangswerte für die Kompartimente $z(t_0)$ für $z \in \mathcal{Z}$ im IDE-Modell her. Diese müssen als Eingabe für den Algorithmus 1 gegeben sein. Die Anfangswerte $E(t_0)$, $C(t_0)$, $I(t_0)$, $H(t_0)$ und $U(t_0)$ können aus den vorausgesetzten Übergängen mittels der Gleichungen (4.50)–(4.54) bestimmt werden. Dafür benötigen wir Werte für die Übergänge $\Sigma_{z_1}^{z_2}(t_m)$ für alle $m \in \mathbb{Z}$ mit $a \leq m \leq 0$. Der Zeitpunkt t_a liegt potentiell weit in der Vergangenheit, sodass eine lange Zeitreihe notwendig ist. Auf der Grundlage der Funktionen $\gamma_{z_1}^{z_2}$ kann der benötigte Zeitraum manchmal eingeschränkt werden. Wählen wir einen Gitterpunkt t_b , sodass

$$b = \inf \{ i : i \in \mathbb{Z}, i \leq 0 \text{ und } \exists (z_1, z_2) \in \mathcal{Z} \times \mathcal{Z} \text{ ein sinnvolles Paar, sodass } \gamma_{z_1}^{z_2}(t_{-i}) > 0 \}$$

gilt. Zur Berechnung von $E(t_0)$, $C(t_0)$, $I(t_0)$, $H(t_0)$ und $U(t_0)$ mittels (4.50)–(4.54) sowie zur Durchführung von Algorithmus 1 genügt es, wenn die $\Sigma_{z_1}^{z_2}(t_m)$ für alle $b \leq m \leq 0$ bekannt sind. Betrachten wir alle Summen aus den Berechnungsvorschriften, entspricht b dem kleinsten Summenindex, für den der entsprechende Summand in einer der Vorschriften ungleich null sein könnte. Dadurch kann die benötigte Anzahl an Anfangswerten eingeschränkt werden, falls t_a weiter in der Vergangenheit liegt als t_b . Dies hängt von den Trägern der Funktionen $\gamma_{z_1}^{z_2}$ ab. Wenn wir $R(t_0)$ mittels Gleichung (4.57) bestimmen möchten, könnten wir den benötigten Zeitraum nicht einschränken, da in dieser Formel keine Funktion $\gamma_{z_1}^{z_2}$ vorkommt. Dasselbe gilt für das Kompartiment D .

Entsprechend übernehmen wir für diese beiden Anfangswerte die Überlegungen für das ODE-Modell aus den Formeln (4.59) und (4.60). Wir initialisieren die Anzahl der Todesfälle somit direkt aus den verschobenen und aufbereiteten Daten des RKIs. Der Wert $R(t_0)$ wird unter Berücksichtigung der bestätigten Fälle und der zuvor für das IDE-Modell berechneten Anfangswerte $I(t_0), H(t_0), U(t_0)$ und $D(t_0)$ gesetzt. Den Anfangswert $S(t_0)$ könnten wir mit Gleichung (4.40) für $m + 1 = 0$ bestimmen, das heißt mit

$$\Sigma_S^E(t_0) = \Delta t \varphi(t_{-1}) S(t_0),$$

unter Verwendung des bekannten Werts $\Sigma_S^E(t_0)$. Zuvor müsste $\varphi(t_{-1})$ berechnet werden, wozu Gleichung (4.39) verwendet werden kann. Eventuell benötigen wir hierzu zusätzlich die Übergänge $\Sigma_{z_1}^{z_2}(t_m)$ für $m = b - 1$. Außerdem muss $D(t_{-1})$ mittels (4.56) sowie $D(t_0)$ und $\Sigma_U^D(t_0)$ ermittelt werden. Diese Berechnungsweise führt unter bestimmten Umständen zu einer Überschätzung von $S(t_0)$, sodass $S(t_0) \leq N_0$ nicht eingehalten wird. Entsprechend berechnen wir auch $S(t_0)$ wie im ODE-Modell mittels der Formel (4.62). Damit haben wir eine mögliche Methode zur Bestimmung der Anfangswerte für die Kompartimente im IDE-SECIR-Modell aus vorgegebenen Übergängen behandelt.

Wenn wir die Anfangswerte für das IDE-Modell aus den Werten für die Übergänge ermitteln und die für das ODE- und das LCT-Modell mittels der Methoden nach Koslow et al. [56, S2 Appendix], sind die Wahlen der Anfangsbedingungen nicht unbedingt konsistent. Die Anfangswertprobleme, die das ODE- und das LCT-Modell definieren, bilden dann jeweils nicht zwingend einen Spezialfall des IDE-Modells mit den dortigen Anfangsbedingungen. Unter diesen Umständen sind die Simulationsergebnisse nicht direkt miteinander vergleichbar. Betrachten wir deshalb, wie wir die Anfangsbedingungen für das LCT-SECIR-Modell auf eine zum IDE-Modell konsistente Weise bestimmen können. Eine passende Wahl für das ODE-SECIR-Modell entspricht dann schlicht der Annahme $n_z = 1$ für alle $z \in \mathcal{A}$ im LCT-Modell. In Kapitel 3.2 haben wir für das LCT-SIR-Modell bereits zum IDE-SIR-Modell passende Anfangswerte definiert, sodass wir uns daran orientieren können. Wir müssen uns nur mit den Anfangswerten $z(t_0)$ für $z \in \mathcal{A}$ beschäftigen, da die Berechnung für die übrigen Kompartimente für das IDE-Modell aus dem LCT- beziehungsweise ODE-Modell übernommen wurde. Die zum IDE-Modell passenden Anfangswerte in $t_0 = 0$ formulieren wir beispielhaft anhand des Kompartiments C . Für die restlichen Zustände gehen wir analog vor. Um geeignete Anfangsbedingungen zu formulieren, haben wir für das LCT-SIR-Modell die Integraldarstellung für die Subkompartimente aus Satz 3.6 in Gleichung (3.14) verwendet. Um hier eine analoge Vorgehensweise anzuwenden, ist es erforderlich, den Zusammenhang zwischen dem IDE-SECIR- und dem LCT-SECIR-Modell zu erfassen. Ausgehend von der Gleichung für C des IDE-Modells in (4.6) wählen wir dazu die Parameter passend, sodass sich die Gleichung auf die des LCT-Modells reduziert. Zuerst nehmen wir an, dass die Verteilung der Aufenthaltszeit im Kompartiment C unabhängig davon ist, ob ein Individuum nach I oder nach R wechseln wird. Es gilt demnach $\gamma_C^L(\tau) = \gamma_C^R(\tau)$. Die Aufenthaltszeit soll Erlang-verteilt sein mit dem Ratenparameter $n_C \gamma_C$ und dem Formparameter n_C . Aus dem Satz 3.6 gemeinsam mit den Ausführungen in Kapitel 3.4.1 wissen wir, dass dies der

Wahl

$$\gamma_C^I(\tau) = \gamma_C^R(\tau) = \frac{1}{n_C \gamma_C} \sum_{j=1}^{n_C} f_{n_C \gamma_C, j}(\tau)$$

entspricht. Setzen wir die Funktionen in die Formel (4.6) des IDE-SECIR-Modells ein, erhalten wir

$$\begin{aligned} C(t) &= \int_{\hat{T}}^t \left(\gamma_C^I(t-x) \mu_C^I + \gamma_C^R(t-x) (1 - \mu_C^I) \right) \sigma_E^C(x) dx \\ &= \int_{\hat{T}}^t \left(\frac{1}{n_C \gamma_C} \sum_{j=1}^{n_C} f_{n_C \gamma_C, j}(t-x) \right) \sigma_E^C(x) dx \\ &= \sum_{j=1}^{n_C} \left(\int_{\hat{T}}^t \frac{1}{n_C \gamma_C} f_{n_C \gamma_C, j}(t-x) \sigma_E^C(x) dx \right). \end{aligned}$$

Eine Integralgleichung für C_j mit $j \in \{1, \dots, n_C\}$ kann durch eine Benennung der Summanden wie im Beweis von Satz 3.6 durch

$$C_j(t) = \int_{\hat{T}}^t \frac{1}{n_C \gamma_C} f_{n_C \gamma_C, j}(t-x) \sigma_E^C(x) dx$$

definiert werden. Analog zu diesem Beweis kann man zeigen, dass die Ableitungen der Integralgleichungen zu den Formulierungen von C_1 in (4.27) und C_j für $j \in \{2, \dots, n_C\}$ in (4.28) des LCT-SECIR-Modells passen. Dazu müssen wir ausnutzen, dass $\sigma_E^C(t) = n_E \gamma_E E_{n_E}(t)$ für das LCT-Modell gilt. Wenn wir auf die Integralgleichung dasselbe numerische Lösungsschema wie für das IDE-Modell in Kapitel 4.2.2 anwenden und den Wert in t_0 betrachten, erhalten wir für $j \in \{1, \dots, n_C\}$ die Anfangswerte

$$C_j(0) = \sum_{i=\max\{a,b\}}^{-1} \frac{1}{n_C \gamma_C} f_{n_C \gamma_C, j}(t_{-i}) \Sigma_E^C(t_{i+1}),$$

wobei $b \in \mathbb{Z}$ auf die Wahl der Überlebensfunktionen $\gamma_{z_1}^{z_2}$, unter der sich das IDE-Modell auf das LCT-Modell reduziert, angepasst sei. Hierbei ist Δt wie für das IDE-Modell ein konstanter Wert und die Definitionen von $\Sigma_{z_1}^{z_2}$ und a werden für die Initialisierung des LCT-Modells aus dem numerischen Lösungsschema des IDE-Modells aus Kapitel 4.2.2 übernommen. Für die anderen Gruppen $z \in \mathcal{A}$ kann eine analoge Vorgehensweise angewendet werden. Für die Subkompartimente dieser Zustände ergeben sich unter Anwendung des numerischen Lösungsschemas des IDE-Modells die Anfangswerte

$$\begin{aligned} E_j(0) &= \sum_{i=\max\{a,b\}}^{-1} \frac{1}{n_E \gamma_E} f_{n_E \gamma_E, j}(t_{-i}) \Sigma_S^E(t_{i+1}) && \text{für } j \in \{1, \dots, n_E\} \\ I_j(0) &= \sum_{i=\max\{a,b\}}^{-1} \frac{1}{n_I \gamma_I} f_{n_I \gamma_I, j}(t_{-i}) \Sigma_C^I(t_{i+1}) && \text{für } j \in \{1, \dots, n_I\} \\ H_j(0) &= \sum_{i=\max\{a,b\}}^{-1} \frac{1}{n_H \gamma_H} f_{n_H \gamma_H, j}(t_{-i}) \Sigma_I^H(t_{i+1}) && \text{für } j \in \{1, \dots, n_H\} \\ U_j(0) &= \sum_{i=\max\{a,b\}}^{-1} \frac{1}{n_U \gamma_U} f_{n_U \gamma_U, j}(t_{-i}) \Sigma_H^U(t_{i+1}) && \text{für } j \in \{1, \dots, n_U\}. \end{aligned}$$

Auf diese Weise können wir für das LCT- und das ODE-SECIR-Modell zum IDE-Modell konsistente Anfangswerte wählen. Entsprechend sind die Ergebnisse von Simulationen, die mit dieser Methode initialisiert werden, vergleichbar. Das Verfahren könnte für das LCT- und das ODE-Modell außerdem den Vorteil haben, dass nicht angenommen wird, dass Individuen immer die erwartete Aufenthaltszeit in einem Kompartiment bleiben. Allerdings hat die Methode basierend auf den Übergängen $\Sigma_{z_1}^{z_2}$ einen erheblichen Nachteil. Obgleich wir den Zeitraum, in welchem die Übergänge bekannt sein müssen, bereits eingeschränkt haben, kann die Anzahl der benötigten Werte erheblich groß werden. Außerdem haben wir ausgelassen, wie die Übergänge $\Sigma_{z_1}^{z_2}$ aus realen Daten bestimmt werden könnten. Mit dieser Thematik wird sich die Publikation [95] befassen.

Da wir die Übergänge zwingend zur vollständigen Initialisierung des IDE-Modells benötigen, werden wir das Modell bei Simulationen von realen Szenarien nicht verwenden. Das ODE- sowie das LCT-Modell werden wir für reale Simulationen mit den beschriebenen Methoden nach Koslow et al. [56, S2 Appendix] initialisieren. Zusätzlich führen wir in Kapitel 5 Simulationen von fiktiven Szenarien durch, bei denen die Anfangsbedingungen synthetische Daten darstellen. Hier definieren wir Anfangsbedingungen für die Übergänge $\Sigma_{z_1}^{z_2}$ für einen ausreichenden Zeitraum vor dem Startzeitpunkt t_0 und wenden die soeben beschriebene Initialisierung auf der Basis der Übergänge auf die drei SECIR-Modelle an. Alle in diesem Abschnitt beschriebenen Methoden werden an geeigneten Stellen für Simulationen verwendet.

Aus den Erläuterungen wird deutlich, dass sich die Initialisierung der verschiedenen Modelle als unterschiedlich herausfordernd darstellt. Wir können hiermit die Behauptung aus Kapitel 3.6 validieren und erweitern. Die notwendigen Anfangsbedingungen für das IDE-Modell schränken die Anwendbarkeit so weit ein, dass wir es nicht mit realen Daten initialisieren. Für das ODE- und das LCT-Modell können wir handhabbare Methoden verwenden, wobei sich die Initialisierung des LCT-Modells je nach verwendetem Verfahren als aufwendiger herausstellt.

4.4 Implementierung des LCT-SECIR-Modells

Nachdem wir die theoretischen Voraussetzungen zur Umsetzung der Modelle in Software untersucht haben, klären wir nun Aspekte der Implementierung. Das ODE-SECIR- und das IDE-SECIR-Modell sind bereits in der Software MEmilio [61] berücksichtigt. Das ODE-Modell ist mit Erweiterungen wie einer Auflösung nach dem Alter oder nach Regionen implementiert. Dabei ist das Modell so flexibel aufgebaut, dass auch Simulationen mit dem Modell aus Kapitel 4.1.1 möglich sind. Das IDE-SECIR-Modell aus Kapitel 4.1.2 wurde ebenfalls in der Software programmiert und es müssen nur kleine Erweiterungen hinzugefügt werden, um die Simulationen in Kapitel 5 durchführen zu können. Beispielsweise müssen die passenden Aufenthaltszeitverteilungen für die Parameter in den Code inkludiert werden. Da das Modell nicht im Fokus der vorliegenden Arbeit steht, gehen wir hier nicht über die Formulierungen aus den vorherigen Abschnitten hinaus. Die Modelle sind in der Programmiersprache C++ implementiert, in welcher wir auch das LCT-Modell umsetzen. Für dieses Kapitel wird ein Vorwissen bezüglich der Programmiersprache vorausgesetzt. Das Projekt verwendet aktuell den Standard C++17. Im folgenden Kapitel

halten wir fest, wie das LCT-SECIR-Modell in den bestehenden Code eingebaut wird. Zuvor wurde im Projekt noch kein LCT-Modell betrachtet, sodass die gesamte Umsetzung des LCT-Modells Teil der vorliegenden Arbeit ist. Einige Konzepte und die Struktur des implementierten ODE-Modells können wir jedoch wiederverwenden. Entsprechend beschreiben wir zunächst, welche Voraussetzungen in der Software MEmilio [61] vorhanden sind. Im Anschluss wird die Eigenleistung im Rahmen der vorliegenden Masterarbeit besprochen.

4.4.1 Vorhandene Voraussetzungen in der Software MEmilio

Wir geben nun die wichtigsten Aspekte des Aufbaus des ODE-SECIR-Modells in der Software MEmilio an, um die Konzepte später auf das LCT-Modell übertragen zu können. Dabei ist zu beachten, dass sich die beschriebenen Klassen zum Teil selbst noch in der Entwicklung befinden. Die Beschreibung bezieht sich auf den Entwicklungsstand der Software MEmilio [61] Anfang Oktober 2023. Der Rahmen der ODE-Modelle wurde erst kürzlich umfangreich verallgemeinert. Die Neuerungen werden hier nicht besprochen. Der Ausgangspunkt der Implementierung ist im Ordner `cpp/models/ode_secir` zu finden. Einige Umsetzungen erben von Basisklassen für Funktionalitäten von Modellen mit ähnlichen Strukturen aus dem Ordner `cpp/memilio/compartments`. Dies ermöglicht beispielsweise das Wiederverwenden von Codekonzepten für ein Modell mit anderen Kompartimenten, wie für die Berücksichtigung von Impfungen in [56]. Wir konzentrieren uns nur auf die Aspekte des ODE-SECIR-Modells, die für die Implementierung des LCT-SECIR-Modells relevant sind.

Das ODE-SECIR-Modell ist in einer Klasse `Model` konkretisiert, die von einer Basisklasse `Compartmentalmodel` mit den Template-Parametern `class Comp`, `class Pop` und `class Params` erbt. Für die Klasse `Model` wird `Comp` durch eine enum class `InfectionState` spezifiziert. Hier werden die Namen der verwendeten Kompartimente aufgezählt und durchnummeriert. Es gibt einen Enumerator mit dem Namen `Count`, der die Anzahl der Kompartimente angibt. Über den Template-Parameter `Pop` wird eine Aufteilung in Altersgruppen spezifiziert. Zudem werden über die Klasse die Anfangswerte definiert. Da wir für das LCT-Modell keine Altersauflösung implementieren, lösen wir diese Thematik dort anders. Der Typ des Template-Parameters `Params` der Klasse `Compartmentalmodel` wird für `Model` als die Klasse `Parameters` definiert. Diese Klasse erbt wiederum von der Klasse `ParameterSet`, die durch ein Template, das beliebig viele Argumente zulässt, alle benötigten Parameter des Modells umfasst. Dazu werden die Modellparameter, wie beispielsweise die durchschnittlich verbrachte Zeit in I oder die Transmissionswahrscheinlichkeit, jeweils in Form eines structs definiert. In dem struct wird der Typ spezifiziert, eine Funktion für die Abfrage eines default-Werts definiert und es kann der Name des Parameters abgerufen werden. Herauszustellen ist der Parameter `ContactPatterns` welcher den Typ `UncertainContactMatrix` verwendet. Mithilfe von dieser Klasse können Veränderungen in der Kontaktrate durch nicht-pharmazeutische Maßnahmen berücksichtigt werden, sodass wir dies für das LCT-Modell übernehmen. Durch die Vererbung von der Klasse `ParameterSet` können wir für alle Parameter getter und setter verwenden und die default-Werte aus den definierten structs werden ohne explizite Aufrufe angewendet.

```

1 template <class Model, class Sim = Simulation<Model>>
2 TimeSeries<ScalarType> simulate(double t0, double tmax, double dt, Model
   const& model, std::shared_ptr<IntegratorCore> integrator = nullptr
   ){
3     model.check_constraints();
4     Sim sim(model, t0, dt);
5     if (integrator) {
6         sim.set_integrator(integrator);
7     }
8     sim.advance(tmax);
9     return sim.get_result();
10 }

```

Codeausschnitt 4.1: Code der Funktion `simulate()` aus der Datei `cpp/memilio/-compartments/simulation.h`.

Die Klasse `Parameters` implementiert zusätzlich beispielsweise eine Funktion namens `check_constraints()`, mit der notwendige Bedingungen überprüft werden können, wie beispielsweise, dass die Wahrscheinlichkeiten zwischen null und eins liegen. In der Basis-Klasse `Compartmentalmodel` mit den spezifizierten Template-Parametern für `Model` gibt es ebenfalls eine Funktion `check_constraints()`, die unter anderem die Funktion mit demselben Namen der Klasse `Parameters` aufrufen kann und die gesamten Bedingungen für `Model` testet. Hier wird auch überprüft, ob alle Anfangswerte für die Kompartimente nicht-negativ gesetzt wurden. Eine weitere wichtige Funktion von `Compartmentalmodel` ist `eval_right_hand_side()`, welche die rechte Seite des gewöhnlichen Differentialgleichungssystems des Modells in Abhängigkeit von der Zeit t und einem Vektor y mit den Anzahlen der Individuen in allen Kompartimenten zu dieser Zeit berechnet. Das Ergebnis wird durch ein Update eines Vektors `dydt` zurückgegeben, welcher per Referenz an die Funktion übergeben wird. Die Klasse `Compartmentalmodel` ruft hier nur eine virtuelle Funktion `get_derivatives()` auf, die in der Klasse `Model` dann auf das ODE-SECIR-Modell mit den in `InfectionState` spezifizierten Krankheitszuständen angepasst wird.

Dies sind die wichtigsten Aspekte der Implementierung des ODE-SECIR-Modells. Wenn wir eine Simulation mit dem Modell ausführen möchten, müssen wir die Funktion `simulate()` aufrufen. Der Aufruf wird an eine allgemeinere Funktion weitergeleitet, deren Code im Codeausschnitt 4.1 dargelegt ist. Die Funktion ist ein Funktions-Template und funktioniert für verschiedene Modelle. Der verwendete Typ für den Template-Parameter `Model` ist hier die definierte Klasse `Model` des ODE-SECIR-Modells. An die Funktion wird zusätzlich zum betrachteten Modell, mit den Definitionen der Parameter, der Infektionszustände und weiteren Informationen, der Anfangszeitpunkt sowie der Endzeitpunkt der Simulation übergeben. Zusätzlich erhält die Funktion die initiale Schrittweite, die vom Löser verwendet werden soll. Außerdem ist als Funktionsparameter ein Zeiger auf eine Instanz der Klasse `IntegratorCore` angegeben. Diese Klasse definiert die konkrete Ausgestaltung der verwendeten numerischen Methode zur Lösung des betrachteten gewöhnlichen Gleichungssystems. Beispielsweise erbt eine Klasse `ControlledStepperWrapper` von `IntegratorCore`, die Löser von der Softwarebibliothek `boost` zulässt. Durch die Klasse können zusätzliche Aspekte definiert werden, wie zum

```

1 Simulation(Model const& model, double t0 = 0., double dt = 0.1)
2   : m_integratorCore(std::make_shared<mio::ControlledStepperWrapper<
3     boost::numeric::odeint::runge_kutta_cash_karp54>>()),
4   m_model(std::make_unique<Model>(model)),
5   m_integrator([&model = *m_model](auto&& y, auto&& t, auto&& dydt) {
6     model.eval_right_hand_side(y, y, t, dydt);}, t0, m_model->
    get_initial_values(), dt, m_integratorCore)
    {
    }

```

Codeausschnitt 4.2: Konstruktor der Klasse `Simulation` aus der Datei `cpp/memilio/-compartments/simulation.h`.

Beispiel eine minimale oder maximale Schrittweite oder Fehlertoleranzen für adaptive Verfahren. Über eine Vererbung von der Klasse `IntegratorCore` können ebenso eigene Löser zur Lösung des gewöhnlichen Differentialgleichungssystems verwendet werden.

Betrachten wir den Methodenrumpf der Funktion `simulate()`. Zunächst wird die besprochene Funktion `check_constraints()` für das Modell aufgerufen, sodass sichergestellt ist, dass alle Bedingungen eingehalten und alle Spezifikationen des Modells sinnvoll gesetzt worden sind. Die Simulation wird mit einer Instanz der Klasse `Simulation` durchgeführt. Das ODE-SECIR-Modell hat eine eigene Klasse `Simulation`, die von einer Basisklasse `Simulation`, mit der Spezifikation des Template-Parameters als die Klasse `Model` des ODE-SECIR-Modells, erbt. Beide Klassen heißen `Simulation` und sind in verschiedenen namespaces definiert. Wenn man das ODE-SECIR-Modell ohne eine örtliche Auflösung anwendet, ist ausschließlich die aufgerufene Funktionalität der Basisklasse relevant. Der Konstruktor dieser Basisklasse ist im Codeausschnitt 4.2 zu finden, da anhand dieses Konstruktors die Funktionalität der Klasse erörtert werden kann. Wir stellen fest, dass die Klasse `Simulation` ebenfalls einen Zeiger auf eine Instanz von `IntegratorCore` hält. Wenn später kein anderer `IntegratorCore` gesetzt wird, wird zur Simulation die Methode `runge_kutta_cash_karp54` aus der Bibliothek `boost` verwendet, wie in Kapitel 4.2.1 beschrieben wurde. Das zugrundeliegende Modell wird gespeichert. Zusätzlich gibt es eine Klassenvariable der Klasse `OdeIntegrator`, der sich die rechte Seite des zu lösenden Systems merkt, hier `eval_right_hand_side()`, sowie die Anfangswerte des Anfangswertproblems, die aus der Instanz von `Model` übernommen werden. Zusätzlich erhält der `OdeIntegrator` den Startzeitpunkt sowie die initiale Schrittweite und den `IntegratorCore`. Er enthält folglich alle für die Lösung des gewöhnlichen Differentialgleichungssystems notwendigen Informationen. Die Funktion `simulate()` ruft die Funktion `advance()` der Instanz der Klasse `Simulation` auf. Diese löst das gewöhnliche Differentialgleichungssystem numerisch bis zum Zeitpunkt `tmax` unter Verwendung des spezifizierten Lösers und mithilfe des `OdeIntegrators`. Hierbei wird die Funktion `eval_right_hand_side()` wiederholt aufgerufen. Weitere Details zur Funktionsweise der Integratoren beziehungsweise des numerischen Lösers lassen wir aus. Das Ergebnis der Simulation wird als Instanz einer Klasse `TimeSeries` zurückgegeben. Darin sind die meist adaptiv gesetzten Zeitpunkte in Verbindung mit den erhaltenen approximativen Werten für alle Kompartimente gespeichert.

Grundsätzlich umfasst die Klasse `Model` den reinen Aufbau des Modells und die Werte der Modellparameter. Die Klasse `Simulation` beschäftigt sich mit der Simulation unter der Verwendung eines solchen Modells und gestaltet die Methode zur Lösung des zugrundeliegenden gewöhnlichen Differentialgleichungssystems. In der Funktion `simulate()` wird das Zusammenspiel der Klasseninstanzen deutlich. Damit haben wir die Grundlagen der Struktur des ODE-SECIR-Modells in der Software `MEMilio` dargelegt. Die Vererbungshierarchien und die Aufteilungen der Funktionalitäten können komplex wirken, jedoch unterstützen sie die Wiederverwendbarkeit von Codeteilen und die konsistente Struktur verschiedener Modelle, die auf gewöhnlichen Differentialgleichungen mit einer festen Anzahl von Kompartimenten beruhen. In der Datei `cpp/models/ode_secir/parameters_io.cpp` sind noch Methoden implementiert, um mittels der Initialisierung nach Koslow et al. [56, S2 Appendix] aus Kapitel 4.3 Anfangswerte für das ODE-SECIR-Modell aus realen Daten zu berechnen.

4.4.2 Einbindung des LCT-SECIR-Modells in die Software

Wir beschreiben nun, wie wir ein LCT-SECIR-Modell in die Software `MEMilio` einbauen. Die neue Funktionalität, die im Rahmen der vorliegenden Masterarbeit entwickelt wurde, ist aktuell auf dem Branch `680-implement-lct-model`¹ zu finden. Eine erste Idee, das LCT-Modell in den Code einzufügen, ist leicht umzusetzen. In der Umsetzung des ODE-Modells ist, wie zuvor beschrieben, die Flexibilität geboten, über Vererbung der Klasse `CompartmentalModel` unkompliziert Modelle mit selbst definierten Kompartimenten zu erstellen und die rechte Seite des Differentialgleichungssystems zu definieren. Für das `CompartmentalModel` sollte als Template-Argument für `Pop` eine abgeleitete Klasse von `Population` gesetzt werden. Für die Klasse `Population` wird erwartet, dass die Krankheitszustände in einer enum class mit einem Enumerator `Count` definiert sind, wie es für das ODE-SECIR-Modell gegeben ist. Für das LCT-Modell ist die Vorgehensweise problematisch, wenn wir verschiedene Anzahlen von Subkompartimenten ausprobieren wollen und keine festen Anzahlen implementieren. Die Variable `Count` der Anzahl der Kompartimente inklusive der Subkompartimente könnte hier nicht durch eine constexpr bestimmt werden, was für die Vererbung von `CompartmentalModel` notwendig wäre. Würden wir beispielsweise immer mit zwei Subkompartimenten für jedes der Kompartimente in \mathcal{A} Simulationen durchführen, wäre die genannte Möglichkeit die einfachste Art der Umsetzung. In unserer Implementierung möchten wir allerdings die Möglichkeit haben, die Anzahl der Subkompartimente in \mathcal{A} variabel zu gestalten und die Anzahlen unabhängig voneinander zu setzen. Dies hat zur Folge, dass wir eine vollständige eigene Klasse für die LCT-Modelle erstellen müssen.

Bei der Implementierung des LCT-SECIR-Modells wird dennoch die Modellstruktur sowie Benennungen von Variablen, Funktionen, Klassen und Ähnliches bestmöglich an die Umsetzung des ODE-SECIR-Modells angepasst. Dies zielt darauf ab, bereits vorhandene allgemeine Erweiterungen auf das LCT-Modell anwenden zu können, wie beispielsweise die örtliche Auflösung. Zudem wird die Projektstruktur beibehalten. Auf den ersten Blick macht dies die Umsetzung des LCT-Modells in die Software komplexer. Der zugefügte

¹Siehe <https://github.com/SciCompMod/memilio/pull/771> (abgerufen am 04.12.2023).

```

1 class Model{
2 public:
3     // Constructor, getter, ...
4     bool check_constraints() const;
5     void eval_right_hand_side(Eigen::Ref<const Eigen::VectorXd> y, double
        t, Eigen::Ref<Eigen::VectorXd> dydt) const;
6     Parameters parameters;
7     InfectionState InfectionStates;
8 private:
9     Eigen::VectorXd m_initial_values;
10    double m_N0;
11 };

```

Codeausschnitt 4.3: Schematische Darstellung der Klasse `Model` aus der Datei `cpp/models/lct_secir/model.h`.

Code umfasst insgesamt rund 6000 Zeilen (inklusive Tests, der Implementierung der Simulationen aus Kapitel 5 und Ähnlichem), sodass wir nicht auf alle Details eingehen. Wir beschränken uns auf die essentiellen Aspekte des Modells und betrachten nur die wichtigste Funktionalität. Für das LCT-Modell definieren wir die Klasse `Model`, die wie beim ODE-Modell nur die Informationen des Modells und die Parameterwahlen beinhaltet. Zusätzlich gibt es eine eigene Klasse `Simulation` zur Simulation mittels dieser Modelle. Auch diese Klasse schreiben wir neu und verwenden nicht die für die ODE-Modelle. In der zuvor beschriebenen Klasse `Simulation` wird für `Model` ein `CompartmentalModel` erwartet und es wird eine Funktion aufgerufen, die dies überprüfen soll. Die Klasse `Simulation` und auch die Funktion `simulate()` werden hier nur so weit angepasst, dass sie LCT-Modelle der Klasse `Model` zulassen. Entsprechend gehen wir nicht näher auf die Implementierung ein. Die numerische Lösungsmethode für das Modellgleichungssystem des LCT-SECIR-Modells, das ebenfalls ein gewöhnliches Differentialgleichungssystem ist, kann analog zum ODE-Modell verändert werden. Der Code des ODE- und des LCT-Modells für diese Klasse `Simulation` könnte zusammengeführt werden, jedoch müsste dafür der bestehende Code des ODE-Modells leicht verändert werden, wovon hier zunächst abgesehen wird.

Das Konzept der Klasse `Model` ist im Codeausschnitt 4.3 dargestellt. Die Klasse verzichtet auf Vererbungen und wird für die festen acht Infektionszustände, wie in Kapitel 4.1 definiert, implementiert. Die Funktionalität bleibt ansonsten dieselbe wie für das ODE-SECIR-Modell für `Model` beschrieben. In der Variable `parameters` sind alle möglichen benötigten Modellparameter berücksichtigt, das heißt beispielsweise die Aufenthaltszeiten sowie die Übergangswahrscheinlichkeiten und die Parameter ρ , ϕ , ξ_C und ξ_I . Die Klasse `Parameters` erbt dafür wieder von `ParameterSet`, sodass wir wie zuvor für jeden Parameter einen struct definieren, in dem beispielsweise ein default-Wert gesetzt wird. Zusätzlich implementiert `Parameters` die Funktion `check_constraints()`. Außerdem hat `Model` eine Klassenvariable der Klasse `InfectionState`. Diese Klasse ist im Codeausschnitt 4.4 schematisch dargestellt. Sie verwendet eine enum class `InfectionStateBase`, in welcher die Kompartimente ohne Berücksichtigung der Subkompartimente spezifiziert sind, das heißt hier die acht Zustände S , E , C , I , H , U , R und D . Die Klasse `InfectionState`

```

1 class InfectionState{
2 public:
3     // Constructors, setter and getter methods, ...
4     int get_number(InfectionStateBase infectionstatebase) const;
5     int get_firstindex(InfectionStateBase infectionstatebase) const;
6     bool check_constraints() const;
7 private:
8     // ...
9     std::vector<int> m_SubcompartmentNumbers;
10    std::vector<int> m_SubcompartmentNumbersindexfirst;
11    int m_Count;
12 };

```

Codeausschnitt 4.4: Schematische Darstellung der Klasse `InfectionState` aus der Datei `cpp/models/lct_secir/infection_state.h`.

speichert die Anzahlen der Subkompartimente für die Kompartimente aus `InfectionStateBase` in einem Vektor `m_SubcompartmentNumbers` und besitzt einige zusätzliche Funktionalitäten. Beispielsweise kann die Anzahl der Subkompartimente eines Enumerators von `InfectionStateBase` abgefragt werden. Häufig werden die Anzahlen der Individuen in den Subkompartimenten in Form eines Vektors gespeichert. Dabei wird die Reihenfolge der Definition in den Modellgleichungen (4.24)–(4.36) eingehalten, das bedeutet, dass sich in dem Vektor zunächst der Wert für S dann von E_j mit $j \in \{1, \dots, n_E\}$ und so weiter befindet. Für Berechnungen müssen wir wissen, an welchem Index ein bestimmtes Subkompartiment im Vektor verortet ist. Dafür könnten wir jedes Mal die Anzahlen der Subkompartimente aller in der Reihenfolge vorhergehender Kompartimente aufaddieren. Um dies zu vermeiden, gibt es den Vektor `m_SubcompartmentNumbersindexfirst` beziehungsweise die Funktion `get_firstindex()`, die den Index des ersten Subkompartiments in einem solchen Vektor für einen `InfectionStateBase` zurückgibt. Es sind zudem eine Variable für die Gesamtzahl aller Subkompartimente und die Funktion `check_constraints()` vorhanden, die beispielsweise prüft, ob für die Kompartimente S , R und D die Anzahl der Subkompartimente auf eins gesetzt wird.

Weiterhin besitzt die Klasse `Model` für das LCT-Modell eine Variable für den Wert N_0 aus der Beschreibung in Kapitel 4.1. Die Anfangswerte aller Kompartimente beziehungsweise Subkompartimente werden in einem Vektor gespeichert, der die Form hat, die bei der Klasse `InfectionState` beschrieben wurde. Den Kern der Klasse bildet die Funktion `eval_right_hand_side()`. Diese berechnet das Ergebnis der rechten Seite des Differentialgleichungssystem (4.24)–(4.36) in Abhängigkeit von der Zeit t und einem Vektor y mit den Werten aller Kompartimente und Subkompartimente zu diesem Zeitpunkt. Hier fließen die definierten Parameter, der Wert N_0 und die Instanz von `InfectionState` zusammen. Die Berechnung der rechten Seite erfolgt so, wie im Differentialgleichungssystem vorgegeben. Werte, die in den Gleichungen mehrfach vorkommen, werden nicht neu berechnet, sondern wiederverwendet. Damit haben wir den Aufbau der Klasse `Model` aus dem Codeausschnitt 4.3 für das LCT-SECIR-Modell besprochen, ohne auf Details der Implementierung einzugehen.

Für eine Initialisierung mittels Übergängen am Vorbild des IDE-Modells, wie in Ka-

Algorithmus 2: Erstellung einer Simulation mit einem LCT-SECIR-Modell.

Erstelle eine Instanz von `Parameters` und setze die Werte der Parameter.
 Erstelle eine Instanz von `InfectionState` mit den Anzahlen der Subkompartimente.
 Definiere Anfangswerte in einem Vektor, beispielsweise mit der Klasse `Initializer`.
 Erstelle ein `LCT-Model`.
 Rufe die Funktion `simulate()` ähnlich zu der aus dem Codeausschnitt 4.1 auf und setze den Löser. Das Ergebnis der Simulation wird zurückgegeben.

pitel 4.3 beschrieben wurde, gibt es eine eigene Klasse `Initializer`. Diese bestimmt Anfangswerte, mit denen eine Instanz der Klasse `Model` initialisiert werden könnte. Die Implementierungsdetails entsprechen denen des IDE-Modells, sodass wir die Klasse nicht im Detail besprechen. Einige Anpassungen für das LCT-Modell, wie die Aufteilung in Subkompartimente, gehen über die Methode des IDE-Modells hinaus, sodass wir die Implementierung der Initialisierung des IDE-Modells nicht wiederverwenden können. Zusätzlich ist in der Datei `cpp/models/lct_secir/parameters_io.cpp` eine zweite Möglichkeit zur Bestimmung der Anfangswerte für das LCT-Modell mit der Methode nach Koslow et al. [56, S2 Appendix] aus Kapitel 4.3 implementiert. Mit dieser Methode können die Anfangswerte aus realen Daten ermittelt werden. Das Verfahren unterscheidet sich von dem für das ODE-Modell, da eine Aufteilung in Subkompartimente vorgenommen werden muss. Für die erforderlichen Zeitpunkte werden keine gerundeten Werte verwendet, sondern es wird linear interpoliert, wie in Kapitel 4.3 beschrieben wurde. Hierdurch erhöht sich der Schwierigkeitsgrad der Implementierung deutlich.

Wir haben nun den groben Aufbau aller Funktionalitäten für das LCT-Modell abgebildet. Im Algorithmus 2 ist dargelegt, wie die beschriebenen Komponenten verwendet werden können, um eine Simulation mittels eines LCT-Modells durchzuführen. Die Ergebnisse der Simulation in Form einer `TimeSeries` speichern wir in einer `.h5` Datei ab. Sollen die Ergebnisse grafisch dargestellt werden, importieren wir die gespeicherten Ergebnisse in einer Python-Datei und erstellen die Plots unter Verwendung der Bibliothek „matplotlib“.

Für das Modell und die zugehörige Funktionalität sind umfangreiche Softwaretests implementiert. Aus der theoretischen Betrachtung des LCT-Modells in Kapitel 3.4.2 wissen wir, dass das LCT-SECIR-Modell mit nur einem Subkompartiment das ODE-SECIR-Modell ergeben sollte. Ein implementierter Softwaretest bestätigt, dass das LCT-SECIR-Modell mit jeweils nur einem Subkompartiment für alle Gruppen in \mathcal{A} dieselben Simulationsergebnisse mit ansonsten identischen Parametern liefert wie das ODE-SECIR-Modell. Dieser Test zeigt noch nicht, dass das Modell mit mehreren Subkompartimenten so funktioniert wie erwartet. Entsprechend wird das Ergebnis der Funktion `eval_right_hand_side()` händisch anhand eines Beispiels berechnet und in einem Unittest mit dem Ergebnis der Funktion verglichen. Ein Softwaretest, der eine Simulation mit einem Referenzdurchlauf vergleicht, prüft, ob unbeabsichtigte Änderungen an der Funktionalität vorgenommen werden. Einzelne Bestandteile und Funktionen werden zusätzlich mit Unittests getestet, sodass die Qualität des implementierten Codes überprüft wird und einige potentielle Fehler ausgeschlossen werden können. Insgesamt erreichen wir eine Patch Coverage von 98.86 % des zugefügten Codes, was noch etwas höher ist als die des

Projekts. Der Code wird umfangreich dokumentiert und kommentiert und erklärende Beispiele werden geschrieben, um die Anwendung des Modells zu vereinfachen.

Bei der Implementierung fällt auf, dass das LCT-Modell wesentlich einfacher in Software umzusetzen ist als das IDE-Modell. Dies liegt vor allem an der Anwendbarkeit von bereits implementierten Lösern aus Softwarebibliotheken für gewöhnliche Differentialgleichungssysteme. Die Umsetzung des IDE-SECIR-Modell ist aufwendiger, da hierfür das numerische Verfahren aus Kapitel 4.2.2 implementiert werden muss.

In diesem Kapitel haben wir die Grundlagen für praktische Anwendungen der drei Modellarten gelegt. Für das LCT-Modell haben wir die Umsetzung in die Software MEMilio ausführlich besprochen. Damit wurden alle Voraussetzungen geschaffen, um das ODE-, das LCT- und das IDE-SECIR-Modell anhand von verschiedenen Simulationen zu erproben. Mit den Überlegungen aus diesem Kapitel über die relevanten Aspekte in der Praxis konnten Hypothesen aus Kapitel 3.6 bestätigt werden, vornehmlich bezüglich der Praktikabilität und Anwendbarkeit der Modelle. Wir haben festgestellt, dass sich die Formulierung des IDE-Modells für größere Modelle als nur mit drei Kompartimenten als deutlich komplizierter erweist, als die eines entsprechenden ODE-Modells. Das LCT-Modell ist herausfordernder herzuleiten als das ODE-Modell, aber bedeutend unkomplizierter als das IDE-Modell. Für die numerische Lösung von gewöhnlichen Differentialgleichungssystemen stehen Methoden aus Softwarebibliotheken zur Verfügung, wohingegen für das IDE-Modell ein eigener Löser entwickelt werden muss. Auch bei der Herleitung geeigneter Anfangswerte und in der Implementierung weist das IDE-Modell eine höhere Komplexität auf als die anderen beiden Modelle. Die Wahl passender Anfangsbedingungen für das IDE-Modell aus realen Daten könnte besonders herausfordernd sein, sodass wir den Fall für den Vergleich mit realen Daten nicht heranziehen. Es stellt sich somit heraus, dass das ODE-Modell am besten in die Praxis umzusetzen ist. Das Verwenden eines LCT-Modells erfordert im Gegensatz zum IDE-Modell einen überschaubaren zusätzlichen Aufwand. Im folgenden Abschnitt werden wir Argumente aus Kapitel 3.6 hinsichtlich der Anpassungsfähigkeit der Modelle und bezüglich der Parameterwahl untersuchen.

5 Simulation und Ergebnisse am Beispiel von SARS-CoV-2 und COVID-19

Im Folgenden vergleichen wir das Verhalten der Modelle anhand verschiedener Szenarien. Die Simulationen führen wir am Beispiel des Erregers SARS-CoV-2 beziehungsweise der Krankheit COVID-19 durch. Wir modellieren die Ausbreitung des Krankheitserregers in ganz Deutschland. Dabei bewerten wir, ob die Flexibilität der Modelle ausreicht, um realistische Prognosen für die spezifische Krankheit zu generieren. Einzelne Aspekte der Anwendbarkeit und Praktikabilität aus Abschnitt 3.6 werden ebenfalls überprüft. Spätestens seit der Deklaration von COVID-19 als Pandemie durch die WHO im März 2020 [24, S. 2], ist die Krankheit ein intensiv diskutiertes Thema in der Gesellschaft, Politik und Wissenschaft [24, 71]. Die epidemischen Ausbrüche führten zu einem Anstieg des Interesses an der mathematischen Modellierung von Epidemien [38, S. 1]. Die Aktualität und das Interesse an der Infektion begründen die Eignung von COVID-19 beziehungsweise von SARS-CoV-2 als Beispiel. In den Simulationen fließen die in Kapitel 4 besprochenen Aspekte zusammen, das heißt, dass wir die erweiterten Modelle, die numerischen Lösungsverfahren sowie die beschriebenen Methoden zur Initialisierung verwenden. Die Implementierung wird genutzt, um die Simulationsergebnisse zu berechnen. Die Anwendung auf den Fall von SARS-CoV-2 und COVID-19 erfolgt durch angemessene Wahlen der epidemiologischen Parameter. Mit der Definition dieser Modellgrößen beschäftigen wir uns in Abschnitt 5.1. Im Anschluss betrachten wir einige fiktive Szenarien, unter anderem, um die Reaktion von den drei Modellarten auf die Veränderung des Parameters der Kontaktrate zu analysieren. Diese Veränderung ist vergleichbar mit dem Beschluss einer nicht-pharmazeutischen Maßnahme oder der Aufhebung einer solchen. Zum Abschluss erproben wir das ODE- und das LCT-Modell an realen Szenarien und vergleichen die Simulationsergebnisse mit echten Daten der COVID-19 Pandemie.

5.1 Verwendete epidemiologische Parameter für SARS-CoV-2 und COVID-19

Wir beschäftigen uns nun mit der Wahl der epidemiologischen Parameter, die für die Simulationen benötigt werden. Die Prognosen sollen sich auf die Ausbreitung der Krankheit COVID-19 beziehungsweise des Krankheitserregers SARS-CoV-2 in Deutschland beziehen, sodass wir die Parameter auf diesen Fall abstimmen. Bei der Definition der Parameter orientieren wir uns an Kühn et al. [60]. An die Publikation wurde bereits das ODE-SEIR-Modell aus Kapitel 4.1.1 angelehnt. Dort werden passende Parameterbereiche für SARS-CoV-2 auf der Basis von Literaturrecherchen sowie einer Studie mit hospitalisierten COVID-19 Patienten bestimmt, siehe [60, Kapitel 2.3 & Table 1]. Auf

Parameter	berechneter Wert		Parameter	berechneter Wert
$\rho^{(0)}$	0.0733271		ξ_C	1
T_E	3.335		ξ_I	0.3
T_C	3.31331		μ_C^I	$1 - 0.206901$
T_I	6.94547		μ_I^H	0.0786429
T_H	11.634346		μ_H^U	0.173176
T_U	17.476959		μ_U^D	0.217177

Tabelle 5.1: Verwendete Werte epidemiologischer Parameter für SARS-CoV-2 und COVID-19. Die Werte basieren auf den Ergebnissen in [60, Table 2]. Die beiden rot markierten Parameter sind nicht nach [60, Table 2] berechnet, sondern an die Parameter des IDE-Modells angepasst. Siehe dafür den Abschnitt über die Parameterwahl für das IDE-SECIR-Modell.

die Größen wird somit nicht anhand von Fallzahlen geschlossen, mit denen die Modelle bewertet werden. Stattdessen werden epidemiologische Studien herangezogen, was die Vorhersagekraft der Modelle unterstützt, siehe auch [63, S. 61]. Die gefundenen Werte für die epidemiologischen Parameter sind in [60, Table 2] zusammengefasst. Aus diesen Angaben leiten wir wie folgt Ergebnisse für unsere benötigten Parameter für das ODE-SECIR-Modell ab.

Zunächst sind für viele Parameter Intervallbereiche angegeben. Wir sind an einzelnen Werten interessiert, sodass wir jeweils den mittleren Wert des Intervalls annehmen. Zudem werden die Parameter $T_{z_1}^{z_2}$ und $T_{z_1}^{z_3}$ anstatt der hier verwendeten Aufenthaltszeit T_{z_1} dargestellt, vergleiche Kapitel 4.1.1. Die Intervallgrenzen des Parameters T_{z_1} berechnen wir, indem wir unter Verwendung von der Mitte des Intervalls für die Übergangswahrscheinlichkeit $\mu_{z_1}^{z_2}$ und von den Intervallgrenzen der getrennten Aufenthaltszeiten $T_{z_1}^{z_2}$ und $T_{z_1}^{z_3}$ das Schema

$$T_{z_1} = \mu_{z_1}^{z_2} T_{z_1}^{z_2} + (1 - \mu_{z_1}^{z_2}) T_{z_1}^{z_3}$$

anwenden. Für T_{z_1} übernehmen wir dann wieder den mittleren Wert des Intervalls. Für altersaufgelöste Werte berechnen wir die Parameter zunächst für jede Altersgruppe nach dem beschriebenen Vorgehen. Im Anschluss gewichten wir die Ergebnisse gemäß dem Anteil der Altersgruppe an der Gesamtbevölkerung, um einen einzelnen Parameterwert zu erhalten. Dafür verwenden wir die Bevölkerungsstatistik mit dem Stichtag 31.12.2020 von [81], da wir uns wie [60] mit Zeiträumen im Jahr 2020 befassen werden. Die Bevölkerungszahl setzen wir gemäß dieser Quelle als $N_0 = 83\,155\,031$. In die Transmissionswahrscheinlichkeit ρ inkludieren wir wie in [60, Kapitel 2.3.1 & 2.3.2] einen Saisonalitätsfaktor, um eine variierende Transmissionswahrscheinlichkeit über das Jahr zu ermöglichen. Unter Verwendung eines Basiswerts $\rho^{(0)}$ legen wir die Transmissionswahrscheinlichkeit durch $s_k(\tilde{t})\rho^{(0)}$ mit der Formel

$$s_k(\tilde{t}) = 1 + k \sin\left(\pi \left(\frac{\tilde{t}}{182.5} + \frac{1}{2}\right)\right)$$

für den Saisonalitätsfaktor fest. Hierbei ist \tilde{t} der Tag des Jahres und k bestimmt den Einfluss der Saisonalität. Für die fiktiven Szenarien vernachlässigen wir die Saisonalität und setzen $k = 0$. Für die realen Szenarien wählen wir $k = 0.2$ gemäß der Mitte des in [60, Table 2] angegebenen Intervalls. Den Basiswert $\rho^{(0)}$ berechnen wir mit dem zuvor beschriebenen Vorgehen aus den angegebenen altersspezifischen Intervallen. Für betrachtete Zeiträume während Herbst und Winter nehmen wir an, dass nur wenige Individuen der Gruppe C detektiert und in Quarantäne geschickt werden, sodass $\xi_C = 1$ angenommen werden kann. Symptomatische Individuen werden häufiger detektiert oder isolieren sich, sodass wir $\xi_I = 0.3$ setzen. Die Bestimmung der Parameter anhand der Angaben in [60] ist in der Datei `cpp/examples/compute_parameters.cpp` dokumentiert. Die Ergebnisse dieser Berechnungen werden in Tabelle 5.1 präsentiert. Zwei Parameter sind farbig markiert, da die Werte auf einer anderen Berechnung basieren. Siehe dafür den Abschnitt über die Parameterwahl für das IDE-Modell. Abgesehen vom Saisonalitätsfaktor werden die beschriebenen Parameter für den jeweiligen Simulationszeitraum als konstant angenommen. Die Kontaktrate ϕ kann sich innerhalb des Zeitraumes ändern, um eine Veränderung des Kontaktverhaltens simulieren zu können, beispielsweise aufgrund des Beschlusses von nicht-pharmazeutischen Maßnahmen. Entsprechend passen wir diesen Parameter an das betrachtete Szenario an und beschreiben die Wahl gemeinsam mit den Szenarien.

Parameterwahl für das LCT-SECIR-Modell

Für das Modell basierend auf dem Linear Chain Trick werden die Parameterwerte des ODE-SECIR-Modells übernommen. Zusätzlich benötigen wir die Anzahlen der Subkompartimente n_z für alle $z \in \mathcal{A}$. Diese werden bei vielen Quellen, die den Linear Chain Trick verwenden, ohne einen epidemiologischen Bezug gesetzt, siehe zum Beispiel [17, 31, 58, 91]. Optimalerweise würden die Anzahlen mittels der passenden Varianz zu den erwarteten Aufenthaltszeiten bestimmt werden, siehe hierzu Abschnitt 3.4.2. Obgleich mögliche Parameterintervalle in [60] angegeben sind, ist die Varianz der epidemiologischen Größen nicht bekannt. Auch aus anderen Quellen sind die benötigten Varianzen der spezifischen epidemiologischen Größen nicht ableitbar. Entsprechend untersuchen wir verschiedene Wahlen der Anzahlen der Subkompartimente. Wir versuchen zum einen, alle Anzahlen auf denselben Wert zu setzen, das heißt $n_z = n$ für alle $z \in \mathcal{A}$. Für n betrachten wir die Werte 3, 10 und 20. In [93, Table 1] sind Schätzungen der Anzahlen der Subkompartimente für E und I für verschiedene Krankheiten angegeben, vergleiche auch Kapitel 3.4.2. Die ermittelten Werte reichen von 2 bis 40, sodass unsere Anzahlen in einem realistischen Bereich liegen. Zudem untersuchen wir die Möglichkeit, die Anzahlen so zu setzen, dass die erwartete Aufenthaltszeit in jedem Subkompartiment etwa einen Tag beträgt. Entsprechend runden wir T_z auf die nächste natürliche Zahl und setzen n_z auf diesen Wert, wobei T_z als die erwartete Aufenthaltszeit im Kompartiment $z \in \mathcal{A}$ definiert war. Diese Methode wird in [52, S. 337 & 342] vorgeschlagen und beispielsweise in [58, S. 7] erwähnt. Wir berücksichtigen folglich insgesamt vier verschiedene Wahlen der Subkompartimente im LCT-Modell. Wie bereits in Kapitel 4.1.3 angeführt, wäre eine weitere Idee, nur für bestimmte Krankheitszustände Subkompartimente einzuführen.

Parameterwahl für das IDE-SECIR-Modell

Das IDE-Modell erlaubt eine anspruchsvollere Parametrisierung. Aus dem ODE-Modell übernehmen wir die Wahl für ρ sowie $\xi_C(t, \tau) = \xi_C$ und $\xi_I(t, \tau) = \xi_I$. Hier besteht Potential für Verbesserungen durch die Option, ξ_C und ξ_I in Abhängigkeit von der verstrichenen Zeit τ seit dem Eintritt in das jeweilige Kompartiment zu variieren. Dies ermöglicht die Berücksichtigung von Veränderungen im Kontaktverhalten im Verlauf der Infektion sowie Veränderungen in der Infektiösität beziehungsweise der Transmissionswahrscheinlichkeit. Wegen dieser verschiedenen Einflüsse und der kritischen Datenlage gestaltet sich die Auswahl geeigneter Funktionen $\xi_C(\tau)$ und $\xi_I(\tau)$ als herausfordernd. Die Effekte wirken zudem vermutlich gegenläufig. Es kann erwartet werden, dass der Anteil der Individuen in Isolation mit einer fortschreitenden Aufenthaltsdauer steigt. Dies würde zu einer Reduktion von $\xi_C(\tau)$ und $\xi_I(\tau)$ mit steigendem τ führen. Eine gegenteilige Auswirkung könnte die Transmissionswahrscheinlichkeit haben, da Individuen erst im Verlauf der Krankheit infektiös werden. In diesem Kontext beschränken wir uns darauf, den Einfluss der übrigen Modifikationen in den Modellannahmen auf die Simulationsergebnisse zu untersuchen. Die Wahrscheinlichkeiten $\mu_{z_1}^{z_2}$ werden ebenfalls aus dem ODE-Modell übernommen.

Nach dem Ergebnis in Bemerkung 3.9, das auf die SECIR-Modelle übertragbar ist, unterscheiden sich die drei Modelle dann ausschließlich in Bezug auf die Auswahl der Funktionen $\gamma_{z_1}^{z_2}(\tau)$. Demzufolge führen wir in den Simulationen einen Vergleich der verschiedenen Möglichkeiten durch, diese Überlebensfunktionen auszuwählen. Das IDE-Modell bietet die größte Flexibilität, sodass die $\gamma_{z_1}^{z_2}(\tau)$ so gewählt werden können, dass sie möglichst realistischen Verteilungen entsprechen. Eine Herausforderung ist, dass nur begrenzt Forschungsarbeiten existieren, die realistische Verteilungen für epidemiologische Parameter untersuchen. Insbesondere fehlen in der aktuellen Literatur solche Verteilungen, die exakt den definierten Bedeutungen unserer Parameter entsprechen. Wenn vollständige epidemiologische Datensätze bekannt wären, können die benötigten Verteilungen daraus bestimmt werden. Für ausgewählte Parameter wird dies in [60, Supplementary material] auf der Basis des zuvor erwähnten Datensatzes von hospitalisierten COVID-19-Patienten umgesetzt. Hier finden sich Ergebnisse für die Funktionen γ_I^H , γ_H^U , γ_H^R , γ_U^D und γ_U^R , welche wir übernehmen können.

Die Verteilungen in [60, Supplementary material] sind in Form von Dichten angegeben, sodass wir sie in Überlebensfunktionen überführen. Zum Anpassen der Dichten an die Daten wird in der Veröffentlichung das Paket „scipy.stats“ der Programmiersprache Python verwendet, siehe hierzu [82]. Entsprechend folgen auch die Definitionen der angegebenen Funktionen und Parameter denen in dem Paket aufgeführten. Wenn für eine Funktion nur zwei Parameter vorgesehen sind, ist der erste aus [60, Supplementary material] zu vernachlässigen und die Reihenfolge der übrigen Parameter folgt der in scipy.stats. Die dort verwendeten Parameter `loc` $\in \mathbb{R}$ und `scale` $\in \mathbb{R}^+$ sind so zu verstehen, dass eine Dichtefunktion f entsprechend der Formel

$$f(x, \text{loc}, \text{scale}) = \frac{1}{\text{scale}} f\left(\frac{x - \text{loc}}{\text{scale}}\right)$$

verschoben und skaliert werden kann. Wenn F die zugehörige Verteilungsfunktion ist,

Parameter	Verteilung	erwartete Aufenthaltszeit
γ_E^C	$1 - \text{ERLANG}(\lambda = n_E \frac{1}{3.335}; \alpha = n_E)$	3.335
γ_C^I	$1 - \text{ERLANG}(\lambda = n_C \frac{1}{1.865}; \alpha = n_C)$	1.865
γ_C^R	$1 - \text{ERLANG}(\lambda = n_C \frac{1}{8.865}; \alpha = n_C)$	8.865
γ_I^H	$1 - \text{ERLANG}(\lambda = n_I \frac{1}{6.30662}; \alpha = n_I)$	6.30662
γ_I^R	$1 - \text{ERLANG}(\lambda = n_I \frac{1}{7}; \alpha = n_I)$	7.0
γ_H^U	$1 - \text{EXPON}(\text{loc} = 0; \text{scale} = 9.36)$	9.36
γ_H^R	$1 - \text{LOGNORM}(\text{s} = 0.76; \text{loc} = -0.45; \text{scale} = 9.41)$	12.110701
γ_U^D	$1 - \text{EXPON}(\text{loc} = 1; \text{scale} = 14.88)$	15.88
γ_U^R	$1 - \text{EXPON}(\text{loc} = 1; \text{scale} = 16.92)$	17.92

Tabelle 5.2: Verwendete Aufenthaltszeitverteilungen für das IDE-SECIR-Modell. Die Verteilungen unterhalb der unteren horizontalen Linie sind aus [60, Supplementary material] übernommen. Die Namen in Großbuchstaben sind als Verteilungsfunktion passend zu den in „scipy.stats“ [82] definierten Dichten zu verstehen und die erwarteten Aufenthaltszeiten werden mit scipy.stats bestimmt. Für die übrigen Verteilungen definieren wir zuerst die erwarteten Aufenthaltszeiten gemäß [60, Table 2] und setzen diese dann in eine passende Erlang-Verteilung ein. Die Notation `ERLANG` steht für die Verteilungsfunktion der Erlang-Verteilung mit den Parameterdefinitionen aus Definition 3.1. Für die Werte n_z mit $z \in \mathcal{A}$ werden in den Simulationen die Wahlen 3 und 10 untersucht.

kann man zeigen, dass

$$F(x, \text{loc}, \text{scale}) = F\left(\frac{x - \text{loc}}{\text{scale}}\right)$$

gilt. Mit dieser Erkenntnis können wir die Überlebensfunktionen aus den angegebenen Dichten errechnen. Die Daten, die zu der Verteilung γ_I^H geführt haben, sind ungenaue Schätzungen und der resultierende Erwartungswert liegt deutlich über den oberen Intervallgrenzen in [60, Table 2]. Entsprechend übernehmen wir diese Verteilung nicht und gehen wie für die fehlenden Verteilungen vor. Für γ_H^R fällt auf, dass ein negativer Parameter `loc` angegeben ist. Dieser negative Wert könnte dazu führen, dass der Träger der Wahrscheinlichkeitsdichte nicht mehr ausschließlich nicht-negative Zahlen einschließt. Folglich könnte die Bedingung $\gamma_H^R(0) = 1$ nicht erfüllt werden. Allerdings ergibt die Auswertung der entsprechenden Überlebensfunktion an der Stelle Null einen Wert nahe eins. Wir behalten die angegebene Verteilung daher in unserer Simulation dennoch bei.

Angesichts der begrenzten Verfügbarkeit von Daten für die übrigen Aufenthaltszeitverteilungen verwenden wir Erlang-Verteilungen und übernehmen die Annahmen für die Anzahlen der Subkompartimente aus dem LCT-Modell. Wir beschränken uns auf die Wahlen $n_z = 3$ und $n_z = 10$ des LCT-Modells für alle $z \in \mathcal{A}$. Die Funktionen $\gamma_{z_1}^{z_2}(\tau)$ dürfen im IDE-Modell von dem Zustand abhängen, in den ein Individuum wechseln wird.

Entsprechend können wir erwartete Aufenthaltszeiten $T_{z_1}^{z_2}$ anstatt T_{z_1} verwenden und diese zur Definition der Erlang-verteilten Aufenthaltszeit annehmen. Wir berechnen dazu die erwarteten Aufenthaltszeiten wie für das ODE-SECIR-Modell beschrieben aus den Werten in [60, Table 2] und lassen die Berechnung von $T_{z_1}^{z_2}$ und $T_{z_1}^{z_3}$ zu einem Wert T_{z_1} weg. Die Verteilungen, die wir zur Simulation verwenden, sind in Tabelle 5.2 dargestellt.

Die erwarteten Aufenthaltszeiten sollten für alle Kompartimente zwischen dem IDE-Modell und den Werten für das ODE- und das LCT-Modell übereinstimmen, damit die Simulationsergebnisse aller Modelle vergleichbar bleiben. Wenn wir wie bei der Parameterwahl des ODE-Modells aus den beiden getrennten Aufenthaltszeiten $T_{z_1}^{z_2}$ und $T_{z_1}^{z_3}$ aus Tabelle 5.2 eine einzelne Aufenthaltszeit berechnen, stimmen die Werte für T_H und T_U nicht mit denen überein, die mit der Methode für das ODE-Modell berechnet wurden. Entsprechend übernehmen wir die Erwartungswerte aus den Verteilungen des IDE-Modells für diese beiden Parameter für das ODE- und LCT-Modell. Anstatt der Werte, die mit der soeben vorgestellten Methode für das ODE-Modell mit den Werten aus [60, Table 2] berechnet werden, verwenden wir somit die Ergebnisse $T_H = 11.634346$ und $T_U = 17.476959$. Dies ist bereits in der Tabelle 5.1 berücksichtigt.

Damit haben wir Werte für alle Parameter besprochen, die für die Simulationen benötigt werden, mit Ausnahme von der Kontaktrate ϕ . Für COVID-19 folgen die Verteilungen vieler bedeutsamer Aufenthaltszeiten in der Realität keiner Exponentialverteilung [21, S. 2]. Die Verteilungen einiger epidemiologischer Größen folgen für die Krankheit eher einer Gamma-Verteilung [9, S. 2]. Entsprechend sind die Annahmen des LCT-Modells für das Beispiel des Erregers SARS-CoV-2 als realistischer einzuschätzen als die des ODE-Modells. Die Wahl der fehlenden Funktionen des IDE-Modells als Erlang-Verteilung ist somit sinnvoll. Die Anpassungsfähigkeit des IDE-Modells wurde genutzt, um feiner aufgelöste Aufenthaltszeiten und vorhandene Schätzungen von Verteilungen zu inkludieren. Dabei konnte die Behauptung aus Kapitel 3.6 verifiziert werden, dass sich die Ermittlung der Parameter für das IDE-Modell als komplex herausstellt. Die Bestimmung der Aufenthaltszeitverteilungen ist so problematisch, dass die Verteilungsannahme des LCT-Modells für einige Aufenthaltszeiten übernommen wurde. Der Vorteil der größeren Flexibilität des IDE-Modells konnte folglich nur teilweise ausgenutzt werden. Dennoch sollten die Annahmen dieses Modells die Realität am genauesten widerspiegeln.

5.2 Simulation fiktiver Szenarien

Nun verwenden wir die epidemiologischen Größen, die an den Erreger SARS-CoV-2 beziehungsweise an die Krankheit COVID-19 angepasst wurden, um fiktive Szenarien zu erstellen. Zuerst untersuchen wir die verschiedenen Möglichkeiten zur Initialisierung des LCT-SECIR-Modells aus Kapitel 4.3. Darüber hinaus möchten wir das qualitative Verhalten des ODE-, des LCT- und des IDE-SECIR-Modells miteinander vergleichen. Dazu müssen wir sicherstellen, dass die Modelle dieselbe Ausgangslage haben und die Simulationsergebnisse vergleichbar sind. Wir halten hierzu das Infektionsgeschehen zu Beginn der Simulationen möglichst konstant und auf dem gleichen Niveau für die Modelle. Um ein gleichbleibendes Infektionsgeschehen zu erreichen, bestimmen wir die Kontaktrate

$\phi(t)$ so, dass die effektive Reproduktionszahl $\mathcal{R}_{\text{eff}}(t)$ etwa eins ist. Vergleiche hierzu auch Bemerkung 2.1. Dazu bestimmen wir für die effektive Reproduktionszahl im ODE-SECIR-Modell eine Formel auf eine naive Weise aus den Definitionen der Parameter. Wie bereits in Bemerkung 2.1 beschrieben wurde, gibt es genauere Vorgehen zur Herleitung der Reproduktionszahlen, jedoch reicht die Heuristik für unseren Zweck aus. Ein Individuum wird im Verlauf seiner Ansteckung im Durchschnitt T_C Tage im Krankheitszustand C verweilen. Pro Tag hat das Individuum in diesem Krankheitszustand $\rho^{(0)} \phi(t) \xi_C$ Kontakte, die zu einer Ansteckung führen, falls sie mit einer Person aus der Gruppe S stattfinden. Für die fiktiven Szenarien verwenden wir den Parameter $k = 0$ für die Saisonalität, sodass wir direkt $\rho^{(0)}$ einsetzen, vergleiche auch den vorigen Abschnitt 5.1. Ein Individuum steckt in seinem Aufenthalt im Krankheitszustand C entsprechend im Durchschnitt $\rho^{(0)} \phi(t) \xi_C T_C \frac{S(t)}{N(t)}$ Personen an. Mit einer Wahrscheinlichkeit von μ_C^I entwickelt die Person Symptome und bleibt eine durchschnittliche Zeit von T_I Tagen in der Gruppe I . Es ist zu erwarten, dass eine symptomatische Person $\rho^{(0)} \phi(t) \xi_I \frac{S(t)}{N(t)}$ Ansteckungen pro Tag verursacht. Entsprechend erhalten wir das Ergebnis, dass ein durchschnittliches Individuum im Zeitraum seiner Infektion

$$\mathcal{R}_{\text{eff}}(t) = \rho^{(0)} \phi(t) \left(\xi_C T_C + \mu_C^I \xi_I T_I \right) \frac{S(t)}{N(t)}$$

Personen mit dem Erreger infizieren wird. Um die Kontaktrate $\phi(t)$ zu bestimmen, für die sich ein konstantes Infektionsgeschehen ergibt, setzen wir die Reproduktionszahl gleich eins und wenden die Parameter aus der Tabelle 5.1 an. Wir vernachlässigen den Einfluss des Faktors $\frac{S(t)}{N(t)}$ und nehmen $S(t) \approx N(t)$ an, da wir nur eine gemeinsame Ausgangslage der Modelle herstellen möchten, wofür eine approximative Größe ausreicht. Dann ergibt sich

$$\begin{aligned} 1 &= \mathcal{R}_{\text{eff}}(t) \approx \rho^{(0)} \phi(t) \left(\xi_C T_C + \mu_C^I \xi_I T_I \right) \\ \Leftrightarrow \phi(t) &\approx \frac{1}{\rho^{(0)} \left(\xi_C T_C + \mu_C^I \xi_I T_I \right)} \stackrel{\text{Tabelle 5.1}}{\approx} 2.746265. \end{aligned}$$

Für das Erreichen eines konstanten Infektionsgeschehens zu Beginn der Simulationen ist außerdem eine passende Wahl der Anfangsbedingungen essentiell. Mit dieser Thematik beschäftigen wir uns im ersten Szenario in Abschnitt 5.2.1. Die Kontaktrate übernehmen wir für alle Modelle, auch wenn wir sie mittels der Formulierung der effektiven Reproduktionszahl anhand der Parameter des ODE-Modells bestimmt haben.

Zur Simulation verwenden wir eine initiale Schrittweite von $\Delta t = 0.1$. Bei den beiden Modellen, die auf gewöhnlichen Differentialgleichungssystemen basieren, wenden wir, wie in Kapitel 4.2.1 beschrieben, eine adaptive Methode zur Lösung an. Allerdings setzen wir zur Erstellung der Plots eine obere Grenze von $\Delta t \leq 0.1$ für die Schrittweite, damit die Reaktion auf die Veränderung der Kontaktrate besser zu vergleichen ist und die Kurven glatt wirken. Neben dem zeitlichen Verlauf der prognostizierten Anzahlen der Individuen in den Kompartimenten betrachten wir zum Vergleich der Simulationen Graphen, welche die Anzahl der Neuansteckungen pro Tag angeben. Wir berechnen hierfür die Anzahl der Personen, die innerhalb eines Tages von S nach E übergehen. Das bedeutet, dass für den

Zeitpunkt t_{n+1} der Wert

$$\frac{|S_{n+1} - S_n|}{t_{n+1} - t_n}$$

abgebildet ist. Die Idee stammt aus einer Publikation von Feng et al. [31, S. 1513f.], die später allerdings Abbildungen für die ebenfalls interessante Größe der kumulierten neuen Ansteckungen präsentieren.

Wir betrachten nun ausgewählte Simulationsergebnisse von fiktiven Szenarien. Dazu werden jeweils zunächst die Voraussetzungen für die einzelnen durchgeführten Simulationen geklärt. Danach präsentieren und beschreiben wir die Resultate kurz, um im Anschluss mögliche Gründe für das Verhalten zu analysieren. Dies führt zu einem besseren Verständnis der Funktionsweisen der Modelle. Zum Schluss zeigen wir jeweils die Konsequenzen des beobachteten Verhaltens in den Szenarien auf. Wir beginnen mit der Untersuchung verschiedener Methoden zur Initialisierung des LCT-SECIR-Modells. Hieraus ergibt sich auch eine geeignete Methode zur Wahl der Anfangsbedingungen für die beiden anderen Modelle, um ein konstantes Infektionsgeschehen zu erzielen. Im Anschluss simulieren wir zwei Fälle, in denen wir die Kontaktrate nach einer Simulationszeit mit einem konstanten Infektionsgeschehen verändern. Hier testen wir die Reaktion der drei Modellarten auf Veränderungen. Als letztes Szenario simulieren wir mit Reproduktionszahlen größer als eins über einen längeren Simulationszeitraum. So können wir das prognostizierte Ausmaß der Epidemie für die verschiedenen Modelle vergleichen.

5.2.1 Vergleich verschiedener Methoden zur Initialisierung des LCT-Modells

Zunächst stellen wir die drei in Kapitel 4.3 vorgestellten Methoden zur Initialisierung des LCT-SECIR-Modells einander gegenüber. Für diesen Abschnitt beschränken wir uns auf das LCT-Modell und führen keinen Vergleich mit anderen Modellen durch. Ziel des Szenarios ist es, anhand von Simulationen zu untersuchen, ob einfachere Methoden zur Bestimmung der Anfangswerte des LCT-Modells bereits ausreichen, um eine adäquate Initialisierung zu gewährleisten. Die Verfahren vergleichen wir unter der Annahme $\mathcal{R}_{\text{eff}}(t) \approx 1$. Die Anfangsbedingungen setzen wir ebenfalls so, dass wir ein gleichbleibendes Infektionsgeschehen in den Simulationen erwarten. Dazu erzeugen wir zunächst passende synthetische Daten. Für die Methoden zur Initialisierung des LCT-Modells wird der konstante Wert $\Delta t = 0.1$ des IDE-Modells übernommen.

Beginnen wir mit der Methode der Übergänge und definieren Werte $\Sigma_{z_1}^{z_2}(t_m)$ für die benötigten Zeitpunkte $t_m = m\Delta t$ für $m \in \mathbb{Z}$ mit $\max\{a, b\} \leq m \leq 0$, sodass wir ein stabiles Infektionsgeschehen erwarten können. Wir möchten erreichen, dass die Anzahl der Neuansteckungen innerhalb eines Zeitintervalls gleich bleibt. Entsprechend sollte $\sigma_S^E(t)$ beziehungsweise $\Sigma_S^E(t)$ über den gesamten benötigten Zeitraum einen konstanten Wert aufzeigen. Unter dieser Annahme lassen sich passende Werte für die übrigen Übergänge $\Sigma_{z_1}^{z_2}$ ableiten. Wenn wir die Gleichung (4.14) für $\sigma_E^C(t)$ betrachten, ergibt sich aus der Annahme $\sigma_S^E(t) = \sigma_S^E$ für alle $t \leq 0$ der Zusammenhang

$$\sigma_E^C(t) = -\sigma_S^E \int_{\hat{T}}^t \gamma_E^{C'}(t-x) dx = -\sigma_S^E \left[-\gamma_E^C(t-x) \right]_{\hat{T}}^t = \sigma_S^E \left(\gamma_E^C(0) - \gamma_E^C(t-\hat{T}) \right) = \sigma_S^E$$

für alle $t \leq 0$. Hier wird $\gamma_E^C(0) = 1$ ausgenutzt und, dass $\gamma_E^C(t - \hat{T}) = 0$ gilt, da \hat{T} ein betragslich großer, negativer Wert ist. Entsprechend wählen wir $\Sigma_E^C(t_m) = \Delta t \sigma_S^E = \Sigma_S^E$ für alle $m \in \mathbb{Z}$ mit $\max\{a, b\} \leq m \leq 0$. Die gleiche Berechnungsweise kann für alle Übergänge fortgeführt werden, sodass sich passende konstante Anfangswerte $\Sigma_{z_1}^{z_2}$ aus einem einzelnen Wert für Σ_S^E ergeben.

Den Wert Σ_S^E übernehmen wir aus der „7-Tage-Inzidenz“ aus dem Lagebericht des RKIs zu COVID-19 vom 15.10.2020 [75]. Wir benötigen für die Anzahl der Neuansteckungen keinen möglichst genauen, sondern lediglich einen realistischen Wert. Deshalb skalieren wir die angegebene Inzidenz auf ganz Deutschland und auf unsere betrachtete Schrittweite und übernehmen das Ergebnis als Anzahl der neuen Ansteckungen. Die Anzahl der Neuansteckungen pro Tag ergibt sich dann als ungefähr 4050. Auch die Todesfälle sowie die Anzahl der gesamten bestätigten Fälle übernehmen wir aus dem besagten COVID-19 Lagebericht. Wenn wir auf diese Weise synthetische Daten für die Übergänge $\Sigma_{z_1}^{z_2}(t_m)$ für benötigte Zeitpunkte t_m mit $m \in \mathbb{Z}$ mit $\max\{a, b\} \leq m \leq 0$ erstellen und $\phi(t)$ wie zuvor wählen, erwarten wir ein konstantes Infektionsgeschehen. In Abschnitt 4.3 haben wir festgestellt, wie wir Anfangsbedingungen für alle drei Modelltypen basierend auf gegebenen Übergängen ableiten können. Diese Methode ist für eine zueinander konsistente Initialisierung der verschiedenen Modelle geeignet. Wenn wir die synthetisch gesetzten Übergänge für die Initialisierung des ODE- und des IDE-Modells in den übrigen Szenarien übernehmen, erwarten wir ebenso ein konstantes Infektionsgeschehen.

Die Vorgehensweise zur Bestimmung der Anfangsbedingungen über die Übergänge für das LCT-Modell vergleichen wir mit den anderen beiden Methoden zur Initialisierung des Modells aus Kapitel 4.3. Für die Methode nach Koslow et al. [56, S2 Appendix] resultiert aus der Voraussetzung der konstanten Anzahl der Neuansteckungen, dass auch die Anzahl der neuen bestätigten Fälle innerhalb eines Zeitschritts gleich bleiben sollte. Für einen beliebigen Zeitpunkt t ergibt sich aus den Annahmen für die Methode, dass

$$F_{\text{MAR}}(t + \Delta t) - F_{\text{MAR}}(t) = \mu_C^I \Sigma_S^E$$

gelten muss. Mit dieser Differenz lassen sich alle Anfangswerte für die Subkompartimente und Kompartimente mittels der in Kapitel 4.3 beschriebenen Formeln bestimmen. Die Anzahl der Todesfälle und der gesamten bestätigten Fälle übernehmen wir wie zuvor direkt aus dem Lagebericht des RKIs [75]. Für die Methode nach Hurtado et al. [45, S. 1836] summieren wir jeweils die Anzahlen in den Subkompartimenten auf, die wir aus dem Ergebnis nach Koslow et al. [56, S2 Appendix] erhalten haben. Die Werte verwenden wir als Initialisierung für das jeweils erste Subkompartiment und setzen die übrigen Subkompartimente auf null. So sind die Kompartimente mit derselben Anzahl an Individuen initialisiert.

Im Folgenden vergleichen wir die Resultate von Simulationen mit den unterschiedlichen Verfahren zur Wahl der Anfangsbedingungen für das LCT-Modell. In der Datei `cpp/-examples/let_secir_initializations.cpp` ist die Implementierung des Szenarios zu finden. Wir präsentieren in jedem Szenario nur die interessantesten Ergebnisse, die zu den größten Erkenntnissen führen könnten. Die Implementierungen können zum Beispiel mit wenig Aufwand so angepasst werden, dass eine andere Anzahl der Subkompartimente ausprobiert werden kann. Es wurden wesentlich mehr Graphen erstellt, als hier vorgestellt werden.

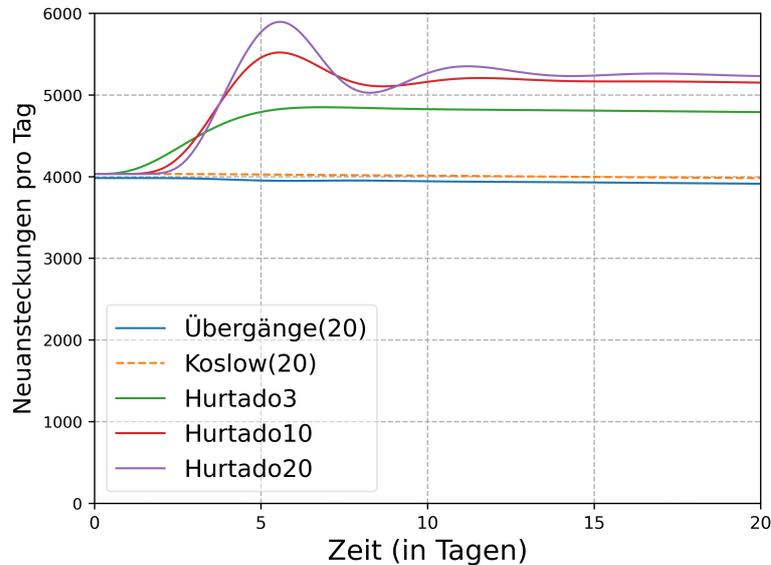


Abbildung 5.1: Tägliche Neuansteckungen für verschiedene Methoden zur Initialisierung des LCT-Modells. Darstellung der Simulationsergebnisse für die neuen Ansteckungen pro Tag im Vergleich für verschiedene Methoden zur Initialisierung des LCT-SECIR-Modells. Für die Initialisierung mittels Übergängen sowie nach der Methode von Koslow et al. wird das Ergebnis für 20 Subkompartimente dargestellt. Andere Wahlen ergeben hier dieselbe Form. Siehe für eine Beschreibung der Methoden auch Kapitel 4.3. Die Zahlen hinter den Methodennamen stehen für die verwendete Anzahl der Subkompartimente n_z für alle $z \in \mathcal{A}$.

In diesem Szenario zur Initialisierung des LCT-Modells betrachten wir beispielsweise kein Ergebnis für die variable Anzahl der Subkompartimente, bei welchen sich in jedem Subkompartiment eine erwartete Aufenthaltszeit von etwa einem Tag ergibt.

Die Simulationsergebnisse zu den täglichen Neuansteckungen sind in Abbildung 5.1 dargestellt. Für die Methode der Initialisierung nach Hurtado et al. [45, S. 1836] sind die Kurven gemeinsam mit der verwendeten Anzahl der Subkompartimente n_z für alle $z \in \mathcal{A}$ bezeichnet. Für die Initialisierung mittels der Übergänge $\Sigma_{z_1}^{z_2}$ und die nach Koslow et al. [56, S2 Appendix] mit den mittleren Aufenthaltszeiten ist nur der Fall von jeweils 20 Subkompartimenten abgebildet. Simulationen mit anderen Anzahlen der Subkompartimente ergeben keine sichtbaren Abweichungen. Für diese beiden Methoden ist eine etwa konstante Anzahl täglicher Neuansteckungen sichtbar. Dementsprechend haben wir die Kontaktrate sowie die Anfangswerte und Übergänge passend zu unserem Ziel, ein konstantes Infektionsgeschehen herbeizuführen, gewählt. Das Ergebnis für die Anzahl der täglichen neuen Ansteckungen zeigt für die Initialisierung nach Hurtado et al. [45, S. 1836] für die drei ausgesuchten Anzahlen der Subkompartimente unterschiedliche Wellenformen auf. Anfangs ist die Anzahl der Neuansteckungen konstant und steigt dann nach einer gewissen Verzögerung an. Mit zunehmender Anzahl der Subkompartimente wird die Wellenform ausgeprägter. Die Amplitude der Wellen wird mit fortschreitender Simulationszeit kleiner. Wenn T_{\max} größer gewählt und über einen längeren Zeitraum

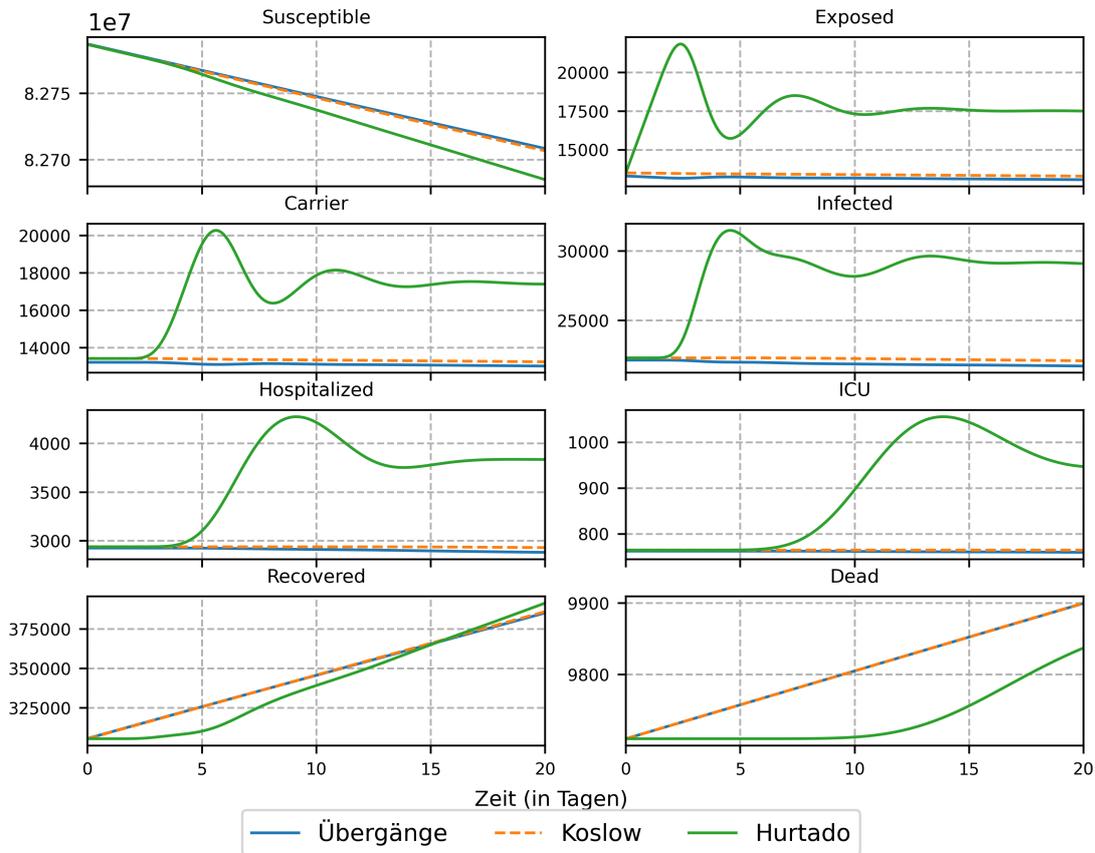


Abbildung 5.2: Kompartmente für verschiedene Methoden zur Initialisierung des LCT-Modells. Darstellung der Simulationsergebnisse für den zeitlichen Verlauf der Anzahlen der Individuen in den verwendeten Kompartmenten im Vergleich für verschiedene Methoden zur Initialisierung des LCT-SECIR-Modells. Siehe für eine Beschreibung der Methoden auch Kapitel 4.3. Alle Simulationen wurden mit $n_z = 20$ Subkompartmenten für alle $z \in \mathcal{A}$ durchgeführt.

simuliert wird, pendeln sich die prognostizierten Anzahlen der Neuansteckungen jeweils auf einem höheren Niveau ein als mit den Methoden nach Koslow et al. [56, S2 Appendix] und mittels der Übergänge. Schwankungen sind nicht mehr sichtbar und die Anzahlen sind weitestgehend konstant. Nach einer gewissen Zeit lässt sich bei allen Methoden zur Initialisierung eine sinkende Anzahl der prognostizierten Neuansteckungen erkennen. Dies ist vermutlich darauf zurückzuführen, dass die effektive Reproduktionszahl mit fortschreitender Simulationsdauer unter eins fällt, da der Faktor $\frac{S(t)}{N(t)}$ sinkt.

In Abbildung 5.2 sind die Ergebnisse der Anzahlen der Individuen in den Kompartmenten dargestellt. Die Anzahlen in den Subkompartmenten werden hier jeweils zu einem Wert für das gesamte Kompartiment aufaddiert. Die Kompartmente sind gemäß der englischen Bezeichnungen betitelt, denen die Buchstaben für die Kompartmente folgen, siehe Kapitel 4.1. Hier werden nur die Ergebnisse für ein LCT-Modell mit 20 Subkompartmenten abgebildet, da für diese Wahl bei der Anzahl der täglichen Neuan-

steckungen das wellenförmige Verhalten am deutlichsten sichtbar ist. Für die Methoden mittels der Übergänge und der nach Koslow et al. [56, S2 Appendix] bleiben die Größen $z(t)$ für $z \in \mathcal{A}$ zeitlich konstant. Mit der Methode nach Hurtado et al. [45, S. 1836] ist für diese Gruppen abermals ein Wellenmuster erkennbar. Bei der Größe $E(t)$ ist unmittelbar nach dem Simulationsbeginn ein Anstieg ersichtlich und es schließen sich Schwankungen an. Für die anderen Kompartimente in \mathcal{A} ist eine unterschiedliche Verzögerungsdauer auffällig, in welcher die Anzahlen in den Krankheitszuständen wie bei den übrigen Initialisierungsmöglichkeiten konstant bleiben. Danach ist jeweils ein Anstieg sichtbar und es ergeben sich verschiedene Schwankungsmuster. Die Anzahlen in D und R bleiben zunächst zeitlich konstant, während die Ergebnisse der anderen Methoden bereits steigende Werte aufweisen. Im Langzeitverhalten ist erkennbar, dass die simulierten Todesfälle mit der Initialisierung nach Hurtado et al. [45, S. 1836] die übrigen Prognosen nach etwa 45 Simulationstagen übertreffen.

Nachfolgend analysieren wir die Gründe für das Verhalten der Ergebnisse mit der Methode nach Hurtado et al. [45, S. 1836]. Sämtliche Individuen im Kompartiment $z \in \mathcal{A}$ haben zum Simulationsbeginn eine erwartete restliche Aufenthaltszeit von T_z , da sich alle im ersten Subkompartiment befinden. Bei den übrigen Verfahren zur Initialisierung gibt es Individuen, die sich in fortgeschrittenen Subkompartimenten befinden und eine kleinere restliche erwartete Aufenthaltszeit haben, siehe auch Abschnitt 3.4.1. Die Kompartimente werden mittels der Methoden nach Hurtado et al. [45, S. 1836] und Koslow et al. [56, S2 Appendix] mit denselben Anzahlen initialisiert, jedoch ergibt sich durch die Verteilung in die Subkompartimente ein unterschiedliches Verhalten. Die Anzahl der Individuen im Kompartiment E steigt bei der Methode nach Hurtado et al. [45, S. 1836] sofort nach dem Simulationsbeginn an. Anfangs verlassen keine Individuen E , da alle Personen zunächst die Kette der Subkompartimente durchlaufen müssen. In E tritt allerdings die gleiche Anzahl der Personen wie bei den übrigen Methoden ein, was in der Grafik für die neuen Ansteckungen erkennbar ist. Deshalb steigt die prognostizierte Anzahl $E(t)$. Die Größe sinkt wieder, wenn viele Individuen die Kette der Subkompartimente durchlaufen haben und mehr Individuen hinaus wechseln, als neu angesteckt werden. Für das Kompartiment C ist eine gewisse Verzögerung des Anstiegs ersichtlich. In der ersten Zeit wechseln weder Personen aus C hinaus noch nach C hinein. Es gilt $T_C \approx T_E$, sodass nach ungefähr dieser Zeit Individuen nach C eintreten und aus C austreten. Hier ist ein Anstieg von $C(t)$ erkennbar, da die steigende Anzahl von Personen in E vom Beginn der Simulation zum Tragen kommt. Im Anschluss fällt $C(t)$, da wieder weniger Individuen aus E kommen und viele Personen ins nächste Kompartiment übergehen. Vergleichen wir die Kurve $C(t)$ mit der Abbildung für die Anzahl der täglichen Neuansteckungen, wirkt der Verlauf der Kurven ähnlich. Dies ist der Fall, da wegen der Wahlen $\xi_C = 1$ und $\xi_I = 0.3$ (vergleiche Tabelle 5.1) der Einfluss von der Gruppe C auf die neuen Ansteckungen größer ist als der Einfluss von I . Nach dem Abfall von $E(t)$ ist abermals eine Erhöhung erkennbar. Hier sind die Personen, die sich anfänglich in E befanden, bereits zum Großteil aus E gewechselt und die Kurve der neuen Ansteckungen steigt, sodass wieder mehr Individuen in E eintreten. Der Anstieg ist nun geringer als der erste, weil nicht nur Individuen nach E kommen, sondern gleichzeitig Individuen nach C fortgehen. Nach einer Verzögerung steigen entsprechend auch $C(t)$ und die Kurve der Neuansteckungen wieder. Dies führt zu einer neuen kleineren

Welle in E . Die Amplitude der Wellen wird immer kleiner, da die Individuen sich nach und nach besser auf die Subkompartimente von E aufteilen und der Zufluss nach E gleichmäßiger wird. Das Verhältnis der Individuen, die nach E übergehen, zu denen, welche die Gruppe verlassen, wird immer ausgeglichener. Die geringeren Ausschläge werden an die übrigen Kompartimente weiter gegeben. Wenn wir den Krankheitszustand I betrachten, erkennen wir eine Zunahme von $I(t)$ nach etwas weniger als T_C Tagen. Hier wechseln die ersten Individuen von C nach I . Da T_I deutlich größer als T_C ist, gehen zu dem Zeitpunkt wenige Personen von I aus in andere Kompartimente über. Die Varianz der Aufenthaltszeitverteilung für I ist wegen des höheren Werts für T_I größer, siehe Gleichung (3.20). Entsprechend verteilen sich die Individuen stärker in den Subkompartimenten. Die Wellenform überträgt sich daher auch in einem anderen Muster auf I . Die Formen von $H(t)$ und $U(t)$ sind mittels derselben Erklärungen verständlich. Anfangs gehen keine Individuen nach R und D über, da alle Individuen zunächst die Ketten der Subkompartimente durchlaufen. Entsprechend bleiben die Größen konstant, während die der übrigen Methoden linear ansteigen. Die nachfolgende Steigung ist größer, zuerst weil viele Individuen im gleichen Zeitraum das Ende der Ketten erreicht haben und im Anschluss, weil die prognostizierte Anzahl der Ansteckungen höher ist. Die Wellenformen in der Abbildung 5.1 sind mit weniger Subkompartimenten nicht so stark ausgeprägt, da Individuen nicht so viele Subkompartimente durchlaufen, bis sie in andere Kompartimente wechseln. Für kleinere Werte n_z ist die Varianz der Aufenthaltszeitverteilungen größer, vergleiche Gleichung (3.20). Entsprechend verteilen sich die Individuen gleichmäßiger auf die Subkompartimente, sodass weniger Schwankungen entstehen.

Die Simulationsergebnisse zeigen, dass die Vorhersagen mit der vereinfachten Methode nach Hurtado et al. [45, S. 1836] erheblich von den Ergebnissen mit den anderen Methoden zur Initialisierung abweichen. Schwankende Ergebnisse sind bei einer konstanten Wahl von $\mathcal{R}_{\text{eff}}(t) \approx 1$ unrealistisch. Daher ist es unzureichend, vereinfachende Annahmen zu treffen und es ist erforderlich, umfassende Überlegungen zur korrekten Initialisierung anzustellen. In Veröffentlichungen wird der Thematik bisher keine große Bedeutung zugesprochen. Zur Simulation realer Fälle ist dies gleichermaßen bedeutsam, da die Prognosen offensichtlich stark von der Wahl der Initialisierungsmethode beeinflusst werden. Die Anfangsbedingungen für die übrigen fiktiven Szenarien werden für alle Modelle mittels der Methode der Übergänge erstellt, wodurch die Initialisierungen der Modelle vergleichbar sind. Für reale Daten ist es eine Herausforderung, die Werte für die benötigten Übergänge herauszufiltern, wie bereits in Kapitel 4.3 festgestellt wurde. Entsprechend führen wir die Simulation realer Szenarien nur mit dem ODE- und dem LCT-SECIR-Modell durch und initialisieren jeweils mit dem Verfahren nach Koslow et al. [56, S2 Appendix].

5.2.2 Vergleich der Modelle bei einer Reduzierung der Kontaktrate

Im folgenden Szenario testen wir die Reaktion der Modelle auf eine Verringerung der Kontaktrate. Dies könnte einem Beschluss einer nicht-pharmazeutischen Maßnahme entsprechen. Nach einer Phase mit gleichbleibendem Infektionsgeschehen zu Beginn der Simulation halbieren wir dazu die Kontaktrate nach einer Simulationszeit von zwei

	ODE	LCT3	LCT10	LCT20	LCTvar	IDE
$T_{\max} = 10$	23	25	45	61	24	101
$T_{\max} = 20$	27	31	59	91	30	201

Tabelle 5.3: Vergleich der benötigten Zeitschritte der Modelle bei einer Verringerung der Kontaktrate. Die Simulationen wurden für zwei verschiedene Simulationsdauern T_{\max} mit einer Halbierung der Kontaktrate nach zwei Tagen und der initialen Schrittweite $\Delta t = 0.1$ durchgeführt. Die Anzahlen sind inklusive der Zeitpunkte 0 und T_{\max} zu verstehen. Die Zahlen bei den Benennungen des LCT-Modells stehen für die verwendeten Anzahlen der Subkompartimente. Die Abkürzung „var“ steht für die variable Anzahl, bei der die erwartete Aufenthaltszeit in jedem der Subkompartimente etwa einen Tag beträgt. Für das IDE-Modell ergibt sich für verschiedene Annahmen dieselbe Anzahl an Zeitschritten, da das Lösungsverfahren nicht adaptiv ist.

Tagen. Da die effektive Reproduktionszahl unter der Annahme $S(t) \approx N(t)$ somit auf $\mathcal{R}_{\text{eff}}(t) \approx 0.5$ abfällt, erwarten wir einen Rückgang des Infektionsgeschehens, vergleiche Bemerkung 2.1. Die Implementierungen der übrigen drei fiktiven Szenarien sind in der Datei `cpp/examples/lct_secir_fictional_scenario.cpp` zu finden.

Wir beginnen unsere Untersuchung mit den erforderlichen Zeitschritten der Modelle als Indikator für die Rechenzeit. Während der Ausführung der Simulationen ist bemerkbar, dass das IDE-Modell mehr Zeit in Anspruch nimmt als die anderen Modelle. Diese Wahrnehmung sollte durch Laufzeitmessungen bestätigt werden. In Tabelle 5.3 sind die benötigten Anzahlen an Zeitschritten bis zur Erreichung von T_{\max} für verschiedene Modelle und zwei verschiedene maximale Zeitpunkte dargestellt. Im Unterschied zu den Simulationsergebnissen, die zur Erstellung der Graphen verwendet werden, wird hier für die adaptiven Verfahren keine maximale Schrittweite definiert. Die Zahlen in der Benennung des LCT-Modells stehen für die verwendeten Anzahlen der Subkompartimente n_z für alle $z \in \mathcal{A}$. Die Abkürzung „var“ steht für die variable Anzahl an Subkompartimenten für die einzelnen Kompartimente, bei der die erwartete Aufenthaltszeit in jedem der Subkompartimente etwa einen Tag beträgt. Mit diesen Bezeichnungen werden wir ebenso die Kurven in den Graphen versehen. Die benötigten Zeitschritte für das IDE-Modell können im Vorhinein bestimmt werden, da die Schrittweite über den Simulationszeitraum konstant gesetzt wird. Die Parameter n_z fließen für einige Kompartimente, wie in Tabelle 5.2 beschrieben, in die Aufenthaltszeitverteilungen im IDE-Modell ein. Für verschiedene Annahmen bezüglich n_z sind die benötigten Anzahlen der Schritte identisch. Für die Graphen kürzen wir die Fälle entsprechend der Anzahl der Subkompartimente mit „IDE3“ und „IDE10“ ab. Die benötigte Anzahl der Zeitschritte erhöht sich mit einer steigenden Anzahl der Subkompartimente für das LCT-Modell. Für den Fall LCTvar liegen die erforderlichen Zeitschritte zwischen dem ODE-Modell und LCT3. Bei der Verwendung einer höheren Anzahl an Subkompartimenten ist eine längere Simulationsdauer zu erwarten. Das IDE-Modell benötigt stets die höchste Anzahl an Zeitschritten, allerdings wird hier die Schrittweite anfangs festgesetzt und könnte auch auf andere Werte fixiert werden. Bei den Modellen, die auf gewöhnlichen Differentialgleichungssystemen basieren, beobachten

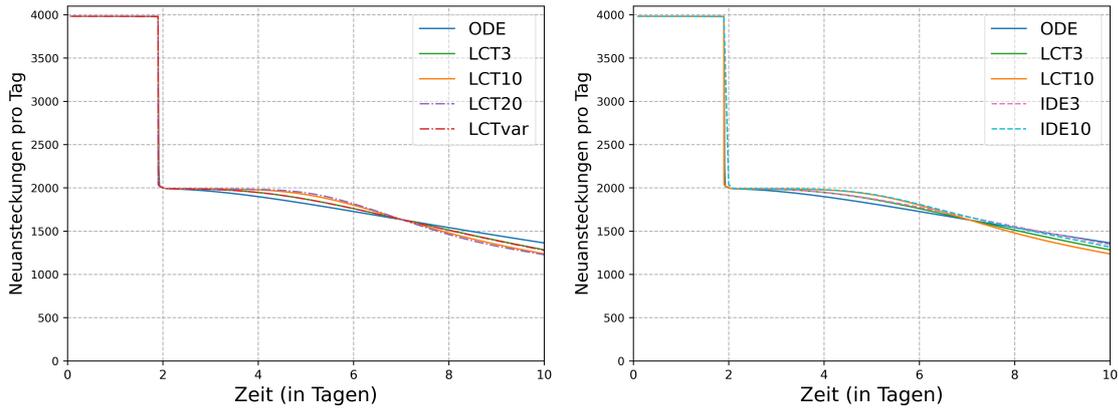


Abbildung 5.3: Tägliche Neuansteckungen bei einer Verringerung der Kontaktrate. Vergleich der Simulationsergebnisse für die täglichen Neuansteckungen des ODE-Modells mit dem LCT-SECIIR-Modell (links) und zusätzlich mit dem IDE-Modell (rechts) unter verschiedenen Annahmen zu den Anzahlen der Subkompartimente und einer Halbierung der Kontaktrate ϕ nach zwei Simulationstagen.

wir, dass sich die Anzahl der benötigten Zeitschritte bei einer Verdopplung der Simulationszeit nicht verdoppelt. Dies ist darauf zurückzuführen, dass die Kontaktrate nach zwei Simulationstagen verändert wird. Das adaptive Verfahren wählt für die Modelle um den Punkt der Veränderung kleine Schrittweiten. Die Anzahl der Zeitschritte lässt nicht direkt auf die Simulationsdauer schließen, sodass diese zusätzlich gemessen werden sollte. Trotzdem bieten die Werte einen ersten Eindruck über den Aufwand der Simulationen.

Im Folgenden werden die Simulationsergebnisse der verschiedenen Modelle vorgestellt. In Abbildung 5.3 sind die täglichen Neuansteckungen dargestellt. Der linke Graph setzt den Fokus auf den Vergleich von LCT-Modellen mit variierender Anzahl von Subkompartimenten. IDE-Modelle werden in der rechten Darstellung in die Betrachtung einbezogen. Die Wahl einer kurzen Simulationsdauer ermöglicht es, die Reaktion der Modelle auf die Veränderung der Kontaktrate deutlich zu erkennen. Die Initialisierung über die Übergänge hat bei sämtlichen Modellen funktioniert, wodurch die Infektionsdynamik in den ersten beiden Tagen konstant bleibt. Nach zwei Simulationstagen wird die Kontaktrate ϕ halbiert. In allen Modellen zeigt sich zu diesem Zeitpunkt eine entsprechende Halbierung der täglichen Neuansteckungen. Wir beginnen den Vergleich mit dem Verhalten der LCT-Modelle in der linken Abbildung. Das ODE-Modell entspricht einem LCT-Modell mit einem Subkompartiment $n_z = 1$ für alle $z \in \mathcal{A}$. Bei einigen Modellen bleibt die Anzahl der Neuansteckungen nach dem Abfall nach zwei Tagen vorerst auf einem konstanten Niveau. Je höher die Anzahl der Subkompartimente ist, desto länger ist die Dauer der Verzögerung bis zu einem weiteren Rückgang der Neuansteckungen. Beim ODE-Modell ist kaum Verzug erkennbar. Nach der Verzögerungsphase ist die Steigung der Kurve der Neuansteckungen umso geringer, je größer die Anzahl der Subkompartimente ist. Für größere n_z sinkt die Anzahl der Neuansteckungen demnach schneller. Das Ergebnis der Simulation bezüglich der täglichen Neuansteckungen von LCTvar entspricht visuell nahezu dem Ergebnis für LCT3. Auch die Simulationsergebnisse von LCT10 und LCT20

	ODE	LCT3	LCT10	LCT20	LCTvar	IDE3	IDE10
$\mathcal{R}_{\text{eff}}(t) \approx 0.5$	–	–5.93 %	–9.40 %	–10.25 %	–6.32 %	–1.35 %	–3.47 %
$\mathcal{R}_{\text{eff}}(t) \approx 2$	–	–0.19 %	–1.15 %	–1.97 %	0.03 %	–6.72 %	–7.86 %

Tabelle 5.4: Relative Abweichungen der Prognosen der Neuansteckungen verschiedener Modelle vom ODE-Modell. Die relative Abweichung der Ergebnisse wird für den Zeitpunkt $T_{\text{max}} = 10$ berechnet. Betrachtet wird der Fall einer Halbierung der Kontaktrate nach zwei Simulationstagen, sodass $\mathcal{R}_{\text{eff}}(t) \approx 0.5$ gilt. Zudem wird der Fall für eine Verdoppelung nach zwei Simulationstagen untersucht, vergleiche dazu auch Abschnitt 5.2.3.

ähneln sich. Dies begründet die Entscheidung, für den Vergleich mit den IDE-Modellen die Fälle mit 3 und 10 Subkompartimenten heranzuziehen. Die Anzahlen der täglichen Neuansteckungen der IDE-Modelle zeigen eine analoge Verzögerung eines weiteren Abfalls der Kurven wie das LCT-Modell mit derselben Annahme bezüglich der Subkompartimente. Nach dem Verzug weist das Ergebnis der IDE-Modelle eine größere Steigung auf, sodass die Kurven nicht so rasch abfallen wie die der LCT-Modelle. In Tabelle 5.4 erfolgt ein Vergleich der prognostizierten Anzahl neuer täglicher Ansteckungen zum Zeitpunkt T_{max} der verschiedenen Modelle mit dem Ergebnis des ODE-Modells. Schon acht Tage nach der Reduktion der Kontaktrate zeigen sich signifikante Unterschiede zwischen den Modellprognosen.

Nun beschreiben wir die prognostizierte zeitliche Entwicklung der Anzahl der Individuen in den einzelnen Krankheitszuständen. In Abbildung 5.4 sind die prognostizierten Anzahlen in den Kompartimenten für verschiedene LCT-Modelle dargestellt. Bei einigen Modellen bleiben die Ergebnisse für die Krankheitszustände C , I , H und U eine gewisse Zeit nach der Verringerung der Kontaktrate konstant. Es dauert einige Tage, bis sich die Veränderung des Kontaktverhaltens in den Größen der Kompartimente bemerkbar macht. Anschließend beginnen die Anzahlen abzunehmen. Ähnlich wie bei den Neuansteckungen ist die Verzögerung kürzer, je geringer die Anzahl der Subkompartimente ist. Im ODE-Modell zeigt sich kaum Verzögerungszeit. Danach ist die Steigung der Kurven wieder umso geringer, je größer die Anzahl der Subkompartimente ist. Bei den Kurven für H und U dauert die Verzögerung so lange an, dass der Abfall erst zum Ende des Simulationszeitraums erkennbar ist. Nur beim ODE-Modell ist rasch ein Abfall festzustellen. Beim Kompartiment E gibt es keine Verzögerung und die Kurve sinkt direkt nach der Reduktion der Kontaktrate. Für die Steigung gilt dieselbe Beobachtung wie für die übrigen Kompartimente. Für die Zustände E , C und I sind die Ergebnisse für LCT3 und LCTvar wieder ähnlich. In Abbildung 5.5 ist die Anzahl der Individuen in den Krankheitszuständen im Vergleich zu den IDE-Modellen abgebildet. Die Kurven der IDE-Modelle zu E unterscheiden sich jeweils nicht wesentlich von denen des entsprechenden LCT-Modells. Für das Kompartiment C zeigt sich der gleiche Trend wie für die Neuansteckungen. Die Verzögerungszeiten der IDE-Modelle sind jeweils ähnlich zu denen der LCT-Modelle. Im späteren Verlauf fällt die Steigung von $C(t)$ bei den IDE-Modellen jedoch nicht so gering aus wie bei den LCT-Modellen. Beim Zustand I ist die Zeitspanne zwischen der Reduzierung der effektiven Reproduktionszahl und dem Abfall der Kurve bei den IDE-

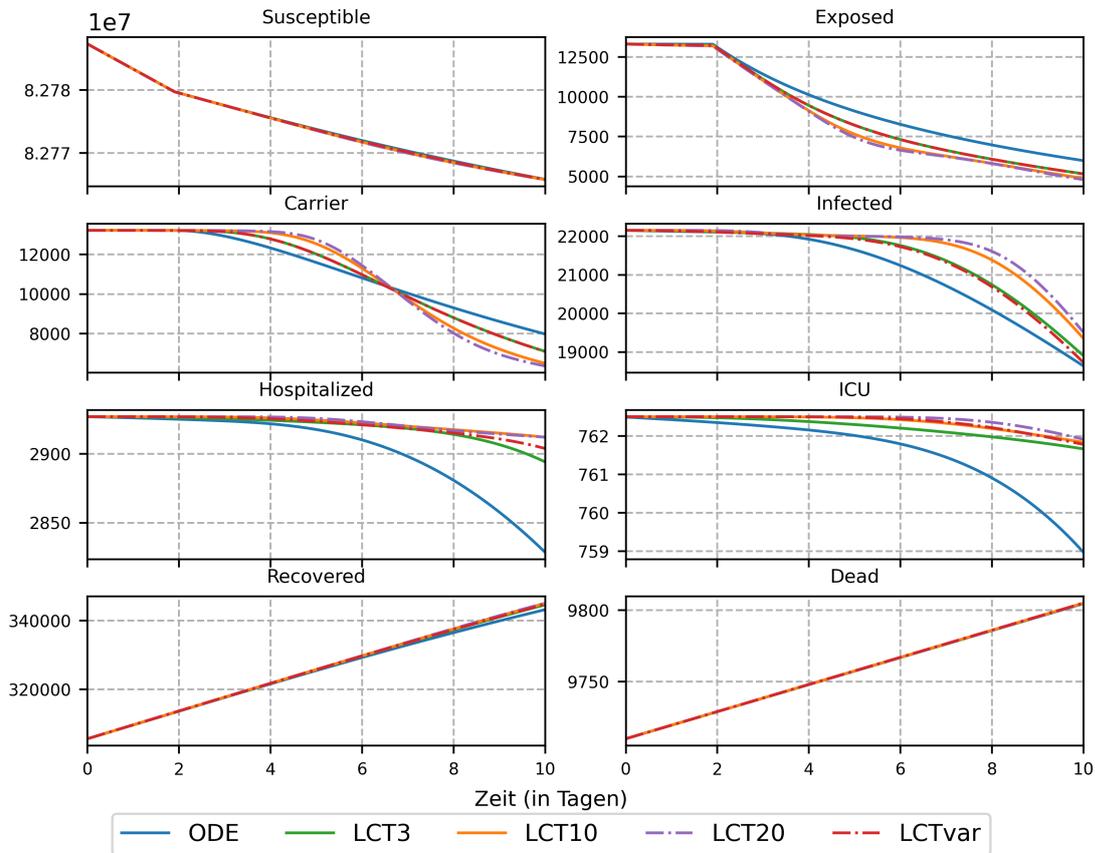


Abbildung 5.4: Kompartmente bei einer Verringerung der Kontaktrate in LCT-Modellen. Darstellung der Simulationsergebnisse für die Anzahl der Individuen in den jeweiligen Kompartimenten des LCT-SECIR-Modells im Vergleich für die Annahmen von 1, 3, 10, 20 und einer variablen Anzahl an Subkompartimenten. Nach zwei Simulationstagen wird die Kontaktrate ϕ halbiert.

Modellen kleiner als bei den LCT-Modellen, während sie bei den Gruppen H und U größer ausfällt.

Wir fangen bei den LCT-Modellen an, um mögliche Gründe für das Verhalten zu erörtern. Die Modelle unterscheiden sich in der Wahl der Anzahl der Subkompartimente und damit in der Varianz der Aufenthaltszeitverteilung in den jeweiligen Krankheitszuständen, vergleiche Gleichung (3.20) und Abbildung 3.2. Je größer n_z gewählt wird, desto geringer ist die Fluktuation um die mittlere Aufenthaltszeit. Nach dem Abfall von ϕ sinkt die Anzahl der Neuansteckungen für alle Modelle unmittelbar, sodass weniger Individuen in die Gruppe E eintreten und $E(t)$ abnimmt. Beim ODE-Modell gibt es einen Anteil der Individuen, die einen Krankheitszustand kurz nach dem Eintritt wieder verlassen, vergleiche Abbildung 3.2. Entsprechend wechseln unmittelbar weniger Individuen nach C , sodass keine Verzögerung sichtbar ist. Diese schnelle Reaktion setzt sich in den übrigen Kompartimenten fort, sodass die Kurven für I , H und U für das ODE-Modell ebenso kaum Verzögerungszeit aufweisen. Je geringer die Varianz ist, desto später wechseln die

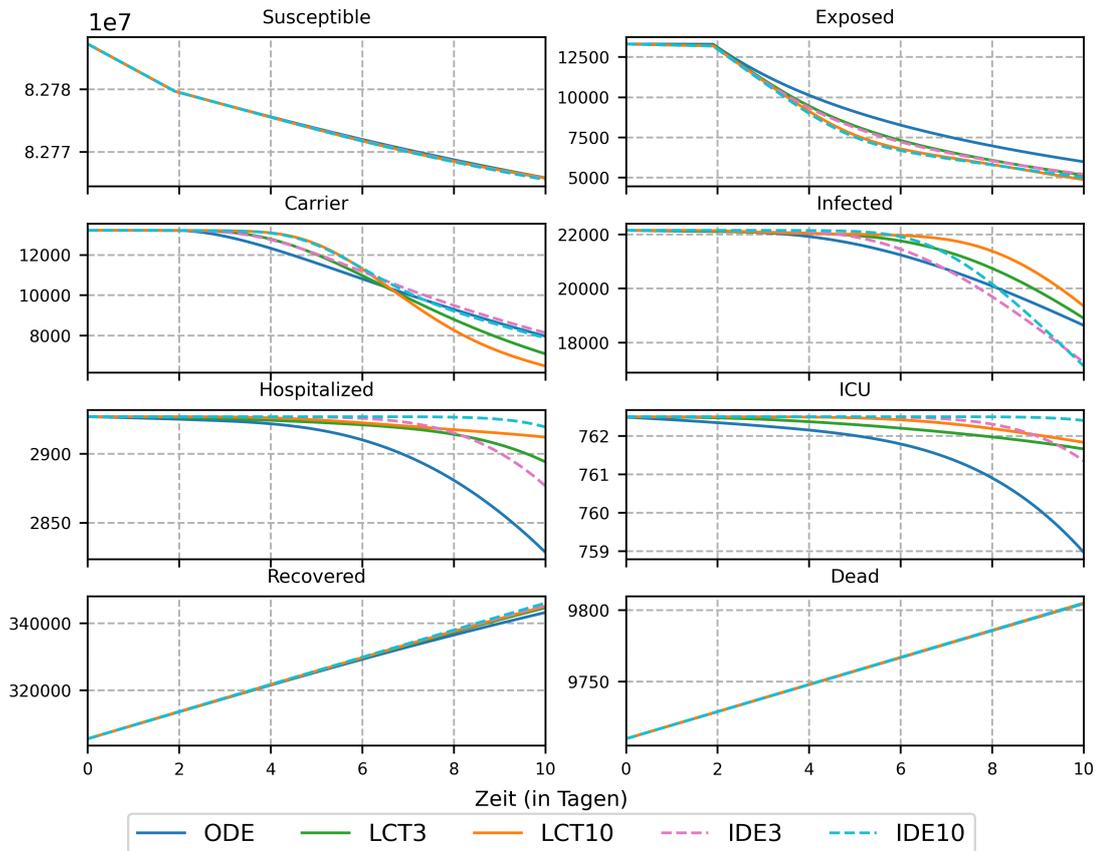


Abbildung 5.5: Kompartmente bei einer Verringerung der Kontaktrate im Vergleich mit IDE-Modellen. Darstellung der Ergebnisse für die Anzahl der Individuen in den jeweiligen Kompartmenten des ODE-SECI_R-Modells im Vergleich mit dem LCT- und IDE-Modell jeweils mit den Annahmen von 3 und 10 Subkompartmenten. Nach zwei Simulationstagen wird die Kontaktrate ϕ halbiert.

ersten Individuen, die nach dem zweiten Tag angesteckt wurden, nach C . Demzufolge ist die Verzögerung umso länger, je größer die Wahl für n_z ist. Dieses Verhalten überträgt sich ebenso auf die Krankheitszustände I , H und U .

Nun analysieren wir einen möglichen Grund dafür, dass die Kurve für $E(t)$ für das ODE-Modell weniger steil abfällt. Im ODE-Modell gibt es aufgrund der höheren Varianz Individuen, die lange nach der erwarteten Aufenthaltszeit noch in E verweilen. Es erfordert eine entsprechende Zeitspanne, bis auch diese Anzahlen der Personen mit einem hohen Infektionsalter abnehmen. Die höhere Anzahl der Neuansteckungen beeinflusst das Kompartiment E demnach über einen längeren Zeitraum. Folglich fällt die Größe $E(t)$ für das ODE-Modell langsamer und die Steigung bleibt größer. Je geringer die Varianz der Aufenthaltszeitverteilung ist, desto geringer ist der Anteil der Individuen, die nach einer langen Dauer noch in dem Kompartiment verweilen und den Abfall der Kurve verlangsamen. Entsprechend ist die Steigung von $E(t)$ geringer, je höher die Anzahl der Subkompartmente gewählt wurde. Diese Erklärung lässt sich auf die übrigen Kompar-

timente übertragen, wobei der Effekt bei den fortgeschrittenen Krankheitszuständen wegen der Simulationsdauer nicht erkennbar ist. Eine Veränderung der Steigung bei den Todesfällen ist vermutlich ebenfalls aufgrund der kurzen Simulationszeit nicht zu erkennen.

Wie bei dem Szenario der unterschiedlichen Initialisierungen im vorigen Abschnitt verhält sich die Kurve der neuen täglichen Ansteckungen nach dem zweiten Tag ähnlich zu der für C . Somit lassen sich die Erklärungen für die unterschiedlichen Verzögerungen und die Steigungen auf die Neuansteckungen übertragen. Die Ergebnisse von LCT3 und LCTvar verhalten sich für die Kompartimente E und C und für die neuen Ansteckungen analog, da die Anzahl der Subkompartimente bei LCTvar für E und C aufgrund der erwarteten Aufenthaltszeit ebenfalls auf drei gesetzt wird. Für die übrigen Kompartimente zeigen sich Unterschiede in den Kurvenverläufen.

Im Folgenden analysieren wir das Verhalten der Ergebnisse, die mit den IDE-Modellen berechnet werden. Die Aufenthaltszeitverteilung γ_E^C ist beim IDE-Modell dieselbe wie beim LCT-Modell mit derselben Annahme bezüglich der Subkompartimente, sodass die Kurven für E jeweils etwa gleich sind. Die Verzögerung des Abfalls von C des IDE-Modells passt zu der Verzögerungszeit des jeweiligen LCT-Modells. Wie zuvor erläutert, ist der Verzug durch die Verteilung der Aufenthaltszeit in E bedingt, weshalb das ähnliche Verhalten zu erwarten ist. Der Unterschied in der Steigung des Abfalls im Kompartiment C ist darauf zurückzuführen, dass im IDE-Modell zwischen den beiden Verteilungen γ_C^I und γ_C^R unterschieden wird. Individuen, die von C nach I übergehen, haben eine kürzere erwartete Aufenthaltszeit als T_C , vergleiche Tabelle 5.2 mit Tabelle 5.1. Entsprechend ist die Verzögerung beim Kompartiment I für die IDE-Modelle kürzer. Die erwartete Aufenthaltszeit T_C^R ist jedoch bedeutend länger als T_C . Somit gibt es beim IDE-Modell mehr Personen mit einer relativ langen Aufenthaltszeit in C , die den Rückgang der Anzahl in der Gruppe hemmen. Daher weisen die Kurven $C(t)$ für die IDE-Modelle eine höhere Steigung auf. Wie bei den LCT-Modellen verhalten sich die neuen täglichen Ansteckungen nach dem zweiten Tag ähnlich zu $C(t)$. Für das Kompartiment I ist die Steigung der IDE-Modelle kleiner als die der LCT-Modelle. Hier ist T_I^R geringfügig größer als T_I . Allerdings ist die erwartete Aufenthaltszeit T_I^H für Individuen, die nach H wechseln werden, bedeutend kürzer. Dies könnte zu einem schnelleren Abfall von $I(t)$ führen. Für die IDE-Modelle ergibt sich ein komplexes Zusammenspiel der verschiedenen Aufenthaltszeitverteilungen. Es könnte zu weiteren Erkenntnissen führen, die zeitlichen Verläufe für die Übergänge $\Sigma_{z_1}^{z_2}$ zu betrachten. Beim Krankheitszustand H sinkt die Kurve der LCT-Modelle früher als die des passenden IDE-Modells. Die erwartete Zeit vom Eintritt in E bis zum Wechsel nach H liegt für die LCT-Modelle bei etwa 13.59 Tagen und für die IDE-Modelle bei ungefähr 11.51 Tagen, wie aus den Tabellen 5.1 und 5.2 berechnet werden kann. Entsprechend liegt die erwartete Zeit des Eintritts in das Kompartiment H für Individuen, die nach dem zweiten Tag angesteckt wurden, außerhalb des Simulationszeitraums. Die erwarteten Aufenthaltszeiten fließen quadratisch in die Varianzen der verwendeten Erlang-Verteilungen ein, siehe Gleichung (3.20). Demzufolge sind die Varianzen für die LCT-Modelle für die Aufenthaltszeiten in den Krankheitszuständen vor dem Eintritt in H bedeutend größer. Deshalb treten die ersten Individuen, die nach dem zweiten Simulationstag angesteckt wurden, trotz der längeren erwarteten

Zeit vom Eintritt in E bis zum Eintritt in H im LCT-Modell früher in das Kompartiment H ein. Deshalb ist die Verzögerungszeit für die LCT-Modelle kürzer als für die IDE-Modelle. Bei IDE3 führt die kleine Wahl der Anzahlen der Subkompartimente zu einer größeren Varianz im Vergleich zu IDE10. Dieser Effekt führt dazu, dass die Verzögerung von IDE3 kürzer ausfällt als die von IDE10. Die Steigungen der Kurven der IDE-Modelle im Anschluss an die Verzögerung sind auf die Wahl der Aufenthaltszeitverteilungen γ_H^R und γ_H^U als logarithmische Normalverteilung beziehungsweise Exponentialverteilung zurückzuführen, siehe Tabelle 5.2. Ähnliche Überlegungen lassen sich auf den Zustand U übertragen.

Wir können festhalten, dass die Reaktionen auf eine Verringerung der Kontaktrate beziehungsweise auf eine Reduktion der effektiven Reproduktionszahl für verschiedene Modelle und Annahmen unterschiedlich sind. Die Ergebnisse variieren hinsichtlich der Verzögerungszeit, bis ein Abfall der Kurven sichtbar ist, sowie bezüglich der Steigung der Kurven. Dey et al. [24] beschäftigen sich mit der Verzögerung zwischen dem Beschluss oder der Lockerung von nicht-pharmazeutischen Maßnahmen und Änderungen in den Daten zu den bestätigten COVID-19 Fällen und Todesfällen in den Vereinigten Staaten. Eine politische Maßnahme führt laut der Publikation im Durchschnitt nach 10 bis 14 Tagen zu einer Veränderung der Daten [24, S. 10]. Das Wissen über die Verzögerungszeit zeigt die Notwendigkeit, dass politische Entscheidungsträger frühzeitig und vorausschauend nicht-pharmazeutische Maßnahmen planen [24, S. 1]. Entsprechend sollten Simulationen zur Unterstützung von Politikern eine solche Verzögerung aufzeigen. Guglielmi et al. [38, S. 3f.] stellen heraus, dass Verzögerungen zwischen der Ansteckung und der Berücksichtigung in den Daten sowohl nicht-natürlicher Art sein können, wie durch Mess- und Meldeverzug, als auch natürlicher Art. Eine Verzögerung tritt demnach durch die Inkubationszeit auf, das heißt die Zeit von der Ansteckung bis zum Auftreten der ersten Symptome der Krankheit, vergleiche [54, S. 76]. Auch sie stellen eine signifikante Verzögerung für Daten aus Italien und der Schweiz fest, die bei der Modellierung von nicht-pharmazeutischen Maßnahmen berücksichtigt werden sollte [38, S. 3 & 17]. Die natürliche Verzögerung sollte in den Simulationsergebnissen ohne weitere Anpassungen sichtbar sein. Beim ODE-Modell konnten wir durch die Annahme der exponentialverteilten Aufenthaltszeiten fast keinen Verzug feststellen. Dies legt die Verwendung eines der anderen Modelle nahe. Durch eine höhere Anzahl an Subkompartimenten im LCT-Modell kann eine längere Verzögerungszeit erreicht werden. Bei einer hohen Anzahl von Subkompartimenten sind die Unterschiede in den Kurven nicht mehr stark ausgeprägt. Hier sollten Kosten und Nutzen abgewogen werden, da die Verwendung höherer Anzahlen der Subkompartimente sowie die Verwendung eines IDE-Modells zu einer größeren benötigten Anzahl an Zeitschritten führen könnte, wie wir zu Beginn des Abschnitts festgestellt haben.

5.2.3 Vergleich der Reaktion auf eine Erhöhung der Kontaktrate

In diesem Szenario vergleichen wir wie im vorangegangenen Abschnitt die Reaktion der Modelle auf eine Veränderung der Kontaktrate. Wir verdoppeln ϕ nach zwei Simulationstagen, was einem Anstieg der effektiven Reproduktionszahl auf $\mathcal{R}_{\text{eff}}(t) \approx 2$ unter der Annahme $S(t) \approx N(t)$ entspricht. Eine Steigerung der Reproduktionszahl könnte durch

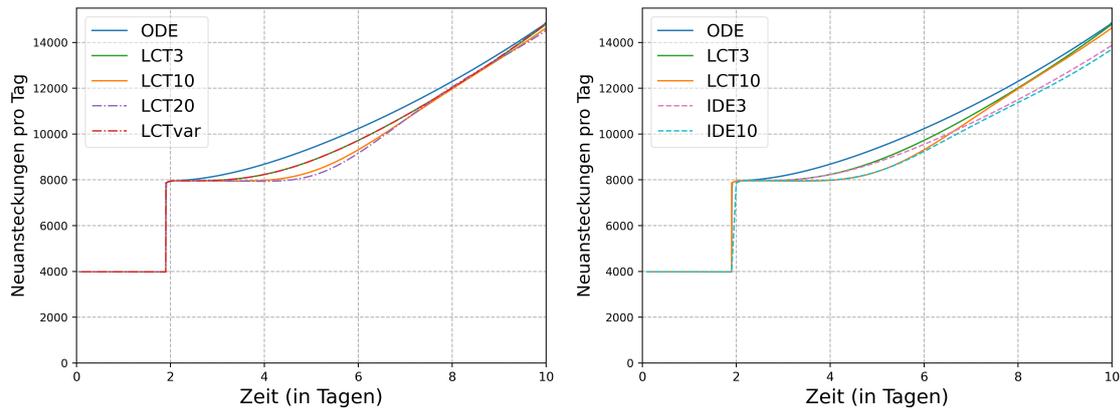


Abbildung 5.6: Tägliche Neuansteckungen bei einer Erhöhung der Kontaktrate. Vergleich der Simulationsergebnisse für die täglichen Neuansteckungen des ODE-Modells mit dem LCT-SECIR-Modell (links) und zusätzlich mit dem IDE-Modell (rechts) mit verschiedenen Annahmen zu den Anzahlen der Subkompartimente und einer Verdopplung der Kontaktrate ϕ nach zwei Simulationstagen.

die Aufhebung einer nicht-pharmazeutischen Maßnahme verursacht werden. Wir erwarten ein intensiviertes Infektionsgeschehen und einen Anstieg der neuen Ansteckungen als Reaktion auf die erhöhte Reproduktionszahl, vergleiche Bemerkung 2.1.

In Abbildung 5.6 sind die prognostizierten täglichen Neuansteckungen abgebildet, wobei die linke Grafik verschiedene LCT-Modelle vergleicht und die rechte die IDE-Modelle einbezieht. Am zweiten Tag der Simulation verdoppeln sich die Anzahlen der Neuansteckungen. Die Kurven zeigen ein analoges Verhalten zum ausführlich besprochenen Fall der Reduktion der Kontaktrate auf. Die Verzögerung des weiteren Anstiegs der neuen Ansteckungen ist länger, je größer die Anzahl der Subkompartimente gewählt wird. Für das ODE-Modell ist fast kein Zeitverzug sichtbar. Im Anschluss zeigen die Modelle mit den höheren Anzahlen eine extremere Steigung auf, was hier einer größeren Steigung entspricht. Für LCT3 und LCTvar ist keine sichtbare Abweichung festzustellen. Zudem unterscheiden sich die Kurven für LCT10 und LCT20 nicht stark. Die Ergebnisse der IDE-Modelle verhalten sich ebenso analog zu denen aus dem vorigen Szenario. Die Verzögerungszeit entspricht etwa der des LCT-Modells mit derselben Annahme bezüglich der Subkompartimente und im Anschluss ist eine weniger extreme Steigung beobachtbar. Die Steigung ist hier demnach kleiner. Die relativen Abweichungen der Ergebnisse von den Prognosen des ODE-Modells zum Zeitpunkt T_{\max} sind gemeinsam mit denen des letzten Szenarios in Tabelle 5.4 zu finden. Die Abweichungen sind hier teilweise nicht so deutlich wie bei der Reduzierung der Kontaktrate. Die Effekte der Verzögerung und der Steigung gleichen sich aus, was zu ähnlichen Prognosen für die LCT-Modelle am Ende der Simulation führt. Die Verläufe der Kurven für die Kompartimente zeigen ebenso gleiche Tendenzen wie zuvor. Entsprechend sind die Abbildungen nicht in die vorliegende Arbeit inkludiert. Die ausführlichen Erklärungen für die Simulationsergebnisse lassen sich aus dem Abschnitt 5.2.2 für den Fall einer steigenden Kontaktrate übernehmen.

Alles in allem konnten die Ergebnisse aus dem vorigen Kapitel bezüglich der Reak-

tionen der verschiedenen Modelle auf Veränderungen im Kontaktverhalten anhand von Simulationen mit einer Steigerung der Kontaktrate bestätigt werden. Das Fehlen einer Verzögerungszeit könnte ein Grund sein, sich gegen ein ODE-Modell für eine Simulation zu entscheiden.

5.2.4 Vergleich des prognostizierten Höhepunkts der Epidemie

In diesem Abschnitt behandeln wir Simulationen, um die prognostizierten Höhepunkte einer Epidemie von verschiedenen Modellen zu vergleichen. Laut [55, S. 1] ist ein Ziel von nicht-pharmazeutischen Maßnahmen, das Maximum der Infektionszahlen zu verringern, um zu verhindern, dass die Kapazität des Gesundheitswesens überschritten wird und um Zeit für die Forschung zu gewinnen. Während der COVID-19-Pandemie verwendeten Politik und Presse den Slogan „flattening the curve“, um die Strategie der Verlangsamung der Ausbreitung zu kommunizieren [68, S. 1]. Aus diesem Grund untersuchen wir, wie die Annahmen verschiedener Aufenthaltszeitverteilungen den prognostizierten Höhepunkt der Epidemie beeinflussen. Verschiedene Ergebnisse könnten zu unterschiedlichen Bewertungen von nicht-pharmazeutischen Maßnahmen auf der Basis von den Simulationen führen.

In der Literatur gibt es bereits Studien zu diesem Themenbereich. Wir skizzieren drei ausgewählte Untersuchungen und deren Resultate kurz und evaluieren die Ergebnisse im Anschluss anhand von eigenen Simulationen. Wearing et al. [93] stellen für ein SIR-Modell mit dem Linear Chain Trick wie das aus Kapitel 3.2 fest, dass die Annahme einer Erlang-Verteilung mit $n = 100$ anstatt einer Exponentialverteilung erhebliche Auswirkungen auf den Verlauf von $I(t)$ hat. Hier wird eine Reproduktionszahl von fünf angenommen. Das ODE-Modell prognostiziert zu Beginn einen geringeren Anstieg von $I(t)$ und die maximale Anzahl an Infizierten ist deutlich geringer. Die Dauer der Epidemie sei unter der Annahme eines LCT-Modells zudem kürzer [93, S. 0622f.]. Blythe et al. [7] betrachten ein komplexeres Modell für die Simulation der Ausbreitung von HIV mit der Einbeziehung von Geburts- und Sterbeprozessen sowie beliebiger Verteilungen für die infektiöse Periode und die Inkubationszeit. Für die Definitionen dieser epidemiologischen Größen siehe auch Seite 7 beziehungsweise Seite 98. Der Höhepunkt der Epidemie wird anhand der Anzahl der neuen Fälle pro Jahr untersucht. Die Studie kommt zu dem Ergebnis, dass die Spitze der Epidemie umso später und höher ist, je geringer die Varianz der Verteilungen ist [7, S. 17]. Blyuss et al. [9] verwenden ein ähnliches Modell wie das LCT-SECIR-Modell für eine Modellierung der Ausbreitung von COVID-19. Allerdings werden die Kompartimente C und I hier zu einem Zustand I zusammengefasst und nur für diesen und den Krankheitszustand E Subkompartimente eingeführt. Nach der Publikation gilt, dass je höher die Anzahl der Subkompartimente in E gewählt wird, desto früher erreicht die Epidemie ihren Höhepunkt. Je größer außerdem die Anzahl der Subkompartimente in I gewählt wird, desto höher sei die maximale Anzahl der Individuen in I und in der Intensivstation [9, S. 9]. Die Ergebnisse der Studien sind bezüglich des Zeitpunkts des Höhepunkts der Epidemie gegensätzlich. Blyuss et al. prognostizieren einen früheren Höhepunkt der Epidemie für eine höhere Anzahl an Subkompartimenten und Blythe et al. einen späteren. Diese Resultate könnten von der verwendeten Modellstruktur abhängig sein. Alle Studien erhalten das Ergebnis, dass die Annahme von mehr Subkompartimenten

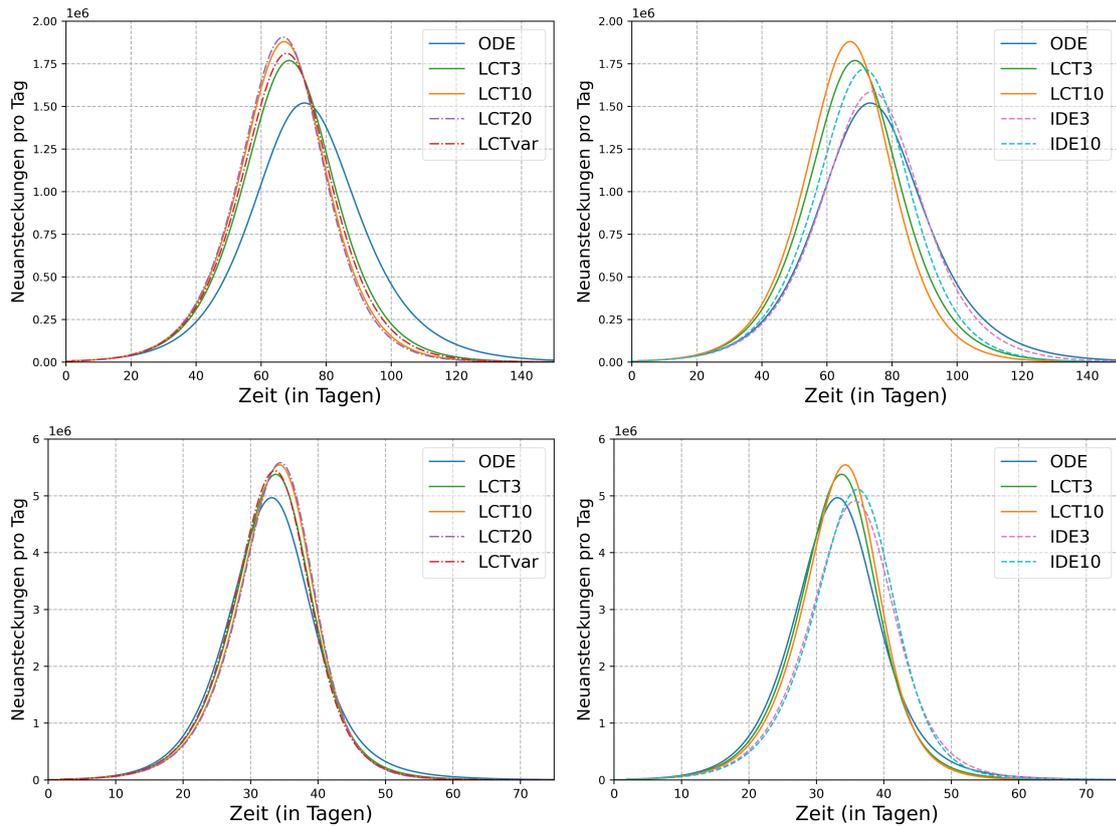


Abbildung 5.7: Tägliche Neuinfektionen zum Vergleich der prognostizierten Höhepunkte der Epidemie. Darstellung der Ergebnisse des ODE-Modells, von LCT-Modellen (oben links) und zusätzlich von IDE-Modellen (oben rechts) mit verschiedenen Annahmen zu den Subkompartimenten und einer Verdopplung der Kontaktrate ϕ nach zwei Tagen. Zudem sind die Ergebnisse derselben Modelle mit einer Vervielfachung von ϕ nach zwei Tagen abgebildet (unten).

zu einem höheren Maximum führt.

Wir überprüfen die Ergebnisse der Publikationen nun anhand von eigenen Simulationen. Dafür testen wir für die effektive Reproduktionszahl mit $S(t) \approx N(t)$ die Werte zwei und vier. Damit die Modelle von einem gleichen Niveau aus beginnen, initialisieren wir die Modelle wie zuvor mit der Annahme von $\mathcal{R}_{\text{eff}}(t) \approx 1$. Nach zwei Simulationstagen verdoppeln oder vervierfachen wir die Kontaktrate ϕ . Die Dauer der Simulation wird jeweils so gewählt, dass der prognostizierte Höhepunkt der Epidemie und der Rückgang des Infektionsgeschehens erkennbar sind. Wir vergleichen den Verlauf der Epidemie für verschiedene Modelle anhand der täglichen Neuinfektionen in Abbildung 5.7. In einigen Publikationen wird der Höhepunkt der Epidemie anhand der Anzahl der Infektiösen untersucht. Wir könnten hier ebenfalls die Kompartimente C oder I oder deren Summe in Betracht ziehen. Die qualitativen Feststellungen, die wir für die täglichen Neuinfektionen treffen, gelten gleichermaßen für die Kurven $C(t)$ und $I(t)$. In Abbildung 5.8 sind die Anzahlen in den Kompartimenten für einige Modelle für den Fall der Verdopplung der

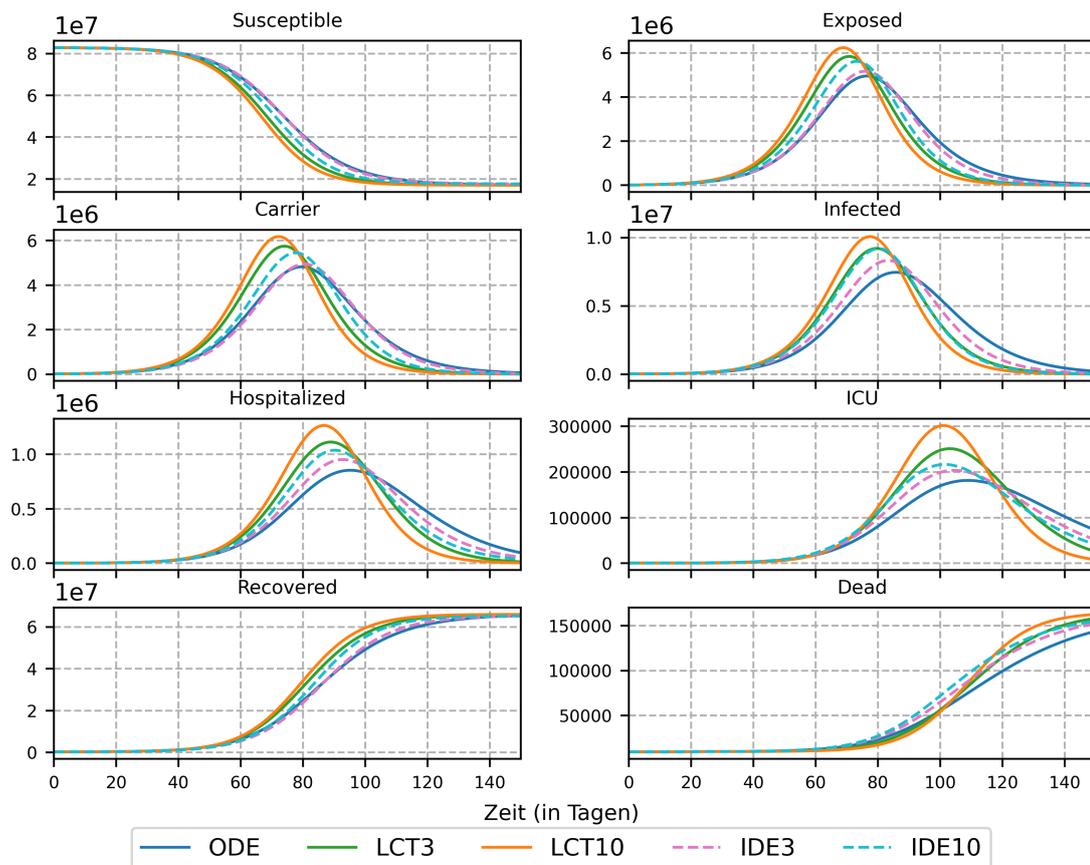


Abbildung 5.8: Kompartimente zum Vergleich der prognostizierten Höhepunkte der Epidemie bei verschiedenen Modellarten. Darstellung der Ergebnisse für die Anzahlen der Individuen in den Kompartimenten des ODE-SECIIR-Modells im Vergleich mit dem LCT- und IDE-Modell jeweils mit 3 und 10 Subkompartimenten. Nach zwei Tagen wird die Kontaktrate ϕ verdoppelt.

Kontaktrate dargestellt. Hier können die Kurven $I(t)$ oder $C(t)$ für die verschiedenen Modelle mit den zugehörigen Neuansteckungen verglichen werden.

Als Indikator für den Höhepunkt der Epidemie betrachten wir somit die Anzahl der täglichen Neuansteckungen. Für beide Reproduktionszahlen ergibt sich im Vergleich der LCT-Modelle, dass der maximale Wert der Kurve umso größer wird, je mehr Subkompartimente gewählt werden. Die Werte der Höhepunkte für die LCT-Modelle mit mehr als einem Subkompartiment sind dabei bedeutend größer als die des ODE-Modells. Die IDE-Modelle prognostizieren jeweils geringere Maxima als das LCT-Modell mit derselben Annahme bezüglich der Subkompartimente. Beim Fall der Reproduktionszahl von vier prognostiziert IDE3 sogar einen geringeren Höhepunkt für die Anzahl der Neuansteckungen und auch für $C(t)$ als das ODE-Modell. Das Verhältnis der Zeitpunkte, zu denen die verschiedenen Modelle das Maximum der Epidemie vorhersagen, unterscheidet sich für die beiden Reproduktionszahlen. Im Fall von $\mathcal{R}_{\text{eff}}(t) \approx 2$ zeigt sich bei den LCT-Modellen,

dass das Maximum umso früher erreicht wird, je größer die Anzahl der Subkompartimente gewählt wird. Für $\mathcal{R}_{\text{eff}}(t) \approx 4$ ergibt sich ein entgegengesetztes Verhalten. Der Höhepunkt wird umso später erreicht, je größer die Anzahl der angenommenen Subkompartimente ist. Für die IDE-Modelle ergibt sich für beide Reproduktionszahlen, dass die Maxima später als das des jeweils passenden LCT-Modells angenommen werden.

Zur Erklärung des Verhaltens können wir die festgestellten Eigenschaften aus Abschnitt 5.2.3 anführen. Für die LCT-Modelle wurde für die Reproduktionszahl von zwei eine längere Verzögerung des Anstiegs der Neuansteckungen als für das ODE-Modell und im Anschluss hieran eine höhere Steigung festgestellt. Für $\mathcal{R}_{\text{eff}}(t) \approx 2$ werden die Hochpunkte deutlich später erreicht als für die höhere Reproduktionszahl. Die Hochpunkte sind so spät, dass die Verzögerungszeit einen geringeren Einfluss hat. Die höhere Steigung der LCT-Modelle führt vermutlich zu einem früheren Erreichen des Maximums. Für $\mathcal{R}_{\text{eff}}(t) \approx 4$ sind die Steigungen der Kurven der LCT-Modelle für die Neuansteckungen ähnlich. Die Verzögerungszeit gemeinsam mit der unterschiedlichen Höhe der Maxima führt wahrscheinlich zu dem späteren Höhepunkt der LCT-Modelle, je größer die Anzahl der Subkompartimente gewählt ist. Die IDE-Modelle haben für die täglichen Neuansteckungen vergleichbare Verzögerungszeiten wie das jeweilige LCT-Modell mit derselben Annahme bezüglich der Subkompartimente. Die geringere Steigung der Neuansteckungen nach der Verzögerung führt zu dem späteren Erreichen des Höhepunkts.

Unsere Simulationen bestätigen das in der Literatur beschriebene Ergebnis eines ausgeprägteren Höhepunkts der Epidemie bei höherer Wahl der Anzahl der Subkompartimente. Für die IDE-Modelle haben wir die Möglichkeit, realistischere Aufenthaltszeitverteilungen zu verwenden als für die LCT-Modelle, siehe Kapitel 5.1. Für diese Modelle wurden die Maxima deutlich geringer prognostiziert als für das passende LCT-Modell. Somit kann infrage gestellt werden, ob das Verhalten der LCT-Modelle realistisch ist. Wir haben zudem festgestellt, dass das Verhältnis der Zeitpunkte des Erreichens des Höhepunkts der Epidemie der verschiedenen Modelle von der effektiven Reproduktionszahl beziehungsweise der Kontaktrate ϕ abhängt. Dies könnte die gegensätzlichen Ergebnisse aus der Literatur erklären.

In [55] wird ein Schwerpunkt darauf gelegt, die Auslastung von Einrichtungen zur Intensivpflege zu prognostizieren. Hier wird festgehalten, dass die Wahrscheinlichkeit einer Überschreitung der Kapazität größer ist, wenn die Aufenthaltszeiten einer Erlang-Verteilung anstatt einer Exponentialverteilung folgen [55, S. 8]. Auch Blyuss et al. [9, S. 9] prognostizieren unter der Annahme der Erlang-Verteilung ein höheres Maximum der Individuen auf der Intensivstation. Anhand der Kurve für U in Abbildung 5.8 können wir dieses Ergebnis bestätigen. Die prognostizierte maximale Anzahl der Personen auf der Intensivstation ist umso höher, je größer die Anzahl der Subkompartimente für das LCT-Modell gewählt wird.

Durch die Simulation fiktiver Szenarien konnten wir einige Einblicke in das Verhalten der Modellarten gewinnen. Die Analysen tragen zu einem besseren Verständnis der Funktionsweise der Modelle bei. Zunächst haben wir festgestellt, dass eine unpassende Initialisierung eines LCT-Modells zu verfälschten Ergebnissen führen kann, sodass diese Thematik bei der Modellierung bedacht werden sollte. Im Anschluss haben wir die Reak-

tion der Modelle auf die Veränderung der Kontaktrate beziehungsweise der effektiven Reproduktionszahl verglichen. Die fehlende Verzögerungszeit beim ODE-Modell haben wir als unrealistisch identifiziert. Außerdem wurden Anzeichen für eine längere Laufzeit der IDE-Modelle gefunden. Zum Schluss haben wir die Prognosen bezüglich des Höhepunkts der Epidemie für verschiedene Modelle und Modellannahmen verglichen. Auch hier ergaben sich deutliche Unterschiede. Insgesamt konnten Indizien dafür gesammelt werden, dass die Modellannahmen des LCT-Modells sinnvoller sind als die des ODE-Modells. Das IDE-Modell ließ die passendsten Annahmen bezüglich der Parameter zu. Die Simulationsergebnisse des LCT-Modells unterscheiden sich teilweise noch deutlich von denen des IDE-Modells. Durch die Ergebnisse wird die Notwendigkeit unterstrichen, die Aufenthaltszeitverteilungen in den Krankheitszuständen genau zu bestimmen, vergleiche [55, S. 8]. Die Prognosen können stark von den Annahmen bezüglich der Verteilungen abhängen, wie bereits in Kapitel 3.6 behauptet wurde. Wenn nicht-pharmazeutische Maßnahmen auf der Basis verfälschter Modellprognosen beschlossen werden, könnten diese zu schwach sein, um die Epidemie zu kontrollieren oder unverhältnismäßig stark, was wirtschaftliche und soziale Auswirkungen haben könnte, vergleiche [93, S. 0624] und [55, S. 9].

5.3 Simulation realer Szenarien

Zuletzt erproben wir die Modelle anhand von realistischen Szenarien im Vergleich zu realen Daten der COVID-19 Pandemie. Zu beachten ist, dass wir nur erste Ergebnisse präsentieren, um Tendenzen bezüglich der Eignung der Modelle für Prognosen von realistischen Ereignissen zu geben. Um die Realität präziser abbilden zu können, könnten in zukünftigen Arbeiten einige Annahmen optimiert und Erweiterungen des Modells vorgenommen werden. Wie bereits erwähnt, beschränken wir uns auf das ODE- und das LCT-Modell, da es eine Herausforderung ist, die benötigten Übergänge für die Initialisierung des IDE-Modells aus gegebenen Daten zu bestimmen. Für das LCT-Modell verwenden wir, wie für den Vergleich mit den IDE-Modellen in den fiktiven Szenarien, nur die Annahmen von $n_z = 3$ und 10 Subkompartimenten für alle $z \in \mathcal{A}$. Die Implementierung der realen Szenarien ist in der Datei `cpp/examples/lct_secir_real_scenario.cpp` zu finden. Wir führen die Simulationen anhand von zwei Zeiträumen im Jahr 2020 durch, einen Zeitraum im Sommer sowie einen im Herbst. Die Zeiträume sind im Jahr 2020 gewählt, da die Parameter aus Kapitel 5.1 an die Veröffentlichung [60] angelehnt und die dortigen Parameter an Szenarien in diesem Jahr angepasst sind. Beispielsweise könnte die Transmissionswahrscheinlichkeit für einen anderen Zeitraum aufgrund einer anderen vorherrschenden Virusvariante angepasst werden, siehe [56, S. 2 & 7]. Zudem treffen wir mit den Zeiträumen eine epidemische Ausbreitung der Krankheit COVID-19. Die Initialisierungen der Modelle für die beiden Zeiträume werden, wie in Kapitel 4.3 beschrieben, mit den Methoden nach Koslow et al. [56, S2 Appendix] anhand von veröffentlichten Daten des RKIs bestimmt. Orientiert an den Ausführungen in [60, S. 12] nehmen wir für das Szenario im Sommer an, dass die Anzahl der Fälle so gering ist, dass die Anzahl der unentdeckten Fälle klein ist. Aus diesem Grund verwenden wir die Parameter $\xi_C = 0.7$ und $\xi_I = 0.2$, da viele Fälle detektiert werden und eine

Kontaktverfolgung stattfinden kann. Im Herbst sind die Fallzahlen höher, sodass auch der Anteil der unentdeckten Fälle größer ist. Entsprechend skalieren wir die Anzahl der bestätigten Fälle für dieses Szenario mit dem Faktor $d = 2$, vergleiche auch Kapitel 4.3. Darüber hinaus übernehmen wir die Parameter $\xi_C = 1$ und $\xi_I = 0.3$ aus Abschnitt 5.1. Wir vergleichen die Resultate der Simulationen mit den realen Daten bezüglich der täglichen Neuansteckungen, der Anzahl der Infizierten I und der Todesfälle D . Die realen Werte extrapolieren wir zu diesem Zweck konsistent zu den Überlegungen in Kapitel 4.3 für die Initialisierung des ODE-Modells nach Koslow et al. [56, S2 Appendix]. Für jeden Simulationstag t berechnen wir die Anzahl der Infizierten I für die realen Daten anhand von Gleichung (4.58). Die Todesfälle werden gemäß Gleichung (4.59) zeitlich verschoben. Für die Anzahl der neuen täglichen Ansteckungen nehmen wir ungeglättete Daten, das heißt wir berechnen nicht wie in Kapitel 4.3 einen gleitenden Durchschnitt mit sieben Tagen. Geglättete Daten würden bewirken, dass Veränderungen in den täglichen Neuansteckungen in den realen Daten unzureichend sichtbar sind. Angepasst an die Formel (4.61) für E berechnen wir die Anzahl der Individuen, die innerhalb des letzten Tages angesteckt wurden, für einen Tag t aus den realen Daten anhand der Formel

$$\frac{1}{\mu_I} \left(F_{\text{RKI}}(t + T_C + T_E) - F_{\text{RKI}}(t + T_C + T_E - 1) \right).$$

Die Notation $F_{\text{RKI}}(t)$ steht hier nun für die ungeglätteten kumulierten bestätigten Fälle, die vom RKI veröffentlicht werden. Auf diese Weise erhalten wir reale extrapolierte Daten, mit denen wir unsere Simulationsergebnisse vergleichen können.

Wie für die fiktiven Szenarien müssen wir eine passende Kontaktrate $\phi(t)$ bestimmen. Hierfür ist es notwendig, nicht-pharmazeutische Maßnahmen für die betrachteten Zeiträume zu implementieren, um Änderungen im Kontaktverhalten darzustellen. Dazu gehen wir abermals wie in der Publikation [60, S. 5f.] vor. Wir teilen die Kontakte in die Kategorien „Zuhause“, „Schule“, „Arbeit“ und „Andere“ ein. Für jede Kategorie definieren wir einen Basiswert der Kontaktrate für das Niveau der Kontakte vor der Pandemie und einen minimalen Wert für die Kontaktrate in einem strikten Lockdown. Wir fügen für die verschiedenen Kontaktraten zur Notation ϕ den Index B oder M für den Basiswert oder den minimalen Wert und für die Kategorien h für Zuhause (engl. *home*), s für Schule (*school*), w für Arbeit (*work*) und o für Andere (*other*) hinzu. Beispielsweise wird die minimale Kontaktrate für die Kategorie Schule demnach mit $\phi_{M,s}$ bezeichnet. Die Kontaktraten sind aus der zitierten Quelle übernommen, wobei altersaufgelöste Kontakte gemäß eines gewichteten Durchschnitts anhand dem Anteil der jeweiligen Altersgruppe an der Gesamtbevölkerung zu einem Wert zusammengefasst werden. Nicht-pharmazeutische Maßnahmen können anhand von Faktoren zur Reduktion der Kontakte eingeführt werden. Da Maßnahmen wie eine Maskenpflicht und eine Empfehlung für Homeoffice gleichzeitig wirken können, führen wir zwei verschiedene Ebenen für die Maßnahmen ein. Wir betrachten eine Hauptebene für Maßnahmen, die sich direkt auf die Anzahl der Kontakte auswirken und eine zweite Ebene für Vorsichtsmaßnahmen wie Abstand halten und Masken tragen [56, S. 9]. Zur Einführung einer Maßnahme passen wir die Faktoren $r_k^\ell \in [0, 1]$ für die Kategorie $k \in \{h, s, w, o\}$ und die Ebene $\ell \in \{1, 2\}$ an. Faktoren vom Wert Null entsprechen der Verwendung des Basiswerts für die Kontaktrate und für den

Wert Eins wird die minimale Kontaktrate genutzt. Schlussendlich erhalten wir für einen Zeitpunkt t die Kontaktrate

$$\phi(t) = \sum_{k \in \{h,s,w,o\}} \left(\phi_{B,k} - \left(1 - \prod_{\ell \in \{1,2\}} (1 - r_k^\ell(t)) \right) (\phi_{B,k} - \phi_{M,k}) \right).$$

Die festgelegten Werte für nicht-pharmazeutische Maßnahmen werden später für die Szenarien spezifiziert. Die genaue Ausgestaltung dieser Faktoren für bestimmte Zeiträume ist eine komplexe Aufgabe. Während der COVID-19 Pandemie wurden in Deutschland vielfältige nicht-pharmazeutische Maßnahmen ergriffen, um die Ausbreitung des Erregers einzudämmen. Die Maßnahmen unterscheiden sich in ihrer Ausprägung, beispielsweise in Form von Schulschließungen oder einer Maskenpflicht und in der Strenge der Maßnahme. Zudem wurden Maßnahmen lokal für Landkreise mit höheren Fallzahlen, für einzelne Bundesländer oder für ganz Deutschland implementiert. Entsprechend ist es wichtig zu untersuchen, inwieweit bestimmte politische Maßnahmen Veränderungen im Kontaktverhalten bewirken. Vergleiche für die vorangegangene Erläuterung auch [59, S. 10ff.]. Diese Komplexität erschwert eine exakte Bestimmung der Kontaktrate beziehungsweise der Faktoren zur Reduktion zu verschiedenen Zeitpunkten der Simulation. Mit diesen Fragestellungen beschäftigt sich die zukünftige Veröffentlichung [94]. Auch psychologische Effekte können für Veränderungen im Kontaktverhalten verantwortlich sein. Beispielsweise könnten niedrigere Infektionszahlen dazu führen, dass die Einhaltung der Maßnahmen nachlässt. Ängste und Sorgen können eine freiwillige Einschränkung von Kontakten auslösen, siehe hierzu [39, S. 678]. Entsprechend müssen Veränderungen in der Kontaktrate in der Realität nicht zwingend mit einer politischen Maßnahme zusammenhängen. Aufgrund der Schwierigkeit der Festlegung realistischer Kontakte skalieren wir die Kontaktrate zusätzlich so, dass die Anzahl der täglichen Neuansteckungen zu Beginn der Simulation den realen Daten entspricht. Als Referenzwert für die realen Daten verwenden wir hier einen extrapolierten Wert für die Anzahl der täglichen Neuansteckungen, der unter Verwendung des gleitenden Durchschnitts von sieben Tagen berechnet wird. So verhindern wir, dass die Kontaktrate mittels eines Ausreißers der extrapolierten täglichen Neuansteckungen skaliert wird. Wir beschreiben nun die Voraussetzungen für die konkreten Szenarien und betrachten die Simulationsergebnisse.

5.3.1 Simulation über einen Zeitraum im Herbst

Die Simulation eines Szenarios im Herbst beginnen wir zum Startdatum 01.10.2020 und wählen eine Simulationsdauer von 45 Tagen. Bei den Werten für die nicht-pharmazeutischen Maßnahmen orientieren wir uns an den Beschreibungen von Kühn et al. [60, S. 10ff.]. Aufgrund der Skalierung der Kontakte zur Reproduktion der täglichen Neuansteckungen zum Simulationsbeginn wirken die Maßnahmen sich in absoluten Werten allerdings nicht identisch zu der Publikation auf die Kontaktrate aus. Wir verändern die Werte und die Zeitpunkte der Maßnahmen an wenigen Stellen, um die realen Daten besser zu reproduzieren. Entsprechend bilden die Überlegungen aus der Veröffentlichung nur die Grundlage. Zu Beginn nehmen wir dieselben Maßnahmen an, wie in der Quelle für den Monat Oktober angegeben. Wenn Bereiche festgelegt sind, in denen die Faktoren

Parameter	01.10.2020	24.10.2020
r_h^1	$0.3(1 - \alpha_1) + 0.5\alpha_1$	0.5
r_w^1	$0.275(1 - \alpha_1) + 0.4\alpha_1$	$0.3(1 - \alpha_2) + 0.4\alpha_2$
r_o^1	$0.1(1 - \alpha_1) + 0.7\alpha_1$	0.7
r_s^1	$0.25\alpha_1$	$0.25\alpha_2$
r_h^2	$0.3(1 - \alpha_1) + 0.7\alpha_1$	$0.3(1 - \alpha_2) + 0.7\alpha_2$
r_w^2	$0.3(1 - \alpha_1) + 0.7\alpha_1$	$0.5(1 - \alpha_2) + 0.7\alpha_2$
r_o^2	$0.3(1 - \alpha_1) + 0.7\alpha_1$	$0.5(1 - \alpha_2) + 0.7\alpha_2$
r_s^2	$0.3(1 - \alpha_1) + 0.7\alpha_1$	0.7

Tabelle 5.5: Überblick über die implementierten nicht-pharmazeutischen Maßnahmen für die Simulation im Herbst. Die Werte für die nicht-pharmazeutischen Maßnahmen sind so zu verstehen, dass sie von dem angegebenen Datum bis zum nächsten spezifizierten Zeitpunkt oder für das letzte Datum bis zum Simulationsende gelten. Die Werte orientieren sich ungefähr an der Spezifikation in [60, S. 10ff.]. Siehe die Erläuterung im Text für die Definition von α_1 und α_2 .

r_k^ℓ liegen könnten, verwenden wir den Mittelwert. Lediglich die Schulferien lassen wir unberücksichtigt. In den Schulferien verlagern sich Kontakte von der Kategorie Schule in andere Kategorien. Zudem sind Schulferien nicht für alle Bundesländer einheitlich, sodass die Implementierung ohne Berücksichtigung einer örtlichen Auflösung erschwert wird. Zum 24.10.2020 nehmen wir eine Verschärfung der Maßnahmen an und verwenden die Werte aus der Publikation für den November. Wir beginnen somit eine Woche früher mit den stärkeren Maßnahmen, da sich in den täglichen Neuansteckungen zu diesem Zeitpunkt bereits eine Stagnation der realen Werte erkennen lässt.

In [60] wird zur Simulation der Ausbreitung von SARS-CoV-2 eine örtliche Auflösung nach Landkreisen implementiert. Es werden zusätzliche lokale, dynamische nicht-pharmazeutische Maßnahmen eingesetzt, falls die Inzidenz über einen bestimmten Wert ansteigt. In diesem Fall werden für 14 Tage härtere Maßnahmen als die Annahmen für November ergriffen. In [60, Fig. 12] werden für eine Simulation im Herbst 371 Überschreitungen dieser Inzidenz in 45 Tagen Simulationszeit angegeben. Wir berechnen hieraus einen Anteil der Landkreise, für den durchschnittlich an einem Tag die harten Maßnahmen gelten. Diesen Anteil inkludieren wir in die Festlegung der Faktoren zur Reduktion für Deutschland, um die lokalen strikteren Maßnahmen nicht unbeachtet zu lassen. In [60, Fig. 12] ist zu erkennen, dass zu Beginn der Simulation der Anteil der Landkreise mit hoher Inzidenz geringer ist und mit fortgeschrittener Simulationsdauer ansteigt. Wir nehmen entsprechend an, dass anfangs nur die Hälfte des berechneten Anteils die strikteren Maßnahmen anwenden. Diesen Anteil bezeichnen wir mit $\alpha_1 \in [0, 1]$. Als Reaktion auf die gestiegene Anzahl der Landkreise mit hoher Inzidenz bei einer fortgeschrittenen Simulationsdauer gelten die strikten Maßnahmen ab dem 24.10.2020 dann für einen höheren Anteil $\alpha_2 \in [0, 1]$. Dieser Anteil entspricht dem 1.5-fachen des anfangs ausgerechneten Werts. Für den Faktor r_s^2 nehmen wir ab dem 24.10.2020 für alle Landkreise den Wert der strikteren Maßnahmen an. Dies könnte beispielsweise durch einen psychologischen Effekt erklärt werden, da

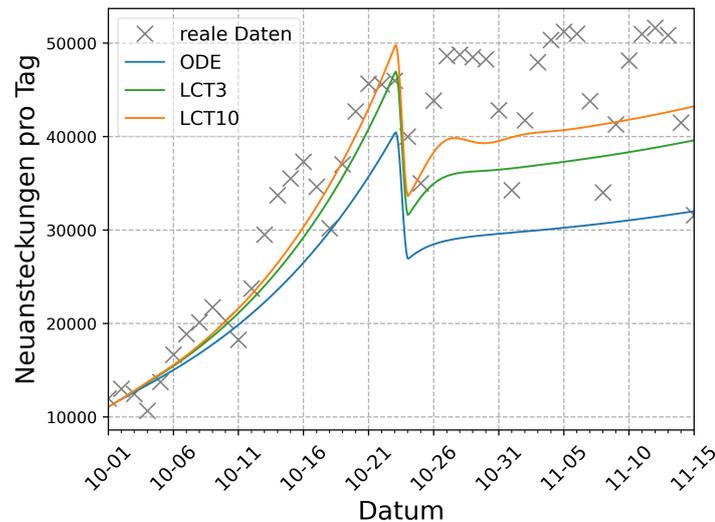


Abbildung 5.9: Tägliche Neuansteckungen mit dem Startdatum 01.10.2020.

Vergleich von extrapolierten realen Daten der täglichen Neuansteckungen mit den Simulationsergebnissen des ODE-SECIR-Modells und von LCT-Modellen unter den Annahmen von 3 und 10 Subkompartimenten.

Schulen eine hohe Aufmerksamkeit in Bezug auf die Ausbreitung von COVID-19 erhalten haben könnten. Die Ergebnisse unserer Überlegungen in Form der verwendeten Faktoren r_k^ℓ für $k \in \{h, s, w, o\}$ und $\ell \in \{1, 2\}$ zur Reduktion der Kontakttrate sind in Tabelle 5.5 angegeben. Zur Inklusion der lokalen, dynamischen nicht-pharmazeutischen Maßnahmen wäre eine örtliche Auflösung angebracht.

Die Simulationsergebnisse für die täglichen neuen Ansteckungen für die drei verwendeten Modelle im Vergleich mit den extrapolierten realen Daten nach dem RKI sind in Abbildung 5.9 dargestellt. Wir können zum Zeitpunkt der Verschärfung der nicht-pharmazeutischen Maßnahmen am 24.10.2020 einen deutlichen Abfall der prognostizierten täglichen Neuansteckungen erkennen. Zu diesem Zeitpunkt zeigen auch die realen Daten einen Rückgang. Der Trend der extrapolierten realen Daten wird von den Prognosen der Modelle zufriedenstellend widerspiegelt. Wie in Abschnitt 5.2.3 fallen die Steigungen der Prognosen der LCT-Modelle extremer aus, was hier realistischer wirkt. Beim ODE-Modell ist das Niveau der Prognosen nach der Einsetzung der strikteren Maßnahmen zu niedrig.

In Abbildung 5.10 sind die Simulationsergebnisse zu den Individuen in der Gruppe I und D im Vergleich mit realen Daten dargestellt. Bei der Anzahl der symptomatisch Infizierten $I(t)$ zeigt sich, dass das ODE-Modell die größten Abweichungen zu den realen Daten aufweist. Die Steigung beim ODE-Modell verringert sich aufgrund der festgestellten fehlenden Verzögerung aus Kapitel 5.2 unmittelbar nach dem 24.10.2020. Sowohl bei den LCT-Modellen als auch bei den realen Daten ist eine größere Verzögerung erkennbar. Die anfängliche Steigung der realen Daten wird von allen Modellen unterschätzt. Das LCT-Modell mit der Annahme von zehn Subkompartimenten weist hier die kleinsten und das ODE-Modell die größten Abweichungen auf. Der Trend der Anzahl der Todesfälle

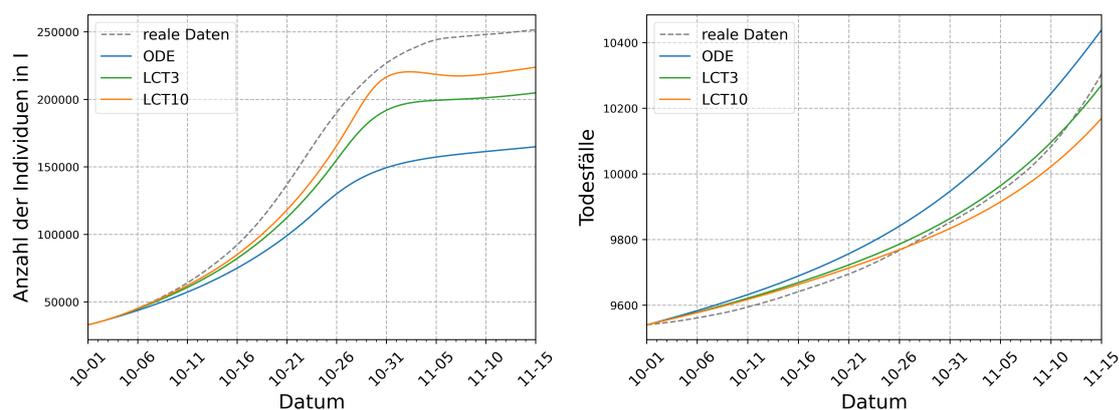


Abbildung 5.10: Infizierte und Todesfälle mit dem Startdatum 01.10.2020.

Vergleich von extrapolierten realen Daten zu den symptomatisch Infektiösen (links) und zu der Anzahl der Todesfälle (rechts) mit den Simulationsergebnissen des ODE-SECIR-Modells und von LCT-Modellen unter den Annahmen von 3 und 10 Subkompartimenten.

wird von allen Modellen relativ genau getroffen. Das Modell LCT10 unterschätzt die echten Zahlen zum Simulationsende hin, das ODE-Modell überschätzt sie und das LCT-Modell mit 3 Subkompartimenten prognostiziert die Zahlen zum Ende der Simulation am zutreffendsten.

Insgesamt treffen alle Modelle den Trend der realen Daten akzeptabel und die Größenordnung der Prognosen entspricht den realen Zahlen. Die nicht-pharmazeutischen Maßnahmen wurden zu diesem Zweck im Vergleich zu den Ausführungen in [60] nur leicht angepasst und basieren auf sinnvollen Annahmen. Das LCT-Modell mit der Annahme von den meisten Subkompartimenten stellt sich als am nächsten an der Realität heraus. Das ODE-Modell erzielt keine zufriedenstellenden Ergebnisse.

5.3.2 Simulation über einen Zeitraum im Sommer

Nun betrachten wir die Ergebnisse eines zweiten Szenarios über einen Zeitraum im Sommer. Die Simulation beginnen wir am 01.06.2020 und führen sie abermals für eine Dauer von 45 Tagen durch. In diesem Zeitfenster überschreiten fast keine Landkreise den Schwellenwert zur Einsetzung härterer Maßnahmen, vergleiche [60, Fig. 6]. Entsprechend lassen wir den Effekt unberücksichtigt. Zum 01.06.2020 implementieren wir fast dieselben nicht-pharmazeutischen Maßnahmen wie in [60] beschrieben. Den Wert r_o^1 nehmen wir geringfügig höher an. Die Beibehaltung dieser Maßnahmen über den Simulationszeitraum führt nicht zu denselben Größenordnungen der Prognosen im Vergleich zu den realen Daten. Die Schwankungen in den realen Daten sind mit der Beibehaltung der Kontaktrate nicht reproduzierbar. Entsprechend führen wir zu verschiedenen Zeitpunkten Maßnahmen ein oder heben sie auf, zu denen Veränderungen in den realen Daten bezüglich der täglichen neuen Ansteckungen erkennbar sind. Die Werte für die Faktoren werden so gewählt, dass sie die Tendenzen in den realen Daten zutreffend abbilden. Die Faktoren

Parameter	01.06.2020	02.06.2020	14.06.2020	03.07.2020
r_h^1	0.1	0	0.1	0
r_w^1	0.275	0.025	0.275	0.025
r_o^1	0.2	0	0.25	0.1
r_s^1	0.5	0.25	0.35	0.2
r_h^2, r_o^2	0.1	0	0.25	0
r_w^2, r_s^2	0.1	0	0.25	0.1

Tabelle 5.6: Überblick über die implementierten nicht-pharmazeutischen Maßnahmen für die Simulation im Sommer. Die Werte für die nicht-pharmazeutischen Maßnahmen sind so zu verstehen, dass sie von dem angegebenen Datum bis zum nächsten spezifizierten Zeitpunkt oder für das letzte Datum bis zum Simulationsende gelten. Die Werte für das erste Datum basieren auf der Spezifikation in [60, S. 10ff.] und die restlichen Werte sind so angepasst, dass sie die Tendenzen der realen Daten der täglichen Neuansteckungen widerspiegeln.

wurden per Augenmaß angepasst, anstatt eine computergestützte Schätzung vorzunehmen. Die Modellentwicklung und Erprobung ist das Hauptziel der vorliegenden Arbeit und Parameterfitting stellt eine separate Aufgabe dar. Die verwendeten Werte für die nicht-pharmazeutischen Maßnahmen mit den passenden angenommenen Zeiträumen sind in Tabelle 5.6 dargelegt. Diese Größen für die Faktoren zur Aufhebung oder Einsetzung von Maßnahmen sind nicht unrealistisch, jedoch korrelieren sie auch nicht zwangsläufig mit politischen Beschlüssen. Die beschriebenen psychologischen Effekte könnten die Schwankungen in den Daten erklären.

In Abbildung 5.11 sind die Prognosen für die täglichen Neuansteckungen verglichen mit den realen Daten für den betrachteten Zeitraum abgebildet. Die Simulationsergebnisse verhalten sich sprunghaft zu den Zeitpunkten, zu denen die Faktoren verändert werden. Die Unterschiede in der Qualität der Modellvorhersagen lassen sich anhand der Abbildung nicht eindeutig beurteilen.

Die Anzahl der symptomatisch Infizierten I und die Anzahl der Todesfälle D sind in Abbildung 5.12 dargestellt. In den ersten Simulationstagen ist deutlich zu erkennen, dass die Steigung der Anzahl der Individuen in I des LCT-Modells mit zehn Subkompartimenten am nächsten an die realen Daten herankommt. Im Anschluss kann dieses Modell die deutlichen Schwankungen der realen Daten am genauesten reproduzieren. Durch die extremeren Steigungen, die in den Abschnitten 5.2.2 und 5.2.3 festgestellt wurden, kann das Modell LCT10 die großen Amplituden der Wellenform der realen Daten zutreffender nachahmen. Wir sehen hier deutlich, dass das ODE-Modell im Vergleich zu beiden LCT-Modellen unrealistischer ist. Die Anzahl der Todesfälle wird von allen Modellen deutlich unterschätzt und es können keine akzeptablen Ergebnisse erzielt werden. Dies spricht möglicherweise für die Einführung einer Altersauflösung, da beispielsweise eine deutlich höhere Wahrscheinlichkeit für ältere Menschen, infolge von COVID-19 zu sterben, in unseren Modellen nicht berücksichtigt werden kann, vergleiche dazu [60, Table 2]. Für ein ähnliches Szenario wird in dieser Publikation unter Verwendung einer Altersauflösung

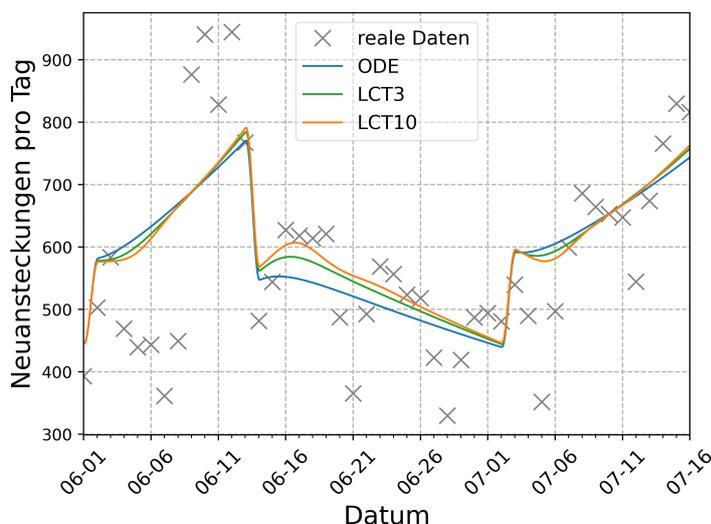


Abbildung 5.11: Tägliche Neuansteckungen mit dem Startdatum 01.06.2020.

Vergleich von extrapolierten realen Daten der täglichen Neuansteckungen mit den Simulationsergebnissen des ODE-SECIR-Modells und von LCT-Modellen unter den Annahmen von 3 und 10 Subkompartimenten.

für die Todesfälle ein genaueres Ergebnis erzielt.

In diesem Szenario mussten wir die Annahmen bezüglich der politischen Maßnahmen aus [60] erheblich erweitern, um die realen Daten ausreichend genau nachzubilden. Dies impliziert, dass Kontaktmuster exakt bestimmt werden müssten, um die Tendenzen der realen Daten zu erklären und genaue Prognosen zu erzielen. Dazu könnte eine Untersuchung der Auswirkung und der Zeiträume von vielfältigen Ausprägungen der politischen Maßnahmen beitragen, was in [94] angestrebt wird. Werden zu den realen Daten passende Werte r_k^ℓ angenommen, ist am Verlauf von $I(t)$ deutlich zu erkennen, dass das LCT-Modell mit zehn Subkompartimenten realistischere Prognosen ermöglicht als das ODE-Modell. Gemeinsam mit den Feststellungen aus dem Szenario im Herbst spricht dies dafür, dass die Annahmen der LCT-Modelle realistischer sind als die des ODE-Modells. Das unterschiedliche Verhalten, das in den fiktiven Szenarien in Kapitel 5.2 gefunden wurde, erlaubt präzisere Vorhersagen der LCT-Modelle. Eine örtliche Auflösung sowie eine Altersauflösung könnten dazu beitragen, die Simulationsergebnisse realistischer zu gestalten.

Wir blicken nun zusammenfassend auf die erlangten Einsichten anhand der Anwendung auf das Beispiel der COVID-19 Pandemie im aktuellen Kapitel zurück. Wir konnten einige Behauptungen aus Kapitel 3.6 bezüglich der Anpassungsfähigkeit und der Parameterwahl untersuchen. Zunächst haben wir festgestellt, dass die Parameter unterschiedlich komplex zu bestimmen sind. Für die epidemiologischen Parameter für das ODE-Modell sind Werte in der Literatur zu finden. Das LCT-Modell wurde anhand von verschiedenen Anzahlen für die Subkompartimente getestet. Die Aufenthaltszeitverteilungen für das IDE-Modell konnten wir nur unvollständig bestimmen, sodass wir zum Teil auf die

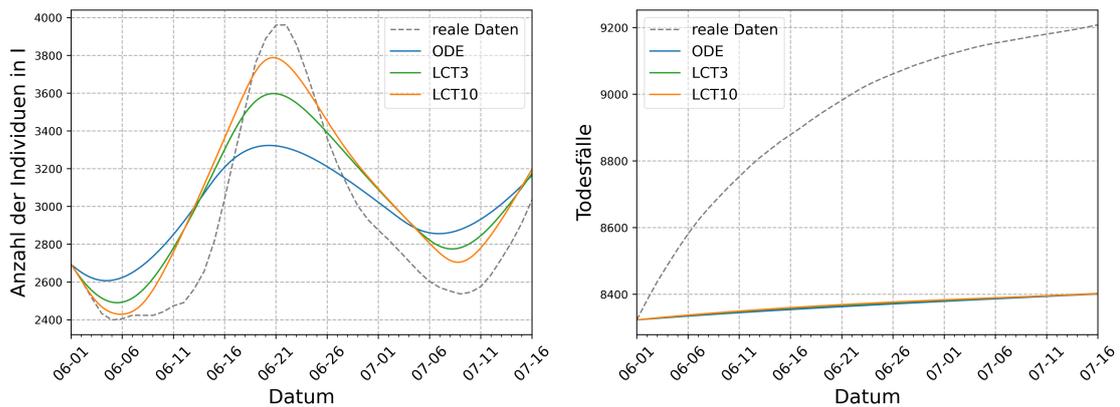


Abbildung 5.12: Infizierte und Todesfälle mit dem Startdatum 01.06.2020.

Vergleich von extrapolierten realen Daten zu den symptomatisch Infektiösen (links) und der Anzahl der Todesfälle (rechts) mit den Simulationsergebnissen des ODE-SECIR-Modells und von LCT-Modellen unter den Annahmen von 3 und 10 Subkompartimenten.

Modellannahmen des LCT-Modells zurückgegriffen haben. Dennoch erlaubt das Modell die realistischste Parameterwahl. Anhand von fiktiven Szenarien haben wir herausgefunden, dass die Initialisierung des LCT-Modells ein essentielles Thema ist. Um Verzerrungen zu vermeiden, sollte nicht auf vereinfachende Annahmen zurückgegriffen werden. Zudem haben wir bemerkt, dass sich die Simulationsergebnisse von verschiedenen Modellen erheblich unterscheiden können. Die Annahmen bezüglich der Aufenthaltszeitverteilung führen zu quantitativen und qualitativen Unterschieden in den Prognosen. Die Modellarten weisen verschiedene Reaktionen auf die Veränderung der Kontaktrate beziehungsweise der effektiven Reproduktionszahl auf. Besonders die fehlende Verzögerungszeit des ODE-Modells konnte als unrealistisch identifiziert werden. Für das IDE-Modell war auffällig, dass, je nach den getroffenen Modellannahmen, eine erhöhte Anzahl von Zeitschritten zu erwarten ist. Dies ist ein Hinweis auf eine längere benötigte Rechenzeit. Die anschließenden realen Szenarien legen nahe, dass sich die LCT-Modelle mit verschiedenen Annahmen bezüglich der Anzahl der Subkompartimente signifikant besser zur Reproduktion der realen Daten eignen als das ODE-Modell. Entsprechend scheint die Anpassungsfähigkeit des LCT-Modells realistischere Modellannahmen zuzulassen, als die des ODE-Modells. Das IDE-Modell lässt viele Freiheiten, die hier wegen der Bestimmung der Anfangsbedingungen nicht umfassend ausgenutzt wurden.

6 Fazit und Ausblick

In dieser Arbeit wurde der Linear Chain Trick im Kontext der epidemiologischen Modellierung eingeführt. Die Funktionsweise von LCT-Modellen sowie deren Beziehung zu einem ODE- und einem IDE-Modell haben wir eingehend analysiert. Die grundlegende Idee des Linear Chain Tricks besteht darin, eine Erlang-verteilte Gesamtaufenthaltszeit in einem Zustand durch eine Kette von Subkompartimenten mit linearen Übergängen zu erzielen. Der Trick basiert auf der Darstellung einer Erlang-Verteilung als Summe von Exponentialverteilungen. Durch die detaillierte Analyse des Linear Chain Tricks für die Gruppe I konnten wir das Prinzip auf weitere Kompartimente übertragen. Sämtliche relevante Konzepte wurden behandelt, um das LCT-Modell für praktische Simulationen anwenden zu können. Die Ergebnisse von einigen Simulationen haben wir mit anderen Modellprognosen und realen Daten zur COVID-19 Pandemie in Deutschland verglichen.

Anhand der Ausführungen wurden verschiedene Aspekte zur Klärung der Frage deutlich, inwieweit das LCT-Modell die Vorzüge des ODE- und des IDE-Modells miteinander vereinen kann. Ein ODE-Modell stellt einen Spezialfall eines LCT-Modells dar, das sich unter einer bestimmten Parametrisierung aus einem IDE-Modell ableitet. Das IDE-Modell erlaubt allgemeine Aufenthaltsverteilungen, während das LCT-Modell Erlang-Verteilungen verwendet und das ODE-Modell die Wahl auf Exponentialverteilungen beschränkt. Die Modelle bieten einen unterschiedlichen Grad an Flexibilität und Anpassungsfähigkeit. Wir haben festgestellt, dass exponentialverteilte Aufenthaltszeiten laut diverser Autoren für die meisten Infektionskrankheiten aus epidemiologischer Perspektive unrealistisch sind. Eine genauere Approximation der realen Aufenthaltszeitverteilung ist mithilfe einer Erlang-Verteilung möglich. Anhand von Simulationen verschiedener Szenarien in Kapitel 5 konnte die Aussage exemplarisch anhand des Erregers SARS-CoV-2 bestätigt werden. Die Flexibilität des LCT-Modells erwies sich als ausreichend, um ein realistischeres Verhalten zu erzeugen, selbst wenn die Anzahlen der Subkompartimente nicht auf epidemiologischen Daten basieren. Dies könnte auf die Eigenschaft der Gedächtnislosigkeit der ODE-Modelle zurückzuführen sein. Im LCT-Modell darf die Entwicklung der Modellgrößen von der Vergangenheit abhängen, anstatt nur von gegenwärtigen Werten. Das IDE-Modell zeichnet sich durch eine hohe Flexibilität aus, die eine adäquate Abbildung der Realität ermöglicht. Jedoch wird diese Anpassungsfähigkeit durch die eingeschränkte Verfügbarkeit von Informationen über reale Prozesse relativiert. Die Datenverfügbarkeit ist ein so limitierender Faktor, dass das IDE-Modell nicht zur Simulation realer Szenarien verwendet wurde und für die Parameter Annahmen des LCT-Modells übernommen wurden. Darüber hinaus verdeutlichten die Simulationen, dass die zugrunde gelegten Verteilungsannahmen die Prognosen entscheidend beeinflussen können. Die gebotene Flexibilität eines Modells ist ausschlaggebend für die Qualität der Ergebnisse. Fehlerhafte Prognosen können bei der Bewertung politischer Maßnahmen zur

Eindämmung des Infektionsgeschehens den Beschluss von unangemessenen Maßnahmen zur Folge haben. Dies kann entweder zu einer unkontrollierten Ausbreitung oder zu übermäßig restriktiven Maßnahmen mit Auswirkungen auf Wirtschaft und Gesellschaft führen.

Hinsichtlich der Praktikabilität zeichnete sich ein entgegengesetztes Bild ab, wobei sich das ODE-Modell als vorteilhaft zeigte. Diesbezüglich wurden in sämtlichen Kapiteln einzelne Aspekte deutlich. Die Formulierung, das Verständnis und die Analyse von ODE-Modellen erwiesen sich als verhältnismäßig unkompliziert. Zur Lösung von gewöhnlichen Differentialgleichungssystemen stehen Softwarebibliotheken zur Verfügung, sodass die Modelle bequem in Software umgesetzt werden können. Die Festlegung von Anfangswerten und benötigten Parametern war unproblematisch. LCT-Modelle basieren ebenfalls auf einem gewöhnlichen Differentialgleichungssystem, jedoch von höherer Dimension als das entsprechende ODE-Modell. Daher sind LCT-Modelle in Bezug auf die genannten Punkte komplexer, bleiben aber zugänglich und anwendbar. Insbesondere können zur Lösung dieser Modellgleichungen ebenso bereits implementierte numerische Löser herangezogen werden. Im Kontrast dazu erwiesen sich Modelle auf der Basis von Integro-Differentialgleichungen als anspruchsvoll, sowohl in der Herleitung als auch in der Anwendung, unter anderem aufgrund des numerischen Lösungsschemas. Die IDE-Modelle zeigten sich in jeder Hinsicht als weniger anwendungsfreundlich, was die Flexibilität des Modells einschränken kann.

Obwohl einige Argumente diskutiert wurden, die diese Feststellungen relativieren, bietet das LCT-Modell hinsichtlich der Anpassungsfähigkeit und der Anwendbarkeit stets einen Mittelweg zwischen dem ODE- und dem IDE-Modell. Der Modellansatz basierend auf dem Linear Chain Trick kombiniert die Vorteile des ODE- und des IDE-Modells und fungiert als Kompromiss zwischen den beiden Typen. Die Fragestellung der vorliegenden Arbeit kann insofern beantwortet werden. Durch den Einsatz eines LCT-Modells anstelle eines ODE-Modells lässt sich mit begrenztem Aufwand eine unrealistische Annahme gegen eine realistischere austauschen. Die Auswahl des geeigneten Modells sollte letztendlich individuell auf die betrachtete Krankheit abgestimmt werden. Unter Berücksichtigung des verfügbaren Wissens über die Infektionskrankheit und vor dem Hintergrund der analytischen und numerischen Einfachheit sollte eine bewusste, begründete Entscheidung getroffen werden, vergleiche [31, S. 1535] und [7, S. 17]. Diese Arbeit hat die Grundlage für eine solche Entscheidungsfindung gelegt.

Ausblick

Die Ergebnisse der vorliegenden Masterarbeit regen dazu an, sich weitergehend mit der Thematik des Linear Chain Tricks auseinanderzusetzen. Zuerst sollten die Laufzeiten der Simulationen der verschiedenen Modelle gegenübergestellt werden, um den Aspekt des Aufwands fundiert bewerten zu können. Einige Erweiterungen für das LCT-Modell wurden in Kapitel 3.5 erörtert, wie beispielsweise die Integration eines Immunitätsverlusts. Eine bedeutende Erweiterung läge in der Annahme einer anderen Aufenthaltszeitverteilung, welche die Erlang-Verteilung als Spezialfall einschließt, für die sich das Modell weiterhin als gewöhnliches Differentialgleichungssystem formulieren lässt. Naheliegend ist die Ein-

beziehung von Aspekten, die für das ODE-Modell in der Software MEmilio [61] inkludiert sind, wie beispielsweise Impfungen. Während der Diskussion realistischer Szenarien wurde deutlich, dass eine räumlich aufgelöste Simulation mit Mobilität und die Differenzierung nach Altersgruppen zu realistischeren Ergebnissen führen könnten. Auf diese Weise müsste für lokale nicht-pharmazeutische Maßnahmen nicht länger ein Durchschnittsverhalten über Landkreise verwendet werden. Um Unsicherheiten in der Bestimmung epidemiologischer Parameter zu berücksichtigen, wäre die Spezifikation von Verteilungen für die betreffenden Größen und die Durchführung mehrerer Simulationsläufe mit jeweils gemäß der Verteilung bestimmten Parametern eine mögliche Vorgehensweise, vergleiche [56, 59]. Sofern alle Aspekte aus MEmilio [61] effizient auf den LCT-Ansatz übertragen werden können, könnte das Modell das ODE-Modell gänzlich ersetzen. Zur Simulation mit exponentialverteilten Annahmen, ist dann jeweils nur ein Subkompartiment anzunehmen.

Mit dem Fortschreiten der Verbreitung von COVID-19 ist ein endemisches Modell mit Geburts- und Sterbeprozessen zur Vorhersage angemessen [49, S. 5379]. Ein derartiges LCT-Modell findet sich in [11, Kapitel 3.6.2]. Für diese Modelle können weitere interessante Eigenschaften analysiert werden. Einige Autoren untersuchen den Einfluss der Wahl der Aufenthaltszeitverteilung auf die Stabilität des endemischen Gleichgewichtszustands, siehe [5, 63].

Die Grundlage für weitere Literaturrecherchen wurde in den Kapiteln 2 und 3 gelegt, indem gängige Benennungen für die Modellansätze aufgeführt sowie grundlegende Quellen herausgearbeitet wurden. Die Funktionsweise des Linear Chain Tricks, die Ergebnisse aus Kapitel 3 und einige praktische Überlegungen aus Kapitel 4 sind nicht auf die Epidemiologie beschränkt und auf andere Anwendungsgebiete übertragbar. Infolgedessen sind die gewonnenen Erkenntnisse von interdisziplinärer Bedeutung.

Literatur

- [1] K. Ahnert und M. Mulansky, *Odeint, solving ODEs in C++ – overview*, 2009 – 2015, zu finden unter https://www.boost.org/doc/libs/1_83_0/libs/numeric/odeint/doc/html/boost_numeric_odeint/getting_started/overview.html (abgerufen am 05.09.2023).
- [2] K. Ahnert und M. Mulansky, *Runge-Kutta Cash-Karp 5(4) Methode aus der Bibliothek Odeint*, 2009 – 2015, zu finden unter https://www.boost.org/doc/libs/1_83_0/libs/numeric/odeint/doc/html/boost_numeric_odeint/runge_kutta_cash_karp54.html (abgerufen am 26.06.2023).
- [3] D. Anderson und R. Watson, „On the Spread of a Disease with Gamma Distributed Latent and Infectious Periods“, *Biometrika*, Jg. 67, Nr. 1, S. 191–198, Apr. 1980. DOI: 10.2307/2335333.
- [4] R. Anderson, S. Blythe, G. Medley und A. Johnson, „Is it possible to predict the minimum size of the acquired immunodeficiency syndrome (AIDS) epidemic in the United Kingdom?“, *The Lancet*, Jg. 329, Nr. 8541, S. 1073–1075, Mai 1987. DOI: 10.1016/S0140-6736(87)90493-4.
- [5] A. Andò, D. Breda und G. Gava, „How fast is the linear chain trick? A rigorous analysis in the context of behavioral epidemiology“, *Mathematical Biosciences and Engineering*, Jg. 17, Nr. 5, S. 5059–5084, Juli 2020. DOI: 10.3934/mbe.2020273.
- [6] M. Bladt, „A Review on Phase-type Distributions and their Use in Risk Theory“, *ASTIN Bulletin: The Journal of the IAA*, Jg. 35, Nr. 1, S. 145–161, Mai 2005. DOI: 10.2143/AST.35.1.583170.
- [7] S. P. Blythe und R. M. Anderson, „Distributed Incubation and Infectious Periods in Models of the Transmission Dynamics of the Human Immunodeficiency Virus (HIV)“, *Mathematical Medicine and Biology: A Journal of the IMA*, Jg. 5, Nr. 1, S. 1–19, März 1988. DOI: 10.1093/imammb/5.1.1.
- [8] S. P. Blythe, R. M. Nisbet und W. S. C. Gurney, „The dynamics of population models with distributed maturation periods“, *Theoretical Population Biology*, Jg. 25, Nr. 3, S. 289–311, Juni 1984. DOI: 10.1016/0040-5809(84)90011-X.
- [9] K. Blyuss und Y. Kyrychko, „Effects of latency and age structure on the dynamics and containment of COVID-19“, *Journal of Theoretical Biology*, Jg. 513, S. 1–10, Jan. 2021. DOI: 10.1016/j.jtbi.2021.110587.
- [10] F. Brauer, „Mathematical epidemiology: Past, present, and future“, *Infectious Disease Modelling*, Jg. 2, Nr. 2, S. 113–127, Mai 2017. DOI: 10.1016/j.idm.2017.02.001.

- [11] F. Brauer, C. Castillo-Chavez und Z. Feng, *Mathematical Models in Epidemiology*, 1. Aufl. Springer New York, Okt. 2019. DOI: 10.1007/978-1-4939-9828-9.
- [12] D. Breda, O. Diekmann, W. F. de Graaf, A. Pugliese und R. Vermiglio, „On the formulation of epidemic models (an appraisal of Kermack and McKendrick)“, *Journal of Biological Dynamics*, Jg. 6, Nr. sup2, S. 103–117, Sep. 2012. DOI: 10.1080/17513758.2012.716454.
- [13] J. Cash und A. Karp, „A Variable Order Runge-Kutta Method for Value Problems with Rapidly Varying Right-hand Sides“, *ACM Trans. Math. Softw.*, Jg. 16, Nr. 3, S. 201–222, Sep. 1990. DOI: 10.1145/79505.79507.
- [14] T. Cassidy, „Distributed Delay Differential Equation Representations of Cyclic Differential Equations“, *SIAM Journal on Applied Mathematics*, Jg. 81, Nr. 4, S. 1742–1766, Aug. 2021. DOI: 10.1137/20M1351606.
- [15] T. Cassidy, M. Craig und A. R. Humphries, „Equivalences between age structured models and state dependent distributed delay differential equations“, *Mathematical Biosciences and Engineering*, Jg. 16, Nr. 5, S. 5419–5450, Juni 2019. DOI: 10.3934/mbe.2019270.
- [16] T. Cassidy, P. Gillich, A. R. Humphries und C. H. van Dorp, „Numerical methods and hypoexponential approximations for gamma distributed delay differential equations“, *IMA Journal of Applied Mathematics*, Jg. 87, Nr. 6, S. 1043–1089, Dez. 2022. DOI: 10.1093/imamat/hxac027.
- [17] D. Champredon, J. Dushoff und D. J. D. Earn, „Equivalence of the Erlang-Distributed SEIR Epidemic Model and the Renewal Equation“, *SIAM Journal on Applied Mathematics*, Jg. 78, Nr. 6, S. 3258–3278, Dez. 2018. DOI: 10.1137/18M1186411.
- [18] D. Champredon, M. Li, B. M. Bolker und J. Dushoff, „Two approaches to forecast Ebola synthetic epidemics“, *Epidemics*, Jg. 22, S. 36–42, Feb. 2018, The RAPIDD Ebola Forecasting Challenge. DOI: 10.1016/j.epidem.2017.02.011.
- [19] G. Ciaravino, A. García-Saenz, S. Cabras, A. Allepuz, J. Casal, I. García-Bocanegra, A. De Koeijer, S. Gubbins, J. Sáez, D. Cano-Terriza und S. Napp, „Assessing the variability in transmission of bovine tuberculosis within Spanish cattle herds“, *Epidemics*, Jg. 23, S. 110–120, Juni 2018. DOI: 10.1016/j.epidem.2018.01.003.
- [20] A. Cintrón-Arias, C. Castillo-Chávez, L. M. A. Bettencourt, A. L. Lloyd und H. T. Banks, „The estimation of the effective reproductive number from disease outbreak data“, *Mathematical Biosciences and Engineering*, Jg. 6, Nr. 2, S. 261–282, Apr. 2009. DOI: 10.3934/mbe.2009.6.261.
- [21] L. Contento, N. Castelletti, E. Raimúndez, R. Le Gleut, Y. Schälte, P. Stapor, L. C. Hinske, M. Hoelscher, A. Wieser, K. Radon, C. Fuchs und J. Hasenauer, „Integrative modelling of reported case numbers and seroprevalence reveals time-dependent test efficiency and infectious contacts“, *Epidemics*, Jg. 43, S. 1–13, März 2023. DOI: 10.1016/j.epidem.2023.100681.

- [22] A. d’Onofrio, „Mixed pulse vaccination strategy in epidemic model with realistically distributed infectious and latent times“, *Applied Mathematics and Computation*, Jg. 151, Nr. 1, S. 181–187, März 2004. DOI: 10.1016/S0096-3003(03)00331-X.
- [23] B. Dawes, D. Abrahams und R. Rivera, *Boost C++ Library*, zu finden unter <https://www.boost.org> (abgerufen am 26.06.2023); siehe auch: B. Dawes, „Proposal for a C++ Library Repository Web Site“, 1998. <https://www.boost.org/users/proposal.pdf> (abgerufen am 26.06.2023).
- [24] T. Dey, J. Lee, S. Chakraborty, J. Chandra, A. Bhaskar, K. Zhang, A. Bhaskar und F. Dominici, „Lag time between state-level policy interventions and change points in COVID-19 outcomes in the United States“, *Patterns*, Jg. 2, Nr. 8, S. 1–15, Aug. 2021. DOI: 10.1016/j.patter.2021.100306.
- [25] O. Diekmann, M. Gyllenberg und J. A. J. Metz, „Finite Dimensional State Representation of Linear and Nonlinear Delay Systems“, *Journal of Dynamics and Differential Equations*, Jg. 30, Nr. 4, S. 1439–1467, Dez. 2018. DOI: 10.1007/s10884-017-9611-5.
- [26] O. Diekmann, M. Gyllenberg und J. A. J. Metz, „On models of physiologically structured populations and their reduction to ordinary differential equations“, *Journal of Mathematical Biology*, Jg. 80, Nr. 1, S. 189–204, Jan. 2020. DOI: 10.1007/s00285-019-01431-7.
- [27] O. Diekmann, H. Heesterbeek und T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton: Princeton University Press, 2013. DOI: 10.1515/9781400845620.
- [28] O. Diekmann, J. A. P. Heesterbeek und J. A. Metz, „The legacy of Kermack and McKendrick“, *Epidemic models: their structure and relation to data*, S. 95–115, Jan. 1995, https://www.researchgate.net/publication/27691377_The_Legacy_of_Kermack_and_McKendrick (abgerufen am 21.06.2023).
- [29] D. Fargue, „Reductibilit e des systemes hereditaires“, *International Journal of Non-Linear Mechanics*, Jg. 9, Nr. 5, S. 331–338, Okt. 1974. DOI: 10.1016/0020-7462(74)90018-3.
- [30] Z. Feng und H. R. Thieme, „Endemic Models with Arbitrarily Distributed Periods of Infection I: Fundamental Properties of the Model“, *SIAM Journal on Applied Mathematics*, Jg. 61, Nr. 3, S. 803–833, Sep. 2000. DOI: 10.1137/S0036139998347834.
- [31] Z. Feng, D. Xu und H. Zhao, „Epidemiological Models with Non-Exponentially Distributed Disease Stages and Applications to Disease Control“, *Bulletin of Mathematical Biology*, Jg. 69, Nr. 5, S. 1511–1536, Juli 2007. DOI: 10.1007/s11538-006-9174-9.
- [32] Z. Feng, Y. Zheng, N. Hernandez-Ceron, H. Zhao, J. W. Glasser und A. N. Hill, „Mathematical models of Ebola – Consequences of underlying assumptions“, *Mathematical Biosciences*, Jg. 277, S. 89–107, Apr. 2016. DOI: 10.1016/j.mbs.2016.04.002.

- [33] O. Forster, *Analysis 1: Differential- und Integralrechnung einer Veränderlichen*, 12. Aufl. Springer Fachmedien Wiesbaden, März 2016. DOI: 10.1007/978-3-658-11545-6.
- [34] H.-O. Georgii, *Stochastik: Einführung in die Wahrscheinlichkeitstheorie und Statistik*, 5. Aufl. Berlin/Boston: Walter de Gruyter GmbH, 2015. DOI: 10.1515/9783110359701.
- [35] W. M. Getz und E. R. Dougherty, „Discrete stochastic analogs of Erlang epidemic models“, *Journal of Biological Dynamics*, Jg. 12, Nr. 1, S. 16–38, Nov. 2017. DOI: 10.1080/17513758.2017.1401677.
- [36] J. Goudriaan, „Boxcartrain methods for modelling of ageing, development, delays and dispersion“, in *The dynamics of physiologically structured populations*, J. Metz und O. Diekmann, Hrsg., Springer Berlin Heidelberg, Aug. 1986, S. 453–473, ISBN: 978-3-662-13159-6, zu finden unter <https://library.wur.nl/WebQuery/wurpubs/fulltext/217642> (abgerufen am 15.06.2023).
- [37] S. Greenhalgh und C. Rozins, „A generalized differential equation compartmental model of infectious disease transmission“, *Infectious Disease Modelling*, Jg. 6, S. 1073–1091, Sep. 2021. DOI: 10.1016/j.idm.2021.08.007.
- [38] N. Guglielmi, E. Iacomini und A. Viguerie, „Identification of time delays in COVID-19 data“, *Epidemiologic Methods*, Jg. 12, Nr. 1, S. 1–20, Mai 2023. DOI: 10.1515/em-2022-0117.
- [39] K. B. Habersaat, C. Betsch, M. Danchin, C. R. Sunstein, R. Böhm, A. Falk, N. T. Brewer, S. B. Omer, M. Scherzer, S. Sah, E. F. Fischer, A. E. Scheel, D. Fancourt, S. Kitayama, E. Dubé, J. Leask, M. Dutta et al., „Ten considerations for effectively managing the COVID-19 transition“, *Nature Human Behaviour*, Jg. 4, Nr. 7, S. 677–687, Juli 2020. DOI: 10.1038/s41562-020-0906-x.
- [40] M. Hanke-Bourgeois, *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*, 3. Aufl. Wiesbaden: Vieweg+Teubner Verlag, 2009. DOI: 10.1007/978-3-8348-9309-3.
- [41] K. Heinz, *Mathematisch-statistische Untersuchungen über die Erlang-Verteilung*, *Forschungsberichte des Landes Nordrhein-Westfalen*. Westdeutscher Verlag, 1969, ISBN: 978-3663063797.
- [42] N. Henze, *Stochastik für Einsteiger, Eine Einführung in die faszinierende Welt des Zufalls*, 13. Aufl. Springer Spektrum Berlin, Heidelberg, Aug. 2021. DOI: 10.1007/978-3-662-63840-8.
- [43] H. W. Hethcote, „The Mathematics of Infectious Diseases“, *SIAM Review*, Jg. 42, Nr. 4, S. 599–653, Mai 2000. DOI: 10.1137/S0036144500371907.
- [44] P. J. Hurtado, „Building New Models: Rethinking and Revising ODE Model Assumptions“, in *An Introduction to Undergraduate Research in Computational and Mathematical Biology: From Birdsongs to Viscosities*, H. Callender Highlander, A. Capaldi und C. Diaz Eaton, Hrsg. Cham: Springer International Publishing, Feb. 2020, S. 1–86. DOI: 10.1007/978-3-030-33645-5_1.

- [45] P. J. Hurtado und A. S. Kiro Singh, „Generalizations of the ‘Linear Chain Trick’: incorporating more flexible dwell time distributions into mean field ODE models“, *Journal of Mathematical Biology*, Jg. 79, Nr. 5, S. 1831–1883, Okt. 2019. DOI: 10.1007/s00285-019-01412-w.
- [46] P. J. Hurtado und C. Richards, „A procedure for deriving new ode models: Using the generalized linear chain trick to incorporate phase-type distributed delay and dwell time assumptions“, *Mathematics in Applied Sciences and Engineering*, Jg. 1, Nr. 4, S. 274–440, Dez. 2020. DOI: 10.5206/mase/10857.
- [47] P. J. Hurtado und C. Richards, „Building mean field ODE models using the generalized linear chain trick & Markov chain theory“, *Journal of Biological Dynamics*, Jg. 15, Nr. sup1, S. 248–272, Apr. 2021. DOI: 10.1080/17513758.2021.1912418.
- [48] O. C. Ibe, „Chapter 4 – Special Probability Distributions“, in *Fundamentals of Applied Probability and Random Processes*, 2. Aufl., Boston: Academic Press, 2014, S. 103–158. DOI: 10.1016/B978-0-12-800852-2.00004-3.
- [49] S. A. Iyaniwura, R. Musa und J. D. Kong, „A generalized distributed delay model of COVID-19: An endemic model with immunity waning“, *Mathematical Biosciences and Engineering*, Jg. 20, Nr. 3, S. 5379–5412, Jan. 2023. DOI: 10.3934/mbe.2023249.
- [50] J. A. Jacquez und C. P. Simon, „Qualitative theory of compartmental systems with lags“, *Mathematical Biosciences*, Jg. 180, Nr. 1, S. 329–362, Juni 2002. DOI: 10.1016/S0025-5564(02)00131-1.
- [51] M. J. Keeling und B. T. Grenfell, „Disease Extinction and Community Size: Modeling the Persistence of Measles“, *Science*, Jg. 275, Nr. 5296, S. 65–67, Jan. 1997. DOI: 10.1126/science.275.5296.65.
- [52] M. J. Keeling und B. T. Grenfell, „Understanding the persistence of measles: reconciling theory, simulation and observation“, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, Jg. 269, Nr. 1489, S. 335–343, Feb. 2002. DOI: 10.1098/rspb.2001.1898.
- [53] W. O. Kermack und A. G. McKendrick, „A Contribution to the Mathematical Theory of Epidemics“, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, Jg. 115, Nr. 772, S. 700–721, Aug. 1927, <http://www.jstor.org/stable/94815> (abgerufen am 21.06.2023).
- [54] W. Kiehl, *Infektionsschutz und Infektionsepidemiologie*, Robert Koch-Institut, Berlin, Feb. 2015. DOI: 10.25646/96.
- [55] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad und M. Lipsitch, „Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period“, *Science*, Jg. 368, Nr. 6493, S. 860–868, Mai 2020. DOI: 10.1126/science.abb5793.

- [56] W. Koslow, M. J. Kühn, S. Binder, M. Klitz, D. Abele, A. Basermann und M. Meyer-Hermann, „Appropriate relaxation of non-pharmaceutical interventions minimizes the risk of a resurgence in SARS-CoV-2 infections in spite of the Delta variant“, *PLOS Computational Biology*, Jg. 18, Nr. 5, S. 1–26, Mai 2022. DOI: 10.1371/journal.pcbi.1010054.
- [57] M. Kreh, R. Goertz und F. Modler, *Tutorium Höhere Analysis*, 1. Aufl. Berlin, Heidelberg: Springer Spektrum, Jan. 2018. DOI: 10.1007/978-3-8274-3004-5.
- [58] O. Krylova und D. J. D. Earn, „Effects of the infectious period distribution on predicted transitions in childhood disease dynamics“, *Journal of Royal Society*, Jg. 10, Nr. 84, S. 1–15, Juli 2013. DOI: 10.1098/rsif.2013.0098.
- [59] M. J. Kühn, D. Abele, S. Binder, K. Rack, M. Klitz, J. Kleinert, J. Gilg, L. Spataro, W. Koslow, M. Siggel, M. Meyer-Hermann und A. Basermann, „Regional opening strategies with commuter testing and containment of new SARS-CoV-2 variants in Germany“, *BMC Infectious Diseases*, Jg. 22, Nr. 1, S. 1–17, Apr. 2022. DOI: 10.1186/s12879-022-07302-9.
- [60] M. J. Kühn, D. Abele, T. Mitra, W. Koslow, M. Abedi, K. Rack, M. Siggel, S. Khailaie, M. Klitz, S. Binder, L. Spataro, J. Gilg, J. Kleinert, M. Häberle, L. Plötzke, C. D. Spinner, M. Stecher, X. X. Zhu, A. Basermann und M. Meyer-Hermann, „Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution“, *Mathematical Biosciences*, Jg. 339, S. 1–19, Sep. 2021. DOI: 10.1016/j.mbs.2021.108648.
- [61] M. J. Kühn, D. Abele, D. Kerkmann, S. A. Korf, H. Zunker, A. C. Wendler, J. Bicker, D. K. Nguyen, M. Klitz, W. Koslow, M. Siggel, J. Kleinert, K. Rack, S. Binder, L. Plötzke, R. Schmieding, P. Lenz et al., *MEMILIO – a high performance Modular Epidemic simuLation software*, Institute for Software Technology of the German Aerospace Center (DLR) and the department of Systems Immunology (SIMM) of the Helmholtz Center for Infection Research (HZI), 2022, zu finden unter <https://github.com/SciCompMod/memilio> (abgerufen am 22.11.2023); siehe auch <https://elib.dlr.de/192140/> (abgerufen am 26.06.2023).
- [62] A. L. Lloyd, „Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods“, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, Jg. 268, Nr. 1470, S. 985–993, Mai 2001. DOI: 10.1098/rspb.2001.1599.
- [63] A. L. Lloyd, „Realistic Distributions of Infectious Periods in Epidemic Models: Changing Patterns of Persistence and Dynamics“, *Theoretical Population Biology*, Jg. 60, Nr. 1, S. 59–71, Aug. 2001. DOI: 10.1006/tpbi.2001.1525.
- [64] A. L. Lloyd, „The dependence of viral parameter estimates on the assumed viral life cycle: limitations of studies of viral load data“, *Proceedings: Biological Sciences*, Jg. 268, Nr. 1469, S. 847–854, Apr. 2001. DOI: 10.1098/rspb.2000.1572.
- [65] J. Ma und D. J. D. Earn, „Generality of the Final Size Formula for an Epidemic of a Newly Invading Infectious Disease“, *Bulletin of Mathematical Biology*, Jg. 68, Nr. 3, S. 679–702, Apr. 2006. DOI: 10.1007/s11538-005-9047-7.

- [66] N. MacDonald, *Time Lags in Biological Models*. Springer Berlin, Heidelberg, Nov. 1978. DOI: 10.1007/978-3-642-93107-9.
- [67] M. Martcheva, *An introduction to mathematical epidemiology*, 1. Aufl. Springer New York, NY, Okt. 2015. DOI: 10.1007/978-1-4899-7612-3.
- [68] A. A. Mat Daud, „Five common misconceptions regarding flattening-the-curve of COVID-19“, *History and Philosophy of the Life Sciences*, Jg. 44, Nr. 3, S. 1–7, Sep. 2022. DOI: 10.1007/s40656-022-00522-x.
- [69] E. Messina, M. Pezzella und A. Vecchio, „A non-standard numerical scheme for an age-of-infection epidemic model“, *Journal of Computational Dynamics*, Jg. 9, Nr. 2, S. 239–252, Apr. 2022. DOI: 10.3934/jcd.2021029.
- [70] P. D. O’Neill, „Introduction and snapshot review: Relating infectious disease transmission models to data“, *Statistics in Medicine*, Jg. 29, Nr. 20, S. 2069–2077, Apr. 2010. DOI: <https://doi.org/10.1002/sim.3968>.
- [71] B. Pell, M. D. Johnston und P. Nelson, „A data-validated temporary immunity model of COVID-19 spread in Michigan“, *Mathematical Biosciences and Engineering*, Jg. 19, Nr. 10, S. 10122–10142, Juli 2022. DOI: 10.3934/mbe.2022474.
- [72] L. Plötzke, „Modellierung epidemischer Infektionskrankheiten auf der Basis von gewöhnlichen und Integro-Differentialgleichungen“, Mathematisches Institut der Universität zu Köln in Kooperation mit dem Deutschen Zentrum für Luft- und Raumfahrt, Bachelorarbeit, Mai 2021, zu finden unter <https://elib.dlr.de/143504/> (abgerufen am 06.11.2023).
- [73] W. H. Press, S. A. T. Teukolsky, W. T. Vetterling und B. P. Flannery, *Numerical Recipes in C, The Art of Scientific Computing*, 2. Aufl. Cambridge University Press, Dez. 1992, ISBN: 978-0-521-43108-8.
- [74] P. Reinecke, L. Bodrog und A. Danilkina, „Phase-Type Distributions“, in *Resilience Assessment and Evaluation of Computing Systems*, K. Wolter, A. Avritzer, M. Vieira und A. van Moorsel, Hrsg. Springer Berlin Heidelberg, Jan. 2012, S. 85–113. DOI: 10.1007/978-3-642-29032-9_5.
- [75] Robert Koch-Institut, „Täglicher Lagebericht des RKI zur Coronavirus-Krankheit-2019 (COVID-19) am 15.10.2020“, Techn. Ber., 2020, https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Okt_2020/2020-10-15-de.pdf?__blob=publicationFile (abgerufen am 15.09.2023).
- [76] Robert Koch-Institut, *SARS-CoV-2 Infektionen in Deutschland*, Datensatzdokumentation, Nov. 2023. DOI: 10.5281/zenodo.10208170, Datensatz ist zu finden unter https://github.com/robert-koch-institut/SARS-CoV-2-Infektionen_in_Deutschland (abgerufen am 27.11.2023).
- [77] M. Roberts, V. Andreasen, A. Lloyd und L. Pellis, „Nine challenges for deterministic epidemic models“, *Epidemics*, Jg. 10, S. 49–53, März 2015, Challenges in Modelling Infectious Disease Dynamics. DOI: 10.1016/j.epidem.2014.09.006.
- [78] S. M. Ross, *A first course in probability*, 8. Aufl. Pearson Prentice Hall, Jan. 2009, ISBN: 978-0-13-603313-4.

- [79] G. Rozhnova, C. H. van Dorp, P. Bruijning-Verhagen, M. C. J. Bootsma, J. H. H. M. van de Wijgert, M. J. M. Bonten und M. E. Kretzschmar, „Model-based evaluation of school- and non-school-related measures to control the COVID-19 pandemic“, *Nature Communications*, Jg. 12, Nr. 1, S. 1–11, März 2021. DOI: 10.1038/s41467-021-21899-6.
- [80] H. Smith, *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, 1. Aufl. Springer New York, NY, Okt. 2010. DOI: 10.1007/978-1-4419-7646-8.
- [81] Statistische Ämter des Bundes und der Länder, *Regionaldatenbank Deutschland: Bevölkerung nach Geschlecht und Altersjahren (79) - Stichtag 31.12. - (ab 2011) regionale Ebenen*, Tabellennummer: 12411-04-02-4-B, verwendeter Stichtag: 31.12.2020, <https://www.regionalstatistik.de/genesis//online?operation=table&code=12411-04-02-4-B&bypass=true&levelindex=1&levelid=1691406873603#abreadcrumb> (abgerufen am 07.08.2023).
- [82] The SciPy community, *Statistical functions (scipy.stats)*, 2008–2023, zu finden unter <https://docs.scipy.org/doc/scipy/reference/stats.html#> (abgerufen am 13.10.2023); siehe auch: P. Virtanen, R. Gommers, T. E. Oliphant u.a. „SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python“, *Nature Methods*, Jg. 17, S. 261–272, Februar 2020. DOI: 10.1038/s41592-019-0686-2.
- [83] H. R. Thieme, „The Transition Through Stages with Arbitrary Length Distributions, and Applications in Epidemics“, in *Mathematical Approaches for Emerging and Reemerging Infectious Diseases: Models, Methods, and Theory*, C. Castillo-Chavez, S. Blower, P. van den Driessche, D. Kirschner und A.-A. Yakubu, Hrsg., Springer New York, 2002, S. 45–84. DOI: 10.1007/978-1-4613-0065-6_5.
- [84] H. R. Thieme, *Mathematics in Population Biology*. Princeton University Press, 2003. DOI: 10.2307/j.ctv301f9v.
- [85] N. T. Thomopoulos, *Probability Distributions, With Truncated, Log and Bivariate Extension*, 1. Aufl. Cham, Schweiz: Springer International Publishing, Apr. 2018. DOI: 10.1007/978-3-319-76042-1.
- [86] A. Vazquez, „Exact solution of infection dynamics with gamma distribution of generation intervals“, *Physical Review E*, Jg. 103, Nr. 4, S. 1–6, Apr. 2021. DOI: 10.1103/PhysRevE.103.042306.
- [87] T. Vogel, „Systèmes déferlants, systèmes héréditaires, systèmes dynamiques“, *Proceedings of the international symposium nonlinear vibrations, IUTAM, Kiev*, Nr. 22, S. 123–130, 1961.
- [88] T. Vogel, „Théorie des Systèmes Évolutifs“, *Traité de physique théorique et de physique mathématique*, Gauthier-Villars, Paris, Nr. 22, 1965, teilweise abrufbar unter <https://excerpts.numilog.com/books/9782706236747.pdf> (abgerufen am 05.06.2023).

- [89] J. Wallinga und M. Lipsitch, „How generation intervals shape the relationship between growth rates and reproductive numbers“, *Proceedings of the Royal Society B: Biological Sciences*, Jg. 274, Nr. 1609, S. 599–604, Nov. 2007. DOI: 10.1098/rspb.2006.3754.
- [90] G. Walz, *Lexikon der Mathematik: Band 3*, 2. Aufl. Berlin, Heidelberg: Springer Spektrum, Dez. 2016. DOI: 10.1007/978-3-662-53502-8.
- [91] X. Wang, Y. Shi, Z. Feng und J. Cui, „Evaluations of Interventions Using Mathematical Models with Exponential and Non-exponential Distributions for Disease Stages: The Case of Ebola“, *Bulletin of Mathematical Biology*, Jg. 79, Nr. 9, S. 2149–2173, Sep. 2017. DOI: 10.1007/s11538-017-0324-z.
- [92] H. Wearing, P. Rohani, T. Cameron und S. Sait, „The Dynamical Consequences of Developmental Variability and Demographic Stochasticity for Host-Parasitoid Interactions“, *The American naturalist*, Jg. 164, Nr. 4, S. 543–58, Okt. 2004. DOI: 10.1086/424040.
- [93] H. J. Wearing, P. Rohani und M. J. Keeling, „Appropriate Models for the Management of Infectious Diseases“, *PLOS Medicine*, Jg. 2, Nr. 7, S. 0621–0627, Juli 2005. DOI: 10.1371/journal.pmed.0020174.
- [94] A. C. Wendler, P. Lenz, M. Klitz, P. Vanella und M. J. Kühn, „Assessment of non-pharmaceutical interventions against SARS-CoV-2 and COVID-19 in Germany“, 2024, In preparation.
- [95] A. C. Wendler, L. Plötzke, H. Tritschak und M. J. Kühn, „Extensions of age of infection models: A SECIR integro differential equation-based model for epidemic outbreaks“, 2024, In preparation.
- [96] World Health Organization (WHO), *Fact sheet: Coronavirus disease (COVID-19)*, [https://www.who.int/news-room/fact-sheets/detail/coronavirus-disease-\(covid-19\)](https://www.who.int/news-room/fact-sheets/detail/coronavirus-disease-(covid-19)) (abgerufen am 21.08.2023), 2023.
- [97] World Health Organization (WHO), *Our work: Communicable and noncommunicable diseases, and mental health*, <https://www.who.int/our-work/communicable-and-noncommunicable-diseases-and-mental-health> (abgerufen am 21.08.2023), 2023.
- [98] World Health Organization (WHO), *Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic*, [https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-coronavirus-disease-\(covid-19\)-pandemic](https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic) (abgerufen am 21.08.2023), Mai 2023.
- [99] P. Yan, „Distribution Theory, Stochastic Processes and Infectious Disease Modelling“, in *Mathematical Epidemiology*, F. Brauer, P. van den Driessche und J. Wu, Hrsg. Springer Berlin Heidelberg, Juni 2008, S. 229–293. DOI: 10.1007/978-3-540-78911-6_10.

Versicherung an Eides statt

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Ort, Datum

Unterschrift