Verifikation sicherheitskritischer Regelsysteme unter Beachtung des Zeitverhaltens einer verteilten Rechenplattform

DISSERTATION

zur Erlangung des akademischen Grades eines
DOKTOR-INGENIEURS
(DR.-ING.)

der Fakultät für Ingenieurwissenschaften,
Informatik und Psychologie der Universität Ulm

von

Dipl.-Ing. Tobias Stefan Bund
aus Ulm

Gutachter:	Prof. Dr.-Ing. Frank Slomka
	Prof. Dr.-Ing. Petru Eles
Amtierender Dekan:	Prof. Dr. rer. nat. Frank Kargl

Ulm, den 07. Juli 2017

Kurzfassung

In vielen Anwendungsgebieten elektrischer Steuer- und Regelsysteme (z.B. Automotive, Avionik, Anlagenbau) sind durch ein instabiles Verhalten des Regelkreises Menschenleben gefährdet. Solche kritische Situationen treten dann auf, wenn das dynamische Verhalten des zu regelnden technischen Systems von den Annahmen des Reglerentwurfs abweicht. Ursachen für das veränderte Verhalten können Zeiteffekte wie Latenzen, Jitter oder verworfene Ereignisse im Regelkreis sein.
Betriebswirtschaftliche Bestrebungen motivieren die kostengünstige Umsetzung solcher Regelsysteme, was unter anderem durch eine größtmögliche Auslastung der Rechenplattform erwirkt wird. Eine hohe Auslastung der verfügbaren Kommunikations- und Rechenkapazitäten begünstigt wiederum Zeiteffekte und steht dadurch in Konflikt zu den Anforderungen an ein sicherheitskritisches Regelsystem. Die Varianz und das Ausmaß der Zeiteffekte erhöht sich zusätzlich bei der Vernetzung von Recheneinheiten zu einer verteilten Rechenplattform.

In dieser Arbeit werden Modelle entwickelt, welche es erlauben Zeiteffekte solch eines verteilten Systems zu beschreiben. Während existierende Modelle die Zeiteffekte anhand einer statischen Grenze spezifizieren, ist es Gegenstand dieser Arbeit, eine formale mathematische Beschreibung zu entwerfen, welche das Auftreten der Zeiteffekte in Relation zu Ereignisintervallen beschränkt. Dichtefunktionen der Latenzen und der verworfenen Ereignisse, welche den Bezug zu Ereignisintervallen herstellen, ermöglichen weniger pessimistische Annahmen bei der Verifikation verteilter Regelsysteme hinsichtlich des Vorkommens von Zeiteffekten. Um die Dichtefunktionen auf nebenläufige und verzweigende Signalpfade einer verteilten Systemarchitektur anzuwenden, sind grundlegend neue Berechnungsvorschriften notwendig.

Basierend auf diesen Zeitmodellen wird in dieser Arbeit der modellbasierte Entwurf verteilter Regelsysteme um neue Verifikationsmethoden erweitert. So sollen für eine Spezifikation des Zeitverhaltens durch eine Dichtefunktion die schlimmstmöglichen Auswirkungen auf ein Regelsystem abgeschätzt werden.
Durch die Implementierung der entwickelten Algorithmen kann gezeigt werden, dass durch Anwendung der Modelle und Methoden eine effizientere Verwendung der Systemressourcen möglich ist. Daraus resultiert unter Garantie der spezifizierten Güteanforderungen an das Regelsystem eine kostenoptimierte Umsetzung der Plattform.

Abstract

In many domains of digital control systems (e.g. automotive, avionics, industrial applications) an instable control behavior may endanger human life. Such unacceptable situations occur, when there is a difference between the dynamic behavior of the technical system and the assumptions in the design phase. The reason for this difference may be timing effects like delay, jitter and dropped samples.
For economic reasons, there is a desire for a cost efficient implementation of the control platform what results in a high utilization of the system resources. However, a high utilization causes timing effects and therefore conflicts with the requirements of a safety critical control system. Additionally, variance and quantity of the timing effects increases for a distributed system architecture.

This thesis introduces an universal model, that is able to describe the platform's time effects in distributed control systems. While existing models only cover the corner case of the timing effects, this work considers their occurrence regarding an interval of consecutive events. This new model of density functions reduces the pessimistic assumptions of the timing effects, when verifying the performance of a control system. To apply these density functions on concurrent and branching signal paths of a distributed system architecture, new calculation methods are necessary.
Based on this model, methods are developed, that bound the effect of these timing effects on a distributed control system. Thus, for such a specification of the timing effects, the worst case impacts on a control systems should be estimated.
An implementation of the developed models and methods presents the beneficial impact on the resource usage while guaranteeing predefined performance specifications.

Symbolverzeichnis

f_s,f_g	Abtastfrequenz, Grenzfrequenz
$T_s = \frac{1}{f_s}$	Abtastzeit
$\tau_{i,k}$	k-te Instanz / k-ter Job des Task τ_i
$t^r_{i,k}$,$t^e_{i,k}$	Zeitpunkt der k-ten Aktivierung bzw. Endzeitpunkt des Task τ_i
$c^{+/-}$	maximale / minimale Nettoausführungszeit
$r_{k,i}$	k-te Antwortzeit des Task τ_i
$r^{+/-}$	maximale / minimale Antwortzeit
P	Ausführungspfad
Δ_t	Zeitspanne
Δ_k	Intervall benachbarter Ereignisse
d_k	Sequenz an verworfenen Ereignissen
D_i	kumulative Summe verworfener Ereignisse
$dD^{+/-}(\Delta_k)$	maximale / minimale Verlustdichte
$iD^{+/-}(n)$	maximales / minimales Intervall an Ereignissen, in welchen n Ereignisse verworfen werden
(Δ^c_k, n^c)	charakteristischer Punkt
(Δ^d_k, n^d)	dominanter Punkt

R_i	kumulative Summe der Antwortzeiten
$dR^{+/-}(\Delta_k)$	maximale / minimale Latenzdichte
F	Frist
$x(t),\ x_k$	kontinuierlicher / zeitdiskreter Zustandsvektor
$\delta^-(k)$	minimale Zeitspanne, in der k Ereignisse auftreten
$\eta^+(\Delta_t)$	maximale Ereignisdichte; maximale Anzahl an Ereignissen in der Zeitspanne (Δ_t)
$\rho^+(\Delta_t)$	maximale Last in der Zeitspanne (Δ_t)
$\|\mathbf{s}(\Delta_k)\|$	Signaldichte

Abkürzungsverzeichnis

CPS	Cyber Physical System (cyber physisches System)
DUT	Device under test (zu testendes Gerät)
GPC	Greedy Processing Component (Greedy Verarbeitungskomponente)
LET	Logical execution time (logische Ausführunszeit)
LTI	Linear, time-invariant (linear zeitinvariant)
QoS	Quality of Service (Dienstgüte)
RTC	Real-Time Calculus (Echtzeitkalkühl)
MPA	Modular performance analysis (Modulare Performanzanalyse)
TTA	Time-triggered architecture (zeitgesteuerte Architektur)
TWCRT	Typical Worst Case Response Time (übliche maximale Ausführungszeit)
WCET	Worst Case Execution Time (maximale Ausführungszeit)
WCRT	Worst Case Response Time (maximale Antwortzeit)

Inhaltsverzeichnis

Kurzfassung iii

Abstract v

Symbolverzeichnis vii

Abkürzungsverzeichnis ix

1 Einleitung 1

2 Motivation und Ziele 5
2.1 Eingebettete Systeme, Echtzeitsysteme und Cyber-Physische Systeme 5
2.2 Verteilte Systemarchitektur 8
2.3 Regelsystem auf verteilter Plattform 10
2.3.1 Digitale und verteilte Regelsysteme 11
2.3.2 Klassifizierung von Regelanwendungen 13
2.4 Ziele der Arbeit 15

3 Modellierung des Zeitverhaltens und des funktionalen Verhaltens . . . 17
3.1 Plattformmodell 18
3.1.1 Taskmodell 18
3.1.2 Ereignismodelle 21
3.1.3 Zeiteffekte der Plattform 24
3.1.4 Verfahren zur Bestimmung der Antwortzeit 26
3.2 Funktionales Modell 29
3.2.1 Zustandsraummodell 30

3.2.2 Diskreter Zustandsregler 31
3.2.3 Störungen im Regelkreis 34

4 Stand der Technik des Regler-/Plattform Co-Designs 37
4.1 Entwurfsziele und Entwurfsparameter 37
4.2 Zeiteffekte im Regelkreis 40
4.2.1 Reglerentwurf unter Beachtung von Zeiteffekten 40
4.2.2 Reglervalidierung unter Beachtung von Zeiteffekten 42
4.2.3 Regler/Plattform Co-Design 43
4.2.4 Zusammenfassung 44
4.3 Bewertung von Regelkreisen 45
4.3.1 Stabilität 45
4.3.2 Regelgüte 47

5 Begrenzung von Zeiteffekten im Intervallbereich 51
5.1 Verworfene und ausgesetzte Ereignisse 53
5.1.1 Definition der Verlustdichte 54
5.1.2 Definition der Intervallfunktion 56
5.1.3 Eigenschaften der Verlustdichte 60
5.1.4 Anwendung der Verlustdichte in verteilten Systemen 63
5.1.5 Aussetzen von Ereignissen zur Lastreduktion 85
5.2 Antwortzeiten und Latenzen 89
5.2.1 Definition der Latenzdichte 90
5.2.2 Eigenschaften der Latenzdichte 91
5.2.3 Ableiten einer Verlustdichte aus einer Latenzdichte 94
5.2.4 Einfluss der Latenzdichte eines Tasks auf die ausgehende Ereignisdichte 96
5.2.5 Bestimmung der Latenzdichte aus den Dichtefunktionen des Echtzeitkalküls 99

6 Analyse von Regelsystemen bezüglich Zeiteffekten 103
6.1 Bewertung der Regelgüte im Intervallbereich 104
6.1.1 Eigenwerte und Matrixnormen 105
6.1.2 Die Signaldichte 106
6.1.3 Mathematische Operationen auf Signaldichten 108
6.1.4 Anwendung der Signaldichte in Regelkreisen 112
6.1.5 Störeinfluss auf die Signaldichte des Zustandsvektors 116

6.2 Verworfene Ereignisse im Regelkreis 120
6.2.1 Auswirkungen verworfener Ereignisse im Zeitbereich 120
6.2.2 Betrachtung im Intervallbereich 124
6.3 Latenzen im Regelkreis 135
6.3.1 Auswirkungen von Latenzen im Zeitbereich 135
6.3.2 Betrachtung im Intervallbereich 138

7 Co-Design und Entwurfsstudie 143
7.1 Modell einer Gleichstrommaschine 143
7.2 Validierung des Störverhaltens 145
7.3 Validierung verworfener Ereignisse 147

8 Zusammenfassung 153

Literaturverzeichnis 155

A Publikationsliste 167

B Formeln 169
B.1 Max-Plus-Algebra 169
B.2 Echtzeitkalkül nach (Wandeler, 2006) 170

C Beweise 171
C.1 Beweis zu Satz 24 171
C.2 Beweis zu Satz 28 173

Abbildungsverzeichnis

2.1 Aufbau eines verteilten Regelsystems mit gemeinsam genutzten Rechen- und Kommunikationsressourcen. 10
2.2 Vergleich der Regelgüte für kontinuierliche Regelung, digitales Regelsystem und verteiltes Regelsystem als Funktion der Abtastrate (Moyne und Tilbury, 2007). 12
2.3 Klassifizierung von Regelsystemen nach Sicherheitskritikalität und Ressourcenverbrauch /-auslastung. 13

3.1 Sichtweisen eines verteilten Regelsystems. 17
3.2 Plattformmodell einer verteilten Systemarchitektur (basierend auf (Künzli et al., 2007)). 19
3.3 Antwortzeit eines Multitasksystems. 20
3.4 Strikt periodisches Ereignismodell. 22
3.5 Periodisches Ereignismodell mit Jitter. 22
3.6 Maximale Anhäufung der Ereignisse für einen Jitter von $j = 2 \cdot p = 10\,ZE$ und $d = 1\,ZE$. 23
3.7 Greedy Processing Component des Echtzeitkalküls nach (Wandeler, 2006). 28
3.8 Hybrides Mehrgrößenregelsystem mit kontinuierlichen und zeitdiskreten Signalen. 31
3.9 Zeitdiskretes Mehrgrößenregelsystem. 34

4.1 Entwurfsparameter und Zusammenhänge des Regler-/Plattform Co-Designs. 38

4.2 Verlauf der Lösung der Differentialgleichungen für den Startwert $x_0 = 1$. 46
4.3 Sprungantwort eines Regelsystems. 48

5.1 Verworfene Ereignisse eines Tasks. 53
5.2 Auflistung einer beispielhaften Sequenz an verworfenen Ereignissen (d_k) mit kumulativer Summe (D_k), maximaler Dichtefunktion $(dD^+(\Delta_k))$ und minimaler Intervallfunktion $(iD^-(n))$. 58
5.3 Vergleich der Spezifikation von verworfenen Ereignissen durch eine Verlustdichte $(dD^+(\Delta_k))$ und einer Rate an verworfenen Ereignissen $(dD^+_{period}(\Delta_k))$. 59
5.4 Abschätzung größerer Intervalle einer endlichen Dichtefunktion $(\Delta^e_k = 14)$. 66
5.5 Treppenfunktionen der charakteristischen Punkte, aus denen die Verlustdichte zusammengesetzt ist. 68
5.6 Charakteristische Punkte und dominanter Punkt einer beispielhaften Dichtefunktion. 71
5.7 Beispielhafte maximale Dichtefunktion einer (m,k)-Beschränkung und deren beider Teilfunktionen, aus deren Minimum sich die Dichtefunktion ergibt. 74
5.8 Beispiel einer Dichtefunktion mit Abschätzung durch eine (m,k)-Begrenzung mit $m = 2$ und $k = 5$ und durch eine feste Rate mit $rD = 3$.
Die charakteristischen Punkte der maximalen Verlustdichte sind $cP = \{(3, 2,), (11, 5)\}$ 76
5.9 Verkettung von Dichtefunktionen. 79
5.10 Zusammengefasste Verlustdichte zweier Tasks mit UND-Vereinigung der ausgehenden Ereignissequenzen. 81
5.11 Vergleich der zusammengefassten Verlustdichte beim Tausch zweier in Serie verknüpfter Tasks. 84
5.12 Historie der zuletzt verworfenen Ereignisse. 87
5.13 Beispielhafte Sequenz von Antwortzeiten eines Tasks und deren Schranken. 90
5.14 Zusammengefasste Latenz bei der UND-Vereinigung zweier simultan aktivierter Signalpfade. 93

5.15 Zusammenhang zwischen Fristüberschreitungen von Tasks und verworfenen Ereignissen. 95
5.16 Konstruktion des kleinsten Intervalls, in dem drei ausgehende Ereignisse auftreten können. 97
5.17 Maximale Ereignisdichte $\eta^+(\Delta_t)$ und minimale Kapazitätskurve $\beta^-(\Delta_t)$, sowie die resultierende Antwortzeit $r^+ = r_3$. 100
5.18 Ereignissequenz, welche den Grenzfall der Ereignisdichte umsetzt. . . 100
5.19 Maximale Latenzdichte $(dR^+(\Delta_k))$ aus den beschränken Kurven des Echtzeitkalkül. 101

6.1 Zweidimensionale Visualisierung der Spektralnorm. 105
6.2 Signaldichte für drei beispielhafte Signalverläufe. 108
6.3 Addition von Signalen, welche durch eine Signaldichte repräsentiert werden. 109
6.4 Testsignal $x_k = \cos(k)$ mit beispielhafter Folge g_k und Faltungsergebnis $|y_k|$, sowie die resultierende Signaldichte $\|y(\Delta_k)\|$ durch direkte Berechnung nach Satz 21. 113
6.5 Norm des Zustandsvektors eines ungestörten Systems mit initialer Auslenkung und korrespondierender Signaldichte. 115
6.6 Die durch ein verworfenes Ereignis verursachte Störung v_k auf die Stellgröße des Regelkreises. 121
6.7 Verlauf der Stellgröße eines asymptotisch stabilen Regelkreises mit einem verworfenen Ereignis. 122
6.8 Zwei Szenarien verworfener Ereignisse mit $iD^-(3) = 7$. 126
6.9 Verlauf der Folgen $\theta_{v,k}$ und $\mu_{v,k}$, sowie deren exponentiell abfallende Hüllkurve. 128
6.10 Verlauf der stabilisierenden und destabilisierenden Anteile des Exponenten aus Gleichung (6.29). 131
6.11 Histogramm der maximalen Störungen durch Auswertung aller Binärkombinationen und valider Grenzwert durch Anwendung der Intervallbetrachtung. 132
6.12 Geltungsbereich der Stellgrößen bei variablen Latenzen. 136

7.1 Ersatzschaltbild des Ankers einer Gleichstrommaschine nach (Stölting und Beisse, 1987). 143
7.2 Signaldichte des Störsignals $(\|z(\Delta_k)\|)$. 146

7.3 Signaldichte der Abweichung der Winkelgeschwindigkeit $\|y_z(\Delta_k)\|$ und spezifizierte Begrenzung $\|\hat{y}(\Delta_k)\|$. 147
7.4 Sensitivitätsfunktion des Regelkreises $(\mu_{v,k})$ bezüglich verworfener Ereignisse. 148
7.5 Signaldichte der Abweichung der Winkelgeschwindigkeit $\|y(\Delta_k)\|$ bei einer Drehmomentstörung und verworfenen Ereignissen. 151

Kapitel 1
Einleitung

Eingebettete Systeme verrichten ihre Aufgabe vorwiegend im Verborgenen und sind zugleich die am meisten verbreiteten Computersysteme (Morales et al., 2015). Dabei müssen diese Computersysteme die geforderten funktionalen Anforderungen unter vorgegebenen Randbedingungen zuverlässig erfüllen. Eine häufig spezifizierte Anforderung ist die Echtzeiteigenschaft, welche erfüllt ist, wenn das korrekte Berechnungsergebnis rechtzeitig vorliegt. Während jede Rechenoperation eine endliche Ausführungszeit benötigt und dadurch eine Latenz bis zur Bereitstellung des Rechenergebnisses verursacht, vergrößert sich diese Zeitspanne ungemein bei verteilten Rechenplattformen. Solche verteilte Systeme ergeben sich durch die räumliche Trennung von Funktionen auf unterschiedliche Computersysteme, bei denen jedoch Abhängigkeiten in der Ausführung bestehen. Um solch ein verteiltes Echtzeitsystem zu verifizieren, wird eine Echtzeitanalyse angewandt, welche durch eine mathematische Untersuchung bestimmt, ob die geforderte Zeitspanne bis zur vollständigen Bearbeitung der Rechenaufgaben eingehalten wird. Diese Vorgehensweise eignet sich besonders für harte Echtzeitsysteme, wie dem rechtzeitigen Auslösen eines Airbags. Dessen lebensrettende Wirkung wird nur bei Einhalten der spezifizierten Zeitfrist voll entfaltet, andernfalls kommt es zu einer Gefährdung von Menschenleben.

Viele Eingebettete Systeme führen eine Aufgabe in wiederkehrender Weise aus und beeinflussen auf der Grundlage ihrer Berechnungen ein physisches System (z.B Regelsysteme). Zeiteffekte einer verteilten Rechenplattform, wie Latenzen, der Verlust von Nachrichten und Jitter, haben einen wesentlichen Einfluss auf das dynamische Verhalten des Regelsystems und können dazu führen, dass die geforderten Güteanforderungen unterschritten werden. Die Begrenzung der Latenzen durch eine Frist wie bei einem harten Echtzeitsystem erweist sich als ungeeignete Kenngröße, da für ein Regelsystem eine größere Latenz tolerierbar sein kann, wenn sie nur selten

und isoliert auftritt.
Ein Ziel der Arbeit ist es daher, Latenzen und den Nachrichtenverlust in Netzwerken durch ausdrucksstärkere und damit weniger pessimistische Modelle zu beschreiben. Dabei werden diese Zeiteffekte durch Dichtefunktionen beschrieben, welche für ein Intervall an benachbarten Ausführungsinstanzen das maximale und minimale Vorkommen der Zeiteffekte beschränken. Um Werte für diese Modelle aus einer gegebenen Rechenplattform ableiten zu können, sind Erweiterungen der bestehenden Methoden zur Echtzeitanalyse notwendig. In dieser Arbeit ist ein Ansatz skizziert, der solch einen neuen Aspekt der Echtzeitanalyse aufzeigt.

Der Entwurf verteilter Regelsysteme ist gegenwärtig so organisiert, dass der Entwurf der verteilten Rechenplattform und der des funktionalen Entwurfs isoliert voneinander betrachtet werden. Dabei beeinflussen Parameter, wie die auftretenden Zeiteffekte die funktionalen Sichtweise und die Plattformsicht. Der Trennung im Entwurf ist es jedoch geschuldet, dass Parameter beider Sichtweisen nicht vollumfänglich modelliert werden, sondern auf ein gemeinsames vereinfachtes Modell reduziert sind. So abstrahiert die maximale Latenz von dem abstrakten Zeitverhalten der Rechenplattform und macht eine tiefgehende Kenntnis des Aufbaus der verteilten Rechenplattform nicht erforderlich. Mit der Begrenzung von Zeiteffekte durch Dichtefunktionen soll eine Separation der Fachdisziplinen weiterhin möglich sein. Ebenso ist es für den systematischen Entwurf verteilter Regelsysteme Voraussetzung, dass die durch diese Modelle repräsentierten Zeiteffekte sich effizient speichern und verarbeiten lassen.

Ein Zielparameter beim Entwurf von eingebetteten Systemen ist eine möglichst hohe Auslastung der verfügbaren Rechen- und Kommunikationsressourcen. Eine hohe Auslastung resultiert in einer Zunahme der Interferenz der ausgeführten Aufgaben und hat im Allgemeinen durch größere Latenzen eine Degradierung des Regelverhaltens zur Folge. Die Bewertung des Regelverhaltens wird für gewöhnlich anhand von Stabilitätsmetriken durchgeführt. Aus der Erfüllung der Stabilitätseigenschaft lässt sich jedoch nicht der Grenzwert der im Regelkreis auftretenden Signale evaluieren. So existiert bisher kein geeignetes Modell, welches das schlimmste Verhalten eines physikalischen Systems in ausdrucksstarker Weise begrenzt. Eine Begrenzung des Regelverhaltens anhand einer Dichtefunktion, ähnlich der Zeiteffekte, ermöglicht so eine elegante Spezifikation des funktionalen Verhaltens und verknüpft durch die in der Arbeit vorgestellten Methoden das Verhalten der Signale in der funktionalen Sichtweise mit den Zeiteffekten der Plattform.

Die Arbeit ist wie folgt aufgebaut. Im nächsten Kapitel werden die verwendeten Begrifflichkeiten eingeführt und das Problem formuliert, welches im Rahmen der Arbeit gelöst wird. Das dritte Kapitel beschreibt die existierenden Modelle und Methoden zur Bestimmung und Beschreibung der unterschiedlichen Zeiteffekte, welche durch die Plattform verursacht werden. Des Weiteren wird das Reglermodell zur Beschreibung des dynamischen Systemverhaltens eingeführt. Im darauf folgenden Kapitel wird eine Übersicht über Arbeiten gegeben, die sich in die Thematik des Regler/Plattform Co-Designs eingliedern. Das fünfte Kapitel führt Modelle ein, welche die Zeiteffekte für Ereignisintervalle begrenzen. Ausgehend von den Eigenschaften der begrenzenden Kurven wird eine effiziente Speicherung dieser Kurven aufgezeigt und Methoden für den Einsatz in verteilten Systemarchitekturen dargestellt. Anschließend wird im sechsten Kapitel in äquivalenter Weise eine Begrenzung des dynamischen Signalverhaltens eingeführt. Diese Modelle bilden die Basis für Algorithmen, die eine funktionale Spezifikation des Regelverhaltens in eine Anforderung der Zeiteffekte einer Plattform transformieren. Beispielhafte Experimente veranschaulichen den Gebrauch und zeigen dabei den vorteilhaften Effekt durch die Nutzung der in der Arbeit vorgestellten Modelle und Methoden beim Entwurf eines verteilten Regelsystems. In der Entwurfsstudie in Kapitel sieben finden die mathematischen Beziehungen in einer kombinierten Betrachtung von Signalstörungen und dem Einfluss von Zeiteffekten auf das Regelsystem Anwendung. Die Arbeit endet mit einer Zusammenfassung.

Kapitel 2
Motivation und Ziele

2.1 Eingebettete Systeme, Echtzeitsysteme und Cyber-Physische Systeme

Ein eingebettetes System ist eine Recheneinheit, welche in einen technischen Kontext integriert ist. Für den Terminus existiert jedoch in der Fachwelt keine einheitliche Definition. So werden in (Lee, 2006) und (Slomka, 2008) eingebettete Systeme durch deren Einsatz in einem physikalischen oder technischen Kontext definiert. In den Definitionen von (Marwedel, 2007) und (Teich, 2013) liegt hingegen der Fokus auf einer spezialisierten Aufgabe, die solch ein System wahr nimmt und die es beispielsweise von der universellen Eignung eines Personal Computer unterscheidet. In Anlehnung an (Lee, 2006) wird hier die Definition eines eingebetteten Systems aufgeführt.

Definition 1 (Eingebettetes System). *Ein eingebettetes System ist ein Computersystem, welches in einen physikalischen Prozess integriert ist.*

Der Aufbau eines eingebetteten Systems umfasst einen Mikroprozessor, Speicher, anwendungsspezifische Sensoren und Aktoren und gegebenenfalls Kommunikationsschnittstellen. Dabei bilden Sensoren das reale Verhalten der Umgebung in eine digitale Repräsentation ab, welche dann von einem Rechner bearbeitet werden kann. Die Berechnungsergebnisse des eingebetteten Systems führen zu einer Reaktion, die über Aktoren auf des Umfeld einwirkt.
Die zunehmende Leistungsfähigkeit der elektronischen Bauteile bei fortschreitender

Miniaturisierung ermöglicht eine immer höhere Integration an Funktionalität auf einem eingebetteten System. Gegenstand der Forschung in den vergangenen Jahrzehnten ist es daher, diese Systemkomplexität unter den vorgegebenen Randbedingungen in einen optimalen Entwurf umzusetzen (Slomka et al., 2011). Anforderungen, aus denen sich Randbedingungen des Systementwurf ergeben sind:

- Echtzeit
- Temperatur / Abwärme
- Leitung / Energie
- Fläche / Bauform / Größe / Gewicht
- Zuverlässigkeit
- Sicherheit
- Kosten

Diese teilweise in Widerspruch stehenden Anforderungen an ein eingebettetes System spannen einen mehrdimensionalen Entwurfsraum auf, aus dem durch eine Mehrkriterienoptimierung Lösungen gesucht werden (Slomka, 2002).
Eine herausragende Stellung unter den oben genannten Anforderungen nimmt die Eigenschaft der Echtzeit ein, da sie für einen Großteil eingebetteter Systeme von Bedeutung ist und deren Verifikation einen eigenständigen Forschungsbereich darstellt . DIN 44300 definiert den Begriff der Echtzeit wie folgt.

Definition 2 (*Echtzeitsystem* (Scholz, 2005)). *„Unter Echtzeit versteht man den Betrieb eines Rechensystems, bei dem Programme zur Verarbeitung anfallender Daten ständig betriebsbereit sind, derart, dass die Verarbeitungsergebnisse innerhalb einer vorgegebenen Zeitspanne verfügbar sind. Die Daten können je nach Anwendungsfall nach einer zeitlich zufälligen Verteilung oder zu vorherbestimmten Zeitpunkten anfallen.“*

Differenziert man darüber hinaus zwischen den Auswirkungen einer überschrittenen Zeitspanne, so ergibt sich die Unterscheidung in weiche und harte Echtzeit. Bei harten Echtzeitsystemen hat eine Fristverletzung desaströse Konsequenzen, während Systeme mit weichen Echtzeitanforderungen mit zunehmender Latenz und Häufigkeit der Fristverletzungen an Leistungsfähigkeit einbüßen (Liu, 2000).

Die Fragestellungen in Zusammenhang mit den Echtzeiteigenschaften eines eingebetteten Systems finden in der Plattformsicht statt.

Definition 3 (Plattform und Plattformsicht). *Der Begriff Plattform bezeichnet die Komponenten, die ein eingebettetes System bereitgestellt, damit dieses die geforder-*

te Funktionalität umsetzen kann. Dieses sind zum einen die Hardwarekomponenten wie Prozessor und Speicher, als auch die Netzwerkkomponenten eines verteilten Rechensystems. Des Weiteren umfasst die Plattform die Infrastrukturkomponenten eines Betriebssystems. In der Plattformsicht werden Aufgaben den Netzerk- und Rechenressourcen zugewiesen und deren Ablauf geplant. Dies geschieht anhand von Modellen, welche von den umgesetzten Funktionalitäten abstrahieren.

Durch die im nächsten Kapitel vorgestellten Modelle der Plattformsicht kann das Zeitverhalten der ausgeführten Aufgaben eines Echtzeitsystems bewertet werden. Einen anderen Blick auf das eingebettete System liefert die funktionale Sichtweise.

Definition 4 (Funktionale Sichtweise). *In der funktionalen Sichtweise werden die auf dem eingebetteten System ausgeführten Algorithmen entworfen und die auftretenden Signalverläufe betrachtet. In dieser Sichtweise findet eine Abstrahierung der Plattform dahingehend statt, dass diese als ideal angenommen wird. Dies bedeutet, dass in der funktionalen Sichtweise unter anderem die Zeiteffekte der Plattform vernachlässigt oder zu einem konstanten Wert angenommen werden.*

Für ein Regelsystem findet in der funktionalen Sichtweise die Entwicklung der Regelalgorithmen und das Auslegen der Regelparameter statt.

Die Trennung der Sichtweisen reduziert die Komplexität beim Entwurf eingebetteter Systeme und ermöglicht die Separation der beim Entwurf involvierten Disziplinen. Für Anwendungen, bei denen eine enge Kopplung zwischen dem funktionalen Verhalten eines physischen Systems und der Plattformsicht besteht, so dass beim Entwurf eine integrierte Betrachtung vorgenommen werden muss, setzt sich zunehmend der Begriff des Cyber-Physischen Systems durch (Lee, 2008).

Definition 5 (Cyber-Physisches System (Lee, 2008)). *Ein Cyber-Physisches System (CPS) ist der Zusammenschluss von digitalen Berechnungen und einem physikalischen Prozess. Eingebettete Computer und Netzwerke überwachen und steuern den physikalischen Prozess, für Gewöhnlich über eine Rückkopplung bei der sich der physikalische Prozess und die Berechnungen gegenseitig beeinflussen.*

Die Abgrenzung zu den eingebetteten Systemen ergibt sich durch die Reduktion eines eingebetteten Systems auf die Betrachtung seiner Rechenkomponenten und damit auf die Plattformsicht. Es existiert jedoch eine sehr weitläufige Verwendung des Begriffs CPS, die weit über die Verwendung in dieser Arbeit hinaus geht. So wird CPS häufig im Zusammenhang mit den Schlagwörtern Smart Grid und Industrie

4.0 genannt, deren Kernbestandteil ein großräumiges Kommunikationsnetzwerk ist. Hierbei liegen die Herausforderungen in der Datenhaltung und -bereitstellung. Im Kontext dieser Arbeit werden Kommunikationsnetzwerke auf deren Zeiteffekte reduziert.

Macht es die räumliche oder logische Separation von Funktionalitäten notwendig, Aufgaben auf mehrere Recheneinheiten aufzuteilen, führt dies zu einer verteilten Systemarchitektur. Bei solch einer verteilten Architektur erhöht sich die Komplexität zur Modellierung und Verifikation eines Systems, da die Ausführung einer Anwendung über ein Kommunikationsnetz hinweg Auswirkungen auf das Verhalten entfernter Recheneinheiten haben kann.

2.2 Verteilte Systemarchitektur

Stand der Technik beim Entwurf komplexer CPS ist die Vernetzung mehrerer unabhängiger Steuergeräte. Eine derartige Architektur hat sich historisch aus der zunehmenden Verknüpfung einzelner Funktionen, als auch der räumlichen Trennung von Sensoren, Aktoren und Recheneinheiten ergeben. Ein gemeinsam genutztes Kommunikationsnetz zur Anbindung der Komponenten reduziert zudem die Anzahl an Leitungen, was zu einer Minderung der Kosten und des Gewichts führt. Durch eine verteilte Systemarchitektur werden die Anforderungen an das System flexibel und kostengünstig umgesetzt. Die Implementierung einer verteilten Systemarchitektur führt zu hohen Anforderungen an den Datendurchsatz des Verbindungsnetzes zwischen den Steuergeräten. So erfordern viele Anwendungen, die auf verteilten Systemarchitekturen ausgeführt werden, Echtzeiteigenschaften, welche die Verbindungsnetze als auch die Recheneinheiten der Steuergeräte erfüllen müssen. Beispielhaft für derartige Systeme sind Anti-Blockier-Systeme, Motorsteuergeräte oder elektrohydraulische Bremssysteme im Automobilbereich (Cullyer, 1991).

Um die Rechtzeitigkeit von Signalen und Nachrichten in verteilten Systemen zu garantieren, werden zeitgesteuerte Zugriffsverfahren wie Time Division Multiple Access (TDMA) eingesetzt. Solche Verfahren gewährleisten eine deterministische Zustellzeit, skalieren jedoch unflexibel in der Nachrichtenanzahl und -länge. Ein grundlegendes Verfahren des zeitgesteuerten Zugriffs ist die Time-Triggered Architecture (TTA). Bei der TTA wird der Ausführungszeitpunkt von Aufgaben über eine globale Zeitbasis gesteuert (Kopetz, 2011) und bildet so durch vorhersagbare Ausführungszeitpunkte eine einfach zu beherrschende Plattform für sicherheitskritische

Systeme. Der Grundgedanke solch einer Implementierung ist die oben beschriebene Trennung von funktionalem Entwurf und Plattformentwurf. Durch die feste Zuordnung von Aufgaben und Nachrichten auf die von der Plattform deterministisch bereitgestellten Zeitschlitze, findet keine zeitliche Beeinflussung ausgeführter Aufgaben statt.
Bei der Bestimmung der Zeitschlitzlänge von Aufgaben, wird deren Dauer so gewählt, dass sie der maximalen Ausführungszeit (worst case execution time, WCET) der Aufgabe entspricht. Wird dabei der für eine Aufgabe bereitgestellte Zeitschlitz nicht in vollem Umfang beansprucht, bleib diese freie Ressource ungenutzt. Daraus folgt eine schlechte Auslastung einer TTA. Des Weiteren führt eine feste Zuordnung bei der TTA dazu, dass die TTA nur in geringem Umfang skaliert und ausschließlich periodische angeregte Funktionen unterstützt werden. Diese Voraussetzung ist für klassische Regel- und Steuersysteme mit periodischer Abtastung gewährleistet. Bei vielen Anwendungen ist diese Annahme jedoch nicht zulässig. So erzeugen beispielsweise rotierende Systeme bei Überschreitung eines bestimmten Winkels ein Ereignis, welches dann eine Messung oder eine Berechnung veranlasst. In solch einem Fall ergibt sich eine von der Winkelgeschwindigkeit abhängige Aktivierung. Auch bei ereignisbasierten Regelsystemen veranlasst der Sensor nur dann eine Neuberechnung des Stellwerts, wenn sich der Sensorwert ausreichend stark verändert (Heemels et al., 2008). Untersuchungen der Anregung solcher Systeme in Abhängigkeit der Reglerdynamik wurden in (Bund et al., 2013) durchgeführt.

Den zeitgesteuerten Echtzeitsystemen stehen die ereignisgesteuerten Systeme gegenüber. In der gängigen Literatur ist es strittig, welcher dieser Ansätze besser geeignet ist, um Echtzeitsysteme zu implementieren. Arbeiten, welche den zeitgesteuerten Ansatz für verteilte eingebettete Systeme bevorzugen sind (Kopetz und Bauer, 2003) und (Kopetz und Grünsteidl, 1993). In den Arbeiten von (Kopetz, 1991) und (Kopetz, 2011) werden die unterschiedlichen Entwurfsparadigmen hinsichtlich ihrer Vorteile und der daraus resultierenden Anwendungen untersucht. In dieser Arbeit wird von einer ereignisgesteuerten oder heterogenen Architektur ausgegangen, da diese im Allgemeinen zu einer höheren Systemauslastung und damit zu einem kostenoptimierten Entwurf führt. Werden sicherheitskritische Applikationen auf solch einer Plattform ausgeführt, ist es zwingend notwendig zu überprüfen, ob die spezifizierten Zeitfristen eingehalten werden. Eine Übersicht der Verifikationsmethoden eines verteilten Echtzeitsystems wird im Kapitel 3.1.4 gegeben.

2.3 Regelsystem auf verteilter Plattform

Eine Anwendung vieler CPS ist die Regelung eines physikalischen Prozesses. Befinden sich Sensoren und Aktoren des Regelkreises in räumlicher Distanz zueinander und sind diese in verschiedene Funktionen eingebunden, ergibt sich daraus ein verteiltes Regelsystem wie in Abbildung 2.1 skizziert. Dargestellt ist die Abstandsregelung eines Fahrzeuges zu einem vorausfahrenden Verkehrsteilnehmer. Solch eine Funktion ist in modernen Oberklassefahrzeugen umgesetzt und übernimmt Teile der monotonen Aufgabe des Fahrens in einer Kolonne. Ein Sensor stellt dem Regler den Abstand (Regelgröße) in digitalisierter Form bereit, so dass der Regler durch Vergleich mit einem geschwindigkeitsabhängigen Sollabstand ein Stellsignal berechnet. Durch Anwendung dieses Stellsignals durch die Aktoren wird eine Änderung der Fahrzeugdynamik veranlasst. Das zu regelnde dynamische System wird als Regelstrecke bezeichnet.
Vergrößert sich der Abstand zum vorausfahrenden Fahrzeug, so beschleunigt die Regelung das Fahrzeug selbstständig bis zu einem vorgegebenen Grenzwert. Verringert sich hingegen der Abstand, wird ein Bremsvorgang eingeleitet um eine Kollision zu vermeiden und den geforderten Sicherheitsabstand aufrecht zu halten. Kommt es bei den vorausfahrenden Verkehrsteilnehmern zu einem unerwarteten Zwischenfall, der

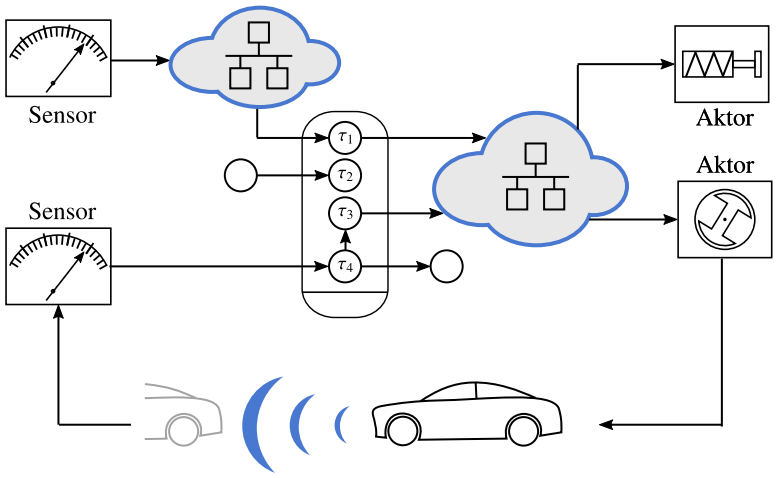


Abb. 2.1: Aufbau eines verteilten Regelsystems mit gemeinsam genutzten Rechen- und Kommunikationsressourcen.

eine Notbremsung erfordert, muss die Regelung einen Aufprall verhindern, woraus sich Echtzeitanforderungen ableiten lassen und sich ein sicherheitskritischer Kontext ergibt.
Der Sensor zur Bestimmung des Abstandes zu einem vorausfahrenden Fahrzeug befindet sich im vorderen Teil des PKW. Die für einen Bremsvorgang zuständigen Aktoren und die Steuergeräte der involvierten Assistenzsysteme sind an unterschiedlichen Stellen im Fahrzeug positioniert. Diese gemeinsame Nutzung von Systemressourcen führt zu Interferenzen in der Ausführung, welche in Zeiteffekten resultieren, die bei der Verifikation des Regelsystems beachtet werden müssen.

2.3.1 Digitale und verteilte Regelsysteme

In der gängigen Regelungstheorie werden idealisierte Annahmen der Plattform, auf welcher die Regelalgorithmen ausgeführt werden, angenommen. Dadurch ist es möglich, von einer konkreten Umsetzung einer Plattform zu abstrahieren und die Regelungstechnik isoliert davon zu betrachten. Dies impliziert die Annahme, dass die Rechen- und Netzwerkressourcen der Reglerfunktionalität exklusiv zur Verfügung stehen. Es wird auch die Ausführungszeit der Algorithmen zu Null angenommen, was einer unendlich großen Rechenleistung entspricht. In den gängigen Lehrbüchern wird darauf verwiesen, eine Plattform zu wählen, bei der sich Latenzen deutlich kleiner als die Abtastzeit der Regelung ergeben und damit vernachlässigt werden können. Die Abtastzeit T_s ist dabei definiert als die Zeitspanne zwischen der Erfassung zweier Sensorwerte.
Die grundlegenden Theorien zum Entwurf von Reglern stammen aus einer Zeit, in der es keine Mikrorechner gab. Entsprechend wurden dynamische Systeme über mechanische und elektrische Komponenten geregelt, die ein zeitkontinuierliches Verhalten aufweisen. Eine weit verbreitete Methode zur Auslegung und Implementierung eines Reglers auf einem Rechensystem ist daher der quasikontinuierliche Reglerentwurf. Dabei werden die zeitlichen Quantisierungseffekte, die bei der zeitdiskreten Abtastung auftreten vernachlässigt. Diese Annahme ist gültig für Systeme, die weit häufiger, als mit der größten im geschlossenen Regelkreis vorkommenden Frequenz abgetastet werden. Üblicherweise wird bei solchen Systemen eine Abtastfrequenz mit 10 bis 30-facher Überabtastung gewählt (Åström und Wittenmark, 2011a). Die Umsetzung solch einer Daumenformel führt offensichtlich zu einem ineffizienten Entwurf, da die Abtastfrequenz einen wesentlichen Einfluss auf den Ressourcen-

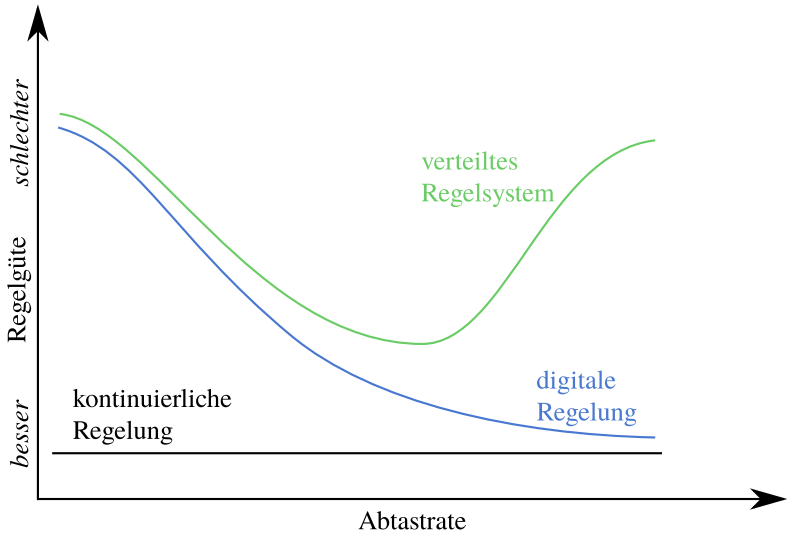


Abb. 2.2: Vergleich der Regelgüte für kontinuierliche Regelung, digitales Regelsystem und verteiltes Regelsystem als Funktion der Abtastrate (Moyne und Tilbury, 2007).

verbrauch hat, indem durch sie die Rate an Nachrichtenübermittlungen und Taskausführungen bestimmt wird.

Durch die mathematische Transformation der Regelstrecke von zeitkontinuierlichen Differentialgleichungen in ein Differenzengleichungssystem über den Mechanismus eines Abtast-Halte-Glieds werden die Signalwerte zu den Abtastzeitpunkten exakt nachgebildet (Franklin et al., 1998). Diese zeitdiskrete Repräsentation der Regelstrecke kann um einen diskreten Regler erweitert werden, um dann wie in (Powell und Katz, 1975) gezeigt, die Auswirkungen der gewählten Abtastzeit hinsichtlich Regelverhalten, Störverhalten und Unsicherheiten der Parameter untersucht werden. Während eine hohe Abtastrate eine hohe Systemauslastung verursacht, nähert sich die Regelgüte des digitalen Regelkreis mit zunehmender Abtastrate dem kontinuierlichen Regelverhalten an. Dieser Zusammenhang ist in Abbildung 2.2 dargestellt. Mit dem Anstieg der Abtastrate eines verteilten Regelsystems muss bei einer Plattform mit begrenzten Ressourcen davon ausgegangen werden, dass durch die Zunahme der Systemauslastung die im vorherigen Abschnitt vorgestellten Zeiteffekte zunehmen. So existiert für verteilte Regelsysteme ein Wendepunkt, ab dem eine weitere

Zunahme der Abtastfrequenz, begründet durch eine zunehmende Latenz, eine Herabsetzung der Regelgüte zur Folge hat. Für ein Regelsystem mit gewählter Plattform kann aus einer Untersuchung des verteilten Regelsystems wie in Abbildung 2.2 die optimale Abtastrate bestimmt werden.

2.3.2 *Klassifizierung von Regelanwendungen*

Cyber-physische Systeme lassen sich abhängig von ihren Anwendungen in verschiedene Bereiche einordnen. In Abbildung 2.3 ist eine Aufteilung bezüglich Sicherheitskritikalität und dem Ressourcenverbrauch /-auslastung der Plattform vorgenommen und für die jeweiligen Anforderungen eine typische Regleranwendung skizziert. Für eine kostenoptimierte Wahl der Plattform muss ein möglichst geringer Verbrauch der Rechen- und Netzwerkressourcen angestrebt werden. Zusätzlich ist es erstrebenswert

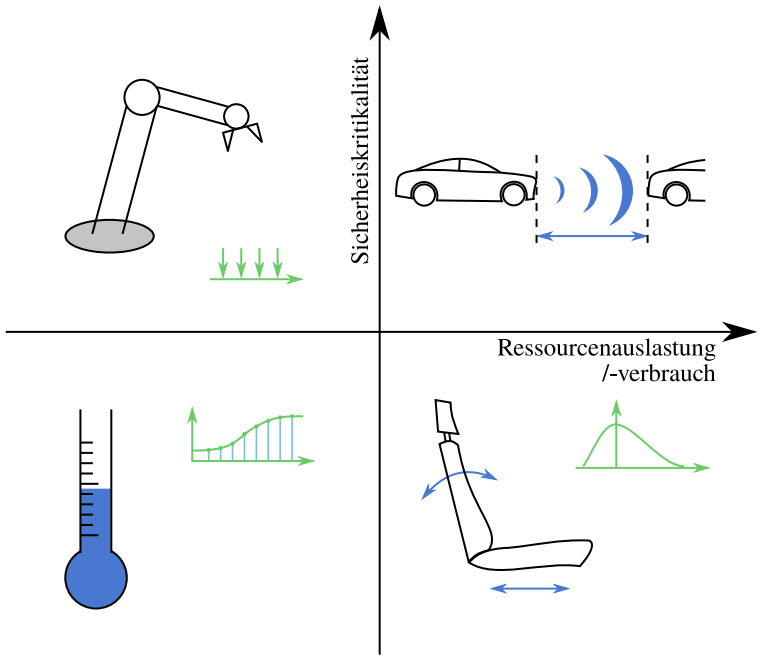


Abb. 2.3: Klassifizierung von Regelsystemen nach Sicherheitskritikalität und Ressourcenverbrauch /-auslastung.

die Auslastung der verfügbaren Ressourcen zu maximieren um die eingesetzte Plattform bestmöglich auszuschöpfen.

In dem unteren linken Quadranten ist die Regelung einer Raumtemperatur mit Überabtastung der Regelgröße angedeutet. Die Überabtastung führt zu einer Robustheit gegenüber Störungen und Zeiteffekten der Plattform. Da hierbei öfter als durch die im vorherigen Abschnitt beschriebene Wahl der Abtastzeit eine Berechnung veranlasst wird, werden mehr Ressourcen als notwendig konsumiert. Bei dem trägen Verlauf der Raumtemperatur ist eine Überabtastung schon bei geringem Ressourcenverbrauch möglich. Auf eine gesonderte Untersuchung der Plattformeffekte wird bei solchen Regelapplikationen verzichtet, da keine sicherheitsrelevanten Aufgaben erfüllt werden.
Ist es hingegen notwendig eine kostenoptimierte Systemplattform zu entwickeln, müssen die auftretenden Zeiteffekte ermittelt und beim Entwurf des Reglers berücksichtigt werden. So ist in dem rechten unteren Quadrant in Abbildung 2.3 eine Komfortanwendung aus dem automobilen Umfeld skizziert. Um eine möglichst hohe Auslastung der Plattform zu erreichen, werden stochastische Werte über das Auftreten von Zeiteffekten in die Entwurfsmethoden des Reglers integriert. Es gibt eine Vielzahl an Arbeiten, die solch eine Vorgehensweise beschreiben. Vertretend sei an dieser Stelle die Forschungsgruppe Automatic Control der Universität Lund genannt. Ihre Arbeiten resultierten in einer MATLAB Toolbox, die Regelkreise unter dem Einfluss von stochastisch modellierten Signallatenzen bewertet (Cervin, 2003; Lincoln und Cervin, 2002). Eine ausführliche Übersicht existierender Arbeiten findet sich in Kapitel 4.

Sind Garantien über das schlimmstmögliche Regelverhalten gefordert, eignet sich eine deterministische Ausführung, wie die im vorherigen Abschnitt vorgestellte TTA. Das Paradigma der TTA für eingebettete Regelsysteme in verteilten Architekturen wird unter anderem in dem an der Universität Berkeley entwickelten Giotto-Framework (Henzinger et al., 2003, 2001) umgesetzt. Durch eine garantierte Ressourcenverfügbarkeit kann jederzeit das geforderte Maß an Regelgüte gewährleistet werden. Jedoch müssen wie oben erläutert die schlechte Skalierbarkeit und die geringe Auslastung in Kauf genommen werden. Ein Beispiel, bei dem solch eine deterministische Ausführung für die Reglerfunktionen gewählt wird, sind industrielle Anwendungen wie im linken oberen Quadranten skizziert. Eine mit der geringen Auslastung einhergehende Kostenerhöhung sowie eine Zunahme von Leistungsaufnahme und des Gewichts des CPS ist für industrielle Anwendungen von untergeordneter Wichtigkeit. Weitere Anwendungsfälle, die sich diesem Quadranten zuordnen

lassen, sind Plattformen zur prototypischen Implementierung und Erprobung von Regelalgorithmen. Dabei steht eine schnelle Umsetzung der Algorithmen und der Test des Regelsystems im Vordergrund. Der Einfluss von Zeiteffekten durch eine geringer dimensionierte Plattform ist in diesem Anwendungsfall unerwünscht.

Anwendungen bei denen ein starker Fokus auf eine kostenoptimierte Umsetzung und dadurch auf eine hohe Plattformauslastung gelegt wird und die zusätzlich sicherheitskritische Anforderungen aufweisen sind Regelungen der Fahrzeugdynamik. So ist im rechten oberen Quadranten in Abbildung 2.3 eine Abstandsregelung skizziert. Auch Anwendungen, die mit dem Schlagwort Industrie 4.0 versehen sind, können diesem Quadranten zugeordnet werden. Bei der zunehmenden Bestrebung konfigurierbare Produktionsanlagen zu entwerfen, erweist sich die statische Struktur einer TTA als ungeeignet. Auch der Trend zur Verwendung von Komponenten der drahtlosen Kommunikation und IP-basierten Netze resultiert in einer ereignisbasierten Ausführung. Zusammen mit dem Anspruch an eine hohe Ressourcenauslastung führt die zunehmende Interferenz in der Ausführung von Berechnungen und Nachrichtenübermittlungen zu nicht vernachlässigbaren Zeiteffekten. Es sind daher für die Verifikation des dynamischen Verhaltens garantierte Grenzen dieser Zeiteffekte zu bestimmen, um anhand dieser den schlimmsten Fall des physikalischen Systemverhaltens abzuleiten. Eine ausführliche Übersicht der existierenden Arbeiten zur Beurteilung eines Regelsystems aus den Grenzen des Zeitverhaltens wird in Kapitel 4 vorgenommen.

2.4 Ziele der Arbeit

Die vorliegende Arbeit lässt sich fachlich in den rechten oberen Quadranten aus Abbildung 2.3 einordnen. Dieser Quadrant repräsentiert das Spannungsfeld aus einer effizienten und kostengünstigen Umsetzung und den deterministischen Anforderungen in einem sicherheitskritischen Umfeld. Der Fokus dieser Arbeit liegt dabei auf der Verifikation einer geforderten Regelgüte unter dem Einfluss der Zeiteffekte einer verteilten Systemarchitektur.
Beim Entwurf eines verteilten Regelsystems ist dieser Schritt am Ende des Entwurfsflusses angegliedert. So kann in den vorausgehenden Schritten der Regler optimal auf das typische Zeitverhalten ausgelegt werden (vgl. Abschnitt 4.2.1).
Ein Beitrag dieser Arbeit für eine effiziente Implementierung ist die Beschränkung des Zeitverhaltens durch eine Kurvenfunktion, die für jedes Ereignisintervall das

maximale Vorkommen angibt (vgl. Kapitel 5). Durch das im Vergleich zu existierenden Beschreibungsformen ausdrucksstärkere Modell wird bei der Verifikation eines Regelsystems (vgl. Kapitel 6) der Pessimismus reduziert. So kann dadurch ein Regelsystem die Güteanforderungen erfüllen, ohne den Reglerkomponenten weitere Ressourcen einzuräumen.

Für den Einsatz in einem sicherheitskritischen Umfeld muss der schlimmste Fall des Regelverhaltens abgeleitet werden. Es existieren Methoden der Echtzeitanalyse, welche garantierte Schranken des Echtzeitverhaltens verteilter Systeme ableiten (vgl. Unterabschnitt 3.1.4). Diese Arbeit erweitert diese Forschungsarbeiten durch ein ausdrucksstärkeres Modell, gewährleistet aber weiterhin eine valide Grenze des Zeitverhaltens. Für die Bewertung des schlimmstmöglichen Regelverhaltens wird in Abschnitt 6.1 eine neue Bewertungsvorschrift eingeführt und in den darauf folgenden Abschnitten der Zusammenhang zu dem Modell der Zeiteffekte hergestellt.

Vorteilhaft an den Modellen ist, dass die Untersuchung des Zeitverhaltens der Plattform und die Verifikation des Regelverhalten isoliert voneinander betrachtet werden können. Jedoch soll aus der Anwendung der entwickelte Modelle und Methoden ein flexiblerer und effizienterer Co-Entwurf resultieren.

Kapitel 3
Modellierung des Zeitverhaltens und des funktionalen Verhaltens

Im vorherigen Kapitel wurde zwischen der Plattformsicht und der funktionalen Sichtweise eines eingebetteten Systems unterschieden. Diese in Abbildung 3.1 für ein Regelsystem dargestellten Sichtweisen ergeben sich aus den am Entwurf beteiligten Fachdisziplinen. Entsprechend den Entwurfsaufgaben, die in der jeweiligen Sichtweisen vorgenommen werden, existieren unterschiedliche Modelle, welche die

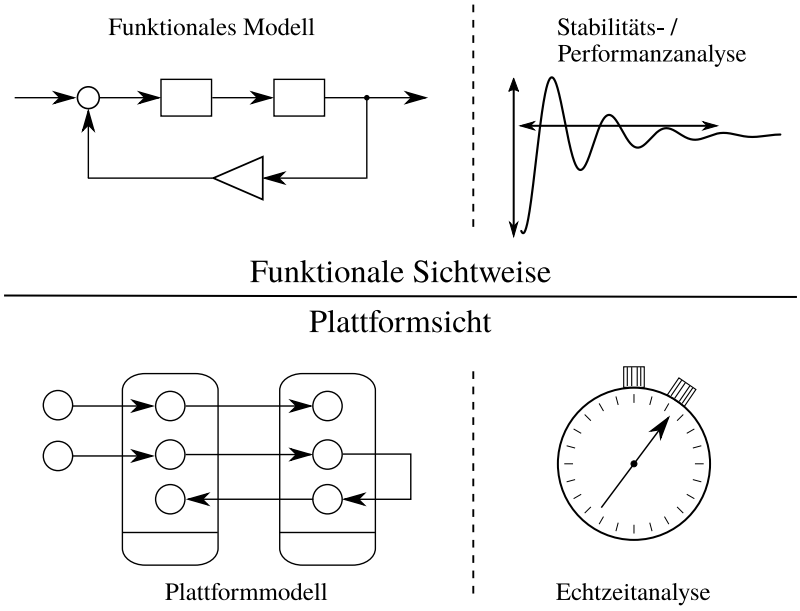


Abb. 3.1: Sichtweisen eines verteilten Regelsystems.

für die Sichtweise relevanten Zusammenhänge darstellt. Für einen unmissverständliche Verwendung der Begrifflichkeiten, werden im Folgenden die Modelle der Plattformsicht und der funktionalen Sichtweise definiert. Zeiteffekte können als Bindeglied dieser Modelle angesehen werden, da sie aus dem Zusammenwirken mit den in der Plattformsicht betrachteten Tasks und Ressourcen resultieren, aber auch das Verhalten der Funktionalen Sichtweise beeinflussen.

3.1 Plattformmodell

Um einen Systementwurf bewerten zu können, benötigt man Modelle die das Zeitverhalten der Plattform beschreiben und anhand derer sich die auftretenden Zeiteffekte evaluieren lassen. In diesem Kapitel werden Plattformmodelle vorgestellt, die dazu von dem konkreten Zeitverhalten einer Hard- und Softwareimplementierung abstrahieren. Basierend auf diesen Modellen werden existierende Methoden zur Bestimmung von Latenzen in verteilten Systemarchitekturen vergleichend dargestellt.

3.1.1 Taskmodell

Wie in Definition 3 festgelegt, besteht die Plattform aus den Infrastrukturelementen, welche zur Ausführung von Rechenaufgaben notwendig sind. Diese Aufgaben werden in Tasks unterteilt, welche dann auf der Pattform ausgeführt werden.

Definition 6 (Task). *Die in Software implementierten Aufgaben werden als Tasks τ_i bezeichnet und bilden eine Ausführungseinheit bestehend aus einer Menge an Rechenoperationen. In der Plattformsicht ist jede Task einer Ressource zugeordnet, auf der sie ausgeführt wird. Ein Rechentask wird auf einer Recheneinheit ausgeführt, während Nachrichten in einem Netzwerk von Kommunikationstasks repräsentiert werden.*

Hinsichtlich des Zeitverhaltens besteht kein Unterschied zwischen einem Rechentask und einem Kommunikationstask, was eine weitere Unterscheidung hinfällig macht und im Weiteren beide als Task bezeichnet werden. Da Tasks für gewöhnlich mehrfach ausgeführt werden, können die jeweiligen Taskinstanzen oder Jobs einzeln benannt werden. Der Ausdruck $\tau_{k,i}$ entspricht somit der k-ten Ausführung des Task τ_i.

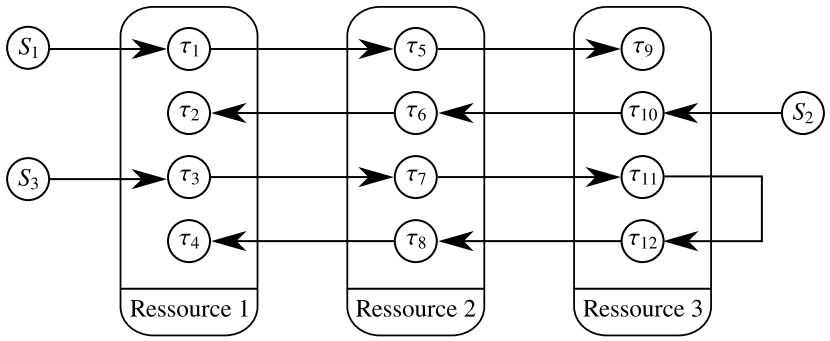


Abb. 3.2: Plattformmodell einer verteilten Systemarchitektur (basierend auf (Künzli et al., 2007)).

Dieser Task τ_i wird zu den Zeitpunkten $t^r_{k,i}$ aktiviert und beendet seine Aufgabe zu den Zeitpunkten $t^e_{k,i}$.

Verfügt die Task τ_i exklusiv über die Ressource, wird die Differenz dieser Zeitpunkte als Nettoausführungszeit $c_{k,i}$ bezeichnet. Die Grenzen dieser Nettoausführungszeit werden minimale und maximale Nettoausführungszeit (engl. best-case / worst-case execution time) (BCET / WCET, c^-, c^+) benannt und können durch eine intrinsische Taskanalyse bestimmt werden (Wilhelm et al., 2008). Für diese Grenzen gilt

$$c_i^- \leqslant \min_k \{c_{k,i}\}, \quad c_i^+ \leqslant \max_k \{c_{k,i}\}.$$

Durch die Mehrfachbelegung von Ressourcen in einem Multitasksystem, wie sie in Abbildung 3.2 skizziert ist, vergrößert sich die Nettoausführungszeit der Tasks durch die Interferenz ihrer Ausführungszeiten. Die Differenz der Zeitpunkte aus Taskterminierung und Taskaktivierung, welche die Unterbrechungszeiten eines Multitasksystems beinhaltet, wird als Antwortzeit $r_{k,i}$ bezeichnet.

Definition 7 (Antwortzeit). *Bezeichnet man mit $t^r_{k,i} \in \mathbb{R}^+_0$ den Zeitpunkt der k-ten Aktivierung des Tasks τ_i und mit $t^e_{k,i} \in \mathbb{R}^+_0$ den Endzeitpunkt (siehe Unterabschnitt 3.1.1), an dem der Task τ_i vollständig abgearbeitet wurde, dann definiert sich Antwortzeit des k-ten Jobs durch*

$$r_{k,i} = t^e_{k,i} - t^r_{k,i}$$

mit $r_{k,i} \in \mathbb{R}^+_0$.

In Abbildung 3.3 ist der Ablaufplan eines Systems mit 2 Tasks skizziert, die eine konstante Nettoausführungszeit aufweisen. Task τ_1 besitzt in dem Beispiel die

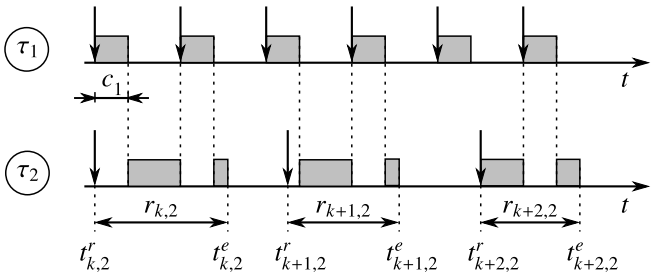


Abb. 3.3: Antwortzeit eines Multitasksystems.

höhere Priorität, so dass er unterbrechungsfrei ausgeführt wird. Seine Antwortzeit entspricht somit seiner Nettoausführungszeit. Für den Task τ_2 ergeben sich variierende Antwortzeiten, da ihm nur die durch Tasks τ_1 verbleibende Rechenzeit zur Verfügung steht. Die Antwortzeit eines Tasks ist daher abhängig von dessen Aktivierungshäufigkeit, der Nettoausführungszeiten und der Zuordnung der Prioritäten. Sind bei einem ereignisgesteuerten Ablauf mehrere Tasks bereit zur Ausführung, wird derjenige Task ausgeführt, welcher die höchste Priorität zugeordnet ist. Methoden, die aus diesen Parametern die Antwortzeiten und deren Grenzwert (r_i^+) (Worst Case Response Time, WCRT) bestimmen, werden im folgenden Abschnitt beschrieben.

Wie in Abbildung 3.2 dargestellt, können in einem verteilten System die Ressourcen und deren zugeordneten Tasks nicht isoliert betrachtet werden, da Pfade den Ausführungszeitpunkt über Ressourcen hinweg steuern.

Definition 8 (Ausführungspfad). *Ein Ausführungspfad P verkettet die Aktivierung von Tasks durch Datenabhängigkeiten. Bei den durch einen Pfad verbundenen Tasks wird bei einer Taskterminierung der darauf folgende Task ausgeführt. Durch sich verzweigende Pfade ist es möglich, Nebenläufigkeit zu modellieren.*

In dem verteilten System aus Abbildung 3.2 sind Ausführungspfade durch Pfeile dargestellt. Für ein verteiltes Regelsystem ergibt sich der Ausführungspfad aus dem Signalpfad des Reglers. Sensoren bilden durch den von ihnen umgesetzten Übergang der analogen zur digitalen Systembetrachtung den Beginn eines Ausführungspfades. Entsprechend der Abtastrate wird ein Sensor als Ereignisquelle modelliert (Bund et al., 2010). Eine Verarbeitung des Sensorwerts beispielsweise in Form einer Filterung kann als angegliederter Task aufgefasst werden.

Definition 9 (Ereignis). *Ein Ereignis bewirkt die Ausführung eines Tasks. Der Zeitpunkt, an dem ein Ereignis eintrifft, entspricht somit dem Zeitpunkt der Aktivierung* $(t^r_{k,i})$*. In einem Ausführungspfad wird zum Zeitpunkt einer Taskaktivierung* $(t^e_{k,i})$ *ein Ereignis generiert.*
Bei der digitalen Signalverarbeitung nimmt ein Task zu Beginn seiner Ausführung Eingabedaten entgegen und stellt das Berechnungsergebnis bei Abschluss der Verarbeitung bereit. Sind in solch einem Fall die Daten an ein Ereignis gebunden, werden sie als Signale bezeichnet. Werden die Daten der Berechnung verworfen, wird dies im Folgenden Signal- oder Ereignisverlust genannt.

Bezogen auf Abbildung 3.2 könnte der Pfad von S_1 bis τ_9 einen Regler repräsentieren. Die Ereignisquelle S_1 spezifiziert die Häufigkeit, mit welcher der Pfad ausgeführt wird. Bei Terminierung des Aktortasks τ_9 wird das physikalische System durch einen Stellwert beeinflusst. Zu diesem Zeitpunkt wurde der Ausführungspfad vollständig durchlaufen. Die Zeitspanne seit Aktivierung des Sensortask wird als Ende-zu-Ende Latenz oder Pfadlatenz bezeichnet. Für eine Ausführung mit nur einem zu betrachtenden Ausführungspfad ist die Latenz durch die Summe aller Antwortzeiten des Pfades gegeben.

3.1.2 Ereignismodelle

Einen wesentlichen Einfluss auf die Antwortzeit eines Tasks hat die Häufigkeit der Taskausführungen. Grundlage der Echtzeitanalyse eines Systems ist daher ein Ereignismodell, welches die Häufigkeit der Taskanregungen begrenzt. So entspricht in diesem Modell ein Ereignis einer Taskaktivierung. Im Folgenden wird eine Auswahl der relevanten Ereignismodelle vorgestellt, mit denen unterschiedlich komplexe Ereignissequenzen beschrieben werden können. Dies ist auch der Entwicklung von Methoden zur Echtzeitanalyse geschuldet. So reduziert sich mit zunehmender Ausdrucksstärke der Ereignismodelle die Überabschätzung der Resultate einer Echtzeitanalyse.

Periodisch

Das periodische Ereignismodell wurde im Rahmen der Planbarkeitsanalyse von Liu und Layland eingeführt (Liu und Layland, 1973). Dabei wird angenommen, dass

alle Tasks des Systems strikt periodisch mit der Periode p angeregt werden (Abb. 3.4). Diese Annahme ist für technische Systeme gültig, bei denen Tasks auf einem Monoprozessor ausgeführt werden und dabei durch einen periodisch abgetasteten Sensorwert angeregt werden. Es ist auch möglich, sporadische Taskausführungen mit dem periodischen Ereignismodell zu modellieren, indem die minimale Zeitspanne zwischen zwei Ereignissen als Periode angenommen wird.

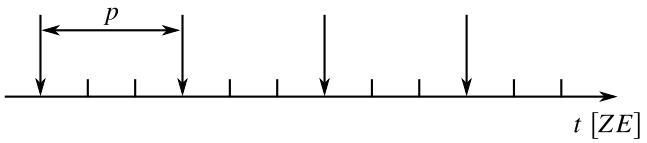


Abb. 3.4: Strikt periodisches Ereignismodell.

Periodisch mit Jitter und minimalem Abstand

Um die mangelnde Ausdrucksstärke des periodischen Modells zu erweitern, wurde dieses in (Tindell und Clark, 1994) um den Jitter und einen minimalen Abstand erweitert. Diese Erweiterung trägt dem Umstand Rechnung, dass die Annahme einer strikt periodischen Taskanregung das reale Verhalten nicht hinreichend beschreibt. Der Jitter j gibt an, in welchem Maße der Ereigniszeitpunkt von dem periodischen Auftreten abweichen kann. Wie in Abbildung 3.5 dargestellt, ist der minimale Abstand zwischen zwei Ereignissen $p - j$.

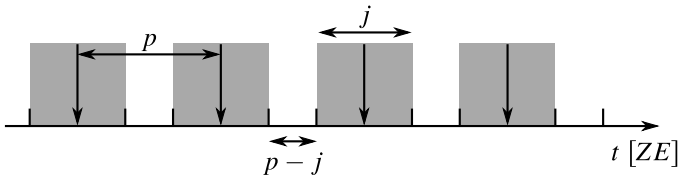


Abb. 3.5: Periodisches Ereignismodell mit Jitter.

Ein gleichzeitiges Auftreten mehrerer Ereignisse ergibt sich für $j > p$. Für die ausgehende Ereignissequenz eines Tasks ist es jedoch nicht möglich, dass mehrere Ereignisse gleichzeitig auftreten. Der minimale Abstand dieser Ereignisse ergibt sich

aus c^-. Um diese Entspannung im Ereignismodell abzubilden wurde dem periodischen Modell mit Jitter der Parameter d hinzugefügt, der den minimalen Abstand zweier Ereignisse spezifiziert. In Abbildung 3.6 ist die maximale Anhäufung einer Ereignissequenz für einen Jitter von $j = 2 \cdot p$ und $d = 1\,ZE$ dargestellt. Ohne einen minimalen Mindestabstand (d) könnten hingegen drei Ereignisse gleichzeitig auftreten.

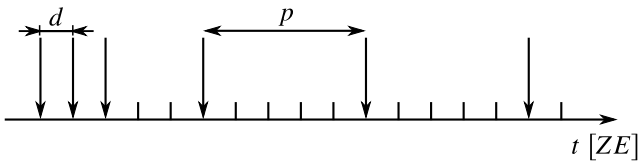


Abb. 3.6: Maximale Anhäufung der Ereignisse für einen Jitter von $j = 2 \cdot p = 10\,ZE$ und $d = 1\,ZE$.

Modelle im Intervallbereich

Das Ereignisstrommodell nach (Gresser, 1993) ist ein universelles Modell zur Beschreibung von Anregungen in einem System. Es ist in der Lage jegliche Ereignissequenz exakt zu beschreiben. Das periodische Ereignismodell mit Jitter und minimalem Abstand lässt sich daher mit geringem Aufwand in das Ereignisstrommodell überführen. Das Ereignisstrommodell beschreibt das maximale Vorkommen von Ereignissen durch eine Menge an Tupeln, bestehend aus einer Periode und einem Offset. Wird dabei die Periode zu Unendlich gewählt, kann ein irreguläres Verhalten modelliert werden, welches einmalig auftritt.

Ein weiteres sehr ausdrucksstarkes Modell zur Beschreibung von Ereignissequenzen sind begrenzende Dichtefunktionen. Diese Ereignisdichten beschränken für jedes Zeitintervall Δ_t die minimale und maximale Anzahl an Ereignissen $\left(\eta^{+/-}(\Delta_t)\right)$, welche in diesem Intervall auftreten. Dieses Modell wird in der Modularen Performanzanalyse nach Wandeler (Chakraborty et al., 2003; Thiele et al., 2000; Wandeler, 2006) eingesetzt und basiert auf den mathematischen Grundlagen des Network Calculus (Le Boudec und Thiran, 2001).

In (Richter, 2005) wurde gezeigt, wie sich die oben beschriebenen Modelle in eine Ereignisdichte abbilden lassen und wie sich die maximale Last $\rho^+(\Delta t)$ durch die

Anregung eines Tasks definiert

$$\rho^{+}(\Delta t) = \eta^{+}(\Delta_t) \cdot c_{\tau}^{+}. \tag{3.1}$$

Mit Dichtefunktionen kann nicht nur das Auftreten von Ereignissen begrenzt werden, es eignet sich auch zur Begrenzung der verfügbaren Berechnungs- und Kommunikationskapazität $\left(\beta^{+/-}(\Delta_t)\right)$. Diese gibt die minimale respektive maximale Anzahl an Ereignissen des Tasks an, die in einem Zeitintervall Δ_t abgearbeitet werden.

3.1.3 Zeiteffekte der Plattform

Die Auswirkungen eingeschränkter Rechen-, Speicher- und Kommunikationsressourcen auf das funktionale Verhalten lassen sich in zwei Gruppen unterteilen. So können in einem realen System nur Informationen mit endlicher Datenbreite übertragen und bearbeitet werden. Durch die Repräsentation realer physikalischer Signale durch eine beschränkte digitale Auflösung entsteht ein fehlerbehaftetes Abbild der Realität. Dieser Quantisierungsfehler kann jedoch in seiner Größe einfach beschränkt werden.
Der zweite negative Einfluss einer Plattform auf das funktionale Verhalten sind Zeiteffekte, welche sich aus der endlichen Rechenzeit ergeben, die für Berechnungen benötigt werden. Hinzu kommen weitere Verzögerungen durch die Interferenz von Aufgaben und Nachrichten welche die selbe Rechen- oder Kommunikationsressource beanspruchen. Durch die Verwendung einer verteilten Plattform, ergeben sich weitere Abhängigkeiten bei der Ausführung von Tasks und eine Aufsummierung der Zeiteffekte entlang eines Signalpfades. Die Betrachtung der im Folgenden aufgeführten Zeiteffekte ist essentiell für einen effizienten integrierten Entwurf des funktionalen Systemverhaltens in einem sicherheitskritischen Umfeld bei gleichzeitig bestmöglicher Ressourcenausnutzung. Diese Zeiteffekte lassen sich thematisch unterteilen in Latenzen, Jitter und verworfene Ereignisse.

3.1.3.1 Latenzen

Der bestimmende Zeiteffekt in der funktionalen Sichtweise, durch dessen Auftreten das dynamische Verhalten eines Regelsystems maßgeblich beeinflusst wird, ist die Signallatenz im Regelkreis. Diese entspricht im Allgemeinen der Pfadlatenz der

Plattformsicht, in der alle Verzögerungen eines Ausführungspfades einer verteilten Systemarchitektur zusammengefasst sind. Bei dem Begriff der Latenz unterscheidet man zwischen einer konstanten und einer variierenden Latenz.
Wenngleich die Annahme einer konstanten Latenz beim Entwurf und der Verifikation eines verteilten Systems häufig Verwendung findet, ist diese für gewöhnlich nicht gegeben und wird durch Anwendung zusätzlicher Maßnahmen erreicht. So wird bei der TTA das Ergebnis einer Berechnung oder einer Kommunikation immer bis zum Ende des dafür vorgesehenen Zeitschlitzes verzögert. Auch bei ereignisbasierten Plattformen kann eine konstante, zeitinvariante Latenz angenommen werden, wenn die Latenzen aller Instanzen bis zu der maximalen Latenz zurückgehalten werden. Diese Methode entspricht dem Logical Execution Time (LET) Modell (Ghosal et al., 2004). So führt der Mechanismus des Zurückhaltens von Berechnungsergebnissen zu einer Verminderung der Systemauslastung zu Gunsten der Vorhersagbarkeit.
Ohne diesen Mechanismus variieren bei einer ereignisbasierten Plattform die Pfadlatenz innerhalb eines Zeitintervalls. Grenzen für dieses Intervall können mit den Methoden der Echtzeitanalyse aus Unterabschnitt 3.1.4.3 bestimmt werden. Für die optimale Auslegung eines Regelsystems kann es des Weiteren vorteilhaft sein die Verteilung der Latenzen über eine Wahrscheinlichkeitsdichtefunktion zu beschreiben. So wurde in (Michiels et al., 2005) gezeigt, dass aus der Kenntnis solch einer stochastischen Beschreibung weniger konservative Stabilitätsbedingungen für ein Regelsystem abgeleitet werden können.

3.1.3.2 Jitter

Durch variierende Latenzen kommt es zwangsläufig zu einem Jitter der Ereignisse, der aufgrund seines Vorkommens in zwei Arten unterteilt wird. Der Sampling Jitter bezeichnet dabei die Variation der Zeitpunkte beim Erfassen eines Sensorwerts. Da jedoch häufig das Aufnehmen und die anschließende A/D-Wandlung einer physikalischen Größe von einer dedizierten Hardwareeinheit vorgenommen wird, besteht keine Beeinflussung durch zeitgleich ausgeführte Rechenaufgaben. Für einen jitterfreien Zeitgeber kann somit der Sampling Jitter vernachlässigt werden und wird in dieser Arbeit nicht weiter betrachtet.
Die Differenz zweier aufeinander folgender Latenzen wird als Input-Output Jitter bezeichnet (Buttazzo und Cervin, 2007). Der Input-Output Jitter ist trivialer Weise durch die Differenz der Grenzerte einer variablen Latenz begrenzt. Im Weiteren wird der Input-Output Jitter schlichtweg als Jitter bezeichnet.

3.1.3.3 Verworfene Ereignisse

Es bestehen mehrere Ursachen für das Aussetzen eines Ereignisses und dem damit verbundenen Informationsverlust. So existiert in einer Vielzahl von Systemen ein Mechanismus, der durch den Abbruch einer Aufgabe einen Ereignisverlust erzwingt, wenn eine vordefinierte Zeitfrist seit der Aufgabenaktivierung überschritten wurde. Dieser Mechanismus wird dann eingesetzt, wenn das Berechnungsergebnis mit fortschreitender Zeit an Wertigkeit verliert. Um das System nicht mit dem veralteten Berechnungsergebnis zu beeinflussen, wird dieser ausgelassen. Durch die Terminierung der Aufgabe werden Ressourcen frei, so dass die Systemauslastung reduziert wird und die darauf folgende Berechnung tendenziell zu einem schnellen Abschluss kommt. Häufig wird für periodisch ausgeführte Berechnungen die Frist äquivalent zur Abtastzeit definiert, so dass keine Änderung in der Ereignisreihenfolge auftreten kann.
Eine weitere Ursache verworfener Ereignisse ist die Kollision in einem Kommunikationsnetzwerk. Ist keine Kollisionsvermeidung implementiert, oder wird bei einer Kollisionsdetektion kein erneutes Senden der Nachricht veranlasst, kommt es zu einem Ereignisverlust.
Schließlich könnte ein intelligenter Sensor mit dem Ziel Systemressourcen zu reduzieren, in einem vorgegebenen Maß die Aktivierung der Berechnungs- und Kommunikationsaufgaben aussetzen. Diese frei werdenden Ressourcen können dann dazu genutzt werden, eine Überlast des Systems zu reduzieren oder Ressourcen für eine Anwendung mit höherer Kritikalität zu allozieren.

3.1.4 Verfahren zur Bestimmung der Antwortzeit

Es existieren verschiedene Methoden zur Bestimmung der Antwortzeit einer Task. Diese unterscheiden sich wesentlich bezüglich ihrer Aussagekraft und dem Aufwand zur Ermittlung der Antwortzeiten.

3.1.4.1 Messung

Eine Möglichkeit zur Bestimmung der Antwortzeiten von Tasks ist die Auswertung von Messungen. Dabei werden die Zeitpunkte des Starts und der Terminierungen von Tasks oder einer Ausführungskette im laufenden Betrieb erfasst. Dies geschieht

anhand eines 'Device Under Test' (DUT). Die Antwortzeiten werden entweder von einer externen Einheit erfasst, wofür unter Umständen weitere digitale Ausgänge des DUT benötigt werden. Alternativ können die Messungen selbst von dem DUT vorgenommen werden. Dies benötigt jedoch Rechenzeit und Speicher der unter Umständen nicht zur Verfügung steht. Des Weiteren wird für das DUT ein Prototyp des Systems benötigt, der zu einer frühen Entwurfsphase noch nicht verfügbar ist.
Durch eine Messung der auftretenden Zeiten erhält man eine realistische Einschätzung des Zeitverhaltens, jedoch kann durch den Betrieb des DUT unter Laborbedingungen im Allgemeinen nicht garantiert werden, dass alle Fälle, oder zumindest der schlimmste Fall abgedeckt ist.

3.1.4.2 Ereignisbasierte Simulation

Durch eine ereignisbasierte Simulation kann das Zeitverhalten des Systems in einer frühen Entwurfsphase bestimmt werden. In solch einer Simulation wird ein abstraktes Systemmodell, basierend auf dem oben beschriebenen Taskmodell modelliert und ausgeführt. Gleich wie die Messung, steigt die Aussagekraft einer Ereignisbasierten Simulation mit der Simulationsdauer. Es ist aber auch hier im Allgemeinen nicht möglich, den Fall in der Simulation zu erfassen, der zu der längsten Ausführungszeit führt (Langer et al., 2007). Auch durch eine begrenzte Detaillierungstiefe bei der Modellierung der Effekte des Echtzeitsystems entstehen Ungenauigkeiten in den Simulationsergebnissen.
Ein ereignisbasierter Simulator, mit dem Fokus der Untersuchung des Echtzeitverhaltens von verteilten eingebetteten Systemen ist ChronSim (Anssi et al., 2012). Mit ChronSim lassen sich sowohl abstrakte Modelle, als auch eine vorhandene C-Implementierung und Effekte wie Uhrendrift modellieren und simulieren.
Durch eine Co-Simulation eines ereignisbasierten Simulators mit einer kontinuierlichen Simulation des angegliederten physikalischen Systems, lässt sich das Gesamtverhalten des CPS simulieren. Solch eine Co-Simulation wird in der MATLAB Toolboxen TrueTime (Cervin und Årzén, 2009; Henriksson et al., 2003) oder mit der kommerziell erhältlichen Toolbox SimEvents (Mathworks, 2010) umgesetzt.

3.1.4.3 Analytische Verfahren

Messungen und simulative Herangehensweisen bestimmen durch eine zeitaufwendige Vorgehensweise das reale Zeitverhalten und erlauben die Generierung einer Häufigkeitsverteilung der auftretenden Antwortzeiten. Durch eine stochastische Echtzeitanalyse wie in der Arbeit (Maxim und Cucu-Grosjean, 2013) lässt sich auch durch eine analytische Berechnung eine Verteilungsfunktion der Antwortzeiten bestimmen.

Ist hingegen im Kontext sicherheitskritischer Systeme eine garantiert obere Schranke der Antwortzeiten gefordert, wird eine deterministische Echtzeitanalyse angewandt. Durch eine deterministische Echtzeitanalyse erfolgt die Bestimmung dieser zeitlichen Grenzen für eine gegebene Hardware-/Software-Lösung eines verteilten Systems. Für einfache verteilte Systeme liefert der Ansatz von Tindell und Clark gute Resultate (Tindell und Clark, 1994). Die dort vorgestellte Antwortzeitanalyse bedient sich eines Betrachtungsfensters, welches über die Zeitlinie geschoben wird, um den schlechtesten Fall zu finden. Mit steigender Komplexität des zu analysierenden Systems verschlechtern sich die Abschätzungen jedoch deutlich. Daher wurden ausgefeiltere Verfahren wie SymTA/S (Richter, 2005) und das Echtzeitkalkül (Chakraborty et al., 2003; Wandeler, 2006) entwickelt, mit deren Hilfe genauere Schranken ermittelt werden konnten.

Das Echtzeitkalkül beruht auf den mathematischen Modellen des Network Calculus (Le Boudec und Thiran, 2001) und spezifiziert die Taskanregungen anhand der im vorherigen Abschnitt vorgestellten Ereignisdichten.

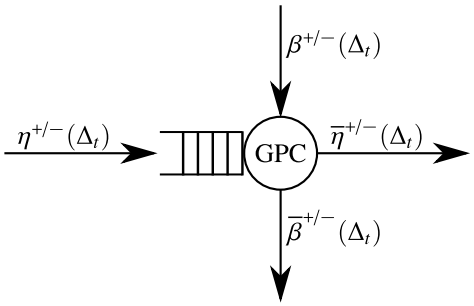


Abb. 3.7: Greedy Processing Component des Echtzeitkalküls nach (Wandeler, 2006).

Das Grundelement des Echtzeitkalküls ist die Greedy Processing Component (GPC) (Abbildung 3.7). Diese Komponente beschreibt einen Task, der durch eine minimale und maximale Ereignisdichte beschränkte Ereignisfolge $\eta^{+/-}(\Delta_t)$ angeregt wird. Zusammen mit der verfügbaren Kapazität $\beta^{+/-}(\Delta_t)$ werden die ausgehenden begrenzenden Kurven der Ereignisfolgen und Kapazität berechnet. Des Weiteren sind die Berechnungsvorschriften in der Lage, die maximale Anzahl zwischengespeicherter Ereignisinstanzen und die maximale Antwortzeit der durch die GPC repräsentierte Task abzuschätzen. Die dafür notwendigen Berechnungsformeln sind in Anhang B.2 aufgeführt. Das Konzept des Echtzeitkalküls erlaubt es mehrere GPC zu verknüpfen und bildet so dass Framework der Modularen Performanzanalyse (MPA). Mit dieser lassen sich verteilte Systeme elegant hinsichtlich der auftretenden Auslastungen und der maximalen Antwortzeit (WCRT) untersuchen.
Erst eine Approximation bei der analytischen Berechnung ermöglicht die Grenzen des Zeitverhaltens in annehmbarer Berechnungszeit zu ermittelt. So entsteht die WCRT aus der pessimistischen Überlagerung von Anregung und Ressourcenverbrauch und tritt für gewöhnlich nicht ein. Ebenso treten Ereignisinstanzen mit hohen Antwortzeiten nur in geringem Umfang auf.
Aus dieser Motivation wurde in (Quinton et al., 2012) die typical worst case response time (TWCRT) eingeführt. Die TWCRT beschreibt eine maximale Antwortzeit, welche für die meisten Fälle eine Grenze darstellt. Des Weiteren werden in der Arbeit die minimalen Intervalle beschrieben, für die diese TWCRT überschritten werden kann, aber die WCRT weiterhin Gültigkeit besitzt.

3.2 Funktionales Modell

Ein CPS nimmt abhängig von dessen Einsatzzweck unterschiedliche Aufgaben wahr. Bei der Überwachung, Steuerung und Regelung eines physikalischen Prozesses bestehen diese Aufgaben überwiegend aus der Verarbeitung von Signalen. Der Entwurf der Algorithmen zur Signalverarbeitung findet in der funktionalen Sichtweise statt, in der idealisierte Annahmen über die ausführende Plattform gemacht werden. Ein gebräuchliches Entwurfswerkzeug ist MATLAB/Simulink (Simulink, 1993), das eine Umgebung zur Modellierung und Simulation von technisch-physikalischen Systemen bereitstellt. Durch die universelle Verwendbarkeit, kann dieses Entwurfswerkzeug applikationsunabhängig eingesetzt werden.
Wie in dem vorherigen Kapitel beschrieben, sind durch die Signalrückkopplung Re-

gelsysteme äußerst sensitiv auf Zeiteffekte des Signalpfades, weshalb sie Gegenstand dieser Forschungsarbeit sind. Es existieren verschieden Möglichkeiten dynamische Systeme im Allgemeinen und Regelsysteme im Speziellen zu modellieren. Neben einer Modellierung des Systemverhaltens als Übertragungsfunktion im Laplace-Bildbereich[1] kann das System im Zustandsraum beschrieben werden. Während eine Laplace-Darstellung sich auf das Ein-/Ausgangsverhalten beschränkt, erlaubt eine Modellierung im Zustandsraum eine Sichtweise auf die inneren Zustände des Systems.
Im Folgenden wird das Modell der Zustandsregler eingeführt, da das Zustandsraummodell ein etabliertes und universelles Modell zur Beschreibung dynamischer Systeme ist. Des Weiteren sind dem Zustandsraummodell eine Vielzahl an bewährten Methoden zur Analyse des Systemverhaltens angegliedert.

3.2.1 Zustandsraummodell

Aus den physikalischen Zusammenhängen des zu regelnden technischen Prozesses lässt sich ein Differentialgleichungssystem aufstellen, welches das dynamische Verhalten des Systems beschreibt (Isermann und Münchhof, 2011). Für ein ungestörtes, lineares, zeitinvariantes System lässt sich dieses Differentialgleichungssystem in eine Matrixdifferentialgleichung erster Ordnung der Form

$$\dot{\mathbf{x}}(t) = A_c \cdot \mathbf{x}(t) + B_c \cdot \mathbf{u}(t) \tag{3.2}$$

überführen. Dabei entspricht $\mathbf{x}(t) \in \mathbb{R}^n$ dem Zustandsvektor, $\mathbf{u}(t) \in \mathbb{R}^r$ dem Eingangssignal mit $A_c \in \mathbb{R}^{n\times n}$ als Systemmatrix und $B_c \in \mathbb{R}^{n\times r}$ als Eingangsmatrix. Um durch einen Regler das System in endlicher Zeit in einen beliebigen Systemzustand zu überführen, wird vorausgesetzt, dass das zu regelnde System vollständig steuerbar ist. Die Regelgröße $\mathbf{y}(t)\in \mathbb{R}^s$, die das Ausgangssignal des Systems beschreibt, berechnet sich dabei über die Ausgangsmatrix $C \in \mathbb{R}^{s\times n}$ aus dem Zustandsvektor durch

$$\mathbf{y}(t) = C_c \cdot \mathbf{x}(t). \tag{3.3}$$

[1] Die Laplace-Transformation überführt über eine Integraltransformation den Zeitbereich in den Spektralbereich. In diesem lassen sich physikalische Systeme, die durch Differentialgleichungen beschrieben werden, vereinfacht darstellen und auswerten.

Auf die Verwendung einer Durchgangsmatrix, die einen proportionalen Zusammenhang zwischen Systemeingang $\mathbf{u}(t)$ und Systemausgang $\mathbf{y}(t)$ herstellt, wird an dieser Stelle verzichtet. Zum Einen handelt es sich bei Systemen mit physikalischem Verhalten nicht um sprunghafte Systeme, so dass sich die Durchgangsmatrix zu Null ergibt Lunze (2008). Zum Anderen hat das Verhalten des Ausgangssignals $\mathbf{y}(t)$ keinen Einfluss auf die Betrachtung von Zeiteffekten, da das Ausgangssignal nicht an der Signalrückführung beteiligt ist. Aus solch einem ungestörten System, wie in Gleichung (3.2) lässt sich für einen gewählten Anfangszustand $\mathbf{x}_0 = \mathbf{x}(0)$ durch Lösung des Differentialgleichungssystems der Zustandsvektor $\mathbf{x}(t)$ zu jedem beliebigen Zeitpunkt eindeutig bestimmen.

3.2.2 Diskreter Zustandsregler

Die in Abbildung 3.8 dargestellte Regelstrecke beinhaltet die graphische Repräsentation der Gleichungen (3.2) und (3.3). Um diese Regelstrecke einem informationsverarbeitenden System zugänglich zu machen, müssen die Größen des Zustandsvektors von einem Sensor erfasst und in eine digitale Repräsentation gewandelt werden. Durch eine periodische Abtastung des Zustandsvektors $\mathbf{x}(t)$ entsteht eine zeitdiskrete Folge $\mathbf{x}_k$ der zu den Zeitpunkten t_k erfassten Zustandsvektoren. Mit dieser Notation ist die Abtastrate, mit welcher der Zustandsvektor durch den Sensor in ein digitales Signal gewandelt wird definiert durch

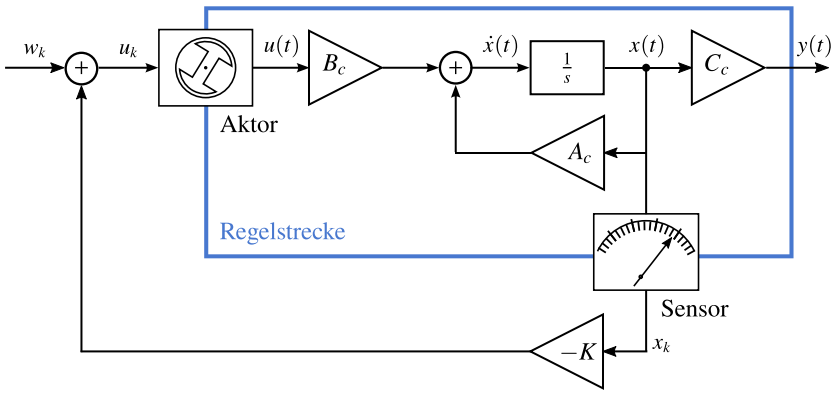


Abb. 3.8: Hybrides Mehrgrößenregelsystem mit kontinuierlichen und zeitdiskreten Signalen.

$$T_s = t_k - t_{k-1},\ k \in \mathbb{N}. \tag{3.4}$$

Eine ausreichende Dimensionierung von Abtastrate und Datenbreite, mit welcher die digitalen Signale abgebildet werden, wird vorausgesetzt. Des Weiteren wird angenommen, dass alle Zustände des Systems messbar sind. Andernfalls können für ein beobachtbares System die nicht messbaren Zustandsgrößen durch eine Beobachterstruktur wie den Luenberger-Beobachter (Luenberger, 1964) geschätzt werden[2]. Die digital abgetastete Folge $\mathbf{x}_k$ verursacht ereignisgesteuert die Berechnung des Stellwertes $\mathbf{u}_k$ und gegebenenfalls die erforderliche Übermittlung über ein Kommunikationsnetzwerk.
Für die Erfassung des Systemzustands wird ein zeitsynchroner Mechanismus angenommen, der alle Zustände des Zustandsvektors zeitgleich erfasst, so dass eine Momentaufnahme des Systems erzeugt wird. Um diese Gleichzeitigkeit der Abtastung bei der Verwendung von verteilten Sensorknoten zu gewährleisten, wird das Abtasten über einen deterministischen Kommunikationskanal initiiert. Andernfalls muss dafür Sorge getragen werden, dass die Uhren an den Sensorknoten sich synchron zueinander verhalten. Ein weiteres Kriterium an den Zeitgeber ist ein drift- und jitterfreier Betrieb, da ansonsten die in Gleichung (3.4) definierte Periodendauer nicht korrekt vom Sensor umgesetzt wird.

Wie in Abbildung 3.8 aufgezeigt wird der Stellwert durch folgende Matrixmultiplikation berechnet.

$$\mathbf{u}_k = K \cdot \mathbf{x}_k \tag{3.5}$$

In der Literatur existiert eine Vielzahl an Verfahren, die Rückführmatrix $K \in \mathbb{R}^{r \times n}$ zu bestimmen. Die geläufigsten sind dabei die Reglerauslegung durch Polvorgabe (Ackermann, 1972) [3] und der Entwurf eines optimalen Reglers anhand eines Gütefunktionals (Willems, 1971). Ein Aktor führt die so bestimmte Folge an Stellgrößen $\mathbf{u}_k$ der Regelstrecke zu. Für gewöhnlich wirkt dabei der Stellwert auf das System ein, bis ein neuer Stellwert vorliegt. Da der Zeitpunkt der Stellgrößenbeaufschlagung in einer verteilten Systemarchitektur verzögert auftritt oder möglicherweise ausgesetzt wird, gilt es im Folgenden diese Effekte auf die Güte des Regelkreis zu untersuchen.

[2] Bei dem Luenberger-Beobachter wird ein Modell der Regelstrecke ausgeführt, dessen Eingang mit dem selben Signal wie die Regelstrecke beaufschlagt wird. Durch einen Abgleich des Modellausgangs mit dem Messwert der Regelstrecke werden Ungenauigkeiten ausgeglichen. Aufgrund des vollständigen Modells ist es möglich, nicht messbare Zustandsgrößen zu schätzen.

[3] Bei der Reglerauslegung durch Polvorgabe wird das dynamische Verhalten des geschlossenen Regelkreises durch dessen Eigenwerte festgelegt. Aus den Eigenwerten und dem Modell der Regelstrecke wird die Rückführmatrix bestimmt.

Kann neben der oben aufgeführten Annahme einer Abtastung des Zustandsvektors zu periodischen Zeitpunkten t_k davon ausgegangen werden, dass gleichzeitig eine Beeinflussung der Regelstrecke durch die Stellgröße stattfindet, dann ist es möglich, ein zeitdiskretes Modell aus der kontinuierlichen Regelstrecke abzuleiten. Die Matrixdifferentialgleichung (3.2) kann so in eine Matrixdifferenzengleichung der Form

$$\mathbf{x}_{k+1} = A \cdot \mathbf{x}_k + B \cdot \mathbf{u}_k \tag{3.6}$$

überführt werden. Die Matrizen des zeitdiskreten Systems ergeben sich dabei nach (Lunze, 2012) zu

$$\begin{aligned} A &= e^{A_c T_S}, \\ B &= \int_0^{T_s} e^{A_c \alpha} d\alpha \cdot B_c. \end{aligned} \tag{3.7}$$

Identisch wie das zeitkontinuierliche Modell verhält sich mit $C = C_c$ die diskrete Ausgangsgleichung

$$\mathbf{y}_k = C \cdot \mathbf{x}_k.$$

Die Signale des so ermittelten diskreten Modells sind dabei an den Zeitpunkten t_k mit denen des zeitkontinuierlichen Modells identisch. In einem geschlossenen Regelkreis nach Abbildung 3.8 ist die Annahme, dass die neue Stellgröße zeitgleich mit der Erfassung des Zustandsvektors einhergeht dann valide, wenn für die Übermittlung der Daten und die im Regelkreis involvierten Berechnungen keine Zeit vergeht. In solch einem idealisierten Fall kann Regler und diskretisierte Regelstrecke zusammengefasst werden zu

$$\mathbf{x}_{k+1} = A_{cl} \cdot \mathbf{x}_k + B \cdot \mathbf{w}_k \tag{3.8}$$

mit

$$A_{cl} = A + B \cdot K.$$

Eine konstante Latenz $r \leqslant T_s$ im Regelkreis kann wie in (Åström und Wittenmark, 2011b) beschrieben durch Erweiterung des Zustandsvektors um einen weiteren Zustand in die Regelstrecke aufgenommen werden. Die graphische Repräsentation dieses vollständig diskretisierten Systems ist in Abbildung 3.9 dargestellt. Das Zustandsraummodell des Regelsystems entspricht dabei der Beschreibung der Regelstrecke in Gleichung (3.6), wenn die Systemmatrix A durch A_{cl}, sowie der Systemeingang $\mathbf{u}_k$ durch die Führungsgröße $\mathbf{w}_k$ ersetzt wird.
In vielen Fragestellungen der Zeitanalyse von Regelsystemen ist die Änderung des

Zustandsvektors von Relevanz. Diese lässt sich auf zwei unterschiedliche Arten darstellen.

$$\begin{aligned}\mathbf{x}'_{k+1} &= \mathbf{x}_{k+1} - \mathbf{x}_k = A_{cl} \cdot \mathbf{x}'_k + B \cdot \mathbf{w}'_k \\ &= (A_{cl} - I) \cdot \mathbf{x}_k + B \cdot \mathbf{w}_k\end{aligned}$$

Diese Gleichungen können direkt aus Gleichung (3.8) abgleitet werden.

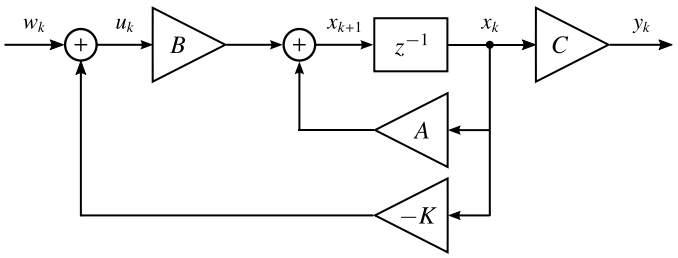


Abb. 3.9: Zeitdiskretes Mehrgrößenregelsystem.

3.2.3 Störungen im Regelkreis

Befindet sich das Regelsystem in einem stationären Zustand, ist eine Signalrückführung durch einen Regler nach obigem Modell nicht erforderlich, da sich für einen gleichbleibenden Zustandsvektor der selbe Stellwert ergibt. Dieser würde aber auch dann auf die Regelstrecke einwirken, wenn kein neuer Stellwert den Aktorknoten erreicht. Insofern wirken sich Zeiteffekte nur dann auf das Verhalten des Regelsystems aus, wenn Änderungen der im Regelkreis involvierten Signale auftreten. So muss für die Untersuchung von Zeiteffekten in einem Regelsystem die Stimulation des Regelkreis durch eine Sollwertänderungen oder Störungen in die Betrachtungen mit einbezogen werden. Bei einer Sollwertänderung besteht die Aufgabe des Regelsystems darin, der neu aufgeschalteten Führungsgröße zu folgen, so dass am Systemausgang $\mathbf{y}(t)$ die eingestellte Führungsgröße $\mathbf{w}_k$ auftritt. In einem ähnlichen Szenario wird bei der Analyse von Regelsystemen die Führungsgröße zu Null angenommen, jedoch eine initiale Auslenkung des Zustandsvektor $\mathbf{x}_0 \neq 0$ eingestellt. Eine weiterer Zweck eines Regelsystems ist das Ausregeln von Störungen im Regelkreis. Das

Einwirken einer Störung kann dabei an unterschiedlichen Stellen auftreten. Ebenso existieren verschieden Modelle, um das Verhalten der Störgröße zu beschreiben. Häufig wird dabei ein stochastisches Modell angenommen, welches es ermöglicht Quantisierungsfehler und Rauschen abzubilden. Über folgende Erweiterung des Zustandsraummodells (3.8) lässt sich die Beeinflussung des Zustandsvektors durch eine diskrete Störgröße z_k in dem geschlossenen Regelkreis darstellen.

$$\mathbf{x}_{k+1} = A_{cl} \cdot \mathbf{x}_k + B \cdot \mathbf{w}_k + E \cdot \mathbf{z}_k \tag{3.9}$$

Dabei wirkt sich über die Matrix E die Störgröße auf den Zustandsvektor aus. Wird jedoch der Stellgröße $\mathbf{u}_k$ eine Störung hinzugefügt, so entspricht die Matrix E der Eingangsmatrix B.

Kapitel 4
Stand der Technik des Regler-/Plattform Co-Designs

In diesem Kapitel werden bestehende Arbeiten vorgestellt, bewertet und eingeordnet, welche in den Entwurf und die Verifikation von Regelsystemen das Zeitverhalten der Plattform einbeziehen.

4.1 Entwurfsziele und Entwurfsparameter

Bei der Implementierung eines Regelsystems auf einer verteilten Plattform ergeben sich konkurierende Entwurfsziele. Auf der einen Seite gibt es das Bestreben ein performantes Regelverhalten zu erwirken. Auf der anderen Seite steht die Wahl einer Plattform mit geringem Ressourcenverbrauch und die Bemühung einer hohen Systemauslastung. In der Übersichtsarbeit (Årzén et al., 2000) werden Möglichkeiten und Forschungsfelder des Co-Designs von Reglerentwurf und der Ablaufplanung von Tasks aufgezeigt, die jedoch überwiegend in Anwendungsbereichen ohne sicherheitskritische Anforderungen anzusiedeln sind.

In Abbildung 4.1 sind die Beziehungen der Einflussparameter auf die Entwurfsziele eingebetteter, sicherheitskritischer Regelsysteme aufgezeigt. In dem oberen Pfad ist der Entwurfsparameter der Abtastrate dargestellt, welcher die Anregungshäufigkeit der im Regelkreis involvierten Tasks festlegt. Wie in Unterabschnitt 2.3.1 erörtert, ist die Wahl der Abtastrate eines Regelsystems eine Einflussgröße, welche die Regelgüte und die Plattformauslastung maßgeblich beeinflusst. Der direkte Zusammenhang einer Anregung in Form einer Ereignisdichte und die dadurch verursachte Last ist in Gleichung (3.1) angegeben. Für konventionelle digitale Regelsysteme ist diese Anregung strikt periodisch. Unter der Randbedingung der Planbarkeit

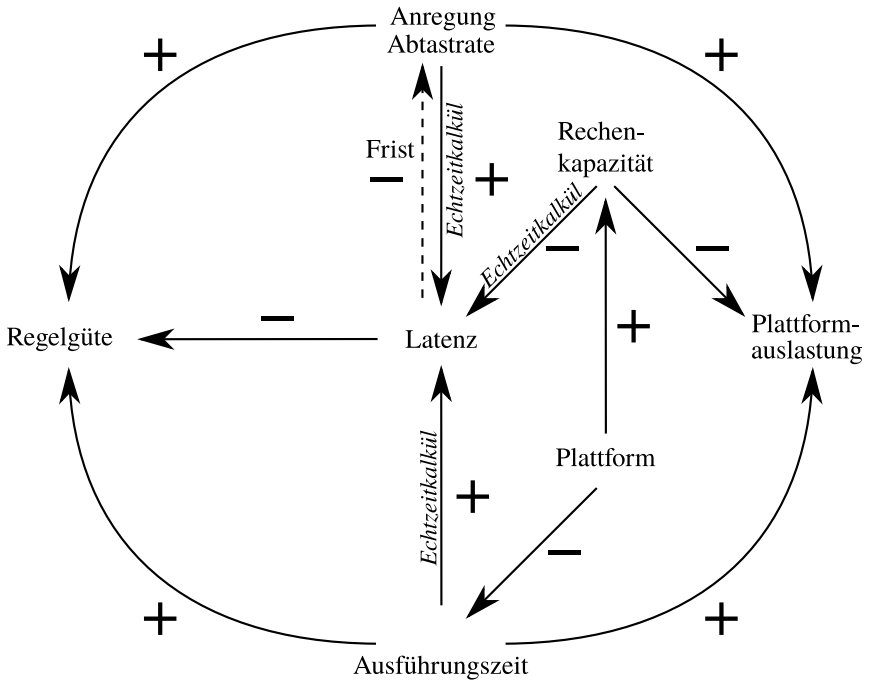


Abb. 4.1: Entwurfsparameter und Zusammenhänge des Regler-/Plattform Co-Designs.

eines Multitasksystems werden in (Seto et al., 1996) die Perioden der Taskausführung so festgelegt, dass die optimale Regelgüte erreicht wird. In (Cervin et al., 2002) wird ein dynamischer Scheduler beschrieben, der zur Laufzeit die Abtastrate der Regler-Task eines Single-Prozessor Systems anpasst, so dass mehrere auf der Plattform ausgeführten Regler die optimale Regelgüte bei hoher Auslastung aufweisen.
Aus der Motivation heraus, Rechenleistung und Datenraten in Kommunikationsnetzen zu reduzieren, wurden neue Methoden der digitalen Regelung entwickelt. In diesen Regelsystemen werden Messereignisse und die daraus resultierenden Nachrichtenübermittlungen und Berechnungen nur dann veranlasst, wenn ein Kriterium zur Bewertung des Systemzustandes einen Schwellwert überschreitet (Heemels et al., 2008). In der Literatur sind verschiedene Mechanismen zur Ereignisgenerierung beschrieben. Ereignisgesteuerte Regelsysteme (event-triggered control systems) (Heemels et al., 2008; Lunze und Lehmann, 2010; Tabuada, 2007) generieren ein Ereignis, wenn der Sensor einen Schwellwert überschreitet. Eine andere Klasse stellen Mechanismen dar, die aus dem gegenwärtigen Systemzustand den nächsten Abtast-

zeitpunkt bestimmen (self-triggered control systems) (Velasco et al., 2003; Anta und Tabuada, 2010). Die Anzahl der letztendlich erzeugten Ereignisse ergibt sich aus der Dynamik des Regelkreises und des eingesetzten Mechanismus zur Ereignisgenerierung. In (Tabuada, 2007) wird ein ungestörtes Regelsystem beschrieben und für seinen vorgestellten ereignisgesteuerten Mechanismus die minimale Zeit zwischen zwei aufeinander folgenden Ereignissen bestimmt. In (Bund et al., 2013) wurde diese Analyse um die Bestimmung einer minimalen Anregung erweitert und der vorteilhafte Effekt bei der Evaluierung einer verteilten Systemarchitektur gezeigt. Auch für self-triggered control systems existieren Arbeiten, welche Anregungsmuster ableiten und dadurch eine effiziente Anwendung der Regelmechanismen in eingebetteten Systemen ermöglichen (Aminifar et al., 2014).

Die Wahl der Abtastzeit beeinflusst über eine weitere Relation die Regelgüte. So ergibt sich über die Formalismen des Echtzeitkalküls (Wandeler, 2006) (Anhang B.2) aus der Anregung die maximale Latenz. Dabei führt grundsätzlich eine Zunahme der Anregung zu einer Vergrößerung der Latenzen, was sich negativ auf die Regelgüte auswirkt. Arbeiten dazu werden im nächsten Unterabschnitt ausführlich erörtert. Auch der umgekehrte Fall, bei der sich eine Latenz auf die Abtastrate auswirkt, ist ein plausibles Szenario. Wird eine Frist festgelegt, bei deren Überschreitung die betreffende Taskausführung abgebrochen wird (siehe Unterabschnitt 3.1.3.3), dann reduziert sich durch diese Fristverletzungen die Anzahl der Ereignisse, welche auf die Regelstrecke einwirken. Die Reduktion der Auslastung ist jedoch bei solch einem Mechanismus marginal, da bis zu dem Überschreiten der Frist Rechenzeit beansprucht wurde.

Über den Zusammenhang des Echtzeitkalküls beeinflusst auch die verfügbare Rechenzeitkapazität $\left(\beta^{+/-}(\Delta t)\right)$ (siehe Unterabschnitt 3.1.4.3) das Ausmaß der auftretenden Latenzen. Die verfügbare Rechenzeitkapazität ergibt sich durch die Wahl der Plattform. So resultiert eine rechenstärkere Plattform in einer Vergrößerung der Rechenkapazitäten, was im Allgemeinen zu einer Verringerung der Latenzen führt. Ist den Tasks des Reglers nicht die höchste Priorität zugewiesen, verbleiben durch die schnellere Abarbeitung der Tasks der selben Ressource mehr Rechenzeit für die den Regler betreffenden Tasks. Des Weiteren sinkt durch eine performantere Plattform die Nettoausführungszeit der Reglertasks.
Gleichung (3.1) zeigt auf, wie sich die Ausführungszeit auf die Auslastung der Plattform auswirkt. Der Einfluss der Ausführungszeit auf die Regelgüte kann über die Wahl der Signalwortbreiten und der Präzision der implementierten Regelalgorithmen hergestellt werden. Wird durch einen iterativen Regelalgorithmus jederzeit ein

Stellwert bereitgestellt, kann diese Präzision zur Systemlaufzeit dynamisch angepasst werden. Solch ein Anytime Control (Bhattacharya und Balas, 2004; Greco et al., 2011) eignet sich jedoch nicht für verteilte Systemarchitekturen, da diese Methode davon ausgeht, dass die Stellwertberechnung am Aktor durchgeführt wird und somit keine weitere Latenz auftreten kann.

4.2 Zeiteffekte im Regelkreis

In Abschnitt 3.1.3 wurden die verschiedenen Zeiteffekte einer Plattform beschrieben. Deren Einbeziehung in den Reglerentwurf, sowie die Untersuchung der Auswirkungen auf ein Regelsystem, wurden in einer Vielzahl an Publikationen untersucht. In diesem Abschnitt wird ein Überblick dieser Arbeiten gegeben und deren Beitrag thematisch geordnet.
Wie in Abschnitt 2.3.2 aufgezeigt, liegt der Fokus dieser Forschungsarbeit auf Regleranwendungen in einem sicherheitskritischen Umfeld, für die ein spezifiziertes Maß an Regelgüte jederzeit erfüllt sein muss. Durch eine Validierung des Regelverhaltens unter möglicherweise variablen Zeiteffekten wird die geforderte Regelgüte garantiert. Dabei ist es für die Validierung notwendig, dass auch die auftretenden Zeiteffekte durch garantierte Grenzwerte beschränkt sind. Dies macht die Verwendung von Resultaten einer Echtzeitanalyse im Zusammenhang mit der Validierung des Regelverhaltens notwendig.

Ein Großteil der Fachaufsätze, die sich mit dem Co-Design von Regelverhalten und dem Zeitverhalten der Plattform befassen, fokussieren den Entwurf von Reglern unter der Beachtung von Zeiteffekten. Dafür wird ein stochastisches Modell der Zeiteffekte angenommen, welches deren durchschnittliches Auftreten beschreibt, da dies mit großer Übereinstimmung dem realen Verhalten entspricht. Würde man dafür die aus einer Echtzeitanalyse ermittelten Werte hernehmen, dann würde der Regler nach einem Grenzfall ausgelegt, welcher möglicherweise stark von dem realen Zeitverhalten abweicht.

4.2.1 Reglerentwurf unter Beachtung von Zeiteffekten

Treten Latenzen unverändert mit dem selben Wert auf, entspricht dies dem zeitgesteuerten Paradigma oder dem LET-Modell. Solch eine Latenz kann wie in Unter-

abschnitt 6.3.1 in das Modell der Regelstrecke aufgenommen werden, wodurch sie beim Reglerentwurf implizit betrachtet wird.
Bei dem Entwurf eines Reglers hinsichtlich der auftretenden Latenzen werden für gewöhnlich stochastische Modelle eingesetzt. So wird in (Xiao et al., 2000) das Vorkommen von zufälligen Latenzen durch eine Markov-Kette (Markov, 1971) beschrieben. Dabei wird die Latenz vereinfacht als Vielfaches der Abtastzeit angenommen. Unter Einbeziehung dieser Kenntnis wird in der Veröffentlichung ein Regler in iterativer Vorgehensweise optimiert.
Eine weitere Artikel wägt den Einfluss einer zeitinvarianten Latenz gegen die Abtastrate des Regelsystems ab. So wird in (Bini und Cervin, 2008) ein Plattformmodell erstellt, welches in Abhängigkeit der Abtastrate eine Antwortzeit abschätzt. Daraus wird ein Kostenfunktional erstellt, welches die optimale Abtastrate des Regelkreises bestimmt. Einen konträren Ansatz verfolgen Arbeiten, welche das nichtdeterministische Auftreten zeitvarinater Latenzen in ein vorhersagbares Zeitverhalten überführen. Diese Mechanismen erfordern jedoch einen höheren Ressourcenverbrauch durch Zeitmessungen und Korrekturberechnungen (Zhang et al., 2001) oder führen wie in (Frehse et al., 2014) und (Kumar et al., 2012) durch eine zusätzliche Verzögerung zu einem ineffizienteren Entwurf.

Die im Folgenden aufgeführten Veröffentlichungen, die sich zu dem Schlagwort "Networked Control Systems" zuordnen lassen, untersuchen den Einfluss verworfener Ereignisse auf ein verteiltes Regelsystem. Die verworfenen Ereignisse sind dabei meist auf ein unzuverlässiges Kommunikationsnetz, wie beispielsweise drahtlose Verbindungen zurückzuführen. Dem Quality of Service (QoS) solch eines Netzes entsprechend wird dabei der Nachrichtenverlust als stochastischer Prozess modelliert (Hespanha et al., 2007; Ling und Lemmon, 2002).
Mit solch einem stochastischen Ausfallmodell kann ein Regler entworfen werden, der für die getroffenen Annahmen ein optimales Regelverhalten aufweist. Auch eine Unterscheidung, ob ein Nachrichtenverlust bei TCP-artigen Protokollen[1] detektiert wird, oder ob dieser wie in UDP-artigen Protokollen[2] nicht erkannt wird, kann in den Entwurf aufgenommen werden (Schenato et al., 2007). Erweiterungen dieses stochastischen Modells sind die kombinierte Betrachtung mit Latenzen (Schenato, 2008) und Quantisierungsfehlern (Trivellato und Benvenuto, 2010). In beiden Ar-

[1] Transmission Control Protocol (TCP) bezeichnet ein Netzwerkprotokoll, bei dem durch Quittierung der übermittelten Nachrichtenpakete durch den Empfänger verloren gegangene Pakete erkannt und gegebenenfalls ein erneutes Senden veranlasst wird. (Postel, 1980b)

[2] User Datagram Protocol (UDP) bezeichnet ein Netzwerkprotokoll, welches keine zuverlässige Zustellung der Nachrichten garantiert. (Postel, 1980a)

beiten werden stochastische Modelle dazu verwendet, einen an die Häufigkeit der verworfenen Ereignisse angepasste Regler zu entwerfen.
Ein ausdrucksstärkeres Modell als die Rate an verworfenen Ereignissen ist eine (m,k)-Begrenzung wie sie in (Ramanathan, 1999) beschrieben wird. Durch eine (m,k)-Begrenzung wird spezifiziert, dass für ein Intervall von k aufeinander folgenden Ereignissen nicht mehr als m Ereignisse verworfen werden. In dem Aufsatz wird ein angepasster Regler vorgestellt, der an solch einen Ereignisverlust angepasst ist.

4.2.2 Reglervalidierung unter Beachtung von Zeiteffekten

Ein Großteil der akademischen Arbeiten zur Verifikation von Regelsystemen beziehen sich auf die Erfüllung der Stabilität des Regelkreises. So wird in (Horowitz, 1963) erstmals ein Stabilitätskriterium vorgestellt, welches den Regelkreis hinsichtlich einer konstanten Latenz untersucht, die nicht beim Entwurf des Reglers berücksichtigt wird. Das grafische Stabilitätskriterium führt den Phasenrand im Bode-Diagramm ein. Totzeiten in einem Regelkreis verändern die Phase des Regelsystems. Überschreitet dabei die Phasenveränderung den Phasenrand, führt dies zu einem instabilen Regelkreis.
Weitere Publikationen wägen den Einfluss einer zeitinvarianten Latenz gegen die Abtastrate des Regelsystems ab. So wird in (Zhang et al., 2001) die Stabilitätsgrenze eines Regelsystems als Funktion von Abtastrate und Verzögerung aufgezeigt.
Die Stabilitätsbetrachtung zeitinvarianter Latenzen, wird in den Fachaufsätzen von (Cervin, 2012; Cervin et al., 2004; Kao und Lincoln, 2004) um den maximal auftretenden Jitter erweitert. Dabei wird die zeitvariante Latenz in eine konstante Latenz und in einen Jitter unterteilt. Durch das Aufspannen eines Koordinatensystems mit den Größen einer konstanten Latenz und einer variablen Verzögerung lässt sich durch die in den Arbeiten vorgestellten Berechnungen eine Stabilitätsregion festlegen.
Die Stabilität eines Regelkreises ist jedoch für Anwendungen, die ein garantiertes Maß an Performanz erfordern, als Gütekriterium nicht ausreichend. So wird in den Veröffentlichungen (Lincoln und Cervin, 2002; Nilsson et al., 1998) die Degradation der Regelgüte über Latenz, Jitter und Abtastzeit betrachtet. Dabei werden mit einem stochastischen Zeitmodell die variierenden Latenzen auf ein quadratisches Performanzkriterium abgebildet, was in einer MATLAB Toolbox zur Analyse von Regelsystemen umgesetzt wurde. Diese Artikel dienen als Grundlage für Methoden eines Regler und Plattform Co-Designs. So findet in (Aminifar et al., 2013b) die-

se Bewertung des Regelsystems Anwendung in einem iterativen Entwurfsfluss, der schrittweise Abtastrate und Priorität des Reglertasks anpasst.
Nachteilig an den stochastischen Zeitmodellen ist, dass sie es nicht ermöglichen Garantien zu gewährleisten, was für sicherheitsrelevante Regelsysteme (Cullyer, 1991) unabdingbar ist.
Es existieren Arbeiten, die solch eine Garantie ermöglichen (Kumar et al., 2012), jedoch wird ein vereinfachtes Modell der auftretenden Latenzen angenommen. So wird eine Grenze der Latenz festgelegt, ab der das Ereignis verworfen wird. Für kleinere Verzögerungen wird der Stellwert bis zu der definierten Grenzlatenz verzögert und dann auf die Regelstrecke angewandt.

Sowohl zur Evaluierung der Stabilität (Hespanha et al., 2007), als auch zur Bestimmung einer Regelgüte (Ling und Lemmon, 2002) von Regelsystemen unter der Einwirkung verworfenen Ereignissen existieren wissenschaftliche Aufsätze. In Verbindung mit den Resultaten einer Echtzeitanalyse wird in weiteren Publikationen die maximale Anzahl an verworfenen Ereignissen aus den Fristüberschreitungen von Tasks bestimmt. So werden daraus in (Soudbakhsh et al., 2013) Begrenzungen der Latenzen für ein stabiles Regelverhalten abgeleitet. In (Frehse et al., 2014) wird hingegen durch eine aufwendige Zustandsraumexploration untersucht, ob das Regelsystem unerlaubte Zustände einnehmen kann.
Werden Regelsysteme hinsichtlich Regelgüte und Stabilität verifiziert, ist es notwendig, die Anzahl an verworfenen Ereignissen begrenzend zu beschreiben. Ein einfaches Modell dafür ist die minimale Rate an verworfenen Ereignissen. So werden in (Goswami et al., 2014) für eine Rate an verworfenen Ereignissen Stabilität und Regelgüte bestimmt. Dieses Modell erlaubt es jedoch nicht, dass zwei benachbarte Ereignisse verworfen werden, selbst wenn anschließend für ein langes Intervall keine Ereignisse verloren gehen.

4.2.3 Regler/Plattform Co-Design

In den bisher aufgeführten wissenschaftlichen Arbeiten werden Zeiteffekte in den Reglerentwurf aufgenommen, oder das Regelverhalten unter Einfluss der Zeiteffekte untersucht. Dabei werden die Freiheitsgrade des Plattformentwurfs und die daraus resultierenden Zeiteffekte als vorgegeben angenommen. Vorteilhaft an dieser Vorgehensweise ist dabei die Trennung der Fachdisziplinen des Plattform- und des Reglerentwurfs.

Existiert jedoch bei unzureichendem Regelverhalten die Freiheit Parameter wie Prioritäten und Ausführungshäufigkeiten zu verändern, ändert sich dadurch das Zeitverhalten der Plattform. Diese Vorgehensweise führt zu einem iterativen Entwurfsfluss, bei dem im Wechsel die Plattform und das Regelsystem untersucht und erneut angepasst werden. So wird in (Samii et al., 2009) durch Anwendung eines evolutionären Algorithmus Abtastrate und Priorität des Reglers verändert und für einen an die veränderten Latenzen angepassten Regler die Regelgüte bewertet. Während diese Forschungsarbeit noch keine Garantien über das schlimmste Regelverhalten bereitstellt, erweitert eine spätere Arbeit diese Anforderung (Aminifar et al., 2012). Für die Anwendung in verteilten Systemen wurde die Herangehensweise um die Zuordnung von Tasks auf Ressourcen erweitert (Aminifar et al., 2013b).

In den Publikationen (Aminifar et al., 2015) und (Fontanelli et al., 2011) wird eine für Regleranwendungen angepasste Ablaufplanung von Tasks vorstellt. Durch einen adaptiven Scheduler wird das Einhalten von Zeitschranken und eine Begrenzung des auftretenden Jitter garantiert. Diese Ablaufplanungen basieren auf dem Scheduling-Verfahren EDF[3] und gewährleisten ein stabiles Regelverhalten.
Eine ähnliche Herangehensweise ist die Reservierung von Ressourcen für Reglerapplikationen (Aminifar et al., 2013a), wodurch neben einer garantierten Regelgüte eine Isolation gegenüber anderen Anwendungen und eine Skalierbarkeit ermöglicht wird.

4.2.4 Zusammenfassung

Modelle, welche die Zeiteffekte für Regelanwendungen beschreiben, unterscheiden sich im Wesentlichen durch ihre Verwendung im Reglerentwurf oder durch die Anwendung bei der Validierung des Regelverhaltens. Beim Entwurf eines Reglers ist man bestrebt, das bestmögliche durchschnittliche Regelverhalten zu erzielen. Dieses setzt das durchschnittliche Zeitverhalten voraus, wie es durch ein stochastisches Modell beschrieben wird.
Bei der Validierung eines Regelsystems wird das schlechteste Regelverhalten für Grenzwerte der Zeiteffekte evaluiert. Grenzwerte, wie die maximale Pfadlatenz, werden jedoch häufig überschätzt. Des Weiteren muss angenommen werden, dass diese

[3] Earliest Deadline First (EDF) bezeichnet ein Planungsverfahren, bei dem jeder Task eine Frist zugeordnet wird. Zur Laufzeit wird derjenige Task ausgeführt, welche den frühesten Zeitpunkt der Fertigstellung aufweist.

Latenz permanent auftritt. In den folgenden Kapiteln soll dieser Grenzwert durch ein weniger pessimistisches Modell der Latenz beschrieben werden. Dieses soll dennoch eine gültige Beschränkung der Latenz darstellen. Die Umsetzung dieses Modells erfordert eine Erweiterung der bestehenden Methoden zur Echtzeitanalyse verteilter Systeme sowie eine neue Herangehensweise bei der Validierung von Regelsystemen. In gleicher Weise soll auch ein neues Modell zur Beschreibung verworfener Ereignisse etabliert werden.
Bestehende Arbeiten und Methoden zum Entwurf von Reglern unter Beachtung von Zeiteffekten werden hiervon nicht beeinflusst und behalten ihre Gültigkeit wie auch ihre Notwendigkeit.

4.3 Bewertung von Regelkreisen

Das Verhalten eines Regelkreises kann mit verschiedenen Metriken bewertet werden. Man unterscheidet dabei zwischen Stabilitätsuntersuchungen und der Bewertung anhand eines Gütekriteriums. In diesem Abschnitt werden diese Begrifflichkeiten mit repräsentativen Forschungsarbeiten beschreiben.

4.3.1 Stabilität

Die fundamentale Eigenschaft, die ein Regelsysteme erfüllen muss, ist die Stabilität. Der Begriff der Stabilität in Mathematik, Physik und Wirtschaftswissenschaften ist geprägt von der Definition nach Lyapunov (Lyapunov, 1992). Nach Lyapunov ist ein ungestörtes dynamische System dann stabil, wenn für einen Startpunkt nahe des Gleichgewichtspunktes, das System in dessen Nähe verweilt.
Definitionen der Stabilität, welche sich besser auf die Theorien von Regelsystemen anwenden lassen, sind die asymptotische Stabilität und die exponentielle Stabilität (Gibson und Annaswamy, 2015), welche jeweils die Lyapunov-Stabilität voraussetzen.

Definition 10 (Asymptotische Stabilität (Kuznetsov, 2008)). *Ein Zustandsraumsystem ist dann global asymptotisch stabil, wenn für einen beliebigen n-dimensionalen initialen Anfangszustand* $\mathbf{x}_0$ *gilt*

$$\lim_{t\to\infty} |\mathbf{x}(t)| = 0. \tag{4.1}$$

Der Ausdruck $|\mathbf{x}(t)|$ in Definition 10 bezeichnet die euklidische Norm [4] des Systemzustandes. Es gilt hierbei die Konvention, dass Vektoren in fetten Kleinbuchstaben geschrieben werden. Solch ein System kann sich beliebig verhalten, wenn für den unendlichen Zeithorizont der Zustandsvektor gegen Null konvergiert. Eine stärkeres Stabilitätskriterium ist das der exponentiellen Stabilität.

Definition 11 (Exponentielle Stabilität (Gibson und Annaswamy, 2015)). *Der Gleichgewichtspunkt ist global exponentiell stabil, wenn positive Parameter c und λ existieren, so dass gilt*

$$|\mathbf{x}(t)| \leqslant c \cdot e^{-\lambda t} |\mathbf{x}_0| . \tag{4.2}$$

Dabei wird der ungestörte Verlauf des normierten Zustandsvektors begrenzt durch eine exponentiell fallende Hüllkurve. Ein exponentiell stabiles System ist immer auch asymptotisch stabil. Zudem sind lineare Zustandsraumsysteme wie in Unterabschnitt 3.2.1 und 3.2.2 mit asymptotischer Stabilität zugleich exponentiell stabil (Gibson und Annaswamy, 2015).

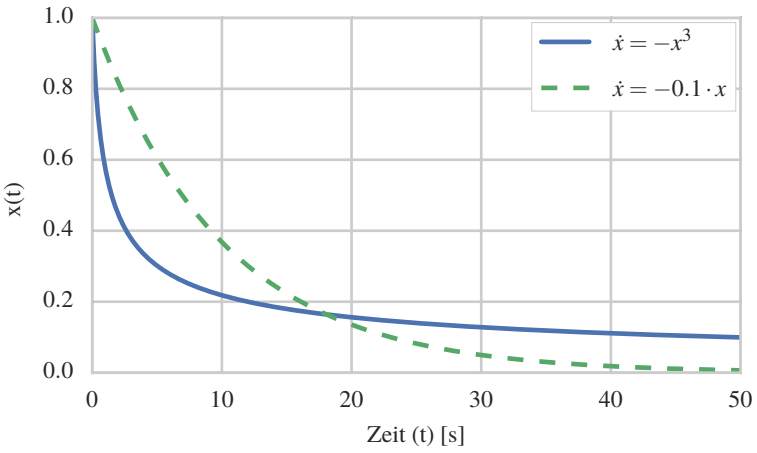


Abb. 4.2: Verlauf der Lösung der Differentialgleichungen für den Startwert $x_0 = 1$.

[4] Für einen n-Dimensionalen Vektor $x(t) \in \mathbb{R}^n$ ergibt sich die euklidische Norm zu $|x(t)| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

In Abbildung 4.2 sind die Verläufe zweier Differentialgleichungen für den Startwert $x_0 = 1$ dargestellt. Während die Lösung der Differentialgleichung $\dot{x} = -0.1 \cdot x$ asymptotische und exponentielle Stabilität aufweist, ist die Lösung von $\dot{x} = -x^3$ ausschließlich asymptotisch stabil. Die Eigenschaft der Stabilität charakterisiert das grundsätzliche Bestreben eines dynamischen Systems hin zu einem Gleichgewichtspunkt. Aussagen über das temporäre Verhalten oder über das Einhalten von Grenzwerten der Regelgröße lassen sich mit der Auswertung einer Stabilitätseigenschaft nicht treffen.
Für eine detailliertere Bewertung des Verhaltens von dynamischen Systemen existieren Verfahren zur Bestimmung der Regelgüte.

4.3.2 Regelgüte

Das Erfüllen einer Stabilitätsanforderung ist für einen Großteil technischer Systeme nicht ausreichend. Ein stabiles System, nach obigen Stabilitätsbedingungen kann ein unbefriedigendes dynamisches Verhalten aufweisen, da der unendliche Zeithorizont in Gleichung (4.1) temporär beliebiges Verhalten zulässt, oder ein kleiner Exponent in Gleichung (4.2) ein langes Einschwingverhalten ermöglicht.
Eine weitere Art der Bewertung des Regelverhaltens sind integrale Gütefunktionale. Ausgangspunkt dieser Bewertungsfunktion ist der Verlauf der Regelabweichung ϵ, welche über ein unendliches Zeitintervall aufintegriert wird (Tavazoei, 2010). Das Integral des absoluten Fehlers (integral of absolute error, IAE) ist definiert durch

$$J_{IAE} = \int_{t=0}^{\infty} |\epsilon(t)| \, dt.$$

Soll dabei eine hohe Regelabweichung verstärkt abgestraft werden, wird in der folgenden Variante der quadrierte Fehler integriert (integral of squared error, ISE).

$$J_{ISE} = \int_{t=0}^{\infty} \epsilon(t)^2 dt$$

In weiteren Ausführungen des integralen Gütefunktionals kann der Zeitfaktor oder einzelne Zustände des Zustandvektors gesondert gewichtet werden. Durch die Betrachtung eines unendlichen Zeithorizonts sind diese Gütekriterien, ähnlich den obigen Definitionen der Stabilität, ungeeignet ein temporäres inakzeptables Verhalten zu begrenzen. Vielmehr werden diese Funktionen als Zielfunktion für den Entwurf

eines optimalen Reglers verwendet. Die Optimierungsalgorithmen suchen dabei die optimalen Reglerparameter, die im gleichem Maße ein Auftreten von Regelabweichung abstrafen, wie es in der Gewichtung des Gütefunktionals spezifiziert wurde.

Eine klassische Bewertungsmethode der Regelgüte ist die Evaluierung des Ausgangsverhaltens auf ein Testsignal am Systemeingang (Piguet und Muellhaupt, 2008). So lassen sich aus der Anstiegszeit, der maximalen Höhe des Überschwingers und der Einschwingdauer einer Sprungantwort eines Regelsystems Rückschlüsse auf die Güte des Regelkreises ziehen. Abbildung 4.3 skizziert diese Metriken anhand

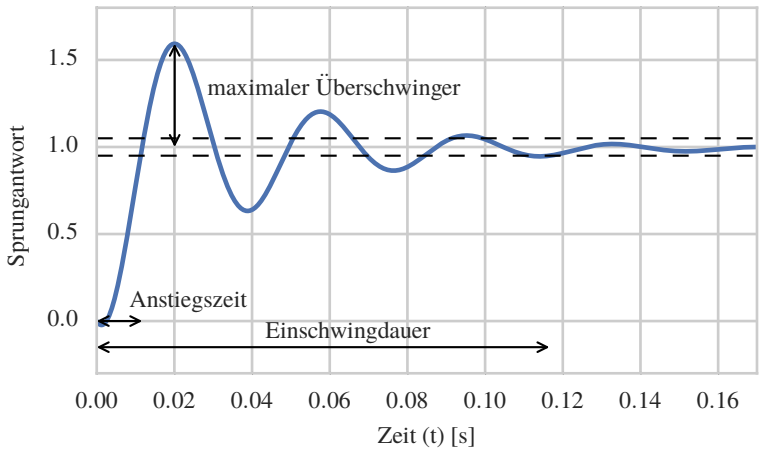


Abb. 4.3: Sprungantwort eines Regelsystems.

einer beispielhaften Sprungantwort. Es lassen sich jedoch auch aus diesen Werten keine Aussagen über das Verhalten des Regelsystems für reale Signalverläufe des Sollwerts oder einer Störgröße machen. Ebenso ist es nicht möglich, aus begrenzten variablen zeitlichen Störeinflüssen oder Störsignalen das schlechteste Regelverhalten abzuleiten. Eine Sprungantwort erlaubt nur die Darstellung des Regelverhaltens für einen konkreten Verlauf der Störeinflüsse. Ob durch variable Störeinflüsse ein unerlaubter Wert der Regelgröße erreicht wird, oder ob der Sollzustand in der geforderten Zeit erreicht wird, kann durch eine Erreichbarkeitsanalyse bestimmt werden (Donzé und Maler, 2007; Frehse et al., 2014; Langer et al., 2007). Dies erfordert jedoch einen erheblichen Rechenaufwand und es kann unter Umständen nicht sicher gestellt werden, dass alle Fälle abgedeckt sind.

Neben diesen gebräuchlichen Bewertungsfunktionen lässt sich für jeden Anwendungsfall ein gesondertes Maß definieren. So wird in (Cervin, 2012) der Multiplikator, mit dem sich eine Störung auf den Systemausgang auswirkt, als Gütekriterium angegeben. In (Kumar et al., 2012) wird der maximale Abstand der Systempole zum Koordinatenursprung als Bewertungmaß festgelegt, während in (Goswami et al., 2014) die Rate mit der eine Störung ausgeregelt wird zur Bewertung des Regelsystems genommen wird.

In dieser Forschungsarbeit wurde eine Bewertungsmethode von Regelsystemen entwickelt, mit der sich das Auftreten beliebiger im Regelkreis involvierter Signale bezüglich der Größe und der Dauer ihres Auftretens beurteilen lässt. Ähnlich der Analyse von Echtzeitsystemen ist das Resultat eine möglicherweise überschätzte aber garantierte Beschreibung des schlechtesten Regelverhaltens.

Kapitel 5
Begrenzung von Zeiteffekten im Intervallbereich

In Systemen bei denen Tasks und Nachrichten um Ressourcen wie Prozessoren und Netzwerkbusse konkurrieren, ergeben sich Antwortzeiten, welche die Nettoausführungszeiten überschreiten. Die daraus resultierenden Zeiteffekte wirken sich im Allgemeinen negativ auf das funktionale Systemverhalten aus. So wurden in Abschnitt 3.1.3 unterschiedliche Zeiteffekte vorgestellt, die das dynamische Verhalten eines Regelsystems beeinflussen. Einen Überblick über existierende Methoden, welche diese Zeiteffekte aus einem Plattformmodell bestimmen, wurde in Abschnitt 3.1.4 gegeben. Den Pessimismus, welchen diese Methoden zur Echtzeitanalyse bei der Bestimmung der maximalen Latenz und der Rate an verworfenen Ereignissen verursachen, wirkt sich auch auf die Ergebnisse bei der Beurteilung von Regelsystemen aus. Für die Validierung einer Funktionalität, wie das rechtzeitige Auslösen eines Airbags, die nur einmalig ausgeführt wird, sind die Resultate einer Echtzeitanalyse ausreichend. Wird jedoch ein dynamisches System wiederkehrend durch Signale beeinflusst, kann ein Modell, das nur Grenzwerte des Zeitverhaltens betrachtet zu einer erheblichen Überschätzung des realen Systemverhaltens führen. So muss bei der Überprüfung des dynamischen Systemverhaltens davon ausgegangen werden, dass die maximale Pfadlatenz dauerhaft auftritt, oder beliebig zwischen minimaler und maximaler Pfadlatenz variiert. Weder die Methoden, noch die zugrunde liegenden Modelle der deterministischen Echtzeitanalyse erlauben die Spezifikation weiterer Schranken neben dem maximalen und minimalen Grenzwert. So können durch das Wissen um geringe Verzögerungen von benachbarten Taskausführungen weniger pessimistische Annahmen getroffen werden.

Diese Einschränkung bestehender Modelle in der Ausdrucksstärke zur Beschreibung der Grenzwerte motiviert die Begrenzung auftretender Zeiteffekte im Intervallbereich. So genannte *Dichtefunktionen* beschränken für eine gegebene Intervalllänge

das maximale respektive minimale Vorkommen von Zeiteffekten. Dies erlaubt eine ausdrucksstarke Beschränkung des Zeitverhaltens der Systemplattform über die gesamte Einsatzdauer. Die Betrachtung im Intervallbereich reduziert dabei den Pessimismus, da für jede Intervalllänge eine Schranke angegeben wird. Des Weiteren werden mit solchen Dichtefunktionen garantierte Schranken spezifiziert, wodurch sie sich für den Einsatz in sicherheitskritischen Anwendungen eignen.

Durch die Verwendung von Dichtefunktionen können die in Abschnitt 2.1 eingeführte Plattformsicht und die funktionale Sichtweise weiterhin getrennt voneinander betrachtet werden. Dadurch lassen sich die Modelle und Methoden in bestehende Entwurfsflüsse einbinden, bei denen zwischen den Fachdisziplinen des Plattformentwurfs und des funktionalen Entwurfs unterschieden wird. So können ausgehend von einem Entwurf des funktionalen Systemverhaltens Dichtefunktionen abgeleitet werden. Die so bestimmten Dichtefunktionen dienen als Spezifikation für den Plattformentwurf. Erfüllt dabei eine Plattform diese Dichtefunktionen, so kann auch ohne Kenntnis des dynamischen Verhaltens garantiert werden, dass das System die funktionalen Performanzanforderungen erfüllt.

In den nachfolgenden Abschnitten werden Dichtefunktionen für verschiedene Zeiteffekte eingeführt und deren Zusammenhänge aufgezeigt. Neben der formalen Definition der Dichtefunktionen werden Eigenschaften abgeleitet, die es erlauben diese Dichtefunktionen in effizienter Weise zu speichern und zu verarbeiten. Auch eine Konvertierung zu bestehenden, weniger effizienten Modellen wird aufgezeigt, um diese vergleichen zu können. In weiteren Abschnitten dieses Kapitels werden Methoden aufgezeigt, die für eine praktische Anwendung der Dichtefunktion in verteilten Systemen essentiell notwendig sind. So ist es für das funktionale Systemverhalten meist irrelevant, an welcher Position im Signalpfad eines Taskgraphen Zeiteffekte auftreten. Von größerer Bedeutung ist es hingegen, die zusammengefassten Zeiteffekte zu ermitteln, so dass Konkatenations- und Vereinigungsoperationen für Dichtefunktionen hergeleitet werden müssen.

In Abschnitt 6.1 wird analog zu den Zeiteffekten die Begrenzung von Signalen im Intervallbereich zur Gütebewertung von dynamischen Systemen eingeführt. So ergibt sich eine enge Verknüpfung von Plattform und funktionaler Sichtweise auf der Modellebene, was zu einem integrierten und effizienten Entwurf des Systems führt. Teile der im Folgenden definierten Modelle und deren Eigenschaften wurden in (Bund und Slomka, 2013, 2014) veröffentlicht.

5.1 Verworfene und ausgesetzte Ereignisse

Wie in Abschnitt 3.1.3.3 beschrieben, existieren mehrere Ursachen, die den Verlust eines Ereignisses zur Folge haben. Betrachtet man verworfene Ereignisse im Kontext der Aktivierung von Taskinstanzen (siehe Definition 9), so ist wie in Abbildung 5.1 dargestellt die Anzahl der generierten Ereignisse eines Tasks geringer als die Anzahl der eingehenden Ereignisse. Der darauf folgende Task wird demnach weniger häufig

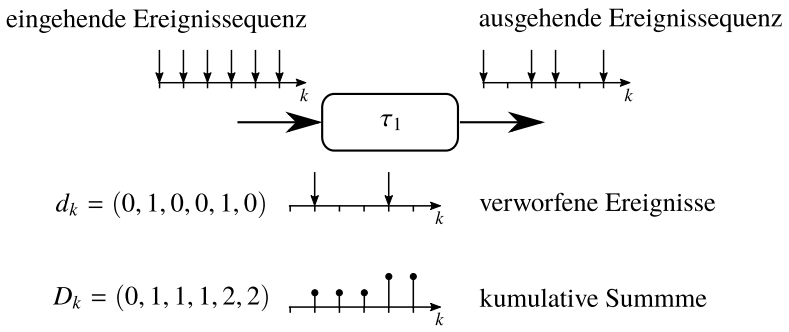


Abb. 5.1: Verworfene Ereignisse eines Tasks.

ausgeführt. So kann aus der Differenz der ausgehenden und eingehenden Ereignissen einer Task die Anzahl an verworfenen Ereignissen ermittelt werden. Werden von i eingehenden Ereignissen D_i Ereignisse verworfen, so verbleiben $i - D_i$ ausgehende Ereignisse. Dabei beschreibt D_i die kumulative Summe der Ereignisse, die während der ersten i Ereignisinstanzen verloren gehen.

Definition 12 (Sequenz an verworfenen Ereignissen). *Mit d_k wird die Sequenz an verworfenen Ereignisse bezeichnet. Hat die k-te Aktivierung der Taskinstanz zur Folge, dass der Task ein Ereignis emittiert, so wird dies mit $d_k = 0$ bezeichnet. Wird jedoch die k-te Ereignisinstanz verworfen, so wird dies mit $d_k = 1$ angegeben.*

Mit dieser Nomenklatur folgt

$$D_i = \sum_{k=1}^{i} d_k \,,\ d_k \in \{0, 1\}. \tag{5.1}$$

Die kumulative Summe der verworfenen Ereignisse D_i erweist sich jedoch als ungeeignet die Anzahl der verworfenen Ereignisse zu begrenzen, da sie an die konkrete

Instanz einer Taskaktivierung gebunden ist. Des Weiteren sind eingebettete Systeme meist so ausgelegt, dass Tasks in endlosen Schleifendurchgängen ausgeführt werden und somit kein Bezug zur Erstaktivierung besteht. So ändert sich die kumulative Summe der verworfenen Ereignisse für unterschiedlich gewählte Startzeitpunkte.

5.1.1 Definition der Verlustdichte

Basierend auf der kumulativen Summe (D_i) lässt sich die maximale und minimale Anzahl an verworfenen Ereignissen anhand einer maximalen respektive minimalen Dichtefunktion beschränken.

Definition 13 (Verlustdichte). *Die maximale Verlustdichte ($dD^+(\Delta_k)$) und die minimale Verlustdichte ($dD^-(\Delta_k)$) beschränkt das Auftreten verworfener Ereignisse in einem Intervall $\Delta_k = (t - s)$ durch*

$$dD^+(t-s) \geqslant D_t - D_s \geqslant dD^-(t-s)\,,\ \forall t \geqslant s \geqslant 0 \tag{5.2}$$

Die Verlustdichte ist gleichermaßen der in Abschnitt 3.1.2 beschriebenen Ereignisdichte eine beschränkende Kurve im Intervallbereich. Jedoch bezieht sich die Verlustdichte in ihrem Definitionsbereich auf diskrete Intervalle von benachbarten Ereignissen (Δ_k), wohingegen die Ereignisdichte nach (Le Boudec und Thiran, 2001) und (Wandeler, 2006) ein kontinuierliches Zeitintervall (Δ_t) als Definitionsbereich annimmt. Einen Bezug der Verlustdichte zu einem kontinuierlichen Zeitintervall lässt sich durch Verkettung mit der Ereignisdichte (vgl. Unterabschnitt 3.1.2) herstellen.

$$dD^+(\Delta_k) = dD^+(\eta^+(\Delta_t))$$

Des Weiteren kann über die Verlustdichte einer Task der Zusammenhang zwischen der eingehenden und der ausgehenden Ereignisdichte dargestellt werden.

Anmerkung 1. Die ausgehende maximale Ereignisdichte $(\bar{\eta}^+(\Delta_t))$ eines Tasks ergibt sich dabei aus der Differenz der eingehenden Ereignisdichte und der minimalen Anzahl an verworfenen Ereignissen, die in diesem Zeitintervall verworfen werden.

$$\begin{aligned} \bar{\eta}^+(\Delta_t) &= \eta^+(\Delta_t) - dD^-(\eta^+(\Delta_t)) \\ \bar{\eta}^-(\Delta_t) &= \eta^-(\Delta_t) - dD^+(\eta^-(\Delta_t)) \end{aligned} \tag{5.3}$$

Die minimale Dichtefunktion ergibt sich analog dazu.

Die Verlustdichte, wie sie in Definition 13 eingeführt wurde, spezifiziert eine Grenze und erlaubt es durch Gleichungen (5.1) und (5.2) zu validieren, ob Sequenzen an verworfenen Ereignissen diese Grenze einhalten. Liegt jedoch eine Sequenz verworfener Ereignisse vor, beispielsweise aus einer Beobachtung, kann über folgenden Zusammenhang die kleinste maximale Verlustdichte abgeleitet werden, welche diese Sequenz begrenzt.

Satz 1. *Für eine gegebene Sequenz an verworfenen Ereignissen d_k lässt sich die kleinste Begrenzung durch eine maximale Verlustdichte $dD^+(\Delta_k)$ wie folgt ermitteln.*

$$\begin{aligned} dD^+(\Delta_k) &= \sup_{i \geqslant 0} \left\{ \sum_{k=1}^{\Delta_k} d_{k+i} \right\} \\ &= \sup_{i \geqslant 0} \left\{ D_{i+\Delta_k} - D_i \right\} \ , \ D_0 \coloneqq 0 \end{aligned} \tag{5.4}$$

Beweis. Die kleinste maximale Verlustdichte, welche die Sequenz an verworfenen Ereignissen begrenzt ergibt sich, indem die Ungleichung (5.2) durch eine Gleichung ersetzt wird und das Supremum über alle Elemente der Sequenz an verworfenen Elementen gebildet wird. Um eine Sequenz an verworfenen Ereignissen, bei der ausschließlich das erste Ereignis verloren geht $(d_1 = 1,\ d_i = 0,\ \forall i > 1)$, korrekt in eine Verlustdichte zu überführen, muss $D_0 = 0$ angenommen werden:

$$\begin{aligned} dD^+(\Delta_k) &\geqslant D_{i+\Delta_k} - D_i \ \forall i \\ &= \sup_{i \geqslant 0} \left\{ D_{i+\Delta_k} - D_i \right\} \ , \ D_0 \coloneqq 0 \\ &= \sup_{i \geqslant 0} \left\{ \sum_{k=1}^{i+\Delta_k} d_k - \sum_{k=1}^{i} d_k \right\} \\ &= \sup_{i \geqslant 0} \left\{ \sum_{k=1}^{\Delta_k} d_{k+i} \right\} \end{aligned}$$

□

Für Gleichung (5.4) im obigen Satz existiert ein äquivalente Darstellung in der Max-Plus Algebra. Durch eine Darstellung der Verlustdichte in Max-Plus und Min-Plus Algebra ergeben sich Eigenschaften, die in einigen Beweisketten der folgenden Abschnitte Anwendung finden. So entspricht auch Gleichung (5.4) einer Min-Plus Entfaltung der kumulativen Anzahl an verworfener Ereignisse mit sich selbst.

$$dD^{+}(\Delta_k) \geqslant (D \oslash D)(\Delta_k)\,,\ D_0 \coloneqq 0$$

Die Definitionen der Faltungs- und Entfaltungsoperationen in Max-Plus und Min-Plus Algebra sind in Anhang B.1 aufgeführt.

5.1.1.1 Max-Plus Algebra

Analog zu den mathematischen Operationen in der Algebra definiert die Max-Plus Algebra korrespondierende mathematische Operationen. So wird die Multiplikation der konventionellen Algebra durch eine Addition ersetzt, die Addition hingegen wird durch das Bilden des Maximums substituiert. Die Max-Plus Algebra ist unter anderem dazu geeignet, das Verhalten diskreter Systeme zu untersuchen (Baccelli et al., 1992; Lunze, 2006). Diese Systeme werden durch Ereignisgraphen modelliert, für welche die Max-Plus Algebra die Zeitpunkte bestimmt, an denen Ereignisse an den Knoten auftreten.
Eine Anwendung der Max-Plus Algebra ist der Network Calculus (Le Boudec und Thiran, 2001). Der Network Calculus beschreibt und untersucht auf der Basis eines mathematischen Gerüsts in der Max-Plus Algebra Kommunikationsnetzwerke. Der Vorgang des Übermittelns von Informationen in einem Computernetzwerk wird dabei durch Datenströme modelliert. Durch Verrechnung mit einer adäquaten Beschreibung der Kommunikationskapazität lassen sich die resultierenden Latenzen berechnen. Die Berechnungen werden im Wesentlichen durch Faltungsoperationen in der Max-Plus und Min-Plus Algebra ausgeführt.

5.1.2 Definition der Intervallfunktion

Bei der Spezifikation von Verlustdichten für eine Task wird jedem Ereignisintervall eine maximale oder minimale begrenzende Anzahl zugeordnet. Es ist jedoch oft die inverse Fragestellung von Interesse, die nach dem minimalen/maximalen Intervall an konsekutiven Ereignissen sucht, in welchem n Ereignisse verworfen werden.

Definition 14 (Intervallfunktion). *Das minimale Intervall an Ereignissen, in dem n Ereignisse verworfen werden ist begrenzt durch*

$$iD^{-}(n) \leqslant \inf\{t-s|\, D_t - D_s \geqslant n\},\ \forall t \geqslant s \geqslant 0. \tag{5.5}$$

Da sich die kumulative Summe für eine weitere Ereignisinstanz um eins erhöht oder auf ihrem Wert verweilt, lässt sich der Ausdruck ändern zu

$$iD^-(n) \leqslant \inf\{t-s|\, D_t - D_s = n\},\ \forall t \geqslant s \geqslant 0.$$

Die minimale Intervallfunktion bildet das Pendant zur maximalen Dichtefunktion. Für eine gegebene Sequenz an verworfen Ereignissen lässt sich eine begrenzende minimale Intervallfunktion bestimmen durch

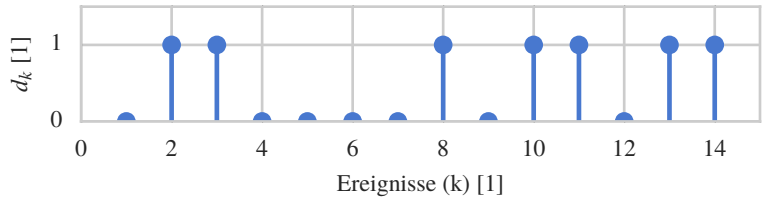
$$iD^-(n) = \inf_{i\in\mathbb{N}_0} \{\Delta_k|\, D_{i+\Delta_k} - D_i = n\,,\ D_0 \coloneqq 0\}\,. \tag{5.6}$$

Für die maximale Intervallfunktion ergibt sich analog zur Definition der minimalen Interallfunktion

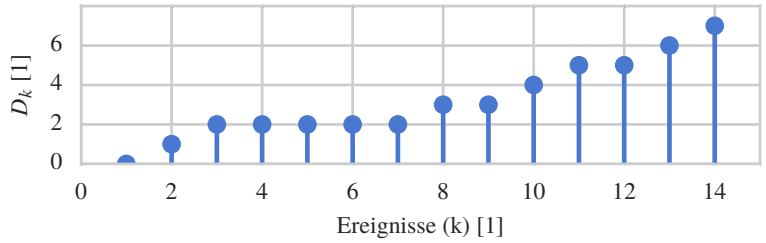
$$iD^+(n) \geqslant \sup\{t-s|\, D_t - D_s = n\},\ \forall t \geqslant s \geqslant 0.$$

Beispiel 1. In Abbildung 5.2a ist eine Sequenz an verworfenen Ereignissen d_k dargestellt, wie sie beispielsweise durch die Datenübertragung über einen unzuverlässigen Kommunikationskanal verursacht wird ($d_k = \{0,1,1,0,0,0,0,1,..\}$). Aus dieser Beobachtung lässt sich anhand Gleichung (5.1) die in Abbildung 5.2b dargestellte kumulative Summe der verworfenen Ereignisse bestimmen. Durch Anwendung von Gleichungen (5.4) lässt sich weiter die maximale Verlustdichte berechnen, welche eine Begrenzung der Sequenz in Abbildung 5.2a darstellt. Die so ermittelte Dichte $dD^+(\Delta_k)$ ist die kleinste obere Schranke, welche die Sequenz d_k in allen Intervallen begrenzt. Anhand Gleichung (5.6) lässt sich die zugehörige minimale Intervallfunktion bestimmen, welche in Abbildung 5.2d dargestellt ist.
Anhand dieses Beispiels lässt sich die Ausdrucksstärke der eingeführten Dichtebetrachtung exemplarisch darstellen. Müsste man das Auftreten der verworfenen Ereignisse in Abbildung 5.2a durch eine Rate an verworfenen Ereignissen bewerten, so resultiert dies in einer Rate von einem verworfenen Ereignis je Ereignisinstanz, da die benachbarten Ereignisse der zweiten und dritten Ereignisinstanz verworfen werden. Die Beschränkung durch eine Rate würde demnach erlauben, dass jedes erzeugte Ereignis verworfen wird. Dies überschreitet das reale Vorkommen verworfener Ereignisse erheblich.

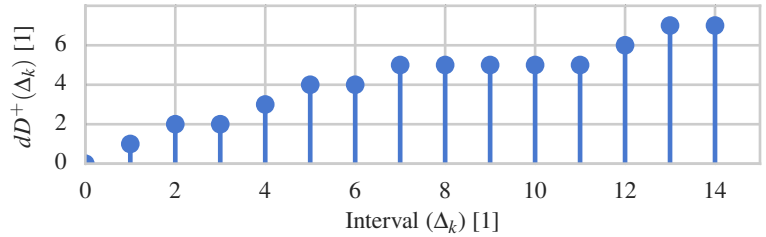
Beispiel 2. Die im vorherigen Beispiel aus einer Sequenz an verworfenen Ereignissen bestimmte maximale Verlustdichte dient nun als Spezifikation der Anzahl an verworfenen Ereignissen. Die Spezifikation ist dann erfüllt, wenn die Verlustdichte,



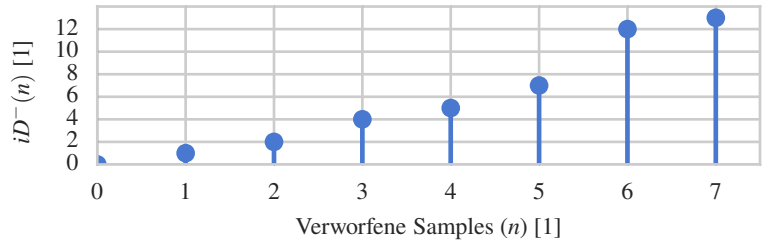
(a) Sequenz an verworfenen Ereignissen d_k.



(b) Kumulative Summe D_k der Sequenz d_k.



(c) Abgeleitete maximale Dichte der verworfenen Ereignisse $dD^+(\Delta_k)$.



(d) Minimale Intervallfunktion $iD^-(n)$.

Abb. 5.2: Auflistung einer beispielhaften Sequenz an verworfenen Ereignissen (d_k) mit kumulativer Summe (D_k), maximaler Dichtefunktion $(dD^+(\Delta_k))$ und minimaler Intervallfunktion $(iD^-(n))$.

welche aus der auftretenden Sequenz an verworfenen Ereignissen gebildet wird, die spezifizierte Verlustdichte in keinem Intervall überschreitet.
Soll diese Spezifikation durch eine Rate an verworfenen Ereignissen beschrieben werden, so ergibt sich eine minimale Rate von einem verworfenen Ereignis je drei Aktivierungen. Eine Rate an verworfenen Ereignissen lässt sich über folgenden Zusammenhang in eine maximale Dichtefunktion überführen

$$dD^{+}_{period}(\Delta_k) = \left\lceil \frac{\Delta_k}{3} \right\rceil .$$

Aus Abbildung 5.3 geht hervor, dass für die meisten Intervalle die Begrenzung durch eine Rate eine geringere Anzahl an verworfenen Ereignissen zulässt, was zu einer pessimistischeren Spezifikation der erlaubten Anzahl an Ereignissen führt. Eine Berechnungsvorschrift zur Bestimmung der minimalen Rate wird in Satz 8 auf Seite 72 gegeben.

Aus dem Beispiel wird ersichtlich, dass das Modell der maximalen Verlustdichte sowohl in der Bewertung als auch in der Spezifikation der Anzahl an verworfenen Ereignissen eine ausdrucksstärkere Beschreibung erlaubt.

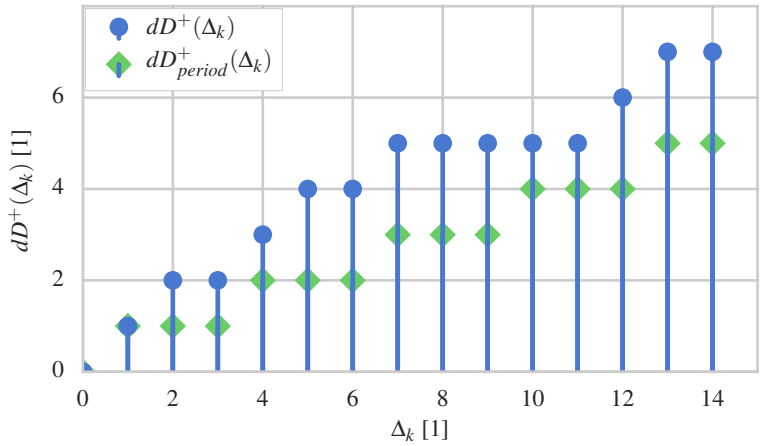


Abb. 5.3: Vergleich der Spezifikation von verworfenen Ereignissen durch eine Verlustdichte $(dD^{+}(\Delta_k))$ und einer Rate an verworfenen Ereignissen $(dD^{+}_{period}(\Delta_k))$.

5.1.3 Eigenschaften der Verlustdichte

Aus den Berechnungsvorschriften der Dichte- und Intervallfunktion aus Abschnitt 5.1.2 und der Darstellung in Max-Plus Algebra resultieren Eigenschaften wie die Monotonie und die Additivität. Aus diesen resultieren Einschränkungen der Verlustdichte, wie sie im folgenden Abschnitt eingeführt werden. Dabei wird im Wesentlichen Bezug auf die maximale Dichtefunktion und die minimale Intervallfunktion genommen. Eigenschaften der weniger relevanten minimalen Dichtefunktion respektive maximalen Intervallfunktion lassen sich analog dazu ableiten.

5.1.3.1 Wertebereiche

Die Dichtefunktion weist jedem Intervall eine maximale Anzahl an verworfenen Ereignissen zu und ist daher eine Abbildung $(dD^+ : \mathbb{N}_0 \rightarrow \mathbb{N}_0)$ mit der Menge an nichtnegativen ganzen Zahlen als Definitions- und Wertebereich. Für ein um eins größeres Ereignisintervall erhöht sich die Dichtefunktion oder verweilt auf dem vorherigen Wert. Da die Dichtefunktion des Weiteren unbeschränkt ist, handelt es sich bei der Dichtefunktion um eine surjektive Funktion (rechtstotal). Sowohl die maximale als auch die minimale Dichtefunktion sind in ihrem Wertebreich begrenzt. So können in einem Intervall von Δ_k konsekutiven Ereignissen mindestens Null und nicht mehr als Δ_k Ereignisse verworfen werden. Die triviale Beschränkung ist demnach

$$0 \leqslant iD^+(\Delta_k) \leqslant \Delta_k. \tag{5.7}$$

Auch die Intervallfunktion ist eine Abbildung mit den nichtnegativen Zahlen als Definitions- und Wertebereich. Jedoch exitiert für jeden möglichen Wert aus dem Wertebereich nicht mehr als ein Element aus dem Definitionsbereich. Somit sind die Intervallfunktionen eine injektive Funktion (linkstotal).

5.1.3.2 Monotonie

Dichte- und Intervallfunktionen sind Funktionen mit monotonem Wachstum. Diese Eigenschaft begründet sich daraus, dass Werte der Sequenz d_k keine negativen Werte annehmen können und somit die kumulative Summe D_k eine monoton wachsende Funktion ist. Durch die Eigenschaft der Monotonie sind Dichte- und Intervall-

funktionen in ihrem Wachstum nach unten beschränkt. So gilt beispielsweise für die maximale Dichtefunktion

$$iD^+(\Delta_{k,b}) \geqslant iD^+(\Delta_{k,a}),\ \forall \Delta_{k,b} \geqslant \Delta_{k,a}.$$

5.1.3.3 Additivität

Die maximale Verlustdichte $(dD^+(\Delta_k))$ ist eine subadditive Funktion.

$$dD^+(\Delta_{k,a} + \Delta_{k,b}) \leqslant dD^+(\Delta_{k,a}) + dD^+(\Delta_{k,b}) \tag{5.8}$$

Diese Eigenschaft lässt sich nach (Le Boudec und Thiran, 2001) aus der Beschreibungsform der maximalen Dichtefunktion durch die Min-Plus Entfaltung herleiten. Anschaulich betrachtet, ergibt sich eine Beschränkung der maximalen Anzahl an verworfenen Ereignissen einer Intervalllänge aus der Begrenzung für kleinere Intervalle. Daher folgt aus der Eigenschaft der Subadditivität eine obere Begrenzung des Wachstums.

Die minimale Verlustdichte $(dD^-(\Delta_k))$ ist hingegen eine superadditive Funktion.

$$dD^-(\Delta_{k,a} + \Delta_{k,b}) \geqslant dD^-(\Delta_{k,a}) + dD^-(\Delta_{k,b})$$

Diese Eigenschaft ist ebenfalls durch die Beschreibung durch eine Max-Plus Entfaltung der kumulativen Summe an verworfenen Ereignissen begründet und ergibt eine untere Beschränkung der minimalen Dichtefunktion.

5.1.3.4 Gegenseitige Beschränkung

Beschränkungen der maximalen und minimalen Dichtefunktion ergeben sich nicht nur durch die isolierte Betrachtung ihrer Eigenschaften. So ergibt sich aus der Festlegung einer maximalen Dichtefunktion eine Einschränkung des Wertebereichs der minimalen Dichtefunktion und umgekehrt. Ein trivialer Zusammenhang ist, dass für alle Intervalllängen die maximale Dichtefunktion nicht kleiner sein darf als die minimale Dichtefunktion.

$$dD^+(\Delta_k) \geqslant dD^-(\Delta_k),\ \forall \Delta_k$$

Eine weitere Beschränkung hinsichtlich des Wachstums der Dichtefunktion ergibt sich durch folgenden Satz.

Satz 2. *Wird die maximale Anzahl an verworfenen Ereignissen für ein Intervall $\Delta_{k,a}$ durch $dD^+(\Delta_{k,a})$ exakt beschränkt, dann werden in einem um $(\Delta_{k,b} - \Delta_{k,a})$ größeren Intervall mehr als weitere $dD^-(\Delta_{k,b} - \Delta_{k,a})$ Ereignisse verworfen.*

$$dD^+(\Delta_{k,b}) - dD^+(\Delta_{k,a}) \geqslant dD^-(\Delta_{k,b} - \Delta_{k,a}),\ \forall \Delta_{k,b} \geqslant \Delta_{k,a}$$

Beweis. Angenommen es existieren o.b.d.A. die Ereignisinstanzen i und j an denen die durch die maximale Dichtefunktion spezifizierten Grenzwerte erreicht, aber nicht überschritten werden, so dass gilt

$$D_{i+\Delta_{k,a}} - D_i = dD^+(\Delta_{k,a})$$

$$D_{j+\Delta_{k,b}} - D_j = dD^+(\Delta_{k,b})$$

Daraus ergeben sich die Ungleichungen

$$D_{i+\Delta_{k,a}} - D_i \geqslant D_{j+\Delta_{k,a}} - D_j$$

und

$$D_{j+\Delta_{k,b}} - D_j \geqslant D_{i+\Delta_{k,b}} - D_i.$$

Ausgehend von der Ereignisinstanz i wird nun die kleinste obere Grenze für ein Intervall $\Delta_{k,b}$ abgeschätzt, da beginnend bei dieser Ereignisinstanz die meisten verworfenen Ereignisse für das kleinere Intervall $\Delta_{k,a}$auftreten. Die letzte Ungleichung lässt sich somit aufteilen zu

$$\begin{aligned} D_{j+\Delta_{k,b}} - D_j &\geqslant D_{i+\Delta_{k,b}} - D_{i+\Delta_{k,a}} + D_{i+\Delta_{k,a}} - D_i \\ dD^+(\Delta_{k,b}) &\geqslant D_{i+\Delta_{k,b}} - D_{i+\Delta_{k,a}} + dD^+(\Delta_{k,a}). \end{aligned}$$

Nach Ungleichung (5.2) in Definition 13 besitzt $D_{i+\Delta_{k,b}} - D_{i+\Delta_{k,a}}$ die untere Schranke $dD^-(\Delta_{k,b} - \Delta_{k,a})$, was den Satz beweist. Die Untersuchung weiterer Aufteilungen des Intervalls $(\Delta_{k,b} - \Delta_{k,a})$ in kleinere Teilintervalle ist nicht notwendig, da die minimale Dichtefunktion, wie im vorherigen Unterabschnitt erläutert, eine superadditive Funktion ist. □

Diese Relation kann auch umgekehrt betrachtet werden. So ist das Wachstum der minimalen Dichtefunktion für ein um $(\Delta_{k,b} - \Delta_{k,a})$ größeres Intervall durch die ma-

ximale Dichtefunktion für das Intervall $(\Delta_{k,b} - \Delta_{k,a})$ nach oben begrenzt.

$$dD^-(\Delta_{k,b}) - dD^-(\Delta_{k,a}) \leqslant dD^+(\Delta_{k,b} - \Delta_{k,a}) \,,\ \forall \Delta_{k,b} \geqslant \Delta_{k,a}$$

5.1.3.5 Bijektivität und Umkehrfunktion

Die maximale Verlustdichte und die minimale Intervallfunktion können jeweils als Umkehrfunktion zueinander angesehen werden. Das Überführen einer Funktion in ihre Inverse setzt jedoch die Eigenschaft der Bijektivität voraus. Da Bijektivität Surjektivität und Injektivität impliziert, können nach Unterabschnitt (5.1.3.1) weder die eingeführte Intervallfuntion noch die Dichtefunktion als bijektive Funktionen bezeichnet werden. Diese Einschränkung nur die Rechtsinverse aus der Dichtefunktion, beziehungsweise die Linksinverse aus der Intervallfunktion bilden zu können kann anhand der Monotonieeigenschaft umgangen werden.

Satz 3. *Aus einer maximalen Dichtefunktion lässt sich die zugehörige minimale Intervallfunktion über folgenden Formalismus ableiten.*

$$iD^-(n) = (dD^+(\Delta_k))^{-1} = \inf\{\Delta_k |\, dD^+(\Delta_k) = n\}$$

Beweis. Der Beweis ergibt sich durch Komposition der Definitionen der Intervallfunktion (Gl. (5.5)) und der Dichtefunktion an verworfenen Ereignissen in Gleichungen (5.2). □

Analog dazu lässt sich aus der minimalen Intervallfunktion die zugehörige maximale Dichtefunktion über folgenden Formalismus ableiten.

$$dD^+(\Delta_k) = \left(iD^-(n)\right)^{-1} = \sup\{n |\, iD^-(n) \leqslant \Delta_k\}$$

5.1.4 Anwendung der Verlustdichte in verteilten Systemen

Aus den Definitionen und Eigenschaften der in den vorherigen Abschnitten eingeführten Intervall- und Dichtefunktionen an verworfenen Ereignissen werden in diesem Abschnitt Methoden entwickelt, die einen effizienten Umgang mit den vorgestellten Modellen ermöglichen. So soll durch eine Reduktion auf die wesentlichen Punkte der Intervall- und Dichtefunktion der Speicheraufwand verringert werden.

Außerdem soll durch die Wahl eines Approximationsgrades der Rechenaufwand bei der Verarbeitung von Intervall- und Dichtefunktionen variabel eingestellt werden. Überdies erfordert es für nebenläufige Signalpfade Berechnungsvorschriften zur Bestimmung von Dichtefunktionen aus vereinigten Signalpfaden.

5.1.4.1 Valide Dichtefunktion

Durch die Anwendung von mathematischen Vorschriften auf Dichtefunktionen, ergeben sich unter Umständen Folgen, die aufgrund ihrer Eigenschaften nicht mehr als gültige Dichtefunktion bezeichnet werden können. Um zu bewerten, ob es sich um eine gültige Dichtefunktion handelt, ist im Folgenden die Definition einer validen Dichtefunktion gegeben.

Definition 15 (Valide Verlustdichte). *Eine Verlustdichte, bestehend aus einer minimalen und maximalen Dichtefunktion, wird als valide bezeichnet, wenn sie den vollständigen ganzzahligen Definitions- und Wertebereich erfasst und dabei die Eigenschaften der Monotonie, sowie der Sub- respektive Superadditivität erfüllt. Eine valide Verlustdichte beinhaltet den Ursprung* $(dD^{+/-}(0) = 0)$ *und ist durch* $(dD^{+/-}(1) \leqslant 1)$ *begrenzt.*

5.1.4.2 Dichtefunktionen endlicher Länge

Im vorherigen Abschnitt wurde die Verlustdichte als Funktion im Wertebereich der natürlichen Zahlen eingeführt. So beschränkt für jede Intervalllänge die maximale Dichtefunktion die maximale Anzahl an verworfenen Ereignissen. Um aus einer Sequenz an Ereignissen eine valide Dichtefunktion mit den natürlichen Zahlen als Definitionsbereich abzuleiten, erfordert dies nach Gleichung 5.4 eine Sequenz mit unendlicher Länge. Es ist jedoch für die praktische Anwendungen nicht möglich alle Werte eines unendlich großen Wertebereichs zu erfassen und für Berechnungen vorzuhalten. Aus einer Sequenz d_k der Länge Δ_k^e ergibt sich aus Gleichung (5.4) eine endliche Verlustdichte, die nur Werte bis zu einer Intervalllänge von $(\Delta_k = \Delta_k^e)$ spezifiziert. Es existiert jedoch eine Berechnungsvorschrift, die solch eine in der Intervalllänge beschränkte Dichtefunktion schrittweise für größere Intervalle abschätzt.

Satz 4. *Existiert bis zu einer Intervalllänge* Δ_k^e *eine vollständige Dichtefunktion, dann ergibt sich folgende Abschätzung für das nächste Element der Dichtefunktion*

$$dD^+(\Delta_k^e + 1) \leqslant (dD_e^+ \otimes dD_e^+)(\Delta_k^e + 1) \tag{5.9}$$

mit

$$dD_e^+(i) = \begin{cases} dD^+(i) & \textit{wenn } i \leqslant \Delta_k^e \\ \infty & \textit{sonst} \end{cases}$$

Beweis. Wie in Abschnitt 5.1.3 beschrieben, ist die maximale Dichtefunktion eine Funktion mit subadditivem Wachstum. Aus dieser Eigenschaft lassen sich obere Grenzen der Dichtefunktion für größere Intervalle aus der gegebenen endlichen Folge abschätzen. Demnach ist der größte Wert des Folgeintervalls, welcher die Eigenschaft der Subadditivität nicht verletzt nach Ungleichung (5.8) begrenzt durch

$$dD^+(\Delta_k^e + 1) \leqslant \inf_{1 \leqslant i \leqslant \Delta_k^e} \{dD_e^+(i) + dD_e^+(\Delta_k^e + 1 - i)\}$$

Wird die endliche Folge der maximalen Dichte um einen unendlichen Wert erweitert $(dD_e^+(\Delta_k^e + 1) = \infty)$, so lässt sich die Gleichung umformen zu

$$dD^+(\Delta_k^e + 1) = \inf_{0 \leqslant i \leqslant \Delta_k^e + 1} \{dD_e^+(i) + dD_e^+(\Delta_k^e + 1 - i)\}$$

Nach Anhang B.1 lässt sich diese Formel als Min-Plus Faltung darstellen. □

Analog zu der Erweiterung einer endlichen maximalen Dichtefunktion lassen sich auch minimale Dichtefunktion und Intervallfunktionen für größere Intervalllängen erweitern.

Beispiel 3. Abbildung 5.4 zeigt die selbe maximale Verlustdichte $dD^+(\Delta_k)$, wie sie bereits in Beispiel 1 und Beispiel 2 verwendet wurde. Die Verlustdichte wurde aus einer endlichen Sequenz an verworfenen Ereignissen bestimmt, so dass sie nur Beschränkungen bis zu dem Intervall $\Delta_k^e = 14$ ermöglicht. Der weitere Verlauf der Verlustdichte kann unter Anwendung des obigen Satzes für größere Intervallängen abgeschätzt werden. Um eine gültige obere Beschränkung für ein Intervall von $\Delta_k = 15$ benachbarten Ereignissen abzuleiten, wird die Folge dD_e^+ aus der endlichen Verlustdichte der Beispiele 1 und 2 bestimmt.

$$dD_e^+ = (0, 1, 2, 2, 3, 4, 4, 5, 5, 5, 5, 5, 6, 7, 7, \infty)$$

Durch Anwendung der Min-Plus Faltung ergibt sich

$$dD^+(\Delta_k^e + 1) \leqslant \inf\{0 + \infty,\ 1 + 7,\ 2 + 7,\ 3 + 6,\ \ldots\} = 8.$$

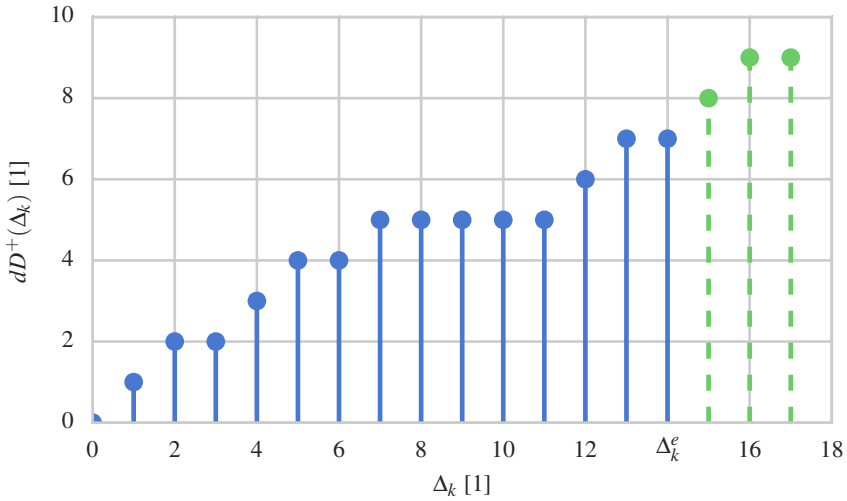


Abb. 5.4: Abschätzung größerer Intervalle einer endlichen Dichtefunktion ($\Delta_k^e = 14$).

Der dadurch bestimmte weitere Verlauf der Verlustdichte zeigt Abbildung 5.4.

Durch wiederholte Anwendung des obigen Satzes lassen sich schrittweise die Folgeelemente der endlichen Verlustdichte mit der Faltung in der Min-Plus Algebra bestimmen. Dabei erhöht sich der Berechnungsaufwand mit zunehmender Intervalllänge und ist ineffizient für die Abschätzung von Werten großer Intervalle. Im Folgenden wird versucht die Dichtefunktion auf relevante Punkte zu reduzieren und so den Aufwand bei der Verarbeitung von Dichtefunktionen zu reduzieren.

5.1.4.3 Effiziente Speicherung und Approximation der Verlustdichte

Werden wie oben beschrieben aus einer endlichen Dichtefunktion Werte für größere Intervalllängen abgeschätzt, so wird durch diese Erweiterung das Vorkommen verworfener Ereignisse nicht weiter eingeschränkt. Es werden nur die bestehenden Schranken auf größere Intervalle angewandt. In diesem Abschnitt wird die Dichtefunktion auf eine Menge an Punkten reduziert, anhand derer sich die ursprüngliche Dichtefunktion rekonstruieren lässt. Solche charakteristischen Punkte können als Stützpunkte der maximalen Dichtefunktion betrachtet werden. Anhand der Menge

an charakteristischen Punkten (cP) lässt sich so eine maximale valide Dichtefunktion bestimmen.

Darauf aufbauend werden Approximationen der Dichtefunktion vorgestellt, die es erlauben Beschränkungen für weniger ausdrucksstarke Modelle abzuleiten. Dadurch kann das Modell der Dichtefunktionen auch in bestehende Analysemethoden integriert werden.

Definition 16 (Charakteristischer Punkt). *Ein charakteristischer Punkt ist ein geordnetes Paar $(\Delta_k^c, n^c) \in \mathbb{N}^2$ und zugleich ein Element der maximalen Dichtefunktion. Die Menge an charakteristischen Punkten ist dabei die kleinste Menge an Punkten, anhand derer sich die ursprüngliche Dichtefunktion rekonstruieren lässt. Dabei wird die maximale valide Dichtefunktion gesucht, die alle charakteristischen Punkte beinhaltet. Existieren keine charakteristischen Punkte $(cP = \{\varnothing\})$, so ergibt sich daraus die initiale maximale valide Dichtefunktion $dD^+(\Delta_k) = \Delta_k$.*

Es lassen sich sowohl aus einer Menge an charakteristischen Punkten eine maximale valide Dichtefunktion bestimmen, als auch umgekehrt aus einer Dichtefunktion die Menge an charakteristischen Punkten extrahieren.

Satz 5. *Aus der Konstruktion einer Treppenfunktion für alle Elemente der Menge an charakteristischen Punkten (cP) lässt sich durch Faltung in der Min-Plus Algebra die Beschränkung der verworfenen Ereignisse für eine beliebig gewählte Inervalllänge bestimmen.*

$$dD^+(\Delta_k) = \left(\left(\bigotimes_{(\Delta_k^c, n^c) \in cP} \left(\left\lceil \frac{i}{\Delta_k^c} \right\rceil \cdot n^c \right)_{i \in \mathbb{N}_0} \right) \otimes (i)_{i \in \mathbb{N}_0} \right) (\Delta_k) \tag{5.10}$$

Beweis. Die obige Gleichung setzt sich aus diskreten Treppenfunktionen der Form

$$s_i = \left(\left\lceil \frac{i}{\Delta_k^c} \right\rceil \cdot n^c \right)$$

zusammen. Diese Treppenfunktionen sind die maximalen subadditiven Funktionen, die den jeweiligen charakteristischen Punkt beinhalten. Eine Min-Plus Faltung dieser Treppenfunktionen erzeugt wiederum eine subadditive Folge, die durch die Einschränkungen aller Treppenfunktionen beschränkt ist. Eine abschließende Min-Plus Faltung mit der Winkelhalbierenden $\left((i)_{i=0,\dots,\Delta_k}\right)$ garantiert dabei, dass nicht mehr als ein Ereignis in einem Intervall von einem Ereignis verworfen werden kann und somit die Bedingung einer validen maximalen Dichtefunktion nach Definition 15 erfüllt ist. □

Um nach obigem Satz die Werte der Dichtefunktion bis zu einer Intervalllänge von Δ_k Ereignissen aus der Menge an charakteristischen Punkten zu ermitteln, ist es ausreichend, Faltungen der Treppenfunktionen mit einer Länge von Δ_k Elementen durchzuführen.

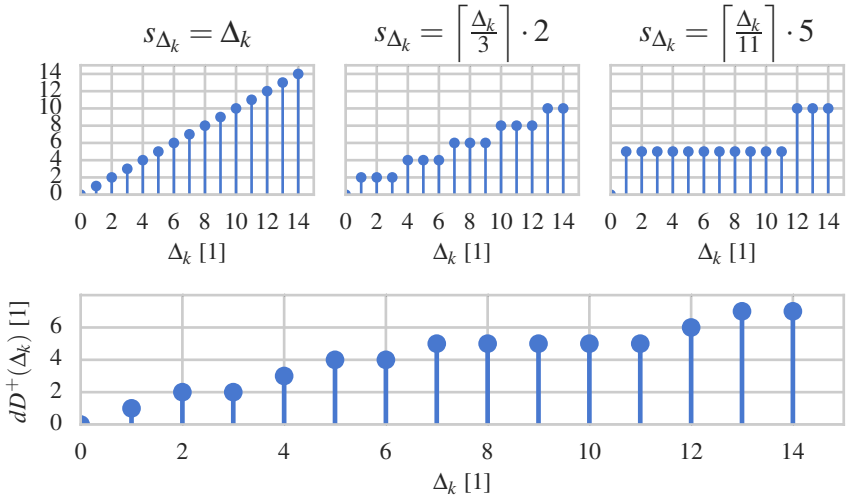


Abb. 5.5: Treppenfunktionen der charakteristischen Punkte, aus denen die Verlustdichte zusammengesetzt ist.

Beispiel 4. Charakteristische Punkte lassen sich in einfacher Weise aus dem Sprachgebrauch ableiten. Wird spezifiziert, dass nicht mehr als zwei Ereignisse in Folge[1] und nicht mehr als fünf Ereignisse je elf Aktivierungen verworfen werden dürfen, ergibt ich daraus folgende Menge an charakteristischen Punkten

$$cP = \{(3,2),(11,5)\}.$$

In Abbildung 5.5 sind die in Satz 5 beschrieben Treppenfunktionen aus den charakteristischen Punkten dargestellt. Durch eine Min-Plus Faltung dieser drei Funktionen ergibt sich die maximale Verlustdichte.

Eine Verlustdichte kann durch das Hinzufügen von zusätzlichen charakteristischen Punkten weiter eingeschränkt werden. Voraussetzung dafür ist, dass die Schranke

[1] Sollen nicht mehr als zwei Ereignisse in Folge verworfen werden, dürfen in einem Intervall von drei benachbarten Ereignissen nicht mehr als zwei Ereignisse verworfen werden.

die bestehende Dichtefunktion an diesem Intervall unterschreitet. Ist dies der Fall, so kann diese weitere Einschränkung als weiterer charakteristischer Punkt aufgefasst werden. Es ist dann zu überprüfen, ob durch die Aufnahme dieses weiteren Punktes zu der Menge an charakteristischen Punkten, bestehende Punkte der Menge obsolet werden.

Satz 6. *Wird eine maximale Verlustdichte $dD^+(\Delta_k)$ mit der Menge an charakteristischen Punkten (cP) durch einen zusätzlichen charakteristischer Punkt $(\overline{\Delta_k^c}, \overline{n^c})$ weiter eingeschränkt, so muss dieser die Eigenschaft $\overline{n^c} < dD^+(\overline{\Delta_k^c})$ aufweisen. Ist dies erfüllt, muss ein Test auf alle bestehenden charakteristischen Punkte angewandt werden, da durch die Aufnahme des neuen charakteristischen Punktes zur Menge bestehende Punkte obsolet werden können. Dabei werden Punkte aus der bestehenden Menge an charakteristischen Punkten (Δ_k^c, n^c) entfernt, die folgende Ungleichung erfüllen:*

$$n^c \geqslant \left\lceil \frac{\Delta_k^c}{\overline{\overline{\Delta_k^c}}} \right\rceil \cdot \overline{n^c},\ \forall (\Delta_k^c, n^c) \in cP$$

Beweis. Wie in Gleichung (5.10) aufgezeigt, lässt sich eine maximale Dichtefunktion aus den Treppenfunktionen ihrer charakteristischen Punkte zusammensetzen. Ein bestehender charakteristischer Punkt (Δ_k^c, n^c) wird dann aus der Menge entfernt, wenn die Treppenfunktion des hinzugefügten Punktes $(\overline{\Delta_k^c}, \overline{n^c})$ die Treppenfunktion des bestehenden Punkte in allen Intervalllängen unterschreitet.

$$\left\lceil \frac{\Delta_k}{\overline{\overline{\Delta_k^c}}} \right\rceil \cdot \overline{n^c} \leqslant \left\lceil \frac{\Delta_k}{\Delta_k^c} \right\rceil \cdot n^c\ \forall \Delta_k \in \mathbb{N}$$

Da n^c durch einen ganzzahligen Wert beschrieben wird, ergibt sich

$$n^c \geqslant \left\lceil \frac{\left\lceil \frac{\Delta_k}{\overline{\overline{\Delta_k^c}}} \right\rceil}{\left\lceil \frac{\Delta_k}{\Delta_k^c} \right\rceil} \right\rceil \cdot \overline{n^c} = \left\lceil \frac{\Delta_k}{\overline{\Delta_k^c} \cdot \left\lceil \frac{\Delta_k}{\Delta_k^c} \right\rceil} \right\rceil \cdot \overline{n^c},\ \forall \Delta_k \in \mathbb{N}.$$

Damit n^c diese Ungleichung für alle Δ_k erfüllt, wird $\left\lceil \frac{\Delta_k}{\Delta_k^c} \right\rceil$ durch $\frac{\Delta_k}{\Delta_k^c}$ abgeschätzt, was obigen Satz beweist. □

Das Hinzufügen eines weiteren Punktes zu der Menge an charakteristischen Punkten führt zu einer restriktiveren maximalen Verlustdichte, die sich durch Anwendung von Satz 5 bestimmen lässt.

Sollen umgekehrt aus einer gegebenen maximalen Verlustdichte die charakteristischen Punkte identifiziert werden, muss folgendermaßen vorgegangen werden: Ausgehend von der leeren Menge wird immer dann ein Punkt der Menge an charakteristischen Punkten hinzugefügt, wenn die aus der Menge resultierende maximale Dichtefunktion die gegebene Verlustdichte überschreitet.

In Unterabschnitt 3.1.3.3 wurden mit der Rate an verworfenen Ereignissen und der (m,k)-Begrenzung weniger aussagekräftige Modelle beschrieben mit denen sich die Anzahl an verworfenen Ereignissen beschränken lässt. Es ist möglich diese Modelle in eine Verlustdichte zu überführen.

Satz 7. *Die maximale Verlustdichte einer Beschränkung der verworfenen Ereignisse durch eine Rate (rD) ergibt sich zu*

$$dD_r^+(\Delta_k) = \left\lceil \frac{\Delta_k}{rD} \right\rceil . \tag{5.11}$$

Beweis. Die Dichtefunktion $dD_r^+(\Delta_k)$, besitzt exakt einen charakteristischen Punkt an der Stelle $cP = \{(rD, 1)\}$. Mit Satz 5 folgt daraus die Verlustdichte

$$\begin{aligned} dD_r^+(\Delta_k) &= \left(\left(\left\lceil \frac{i}{rD} \right\rceil \cdot 1 \right)_{i \in \mathbb{N}_0} \otimes (i)_{i \in \mathbb{N}_0} \right) (\Delta_k) \\ &= \inf_{0 \leqslant s \leqslant \Delta_k} \left\{ \left\lceil \frac{\Delta_k - s}{rD} \right\rceil + s \right\} \\ &= \left\lceil \frac{\Delta_k}{rD} \right\rceil . \end{aligned}$$

Der letzte Schritt gilt für $rD \geqslant 1$, was für reale Systeme erfüllt ist. □

Wird umgekehrt aus einer gegebenen maximalen Dichtefunktion die minimale Rate an verworfenen Ereignissen gesucht, welche die Beschränkung der Dichtefunktion in keinem Intervall verletzt, so wird dazu der dominante charakteristische Punkt benötigt.

Definition 17 (Dominanter charakteristischer Punkt). *Als dominanter charakteristischer Punkt (Δ_k^d, n^d) wird derjenige Punkt aus der Menge an charakteristischen Punkten bezeichnet, der den geringsten Quotienten aus Anzahl verworfener Ereignisse und Intervalllänge aufweist.*

$$(\Delta_k^d, n^d) = (\Delta_k^c, n^c) \mid \inf \left\{ \frac{n^c}{\Delta_k^c} \right\} \forall \, (\Delta_k^c, n^c) \in cP$$

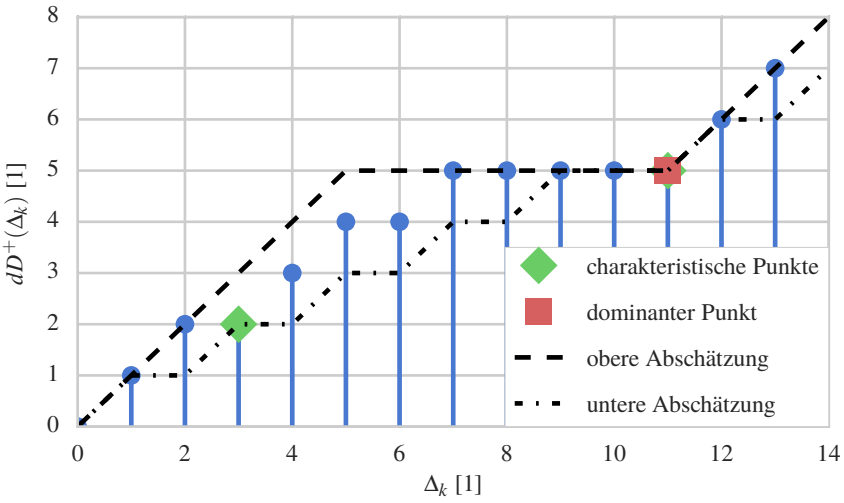


Abb. 5.6: Charakteristische Punkte und dominanter Punkt einer beispielhaften Dichtefunktion.

In Abbildung 5.6 ist die maximale Verlustdichte der vorherigen Beispiele mit den zugehörigen charakteristischen Punkten $cP = \{(3,2),(11,5)\}$, sowie der dominante Punkt $(\Delta_k^d, n^d) = (11,5)$ aufgezeigt. Da der dominante Punkt den geringsten Quotienten aus Anzahl verworfener Ereignisse und Intervalllänge aufweist, befindet sich kein Wert der Verlustdichte unterhalb der Geraden durch den dominanten Punkt und den Ursprung. Diese Gerade lässt sich durch die Funktion

$$g(\Delta_k) = \frac{n^d}{\Delta_k^d}\Delta_k$$

ausdrücken.

Da die Verlustdichte wie in Unterabschnitt 5.1.3.1 beschrieben, auf einen ganzzahligen Wertebereich abbildet, gilt diese Abschätzung auch für die aufgerundete Funktion, so dass sich eine maximale Dichtefunktion durch folgende zwei Funktionen eingrenzen lässt.

$$\left(\left(\left\lceil \frac{i}{\Delta_k^d} \right\rceil \cdot n^d\right)_{i\in\mathbb{N}_0} \otimes (i)_{i\in\mathbb{N}_0}\right)(\Delta_k) \geqslant dD^+(\Delta_k) \geqslant \left\lceil \frac{n^d \cdot \Delta_k}{\Delta_k^d} \right\rceil \tag{5.12}$$

Die obere Abschätzung entspricht dabei nach Satz 5 der Verlustdichte für nur einen charakteristischen Punkt, der dem dominanten Punkt entspricht. An dem Verlauf der oberen und unteren Abschätzung in Abbildung 5.6 ist ersichtlich, dass an dem dominanten Punkt die Abschätzungen den Wert der Verlustdichte exakt wiedergeben. So ist es anhand des dominanten Punktes möglich, effizient die Dichte- und Intervallfunktion für große Intervalle zu bestimmen. So gilt für Vielfache des Intervalls Δ_k^d

$$dD^+(m \cdot \Delta_k^d) = m \cdot n^d \,,\ \forall m \in \mathbb{N} \tag{5.13}$$

und

$$iD^-(m \cdot n^d + 1) = m \cdot \Delta_k^d + 1 \,,\ \forall m \in \mathbb{N},\ dD^+(1) = 1. \tag{5.14}$$

Die Abschätzung der maximalen Verlustdichte durch den dominanten Punkt approximiert deren Verlauf und ermöglicht so die Abbildung auf ein weniger ausdrucksstarkes Modell.

Satz 8. *Die minimale Rate an verworfenen Ereignissen, die eine Begrenzung durch eine maximale Verlustdichte für keine Intervalllänge verletzt und somit unterhalb dieser liegt, ergibt sich zu*

$$rD \geqslant \left\lceil \frac{\Delta_k^d}{n^d} \right\rceil \tag{5.15}$$

Beweis. Nach Formel (5.11) und (5.12) wird eine gegebene maximale Verlustdichte in keinem Intervall durch eine Rate an verworfenen Ereignissen überschritten, wenn gilt

$$dD_r^+(\Delta_k) \geqslant \left\lceil \frac{n^d \cdot \Delta_k}{\Delta_k^d} \right\rceil \geqslant \left\lceil \frac{\Delta_k}{rD} \right\rceil .$$

Substituiert man den Quotienten $\frac{\Delta_k^d}{n^d}$ durch q, dann ergibt sich der Zusammenhang

$$\left\lceil \frac{\Delta_k}{q} \right\rceil \geqslant \left\lceil \frac{\Delta_k}{rD} \right\rceil .$$

Diese Ungleichung ist für alle $rD \geqslant q$ wahr. Durch Resubstitution von q und dem Umstand, dass die Rate (rD) durch die Relation auf diskrete Ereignisse nur ganzzahlige Werte annehmen kann, ergibt sich aus dieser Ungleichung der zu beweisende Satz. □

Ausdrucksstärker als die Rate an verworfenen Ereignissen ist eine (m,k)-Beschränkung. Solch eine Begrenzung erlaubt nicht mehr als m verworfene Ereignisse in einem Intervall von k aufeinander folgenden Ereignissen.

Satz 9. *Die maximale Verlustdichte einer (m,k)-Beschränkung ergibt sich zu*

$$dD^{+}_{m,k}(\Delta_k) = \inf\left\{\left\lceil\frac{\Delta_k}{k}\right\rceil\cdot m,\ \left\lfloor\frac{\Delta_k}{k}\right\rfloor\cdot(m-k)+\Delta_k\right\},\ cP=\{(k,m)\}. \tag{5.16}$$

Beweis. Auch eine (m,k)-Beschränkung lässt sich durch einen einzigen charakteristischen Punkt mit $cP = \{(m,k)\}$ beschreiben. Demnach gilt nach Satz 5

$$\begin{aligned} dD^{+}_{m,k}(\Delta_k) &= \left(\left(\left\lceil\frac{i}{k}\right\rceil\cdot m\right)_{i\in\mathbb{N}_0}\otimes(i)_{i\in\mathbb{N}_0}\right)(\Delta_k)\\ &= \inf_{0\leqslant s\leqslant\Delta_k}\left\{\left\lceil\frac{\Delta_k-s}{k}\right\rceil\cdot m+s\right\}\end{aligned}$$

Der Index s bewirkt immer dann eine Reduktion der Aufrundungsfunktion wenn gilt

$$s=\Delta_k \mod k = \Delta_k - k\cdot\left\lfloor\frac{\Delta_k}{k}\right\rfloor.$$

Es ist daher ausreichend, das Infimum für die Stellen $s=0$ und $s=\Delta_k-k\cdot\left\lfloor\frac{\Delta_k}{k}\right\rfloor$ zu bilden. Eingesetzt in den obigen Formalismus ergibt

$$\begin{aligned} dD^{+}_{m,k}(\Delta_k) &= \inf\left\{\left\lceil\frac{\Delta_k}{k}\right\rceil\cdot m,\ \left\lceil\frac{\Delta_k-\Delta_k+k\cdot\left\lfloor\frac{\Delta_k}{k}\right\rfloor}{k}\right\rceil\cdot m+\Delta_k-k\cdot\left\lfloor\frac{\Delta_k}{k}\right\rfloor\right\}\\ &= \inf\left\{\left\lceil\frac{\Delta_k}{k}\right\rceil\cdot m,\ \left\lceil\left\lfloor\frac{\Delta_k}{k}\right\rfloor\right\rceil\cdot m+\Delta_k-k\cdot\left\lfloor\frac{\Delta_k}{k}\right\rfloor\right\}\\ &= \inf\left\{\left\lceil\frac{\Delta_k}{k}\right\rceil\cdot m,\ \left\lfloor\frac{\Delta_k}{k}\right\rfloor\cdot(m-k)+\Delta_k\right\}\end{aligned}$$

□

Formel (5.16) erschließt sich auch durch die graphische Repräsentation der Dichtefunktion $dD^{+}_{m,k}(\Delta_k)$. Abbildung 5.7 zeigt die Verlustdichte und die beiden Teilfunktionen, aus deren Minimum sich die Dichtefunktion ergibt. Dabei wurde in der beispielhaften Abbildung eine (m,k)-Beschränkung von zwei verworfenen Ereignissen je fünf aufeinander folgenden Ereignissen gewählt. Da eine (m,k)-Beschränkung keine Aussagen über das Auftreten verworfener Ereignisse für weitere Intervalllängen macht, muss für die maximale Dichtefunktion davon ausgegangen werden, dass die verworfenen Ereignisse gehäuft auftreten. So ergibt sich zu Beginn der Dichtefunktion ein linearer Anstieg, da es möglich ist, dass in einem Intervall von m aufeinander folgenden Ereignissen m Ereignisse verworfen werden. Die Dichtefunk-

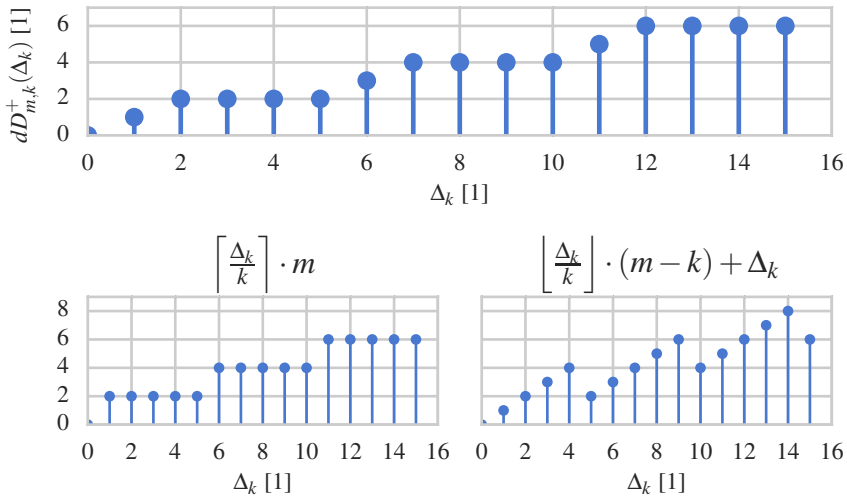


Abb. 5.7: Beispielhafte maximale Dichtefunktion einer (m,k)-Beschränkung und deren beider Teilfunktionen, aus deren Minimum sich die Dichtefunktion ergibt.

tion verbleibt auf diesem Wert, bis zu einer Intervalllänge von k Ereignissen. Im Anschluss setzt sich die Dichtefunktion in wiederholender Weise fort. Die linke der beiden Teilfunktionen beschreibt eine Treppenfunktion und gewährleistet die geforderte Einschränkung für die Intervalllänge k. Die rechte Teilfunktion in Abbildung 5.7 garantiert, dass die resultierende Dichtefunktion im Wachstum beschränkt ist und damit dem Kriterium einer validen Dichtefunktion genügt.
Für mehr als einen charakteristischen Punkt ist die direkte Bestimmung einer maximalen Dichtefunktion wie in Gleichung (5.16) mit erhöhtem Aufwand verbunden, so dass Gleichung (5.10) anzuwenden ist.

Umgekehrt lässt sich aus einer maximalen Dichtefunktion mindestens eine (m,k)-Beschränkung ableiten, welche die maximale Dichtefunktion in keinem Intervall verletzt.

Satz 10. *Eine (m,k)-Begrenzung, die eine maximale Dichtefunktion* $(dD^+(\Delta_k))$ *mit deren charakteristischen Punkten cP in keinem Intervall überschreitet, muss folgende Eigenschaften erfüllen*

$$1 \leqslant m \leqslant dD^{+}(\inf\left\{\Delta_k^c\right\}),\ (\Delta_k^c, n^c) \in cP$$
$$k \geqslant \sup_{(\Delta_k^c, n^c) \in cP} \left\{ \left\lceil m + \frac{\Delta_k^c - n^c}{\left\lfloor \frac{n^c}{m} \right\rfloor} \right\rceil \right\}. \tag{5.17}$$

Beweis. Die erste Ungleichung beschränkt die maximale Anzahl der in Folge verworfenen Ereignissen. Diese Anzahl ergibt sich direkt aus dem charakteristischen Punkt der Dichtefunktion mit der kleinsten Intervalllänge. Für $m = 1$ ist die (m,k)-Begrenzung identisch mit einer Beschränkung durch eine feste Rate an verworfenen Ereignissen.

Über die zweite Ungleichung lässt sich aus dem zuvor bestimmten Parameter m der dazugehörigen Wert k bestimmen. Ausgangspunkt ist dabei die Tatsache, dass sich die Dichtefunktion der (m,k)-Begrenzung mit einem einzigen charakteristischen Punkt in wiederholender Weise fortsetzt, so dass für ein Vielfaches der Intervalllänge des charakteristischen Punktes gilt

$$dD_{m,k}^{+}(a \cdot k) = a \cdot m \ \forall \ \mathrm{a} \in \mathbb{N}.$$

Für einen weiteren Parameter $b \in \mathbb{N}_0$ mit $b \leqslant m$ gilt des Weiteren nach Formel (5.16)

$$dD_{m,k}^{+}(a \cdot k + b) = a \cdot m + b. \tag{5.18}$$

Ziel ist es Parameter m und k zu finden, so dass die Dichtefunktion der (m,k)-Begrenzung $(dD_{m,k}^{+}(\Delta_k))$ für keine Intervalllänge die ursprüngliche Dichtefunktion überschreitet. Dies ist dann der Fall, wenn die resultierende Dichtefunktion der (m,k)-Begrenzung die Werte der charakteristischen Punkte n^c nicht schon an Intervallen kleiner Δ_k^c annimmt. Es ist daher ausreichend, zu überprüfen ob gilt

$$dD_{m,k}^{+}(\Delta_k) = n^c \ f\ddot{u}r\ \Delta_k \geqslant \Delta_k^c,\ \forall\ (\Delta_k^c, n^c) \in cP.$$

Setzt man den Zusammenhang (5.18) in diese Gleichung ein, so ergeben sich die beiden Formeln für einen charakteristischen Punkt der ursprünglichen Dichtefunktion

$$a \cdot m + b = n^c \tag{5.19}$$

$$a \cdot k + b = \Delta_k \geqslant \Delta_k^c \tag{5.20}$$

Der Parameter b muss dabei unter der Einschränkung $(0 \leqslant b \leqslant m)$ betrachtet werden. Für größere Werte $(m < b \leqslant k)$ gilt

$$dD^+_{m,k}(a \cdot k + b) = dD^+_{m,k}((a+1) \cdot k).$$

Unter der Einschränkung $(0 \leqslant b \leqslant m)$ ergeben sich für die Parameter aus Gleichung (5.19)

$$a = \left\lfloor \frac{n^c}{m} \right\rfloor, \; b = n^c - \left\lfloor \frac{n^c}{m} \right\rfloor \cdot m, \; \forall \, (\Delta^c_k, n^c) \in cP.$$

Schließlich lässt sich aus Ungleichung (5.20) der gesuchte k-Parameter der (m,k)-Begrenzung bestimmen durch

$$k \geqslant \frac{\Delta^c_k - b}{a}, \; \forall \, (\Delta^c_k, n^c) \in cP.$$

Durch Einsetzten der zuvor bestimmten Parameter a und b ergibt sich die Formel in Gleichung (5.17). □

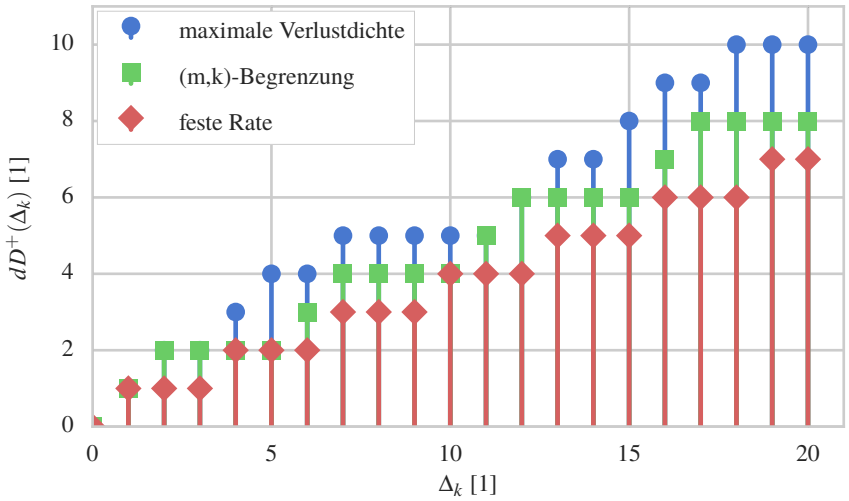


Abb. 5.8: Beispiel einer Dichtefunktion mit Abschätzung durch eine (m,k)-Begrenzung mit $m = 2$ und $k = 5$ und durch eine feste Rate mit $rD = 3$. Die charakteristischen Punkte der maximalen Verlustdichte sind $cP = \{(3,2,), (11,5)\}$.

Beispiel 5. Abbildung 5.8 zeigt die Verlustdichte der Beispiele 1-4 mit den charakteristischen Punkten $cP = \{(3,2), (11,5)\}$ sowie eine (m,k)-Beschränkung und eine Beschränkung der verworfenen Ereignisse durch eine feste Rate, wel-

che diese Verlustdichte in keinem Intervall überschreiten. Der dominante Punkt ist $(\Delta_k^d, n^d) = (11, 5)$, wodurch sich aus Satz 8 die Rate $rD = 3$ ergibt. Die (m,k)-Beschränkung ist mit $m = 2$ in dem Beispiel so gewählt, dass die maximale Anzahl der in Folge verworfener Ereignisse in Anspruch genommen wird. Satz 10 erlaubt auch eine (m,k)-Begrenzung, die weniger in Folge verworfene Ereignisse zulässt. Ist nur ein einziges verworfenes Ereignis in Folge erlaubt, so entspricht die (m,k)-Begrenzung einer Beschränkung durch eine feste Rate an verworfenen Ereignissen, welche ebenfalls in Abbildung 5.8 abgebildet ist. Der Parameter k berechnet sich aus Ungleichung 5.17 in Satz 10 zu

$$\begin{aligned} k &\geqslant \sup_{(\Delta_k^c, n^c) \in cP} \left\{ \left\lceil m + \frac{\Delta_k^c - n^c}{\left\lfloor \frac{n^c}{m} \right\rfloor} \right\rceil \right\} \\ &\geqslant \sup \left\{ \left\lceil 2 + \frac{3-2}{\left\lfloor \frac{2}{2} \right\rfloor} \right\rceil, \left\lceil 2 + \frac{11-5}{\left\lfloor \frac{5}{2} \right\rfloor} \right\rceil \right\} \\ &\geqslant \sup \{3, 5\} = 5. \end{aligned}$$

Obwohl in dem gewählten Beispiel die (m,k)-Begrenzung mit $m = 2$ die Verlustdichte besser annähert als die Begrenzung durch eine feste Rate, lässt sich nicht a priori festlegen für welchen m-Parameter die (m,k)-Begrenzung den geringsten Pessimismus aufweist.

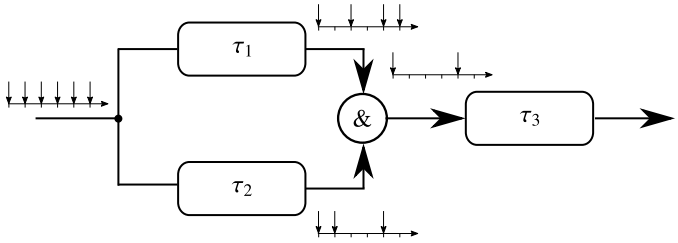
Die in diesem Abschnitt vorgestellten charakteristischen Punkte ermöglichen es, die unendliche Folge einer Verlustdichte anhand weniger Punkte effizient zu repräsentieren. Dies kommt zum einem dem Ressourcenverbrauch der Algorithmen beim Entwurf sicherheitskritischer Systeme zugute. Zum anderen eignen sich die charakteristischen Punkte auch zur Ableitung in eine Rate verworfener Ereignisse und einer (m,k)-Beschränkung, welche weniger Freiheitsgraden aufweisen. Diese Modelle können als Approximation der Dichtefunktion aufgefasst werden und ermöglichen so die Integration in existierende Analysemethoden und Entwurfsflüsse. Verglichen mit der Dichtefunktion entsteht dabei im Allgemeinen immer dann ein Pessimismus in Bezug auf die Anzahl an verworfenen Ereignissen, wenn die Verlustdichte durch mehr als einen charakteristischen Punkt beschrieben wird.

5.1.4.4 Vereinigung von Dichtefunktionen

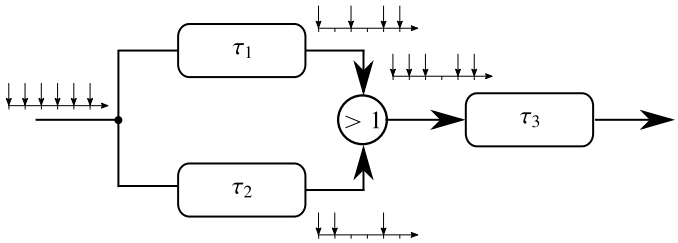
In verteilten Systemen sind entlang eines Signalpfades mehrere Kommunikations- und Berechnungsaufgaben beteiligt. Dabei existieren zwischen vielen Aufgaben Abhängigkeiten, so dass die Reihenfolge ihrer Ausführung verkettet ist. So wird wie in Unterabschnitt (3.1.1) beschrieben, bei einem Signalpfad immer dann die nächste Aufgabe ausgeführt, wenn die vorherige abgeschlossen ist.
Um Nebenläufigkeiten in der Ausführung von Aufgaben abzubilden, werden Signalpfade verzweigt und an einer späteren Stelle wieder vereinigt. Abschätzungen hinsichtlich der Ereignisdichten von vereinigten Ereignissequenzen wurden in (Wandeler, 2006) und (Perathoner et al., 2010) gemacht. Dabei unterscheidet man zwischen einer UND- und einer ODER-Vereinigung der Signalpfade. Bei einer UND-Vereinigung wird dann die Ausführung der nächsten Aufgabe initiiert, wenn alle Pfade ihre Ausführung abgeschlossen haben. Dies wird dann angewandt, wenn Ergebnisse von verschiedenen nebenläufigen Signalpfaden für eine anschließende Berechnung vorliegen müssen. Des Weiteren kann die UND-Vereinigung dazu verwendet werden, verschiedene nebenläufige Prozesse zu synchronisieren.
Werden hingegen Signalpfade durch eine ODER-Vereinigung verknüpft, wird immer dann die Ausführung der folgenden Aufgabe gestartet, wenn mindesten ein Pfad ausgeführt wurde.
Wurden in (Perathoner et al., 2010; Wandeler, 2006) für mehrere eingehende Ereignisdichten die ausgehende Ereignisdichte nach einer UND-/ODER-Vereinigung bestimmt, werden in dieser Arbeit mehrere Tasks durch eine Ereignissequenz angeregt und anschließend zusammengeführt. Des Weiteren wird im Folgenden nicht die ausgehende Ereignissequenz evaluiert, sondern die zusammengefasste Verlustdichte bestimmt. Abbildung 5.9 stellt die hier betrachteten Vereinigungen und Verkettung von Ereignissen anhand exemplarischer Ereignissequenzen dar. Können dabei den Aufgaben τ_1 und τ_2 maximale Verlustdichten zugewiesen werden, so stellt sich die Frage nach der zusammengefassten maximalen Verlustdichte. Für die UND-Vereinigung von Ereignissen eines Signalpfades gilt

$$d_{k,AND} = d_{k,\tau_1} \vee d_{k,\tau_2} \vee d_{k,\tau_3} \cdots = \inf_j \left\{ \sum_j d_{k,j}, 1 \right\} \tag{5.21}$$

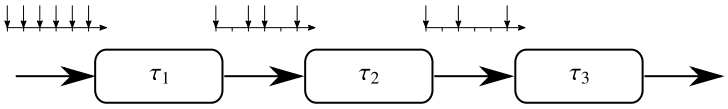
So folgt aus einer UND-Vereinigung der Ereignisse mehrerer Signalpfade eine ODER-Verknüpfung der verworfenen Ereignisse einer zusammengefassten Verzweigung.



(a) UND-Verknüpfung zweier Signalpfade.



(b) ODER-Verknüpfung zweier Signalpfade.



(c) Konkatenation von Dichtefunktionen.

Abb. 5.9: Verkettung von Dichtefunktionen.

Satz 11. *Die resultierende maximale Verlustdichte einer UND-Verknüpfung von j parallelen Signalpfaden ergibt sich zu*

$$dD^{+}_{AND}(\Delta_k) \leqslant \left(\left(\sum_j dD^{+}_j \right) \otimes (i)_{i \in \mathbb{N}_0} \right) (\Delta_k).$$

Beweis. Setzt man den Zusammenhang der UND-Vereinigung aus Gleichung (5.21) in die Vorschrift zur Konstruktion der maximalen Dichtefunktion aus Gleichung (5.4) auf Seite 55, dann erhält man

$$
\begin{aligned}
dD^{+}_{AND}(\Delta_k) &= \sup_i \left\{ \sum_{k=1}^{\Delta_k} \inf \left\{ \sum_j d_{k+i,j}, 1 \right\} \right\} \\
&= \sup_i \left\{ \inf \left\{ \sum_j d_{i+1,j}, 1 \right\} + \inf \left\{ \sum_j d_{i+2,j}, 1 \right\} + \inf \left\{ \sum_j d_{i+3,j}, 1 \right\} \right. \\
&\quad \left. + \cdots + \inf \left\{ \sum_j d_{i+\Delta_k,j}, 1 \right\} \right\}.
\end{aligned}
$$

Zur Abschätzung der UND-Verknüpfung werden alle Kombinationen aus aufsummierten verworfenen Ereignissen und Einsen betrachtet und deren Infimum gebildet. Die Sequenzen an verworfenen Ereignissen werden dann durch die maximale Verlustdichte mit entsprechender Intervalllänge abgeschätzt. Unter Zuhilfenahme der Eigenschaft der Subadditivität (Gl. (5.8)) gilt

$$n \cdot dD^{+}(1) \geqslant dD^{+}(n).$$

Dadurch reduziert sich die Anzahl der Kombinationen zu folgender Ungleichung.

$$
\begin{aligned}
dD^{+}_{AND}(\Delta_k) \leqslant \inf \left\{ \sum_j dD^{+}_j(\Delta_k), \sum_j dD^{+}_j(\Delta_k - 1) + 1, \ldots, \right. \\
\left. \ldots, \sum_j dD^{+}_j(1) + \Delta_k - 1, \Delta_k \right\}
\end{aligned}
$$

Diese Abschätzung entspricht nach Anhang B.1 der Min-Plus Faltung aus dem obigen Satz. Die so erzeugte maximale Dichtefunktion kann nach Definition 15 als valide bezeichnet werden, da wie in (Le Boudec und Thiran, 2001) gezeigt, sowohl die Addition als auch Min-Plus Faltung von subadditiven Funktionen wieder eine Funktion mit subadditivem Wachstum ergeben. Die abschließende Min-Plus Faltung mit der Folge $(i)_{i=\mathbb{N}_0}$ garantiert zudem, dass nicht mehr als ein verworfenes Ereignis pro generiertem Ereignis auftritt. □

Die Abschätzung der verworfenen Ereignisse bei der UND-Vereinigung durch Satz 11 führt einen Pessimismus ein, der dadurch begründet ist, dass Dichtefunktionen es nicht erlauben Korrelationen von Ereignissen darzustellen. Dichtefunktionen beschreiben die maximale Anzahl von verworfenen Ereignissen für ein Intervall an aufeinander folgenden Ereignissen und abstrahieren dadurch von einer konkreten Ereignisinstanz. Existiert eine Korrelation, so dass verworfene Ereignisse in den Si-

gnalpfaden nicht gleichzeitig auftreten können, induziert dies einen Pessimismus in der Abschätzung des obigen Satzes.

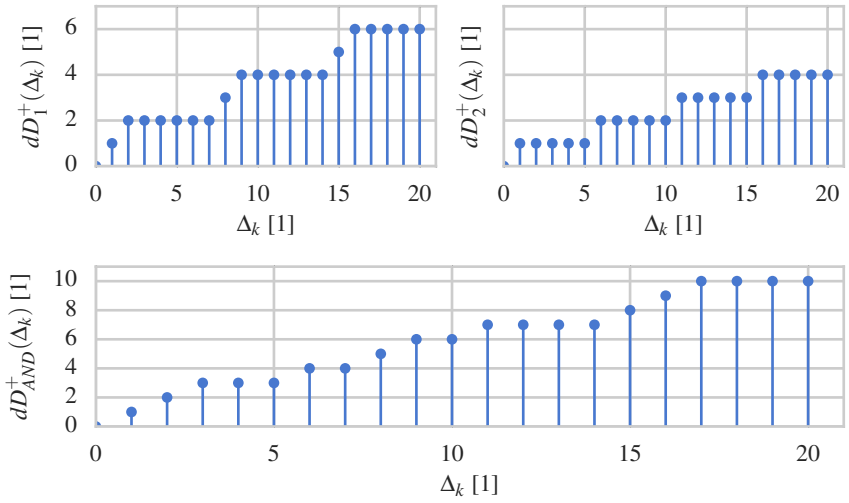


Abb. 5.10: Zusammengefasste Verlustdichte zweier Tasks mit UND-Vereinigung der ausgehenden Ereignissequenzen.

Beispiel 6. In diesem Beispiel nehmen wir die UND-Verknüpfung von Ereignissen zweier parallel aktivierten Tasks wie in Abbildung 5.9a an. Eine der Tasks verwirft maximal zwei Ereignisse in einem Intervall von sieben aufeinander folgenden Aktivierungen ($cP_1 = \{(7,2)\}$), während die andere Task nicht mehr als jedes fünfte Ereignis verwirft ($cP_2 = \{(5,1)\}$). Die beiden Verlustdichten $dD_1^+(\Delta_k)$ und $dD_2^+(\Delta_k)$ sind in Abbildung 5.10 dargestellt. Mit der Berechnungsvorschrift aus Satz 11 lässt sich die zusammengefasste Verlustdichte der UND-Verknüpfung abschätzen, wie sie im unteren Teil von Abbildung 5.10 dargestellt ist.

Bei der Betrachtung der ODER-Verknüpfung von Ereignissen wird immer dann die darauffolgende Aufgabe gestartet, wenn das Ereignis der jeweiligen Instanz mindestens einen der Signalpfade durchlief. So muss für ein verworfenes Ereignis der Verzweigung mit anschließender ODER-Vereinigung in allen Signalpfaden das jeweilige Ereignis verworfen werden. Demnach gilt

$$d_{k,OR} = d_{k,\tau_1} \wedge d_{k,\tau_2} \wedge d_{k,\tau_3} \cdots = \prod d_{k,j} = \inf_j \{d_{k,j}\} \tag{5.22}$$

Analog zu der vorgestellten UND-Verknüpfung erfolgt die Bestimmung der verworfenen Ereignisse einer ODER-Verknüpfung über den binären UND-Operator.

Satz 12. *Die resultierende maximale Verlustdichte einer ODER-Verknüpfung von j Signalpfaden ergibt sich zu*

$$dD^{+}_{OR}(\Delta_k) \leqslant \left(\inf_j \left\{ dD^{+}_j \right\} \otimes \inf_j \left\{ dD^{+}_j \right\} \right) (\Delta_k)$$

Beweis. Setzt man den Zusammenhang der ODER-Vereinigung aus Gleichung (5.22) in die Vorschrift zur Konstruktion der maximalen Dichtefunktion aus Gleichung (5.4) auf Seite 55, dann erhält man

$$\begin{aligned} dD^{+}_{OR}(\Delta_k) &= \sup_i \left\{ \sum_{k=1}^{\Delta_k} \inf_j \{ d_{k+i,,j} \} \right\} \\ &\leqslant \sup_i \left\{ \inf_j \left\{ \sum_{k=1}^{\Delta_k} d_{k,+i,j} \right\} \right\} \\ &\leqslant \inf_j \left\{ \sup_i \left\{ \sum_{k=1}^{\Delta_k} d_{k,+i,j} \right\} \right\} \\ &= \inf_j \left\{ dD^{+}_j(\Delta_k) \right\} \end{aligned}$$

Da das Infimum mehrerer subadditiver Funktionen nicht zwangsläufig eine subadditive Funktion ergibt, muss im Anschluss noch eine Min-Plus Faltung angewandt werden. Diese ergibt nach (Le Boudec und Thiran, 2001) einen subadditiven Abschluss, wodurch die Subadditivität der resultierenden Funktion sichergestellt ist. □

Kann bei einer Aneinanderreihung von Tasks wie in Abbildung 5.9c jeder Task entlang des Signalpfads eine Verlustdichte zugewiesen werden, ergibt sich auch hierbei die Fragestellung nach der zusammengefassten Verlustdichte. Wie in Unterabschnitt 5.1.3.1 ausgeführt, besitzt die Verlustdichte als Definitionsbereich eine Anzahl an benachbarten Ereignissen. Bei der Bestimmung der zusammengefassten Dichtefunktion $(dD^{+}_{SER}(\Delta_k))$ einer seriellen Verkettung von Tasks muss beachtet werden, dass sich die angegliederte Dichtefunktion auf eine um die vorausgehenden verworfenen Ereignisse reduzierte Ereignisfolge bezieht.

Satz 13. *Die resultierende maximale Verlustdichte $(dD^{+}_{SER}(\Delta_k))$ einer Konkatenation von j Tasks mit korrespondierender maximaler Verlustdichte $dD^{+}_j(\Delta_k)$ ergibt sich über folgende Rekursion.*

$$\begin{aligned} a_0(\Delta_k) &= 0 \\ a_j(\Delta_k) &= dD_j^+(\Delta_k - a_{j-1}(\Delta_k)) + a_{j-1}(\Delta_k) \\ dD_{SER}^+(\Delta_k) &= a_j(\Delta_k) \end{aligned}$$

Beweis. Die maximale Anzahl der in einem Intervall verworfenen Ereignisse seriell verketteter Tasks ergibt sich aus der Summe der verworfenen Ereignisse der Einzelaufgaben. Der Maximalwert an verworfenen Ereignissen der ersten Task ergibt sich aus der maximalen Dichtefunktion $dD_1^+(\Delta_k)$. Für ein zu betrachtendes Intervall von Δ_k benachbarten Ereignissen verbleibt dabei für den zweiten Task ein reduziertes Intervall von $\Delta_k - dD_1^+(\Delta_k)$. Entsprechend ergibt sich für den zweite Task der maximale Verlust unter der Annahme möglichst wenig verworfener Ereignisse des ersten Tasks zu $dD_2^+(\Delta_k - dD_1^-(\Delta_k))$.
Aufgrund dieses Widerspruchs muss zur Bestimmung der zusammengefassten Verlustdichte zweier in Serie geschalteter Tasks folgende Gleichung ausgewertet werden.

$$\begin{aligned} dD_{SER}^+(\Delta_k) = \sup_{n\in\mathbb{N}_0} \{ & dD_1^+(\Delta_k) - n + dD_2^+(\Delta_k - dD_1^+(\Delta_k) + n) \\ & \mid n \leqslant (dD_1^+ - dD_1^-)(\Delta_k) \} \end{aligned}$$

Aus der Eigenschaft der Subadditivität maximaler Verlustdichten folgert

$$dD_2^+(\Delta_k - dD_1^+(\Delta_k) + n) \leqslant dD_2^+(\Delta_k - dD_1^+(\Delta_k)) + dD_2^+(n) \leqslant D_2^+(\Delta_k - dD_1^+(\Delta_k)) + n,$$

woraus sich das Maximum der obigen Gleichung abschätzen lässt zu

$$dD_{SER}^+(\Delta_k) \leqslant dD_1^+(\Delta_k) + dD_2^+(\Delta_k - dD_1^+(\Delta_k)). \tag{5.23}$$

Für mehr als zwei aufeinander folgende Tasks ergibt sich der zu beweisende Satz aus der rekursiven Anwendung dieser Gleichung. □

Die Berechnungsvorschriften zur Bestimmung der Verlustdichte für UND- und ODER-Vereinigungen von Ereignissen sind trivialer Weise kommutativ. Die Berechnung der zusammengefassten Verlustdichte der seriellen Verkettung ist hingegen nicht kommutativ. Eine Veränderung der Aufgabenreihenfolge und der zugehörigen Dichtefunktionen kann daher zu einem veränderten Resultat führen.

Beispiel 7. In diesem Beispiel wird angenommen, dass ein Signalpfad durch zwei seriell verkettete Tasks repräsentiert wird, denen jeweils eine maximale Verlustdichte

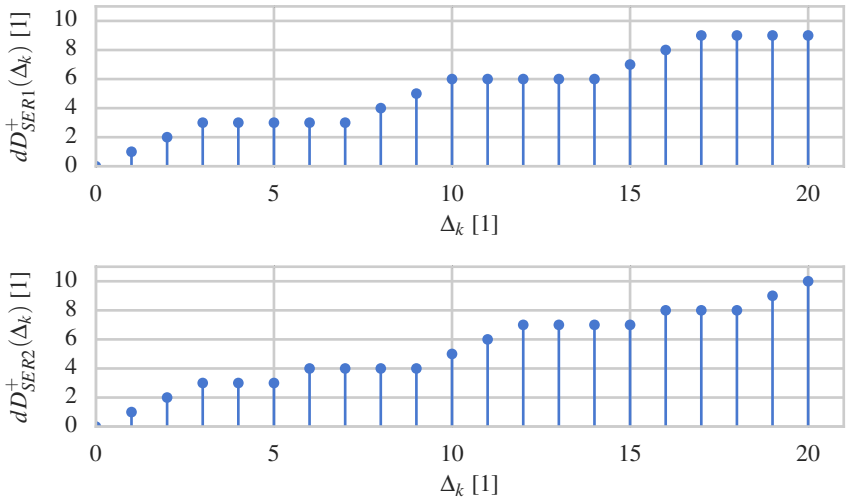


Abb. 5.11: Vergleich der zusammengefassten Verlustdichte beim Tausch zweier in Serie verknüpfter Tasks.

zugewiesen ist (siehe Abb. 5.9c). Die Verlustdichten dieser Tasks ergeben sich wie in Beispiel 6 aus den charakteristischen Punkten $(cP_1 = \{(7,2)\})$ und $(cP_2 = \{(5,1)\})$. Im ersten Szenario wird eine Ereignissequenz durch die Verlustdichte $dD^+_1(\Delta_k)$ und im Anschluss durch die Verlustdichte $dD^+_2(\Delta_k)$ spezifizierte Anzahl an Ereignissen reduziert. Die zusammengefasste Verlustdichte der in Serie verknüpften Tasks berechnet sich aus Satz 13 und ist im oberen Teil von Abbildung 5.11 abgebildet. Wird die Reihenfolge der den Tasks zugeordneten Verlustdichten getauscht, ergibt sich das untere Diagramm im Abbildung 5.11. Die zusammengefassten Verlustdichten $dD^+_{SER1}(\Delta_k)$ und $dD^+_{SER2}(\Delta_k)$ stimmen nur in wenigen Intervalllängen überein. So berechnet sich die Verlustdichte $dD^+_{SER1}(\Delta_k)$ für ein Intervall von $\Delta_k = 6$ zu

$$dD^+_{SER1}(6) = dD^+_2(6 - dD_1(6)) + dD^+_1(6) = dD^+_2(6-2) + 2 = 3.$$

Werden die Verlustdichten der Tasks getauscht ergibt sich hingegen

$$dD^+_{SER2}(6) = dD^+_1(6 - dD_2(6)) + dD^+_2(6) = dD^+_1(6-2) + 2 = 4.$$

Mit den eingeführten Brechungsvorschriften lassen sich beliebig verschachtelte Systeme zu einer resultierenden maximalen Dichtefunktion zusammenfassen, indem die

drei aufgeführten Sätze angewandt werden. Dadurch lässt sich aus den Dichtefunktionen der Elemente eines verteilten Systems die Verlustdichte des Gesamtsystems abschätzen.

5.1.5 Aussetzen von Ereignissen zur Lastreduktion

In einer Vielzahl eingebetteter Systeme werden sicherheitskritische Anwendungen zusammen mit Applikation geringer Sicherheitskritikalität ausgeführt (*mixed-critical systems*). Treten Ereignisse, welche die Ausführung sicherheitskritischer Anwendungen erfordern nicht permanent, sondern in sporadischen zeitlichen Abständen auf, wäre es ineffizient für diese Anwendungen permanent dedizierte Ressourcen bereitzustellen. Erlauben es hingegen zeitgleich ausgeführte Anwendungen ohne maßgeblichen Verlust ihrer Funktionalität, temporär auf Ressourcen zu verzichten, können die vorhandenen Systemressourcen effizienter ausgenutzt werden.
Besitzt eine Anregungsquelle, welche einen Signalpfad durch eine Ereignissequenz anregt, Kenntnis über den Ausführungsstatus der kritischen Anwendung, kann sie die Ausführung des Signalpfades zu Gunsten der sporadisch auftretenden Taskausführung aussetzen. Wird dabei die Ausführung des Signalpfades nicht häufiger als durch eine maximale Dichtefunktion $dD^+(\Delta_k)$ erlaubt ausgelassen, dann wird die zugehörige Performanzanforderung nicht unterschritten. Existiert weiterhin durch eine Plattformanalyse eine valide Schranke $(dD^+_{val}(\Delta_k))$ der verworfenen Ereignisse, die ohnehin in dem betroffenen Signalpfad auftreten, wird wie folgt die erlaubte maximale Dichtefunktion der durch die Anregungsquelle verursachte Ereignisverlust $(dD^+_{quelle}(\Delta_k))$ bestimmt.

Satz 14. *Ist die validierte maximale Verlustdichte des betrachteten Signalpfades kleiner als die erlaubte Verlustdichte $dD^+_{val}(\Delta_k) \leqslant dD^+(\Delta_k)$, so ergibt sich der maximal erlaubte Ereignisverlust in der Anregungsquelle zu*

$$dD^+_{quelle}(\Delta_k) = \left(\overline{n \otimes n}\right)(\Delta_k)$$

mit

$$n(\Delta_k) \leqslant \left(\left(\sup\left\{n \mid n + dD^+_{val}(i-n) \leqslant dD^+(i)\right\}\right)_{i\in\mathbb{N}_0} \overline{\oslash} 0\right)(\Delta_k).$$

Beweis. Die Anregungsquelle und der Signalpfad kann als Konkatenation von Tasks betrachtet werden. Für diese Konkatenation ist mit $dD^+(\Delta_k)$ die maximal erlaub-

te Verlustdichte spezifiziert. Gesucht ist die maximal erlaubte Verlustdichte für die Anregungsquelle, so dass die Verlustdichte der Konkatenation gemäß Satz 13 die spezifizierte Verlustdichte nicht überschreitet. Es wird somit für jede Intervalllänge der größte Wert gesucht, den $dD^{+}_{quelle}(\Delta_k)$ in der Ungleichung aus (5.23) annehmen kann.

$$dD^{+}(\Delta_k) \geqslant dD^{+}_{quelle}(\Delta_k) + dD^{+}_{val}(\Delta_k - dD^{+}_{quelle}(\Delta_k))$$

Die so ermittelte Folge kann jedoch im Allgemeinen nicht als valide Dichtefunktion bezeichnet werden, da sowohl die Eigenschaft der Monotonie als auch der Subadditivität nicht erfüllt sind. So garantiert eine Max-Plus Entfaltung mit der Nullfolge eine monoton steigende Funktion, indem nach der Definition der Max-Plus Entfaltung das Minimum des Wertes und den Folgeelementen gewählt wird. Der Ausdruck $\left(\overline{n \otimes n}\right)$ bezeichnet den subadditiven Abschluss (siehe 'subadditive closure' in Le Boudec und Thiran (2001)), welcher durch wiederholtes Anwenden der Min-Plus Faltung zu einer subadditiven Folge konvergiert. □

Um nun aus der Verlustdichte $dD^{+}_{quelle}(\Delta_k)$ zu bestimmen, ob das anstehende Ereignis verworfen werden kann, wird die Historie der Sequenz an verworfenen Ereignissen (d_k) gespeichert. Anhand dieser kann bewertet werden, in welchem Maße Ereignisse in der Vergangenheit ausgelassen wurden. Es ist dabei ausreichend, die Sequenz an verworfenen Ereignissen bis zu dem Intervall $\Delta^{c,max}_{k} - 1$ zu speichern. $\Delta^{c,max}_{k}$ entspricht dabei der maximalen Intervalllänge, der charakteristischen Punkte der Verlustdichte $dD^{+}_{quelle}(\Delta_k)$.
Für die Entscheidung, ob das gegenwärtige Ereignis ausgelassen werden kann, wird wie in Abbildung 5.12 aufgezeigt, für jede Intervalllänge der charakteristischen Punkte $(\Delta^{c}_{k,1}, \Delta^{c}_{k,2}, ..)$ die Summe der zuletzt verworfenen Ereignisse $(D_1, D_2, ..)$ gebildet. Dabei wird angenommen, dass das anstehende Ereignis verworfen wird $(d_0 = 1)$. Befinden sich dabei alle kumulativen Summen unterhalb der durch die Werte der charakteristischen Punkte festgelegten Grenzwerte, so ist es zulässig, dass das anstehende Ereignis ausgelassen werden kann, ohne dass die Dichtefunktion verletzt wird. Mathematisch betrachtet, kann eine Anregung durch eine Anregungsquelle ausgesetzt werden, wenn für alle charakteristischen Punkte die folgende Ungleichung erfüllt ist.

$$\left(\sum_{k=1-\Delta_c}^{-1} d_k\right) + 1 \leqslant n_c \;\forall(\Delta_c, n_c) \in cP \tag{5.24}$$

Im Kontext der sporadischen Ressourcenanforderung in Systemen mit unterschiedlichen Kritikalitäten ist es nicht ratsam, immer dann ein Ereignis in der Anregungs-

quelle auszulassen, wenn es Ungleichung 5.24 zulässt. Das System kann nur dann flexibel auf Ressourcenanforderungen höherpriorer Aufgaben eingehen, wenn in der Vergangenheit möglichst wenige Ereignisse verworfen wurden.

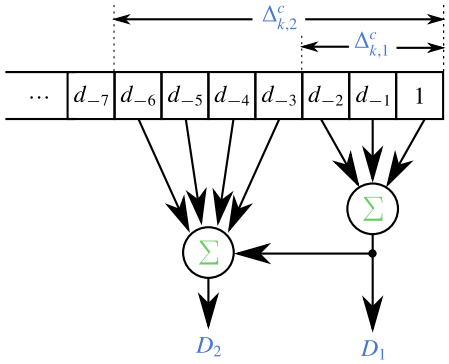


Abb. 5.12: Historie der zuletzt verworfenen Ereignisse.

Durch das Auslassen eines Ereignisses in der Anregungsquelle reduziert sich die Last des Systems um die Ausführungszeit der Aufgaben in dem betroffenen Signalpfad. Im Folgenden wird die garantierte Lastreduktion bestimmt, die ein Auslassen der Ereignisse um die maximale Verlustdichte in der Anregungsquelle verursacht. Dies ist zugleich die minimale garantierte Last, den höherpriore Aufgaben einnehmen können.

Nach (Redell und Sanfridson, 2002) ist die minimale Last, die ein Task erzeugt definiert durch

$$\rho^-_\tau(\Delta t) = \eta^-(\Delta_t) \cdot c^-_\tau$$

Dabei ist c^-_τ die minimale Nettoausführungszeit des Tasks τ und $\eta^-(\Delta_t)$ die minimale Ereignisdichte, mit welcher der Task angeregt wird. Da ein Ereignis der Anregungsquelle die Ausführung aller Tasks des Signalpfades aktiviert, kann die Last des Tasks zu einer Pfadlast eines Pfades P zusammenfasst werden.

$$\rho^-(\Delta_t) = \eta^-(\Delta_t) \cdot c^-, \quad c = \sum_{\tau \in P} c^-_\tau$$

Gleichung (5.3) auf Seite 54 stellt den Zusammenhang einer Verlustdichte zu der verringerten ausgehenden Ereignisdichte eines Tasks her. Über diesen Zusammenhang verringert sich durch den Verlust von Ereignissen die minimale Ereignisdichte

in obiger Gleichung, so dass sich für die reduzierte Last ergibt

$$\rho^{-}_{red}(\Delta_t) = (\eta^{-}(\Delta_t) - dD^{+}(\eta^{-}(\Delta_t)) \cdot c^{-}.$$

Die Last, die ein höherpriorer Task garantiert einnehmen kann, ist die minimale freiwerdende Last durch die Aussetzung des Signalpfades, bestimmt durch

$$\begin{aligned}\rho^{+}_{hp}(\Delta_t) &= \rho^{-}(\Delta_t) - \rho^{-}_{red}(\Delta_t) \\ &= dD^{+}(\eta^{-}(\Delta_t)) \cdot c^{-}.\end{aligned}$$

Aus der obigen Formel geht hervor, dass die minimale freiwerdende Last in direktem Zusammenhang mit der maximalen Dichte an verworfenen Ereignissen steht. Generiert dabei die Anregungsquelle eine periodische Abfolge an Ereignissen, kann nach (Richter, 2005) für die minimale Ereignisdichte folgender Zusammenhang eingesetzt werden.

$$\eta^{-}(\Delta_t) = \left\lfloor \frac{\Delta_t}{p} \right\rfloor$$

Die Betrachtung der Lastreduktion bezieht sich hierbei auf einen Signalpfad, dessen Tasks in einer verteilten Systemarchitektur unterschiedlichen Systemressourcen zugeordnet sind. Verursacht dabei die sporadisch auftretende höherpriore Aufgabe nur eine Last auf einer Systemressource, kann sie demnach nicht die gesamten Lastreduktion des Signalpfades beanspruchen. Die oben beschriebene Lastreduktion muss in solch einem Fall auf die Tasks des Signalpfades der entsprechenden Systemressource eingeschränkt werden.

5.2 Antwortzeiten und Latenzen

Wurde im vorherigen Abschnitt die Anzahl an verworfener Ereignisse durch eine Dichtefunktion im Intervallbereich begrenzt, so soll im Folgenden in ähnlicher Weise eine Spezifikation der auftretenden Latenzen im Intervallbereich definiert werden. Grundsätzlich lässt sich jeder Taskausführung eine Latenz in Form einer Antwortzeit zuordnen (Definition 7, Seite 19), die im Allgemeinen in bestimmten Grenzen variiert (siehe Unterabschnitt 3.1.1). Verändert sich durch die Latenzen das Systemverhalten, müssen für die Verifikation des Systemverhaltens diese Grenzen der Latenzen bestimmt werden.
Grundlage der im Weiteren definierten Latenzdichte ist die in Abschnitt 3.1.1 definierte Antwortzeit. Diese ist trivialer Weise immer ein Element der Menge an positiven reellen Zahlen. Aus der Sichtweise des funktionalen Verhaltens entspricht die Antwortzeit (r_k) der maximalen Zeitspanne, nach deren Verstreichen das Ergebnis der veranlassten Rechen- oder Kommunikationsaufgabe verfügbar ist. Diese Latenz resultiert aus der Rechenzeit der gestarteten Aufgabe zuzüglich der Interferenzen durch Tasks, die der selben Ressource zugeordnet sind. Eine triviale Begrenzung des Zeitverhaltens eines Tasks sind die Grenzwerte der Antwortzeit.

$$\begin{aligned} r_i^+ &\geqslant \sup_k \{r_{k,i}\} \\ r_i^- &\leqslant \inf_k \{r_{k,i}\} \end{aligned} \tag{5.25}$$

Solch eine Abschätzung der maximalen und minimalen Antwortzeit, wie beispielhaft in Abbildung 5.13 dargestellt, kann mit den in Abschnitt 3.1.4 aufgelisteten existierenden Methoden der Echtzeitanalyse bestimmt werden. Zur Reduktion des Rechenaufwands bei der Bestimmung dieser Grenzwerte und aufgrund teilweise pessimistischen Annahmen der Nettoausführungszeit, überschätzen die maximale und minimale Antwortzeit das reale Zeitverhalten. Des Weiteren kann durch diese Methoden keine Bewertung der Häufigkeit gemacht werden, mit der diese Grenzwerte auftreten können. So muss bei der Validierung von sicherheitskritischen Systemen unter anderem davon ausgegangen werden, dass die maximale respektive minimale Latenz permanent auftritt.
Dies manifestiert sich insbesondere bei der Bewertung von Regelsystemen. Die Regelgüte verringert sich nur im geringen Maße, wenn eine überdurchschnittlich große Latenz selten und in einer Umgebung von geringen Latenzen auftritt.

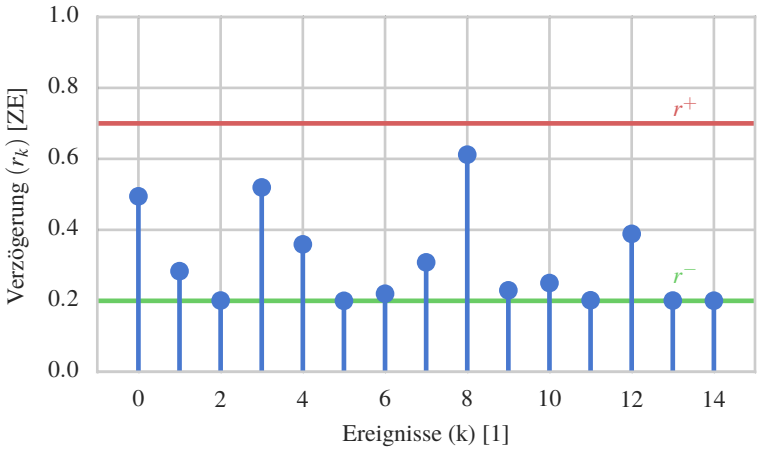


Abb. 5.13: Beispielhafte Sequenz von Antwortzeiten eines Tasks und deren Schranken.

In diesem Abschnitt wird daher eine neue Beschreibung der auftretenden Latenzen eingeführt, die ähnlich der im vorherigen Abschnitt eingeführten Verlustdichte die maximale und minimale kumulativen Latenz für ein Intervall an benachbarten Ereignissen begrenzt. Ein wesentlicher Unterschied, im Vergleich zu der Dichte an verworfenen Ereignissen, ist die Tatsache, dass eine Latenz einen beliebigen positiven reellen Wert annimmt, während für eine Verlustdichte nur ganzzahlige Werte zulässig sind (siehe Abschnitt 5.1.3.1). Da die im Folgenden eingeführte Latenzdichte und die Verlustdichte gleichermaßen das Auftreten von verworfenen Ereignissen und das Vorkommen von Latenzen über einen Grenzwert je Anzahl an benachbarter Ereignisse beschreiben, existiert eine große Gleichartigkeit der Definitionen. Aus diesem Grund werden nachfolgend Eigenschaften der Verlustdichte, die auch für die Latenzdichte gültig sind, nur grundlegend aufgeführt, Unterschiede jedoch detailliert herausgestellt. Das Modell der Latenzdichte wurde erstmals in (Bund und Slomka, 2013) vorgestellt.

5.2.1 Definition der Latenzdichte

Wie in Unterabschnitt 3.1.1 eingeführt, lässt sich jeder Taskaktivierung eine Antwortzeit zuordnen. Für die aufsummierte Antwortzeit der ersten i Taskausführungen

ergibt sich die kumulativen Latenz zu

$$R_i = \sum_{k=1}^{i} r_k,\ r_k \in \mathbb{R}_0^+.$$

Mit dieser streng monoton wachsenden Folge sind die Begrenzungen der Latenzen wie folgt definiert.

Definition 18 (Latenzdichte). *Die maximale Latenzdichte* $(dR^+(\Delta))$ *und die minimale Latenzdichte* $(dR^-(\Delta_k))$ *sind Funktionen* $f : \mathbb{N} \to \mathbb{R}_0^+$*, welche das Auftreten von Latenzen in einem Intervall* $\Delta_k = t - s$ *beschränken durch*

$$dR^+(t-s) \geqslant R_t - R_s \geqslant dR^-(t-s),\ \forall t \geqslant s \geqslant 1$$

Mit dem Zusammenhang aus Ungleichung (5.25) ergibt sich für eine Intervalllänge von einem Sample $(\Delta_k = 1)$

$$\begin{aligned} dR^+(1) &= r^+ \\ dR^-(1) &= r^-. \end{aligned} \tag{5.26}$$

5.2.2 Eigenschaften der Latenzdichte

Analog zu der eingeführten Dichte an verworfenen Ereignissen lässt sich die Latenzdichte durch Faltungsoperationen in der Min-Plus Algebra beschreiben.

$$dR^+(\Delta_k) \geqslant (R \oslash R)(\Delta_k)\,,\ R_0 \coloneqq 0$$

Demnach gelten die in den Unterabschnitten 5.1.3.2 und 5.1.3.3 eingeführten Eigenschaften der Monotonie und Subadditivität auch für die maximale Latenzdichte. Des Weiteren existiert eine gegenseitige Beschränkung der minimalen und maximalen Latenzdichte wie sie in Unterabschnitt 5.1.3.4 für die Dichte an verworfenen Ereignissen eingeführt wurde mit

$$dR^+(\Delta_{k,b}) - dR^+(\Delta_{k,a}) \geqslant dR^-(\Delta_{k,b} - \Delta_{k,a})\,,\ \forall \Delta_{k,b} \geqslant \Delta_{k,a}.$$

Beispiel 8. Angenommen, es ist möglich, dass in zwei aufeinander folgenden Taskinstanzen die minimale Latenz r^- auftreten kann, so gilt $dR^-(2) = 2 \cdot r^-$.

Mit $\Delta_{k,a} = 1$, $\Delta_{k,b} = 3$ und dem Faktum $dR^+(1) = r^+$ ist die maximale Latenzdichte für ein Intervall von drei benachbarten Ereignissen nach obiger Ungleichung begrenzt durch $dR^+(3) \geqslant r^+ + 2 \cdot r^-$.

Zur effektiven Speicherung und Verarbeitung der Latenzdichte lässt sich auch für die Latenzdichte das Konzept der charakteristischen Punkte anwenden (vgl. Abschnitt 5.1.4.3). Es ist jedoch zu beachten, dass im Gegensatz zu der Dichte an verworfenen Ereignissen keine Einschränkung hinsichtlich der Werte existiert. Konnte die Dichte an verworfenen Ereignissen für ein Intervall von einem Ereignis maximal den Wert 1 annehmen ($d_k \in \{0, 1\}$), ist für die Latenz einer Ereignisinstanz ein beliebiger positiver rationaler Wert erlaubt $\left(r_k \in \mathbb{R}_0^+\right)$. Daher muss zur Anwendung von Satz 5 auf die Latenzdichte die abschließende Min-Plus Faltung in Gleichung (5.10) ausgelassen werden, so dass gilt

$$dR^+(\Delta_k) = \left(\bigotimes_{(\Delta_k^c, n^c) \in cP} \left(\left\lceil \frac{i}{\Delta_k^c} \right\rceil \cdot n^c \right)_{i \in \mathbb{N}_0} \right)(\Delta_k).$$

5.2.2.1 Vereinigung von Latenzdichten

Ein wesentlicher Unterschied der Latenzdichte gegenüber der Dichte an verworfenen Ereignissen ergibt sich bei der Betrachtung zusammengesetzter Dichtefunktionen durch vereinigte Signalpfade. Dabei wird angenommen, dass sich das Zeitverhalten der Signalpfade durch eine maximale Latenzdichte beschreiben lässt und die Pfade durch eine vorausgehende Verzweigung zum gleichen Zeitpunkt aktiviert werden. Bei solch einer UND-Verknüpfung nach Abbildung 5.14 wird ein anschließender Task dann aktiviert, wenn alle Signalpfade ihre Berechnungen abgeschlossen haben. Aus diesem Zusammenhang lässt sich die zusammengefasste maximale Latenzdichte über folgenden Satz abschätzen.

Satz 15. *Die resultierende maximale Latenzdichte einer UND-Verknüpfung von j parallelen Signalpfaden ergibt sich zu*

$$dR^+_{AND}(\Delta_k) \leqslant \left(\sum_j dR^+_j \otimes \left(\sup_j \left\{ dR^+_j(1) \right\} \cdot i \right)_{i \in \mathbb{N}_0} \right)(\Delta_k).$$

Beweis. Der UND-Operator zur Vereinigung von Ereignissen bildet für jede Ereignisinstanz das Supremum der Latenz aller Signalpfade ($r_{k,AND} = \sup_j \{r_{k,j}\}$).
Für die Latenzdichte der zusammengefassten Signalpfade ergibt sich äquivalent zu

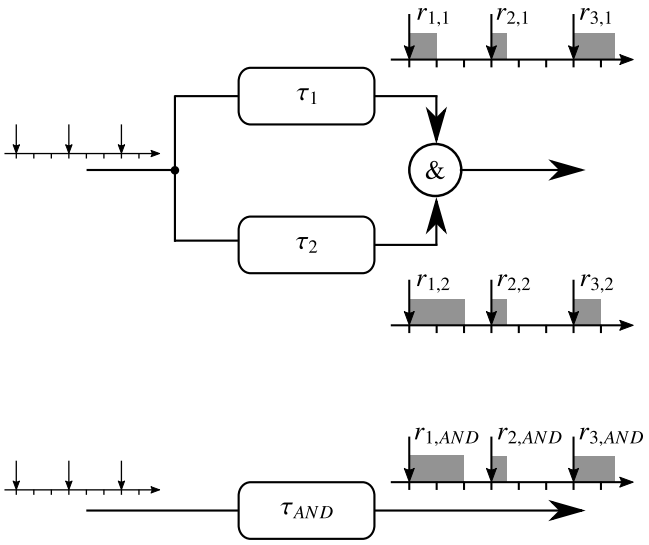


Abb. 5.14: Zusammengefasste Latenz bei der UND-Vereinigung zweier simultan aktivierter Signalpfade.

Gleichung (5.4) auf Seite 55

$$dR^+_{AND}(\Delta_k) = \sup_{i \geqslant 0} \left\{ \sum_{k=1}^{\Delta_k} \sup_j \{r_{k+i,j}\} \right\}.$$

Daraus lassen sich folgende Abschätzungen ableiten.

$$\begin{aligned} dR^+_{AND}(\Delta_k) &\leqslant \sup_{i \geqslant 0} \left\{ \sum_{k=1}^{\Delta_k} \sum_j r_{k+i,j} \right\} \\ &\leqslant \sum_j \sup_{i \geqslant 0} \left\{ \sum_{k=1}^{\Delta_k} r_{k+i,j} \right\} = \sum_j dR^+_j(\Delta_k) \end{aligned}$$

$$\begin{aligned} dR^+_{AND}(\Delta) &\leqslant \sup_i \left\{ \Delta_k \cdot \sup_j \{r_{i,j}\} \right\} \\ &= \sup_j \left\{ \Delta_k \cdot \sup_i \{r_{i,j}\} \right\} = \sup_j \left\{ dR^+_j(1) \right\} \cdot \Delta_k \end{aligned}$$

Die maximale valide Latenzdichte ergibt sich aus der Min-Plus Faltung der beiden Abschätzungen. □

Bei der Implementierung einer UND-Verknüpfung von verzögerten Ereignissen muss darauf geachtet werden, dass vor der Pfadvereingung Puffer vorgehalten werden. Diese werden dann benötigt, wenn folgende notwendige, aber nicht hinreichende Bedingung erfüllt ist.

$$\delta^-(2) < dR^+_{AND}(1) - dR^-_{AND}(1)$$

Werden Tasks, denen eine Latenzdichte zugeordnet werden kann, sequentiell ausgeführt, lässt sich die zusammengefasste Latenzdichte über folgenden Satz abschätzen.

Satz 16. *Die resultierende maximale Latenzdichte einer seriellen Verkettung von j Tasks ergibt sich zu*

$$dR^+_{SER}(\Delta_k) \leqslant \sum_j dR^+_j(\Delta_k).$$

Beweis. Da die maximale Latenzdichte den jeweiligen Extremwert für ein entsprechendes Intervall darstellt, ist die Summe der seriell verketteten Latenzdichten eine pessimistische, aber gültige Schranke. Durch die Addition werden die notwendigen Eigenschaften Subadditivität und Monotonie einer gültigen Latenzdichte nicht verletzt. □

5.2.3 Ableiten einer Verlustdichte aus einer Latenzdichte

Viele Betriebssysteme implementieren den Mechanismus eine Taskinstanz bei der Überschreitung einer Zeitfrist zu beenden. Insbesondere bei verzögerungssensitiven Systemen kann es vorteilhaft sein, das Ergebnis einer Berechnung zu verwerfen, wenn eine vorgegebene Zeitspanne seit der Aktivierung des Tasks verstrichen ist. So ist es bei Regelsystemen mit fester Abtastrate üblich, eine Frist gleich der Periodenlänge zu wählen. Eine derart gewählte Frist garantiert, dass ein aktuell berechneter Stellwert die Regelstrecke nicht vor einem älteren Stellwert beeinflusst. Durch einen bei Verlassen der Anregungsquelle beigefügten Zeitstempel, kann die Pfadlatenz jederzeit ausgewertet werden. Wird dabei die Frist während der Ausführung des Signalpfades überschritten, kann durch Terminierung des betreffenden Jobs die Systemlast reduziert werden.

Die folgende Berechnungsvorschrift zur Bestimmung der maximalen Dichte an verworfenen Ereignissen aus der Latenzdichte und einer Frist bezieht sich auf einen Task und kann durch Aneinanderreihen der Latenzdichten wie in Unterabschnitt 5.2.2.1 auch auf Pfadlatenzen angewandt werden.
Der Formalismus zur Bestimmung eines verworfenen Ereignisses im Zeitbereich ergibt sich für eine Frist F und Latenzen $r_k > 0$ zu

$$d_k = \begin{cases} 0 & \text{wenn } r_k \leqslant F \\ 1 & \text{wenn } r_k > F \end{cases} \tag{5.27}$$

$$= \inf \left\{ \left\lceil \frac{r_k}{F} \right\rceil - 1, 1 \right\} \tag{5.28}$$

und ist beispielhaft in Abbildung 5.15 skizziert.

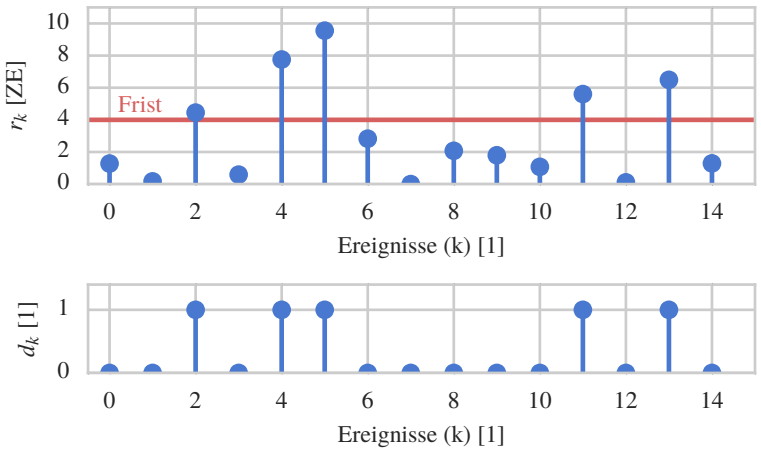


Abb. 5.15: Zusammenhang zwischen Fristüberschreitungen von Tasks und verworfenen Ereignissen.

Satz 17. *Aus einer maximalen Latenzdichte* $(dR^+(\Delta_k))$ *und einer gewählten Frist* F *ergibt sich die maximale valide Verlustdichte* $(dD^+(\Delta_k))$ *durch die Berechnungsvorschrift*

$$dD^+(\Delta_k) \leqslant \inf \left\{ \left\lceil \frac{dR^+(\Delta)}{F} \right\rceil - 1, \Delta \right\}.$$

Beweis. Die Berechnung der maximalen Verlustdichte entspricht der Adaption von Gleichung 5.27 für beliebige Intervalle unter Anwendung von Gleichung (5.4) auf

Seite 55:

$$
\begin{aligned}
dD^+(\Delta) &= \sup_i \left\{ \sum_{k=1}^{\Delta} \left(\inf \left\{ \left\lceil \frac{r_{k+i}}{F} \right\rceil - 1, 1 \right\} \right) \right\} \\
&\leqslant \sup_i \left\{ \inf \left\{ \sum_{k=1}^{\Delta} \left\lceil \frac{r_{k+i}}{F} \right\rceil - 1, \Delta \right\} \right\} \\
&\leqslant \inf \left\{ \sup_i \left\{ \sum_{k=1}^{\Delta} \left\lceil \frac{r_{k+i}}{F} \right\rceil - 1 \right\}, \Delta \right\} \leqslant \inf \left\{ \left\lceil \frac{dR^+(\Delta)}{F} \right\rceil - 1, \Delta \right\}
\end{aligned}
$$

□

5.2.4 Einfluss der Latenzdichte eines Tasks auf die ausgehende Ereignisdichte

Verzögert ein Task die eintreffenden Ereignisse durch eine variable Latenz, so kommt es zu einer Verdichtung der Ereignisfolgen am Ausgang des Tasks gegenüber der eingehenden Ereignisfolge. Kann dabei dem Task eine wie in Definition 18 eingeführte maximale und minimale Latenzdichte zugewiesen werden, lässt sich diese Verdichtung abschätzen. Der folgende Satz bezieht sich dabei auf die Intervallfunktion der Ereignisse, welche als Umkehrfunktion der Ereignisdichte aufgefasst werden kann. Nach (Gresser, 1993) gilt für die minimale Intervallfunktion

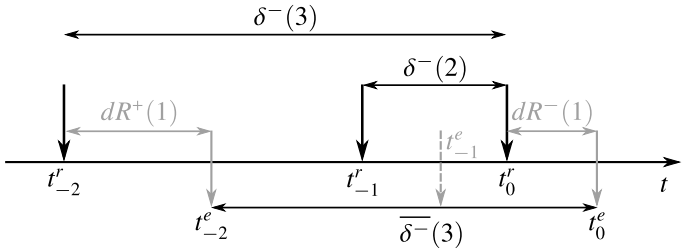
$$\delta^-(n) = \inf \{ \Delta_t \mid \eta^+(\Delta_t) = n \}$$

Satz 18. *Die ausgehende minimale Intervallfunktion $\bar{\delta}^-(n)$ eines Tasks ergibt sich aus der minimalen und maximalen Latenzdichte aus der eingehenden minimalen Intervallfunktion $\delta^-(n)$ durch*

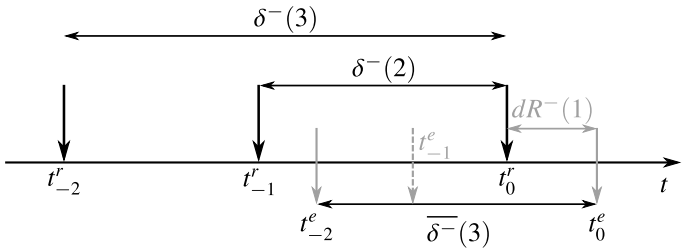
$$\overline{\delta^-}(n) = \sup \left\{ 0, dR^-(1) + \max_{0<l<n} \left\{ \frac{1}{l} \left(\left(\sum_{k=0}^{l-1} \delta^-(n-k) \right) - dR^+(l) \right) \right\} \right\}. \quad (5.29)$$

Beweis. Zur Bestimmung der minimalen ausgehenden Intervallfunktion wird eine eingehende Ereignissequenz angenommen, in der basierend auf dem Releasezeitpunkt (t_0^r) die vorausgehenden Ereignisse in minimalen zeitlichen Abstand zu diesem vorkommen.

$$t_0^r - t_{-n+1}^r = \delta^-(n)$$



(a) Ereignissequenz, wenn Ungleichung (5.30) erfüllt ist.



(b) Ereignissequenz, wenn Ungleichung (5.30) nicht erfüllt ist.

Abb. 5.16: Konstruktion des kleinsten Intervalls, in dem drei ausgehende Ereignisse auftreten können.

Zur Konstruktion des minimalen Intervalls von n ausgehenden Ereignissen, wird der Endzeitpunkt der Ereignisse am Interallbeginn durch Anwendung der maximalen Latenz möglichst weit in die Zukunft verschoben. Im Gegenzug dazu wird das letzte Ereignis in dem zu betrachtenden Intervall möglichst gering verzögert. So ergibt sich, wie in Abbildung 5.16a dargestellt, für den Endzeitpunkt

$$t^e_0 = t^r_0 + dR^-(1)$$

Bei der Untersuchung für ein Intervall von n ausgehenden Ereignissen unterscheidet man zwei Fälle. Für den Fall, dass die Ungleichung

$$dR^+(1) \leqslant t^r_{-n+2} + dR^-(1) - t^r_{-n+1} \tag{5.30}$$
$$\leqslant \delta^-(n) + dR^-(1) - \delta^-(n-1) \tag{5.31}$$

gilt, ist es nicht möglich, dass das Ereignis zu Beginn des Intervalls den Endzeitpunkt des darauf folgenden Ereignisses überschreitet ($t^e_{-n+1} \leqslant t^e_{-n+2}$) und es gilt für das

minimale Intervall von n ausgehenden Ereignissen

$$\overline{\delta^-}(n) = \delta^-(n) - dR^+(1) + dR^-(1). \tag{5.32}$$

Dieser Fall ist in Abbildung 5.16a dargestellt und entspricht für $l = 1$ Gleichung (5.29).

Ist hingegen Ungleichung (5.30) nicht erfüllt, so muss das Zeitverhalten weiterer Ereignisse und demnach die maximale Latenzdichte für größere Intervalllängen betrachtet werden. Bei der Konstruktion des minimalen Intervalls wird dabei der Endzeitpunkt t^e_{-n+1} bis zu dem Releasezeitpunkt t^r_{-n+2} verzögert und anschließend die verbleibende maximale Latenz zweier Ereignisse $(dR^+(2))$ gleichermaßen auf die beiden Ereignisse aufgeteilt. Durch diese Vorgehensweise verlassen die beiden ersten Ereignisse zum selben Zeitpunkt den Task $(t^r_{-n+1} = t^r_{-n+2})$ und das minimale Intervall der ausgehenden Ereignisse berechnet sich zu

$$\begin{aligned}\overline{\delta^-}(n) &= \delta^-(n-1) - \frac{1}{2}\left(dR^+(2) - \left(\delta^-(n) - \delta^-(n-1)\right)\right) + dR^-(1)\\ &= \frac{1}{2}\left(\delta^-(n) + \delta^-(n-1) - dR^+(2)\right) + dR^-(1)\end{aligned}$$

Dies entspricht Gleichung (5.29) für $l = 2$ und ist für ein beispielhaftes Zeitverhalten in Abbildung (5.16b) skizziert. Um bei dieser Konstruktion des kleinsten Intervalls nicht die Beschränkung der Latenzdichte für kleinere Intervalle außer Acht zu lassen, wird das Maximum aus dieser Gleichung mit Gleichung (5.32) gewählt. Gleichung (5.29) erweitert diese Methode für beliebige Anzahlen an Ereignissen und beschränkt die resultierende Intervalllänge durch Null, da andererseits für Latenzen größer als das Anregungsintervall negative Intervalllängen resultieren würden. □

5.2.5 Bestimmung der Latenzdichte aus den Dichtefunktionen des Echtzeitkalküls

Gegenwärtig existiert keine analytische Vorgehensweise, um eine Latenzdichte aus einem Plattformmodell abzuleiten. In diesem Unterabschnitt soll eine Methode skizziert werden, welche dies auf der Grundlage des in Unterabschnitt 3.1.4 eingeführten Echtzeitkalküls umsetzt. Mit diesem Ansatz, welcher im Rahmen einer Masterarbeit untersucht wurde (Naumova, 2016), wird eine Verbindung zwischen den Algorithmen der Echtzeitanalyse und der in diesem Abschnitt vorgestellten Latenzdichte hergestellt.

Wie in Unterabschnitt 3.1.4 beschrieben, sind die elementaren Parameter des Echtzeitkalküls die minimale sowie maximale Ereignisdichte $(\eta^{+/-}(\Delta_t))$ mit der ein Task angeregt wird und die minimale respektive maximale Kapazitätskurve $(\beta^{+/-}(\Delta_t))$, welche die zur Verfügung stehende Rechenzeit begrenzt. Die größte Antwortzeit tritt dann auf, wenn die maximale Anregung unter gleichzeitigem Mangel an Rechenzeit auftritt. So wird für die Bestimmung der maximalen Antwortzeit nach dem Network Calculus (Le Boudec und Thiran, 2001) der maximale horizontale Abstand der maximalen Ereignisdichte und der minimalen Kapazitätskurve gesucht. Für ein System mit diskreter Anregung ergibt sich die maximale Antwortzeit zu

$$r^+ = \sup\{r_n\}$$

mit

$$r_n = \sup\left\{b \,|\, \beta^-(b) \leqslant n\right\} - \inf\left\{a \,|\, \eta^+(a) \geqslant n\right\},\ n \in \mathbb{N} \tag{5.33}$$

Beispiel 9. Angenommen ein Task, der durch ein Greedy Processing Component (siehe Abb. 3.7 auf Seite 28) repräsentiert wird, hat eine maximale Nettoausführungszeit von $c^+ = 1\,ms$ und wird durch eine Ereignisquelle mit einer Periode von $p = 5\,ms$ und einem Jitter von $j = 9\,ms$ angeregt. Der Task wird in einem Zeitschlitzverfahren für $2,5\,ms$ in einem Zyklus von $6\,ms$ zur Ausführung gebracht. Die resultierenden Dichten sind in Abbildung 5.17 dargestellt. Für das Beispiel ergibt sich der maximale horizontale Abstand nach obiger Gleichung zu $r^+ = r_3 = 9\,ms$. Eine weitere Skizze (Abb. 5.18) veranschaulicht das Zustandekommen der maximalen Antwortzeit im Zeitbereich.

Mit der so bestimmten maximalen Antwortzeit ergibt sich die triviale Latenzdichte von $dR^+(\Delta_k) = r^+ \cdot \Delta_k$, was ein permanentes Auftreten der maximalen Antwortzeit

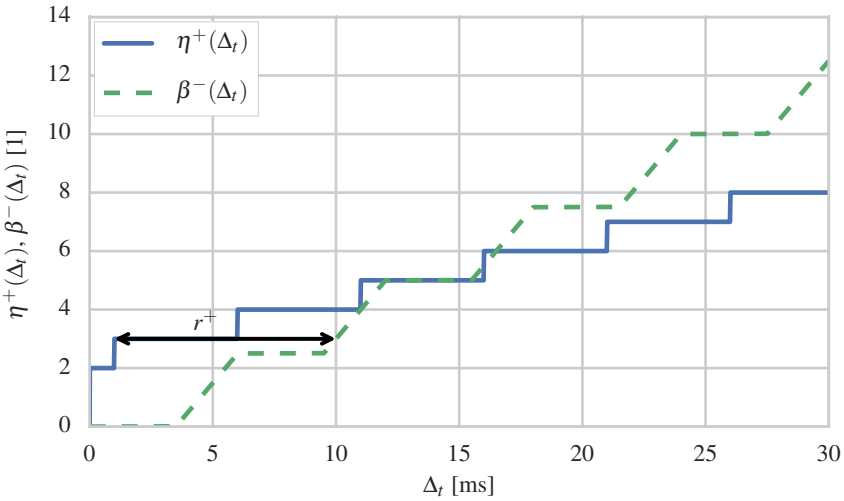


Abb. 5.17: Maximale Ereignisdichte $\eta^+(\Delta_t)$ und minimale Kapazitätskurve $\beta^-(\Delta_t)$, sowie die resultierende Antwortzeit $r^+ = r_3$.

impliziert. In Bezug auf die Latenzdichte ergibt sich die Fragestellung, ob die maximale Verzögerung r^+ an zwei aufeinander folgenden Aktivierungen auftreten kann. Ist dies nicht der Fall, ergäbe sich eine geringere Latenzdichte von $dR^+(2) < 2 \cdot r^+$. Angenommen die maximale Latenz r^+ tritt in einer durch die maximale Ereignisdichte aus Abbildung 5.17 beschränkten Ereignisfolge zur Ereignisinstanz $k = i$ auf $(r_i = r^+ = 9\,ms)$. Es kann mit Sicherheit bestimmt werden, dass die beiden vorausgehenden Antwortzeiten durch die Antwortzeiten aus Abbildung 5.18 mit

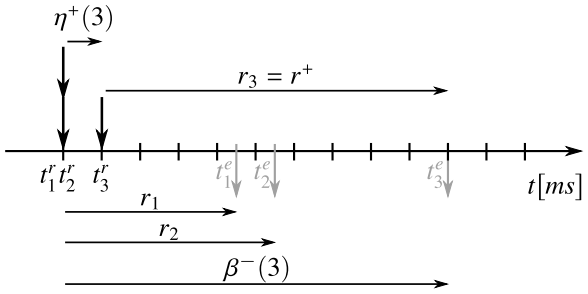


Abb. 5.18: Ereignissequenz, welche den Grenzfall der Ereignisdichte umsetzt.

$r_{i-1} \leqslant r_2 = 5,5\,ms$ und $r_{i-2} \leqslant r_1 = 4,5\,ms$ begrenzt sind.
Die Bestimmung der Antwortzeit r_{i+1} des Ereignisses, welches nach der maximalen Antwortzeit auftritt, ergibt sich aus der folgenden Abwägung. Zum einen kann angenommen werden, dass das Ereignis zum frühestmöglichen Zeitpunkt eintrifft. Für diese Annahme ergibt sich aus Gleichung 5.33 $r_{i+1} \leqslant r_4 = 5\,ms$. Eine weitere Möglichkeit besteht darin, dass das Ereignis, welchem die Antwortzeit r_{i+1} zugewiesen ist, für unbestimmte Zeit pausiert, so dass erneut eine Anhäufung der Ereignissequenz stattfinden kann. Darüber hinaus kann angenommen werden, dass zu diesem späteren Zeitpunkt die Rechenkapazität entsprechend der minimalen Kapazitätskurve zur Verfügung steht. Für diesen Fall wiederholt sich die Sequenz an Latenzen alle drei Ereignisse $r_k = \{4.5,\ 5.5,\ 9,\ 4.5,\ 5.5,\ 9,\ 4.5\ldots\}[ms]$. Bei der zuvor beschrieben Alternative findet eine Wiederholung erst ab dem vierten Ereignis statt $r_k = \{4.5,\ 5.5,\ 9,\ 5,\ 4.5,\ 5.5,\ 9,\ 5,\ 4.5\ldots\}[ms]$.

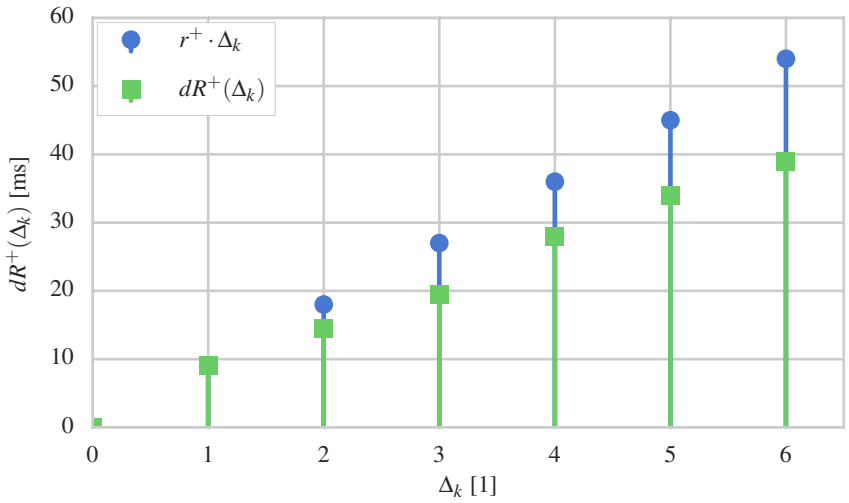


Abb. 5.19: Maximale Latenzdichte $(dR^+(\Delta_k))$ aus den beschränken Kurven des Echtzeitkalkül.

Für beide Szenarien ist es nicht möglich, dass die maximale Antwortzeit an zwei aufeinander folgenden Ereignisinstanzen auftritt. Die maximale Latenzdichte nach Definition 18 auf Seite 91 ist zu Beginn für beide Sequenzen gleich mit $dR^+(1) = 9\,ms$ und $dR^+(2) = 14,5\,ms$. Für ein Intervall von drei aufeinander folgenden Ereignissen ergibt sich die maximale Latenzdichte aus der Wiederholung nach dem vierten

Ereignis mit $dR^+(3) = 19,5\,ms$. Für ein Intervall von vier Ereignissen erreicht die Sequenz mit der Wiederholung nach dem dritten Ereignis die höhere Latenzdichte mit $dR^+(4) = 28\,ms$. Den Wert der maximalen Latenzdichte für größere Intervalllängen kann analog zu Satz 4 auf Seite 64 aus der Eigenschaft der Subadditivität abgeschätzt werden.

Für das Beispiel 9 ergibt sich somit eine Latenzdichte wie in Abbildung 5.19 dargestellt. Ein Vergleich mit Annahme einer permanent auftretenden maximalen Antwortzeit ergibt eine erhebliche Reduktion des Pessimismus. Durch die Kenntnis von $dR^+(2) = 14,5\,ms$ lässt sich die Aussage treffen, dass für eine Taskinstanz mit der Verzögerung von r^+ eine benachbarte Taskinstanz maximal eine Latenz von $14,5\,ms - 9\,ms = 5,5\,ms$ verursacht. Dies entspricht einer Reduktion um $38,9\%$ gegenüber der Annahme eines wiederholten Auftretens von r^+.

Kapitel 6
Analyse von Regelsystemen bezüglich Zeiteffekten

Mit der in Kapitel 5 beschriebenen Verlust- und Latenzdichte wurden ausdrucksstarke Modelle eingeführt, mit denen sich Zeiteffekte in verteilten Systemen effizient beschränken lassen. Solch eine Beschränkung in Form einer Dichtefunktion eignet sich besonders für Systeme, bei denen Signale in wiederkehrender Weise verarbeitet und übermittelt werden. Werden Signale hingegen nur singulär angefordert, wäre eine Betrachtung basierend auf Ereignisintervallen obsolet. Anwendungen, die Signale in wiederkehrender Weise bearbeiten, sind digitale Regelsysteme. Durch die Rückkopplung der im Regelkreis involvierten Signale, wirken sich Zeiteffekte nicht nur auf die gegenwärtige Ausführungsinstanz, sondern auch auf den weiteren Signalverlauf aus. Bei dauerhaftem oder wiederkehrendem Auftreten von Zeiteffekten können sich diese in verstärkender Weise degradierend auf die Regelgüte auswirken.

In diesem Kapitel werden die Auswirkungen von Zeiteffekten auf einen digitalen Regelkreis untersucht. Dazu wird ein Gütekriterium eingeführt, das ähnlich der Dichtefunktionen des vorherigen Kapitels, eine Bewertung des dynamischen Verhaltens eines Regelsystems im Intervallbereich vornimmt.
Die in Unterabschnitt 4.3.1 und 4.3.2 vorgestellten Stabilitäts- und Gütekriterien beziehen sich jeweils auf einen unendlichen Zeithorizont (siehe Definition 10 auf Seite 45) oder beurteilen die Regelgüte anhand von Charakteristika bei der Anregung mit einem Testsignal (vergleiche Abbildung 4.3 auf Seite 48). Im Folgenden wird hingegen ein Gütekriterium im Intervallbereich beschrieben, welches eine Bewertung des dynamischen Verhaltens sowohl für kurzzeitige Vorgänge, als auch für lange Intervalle an Ausführungsinstanzen des digitalen Reglers vornimmt. Dieses Modell zur Bewertung des Regelverhaltens im Intervallbereich kann zugleich auch als Spezifikation verwendet werden, aus der sich eine maximale Verlust- oder Latenzdichte der Plattform ableiten lässt. Durch die zusammengefasste Betrachtung der Signale

eines Regelkreises und den Zeiteffekten im Intervallbereich findet eine ganzheitliche Validierung des Regelverhaltens statt, in der Störsignale und destruktive Zeiteffekte abgewägt werden können. Aufgrund der Beschreibung des Signal- und Zeitverhaltens durch valide Schranken, eignen sich die im Folgenden vorgestellten analytischen Methoden für Systeme mit sicherheitskritischen Anforderungen.
Teile dieser Verfahren und Algorithmen wurden in (Bund und Slomka, 2015) veröffentlicht.

6.1 Bewertung der Regelgüte im Intervallbereich

Äquivalent zu der Verlust- und Latenzdichte wird im Folgenden eine Begrenzung der im Regelkreis involvierten Signale im Intervallbereich definiert. Die Einführung solch einer Signaldichte begründet sich durch eine ähnliche Motivation wie die Dichtefunktion für Zeiteffekte. Tritt beispielsweise ein Störsignal nur in großen zeitlichen Abständen auf, wäre es pessimistisch, die Störung nur durch ihren Maximalwert zu beschreiben. Durch die Auswertung des Signals in Relation zu einem Ereignisintervall kann jedoch ein zeitlicher Bezug hergestellt werden. Damit stellt die Signaldichte der Regeldifferenz ein ausdrucksstarkes Maß der Regelgüte eines Regelkreises dar. So kann über eine Begrenzung für kleine Intervalllängen die maximale Regeldifferenz und deren zeitliches Vorkommen limitiert werden. Entsprechend begrenzt eine Schranke für große Ereignisintervalle das Verhalten des Regelsystems in langen Zeithorizonten analog zu der Einschwingdauer einer Sprungantwort.

Nach Gleichung (4.3.1) auf Seite 35 legt der Zustandsvektor eines Regelsystems mit fortschreitender Anzahl an Ausführungsinstanzen eine Trajektorie im n-dimensionalen Raum zurück. Werte, die der Zustandsvektor dabei einnimmt, ergeben sich aus dem Anfangszustand $(\mathbf{x}_0)$, sowie Störungen und Zeiteffekten. Für die Beurteilung des schlechtesten Regelverhaltens solch eines Systems ergibt sich ein hoher Rechenaufwand, da alle erdenklichen Kombinationen der Einflussnahme auf das Regelsystem betrachtet werden müssen. Diese Berechnungskomplexität ergibt eine weitere Motivation der Bewertung des Regelverhaltens durch eine Begrenzung im Intervallbereich. Durch den Zusammenhang der auf ein Regelsystem einwirkenden Störungen und Zeiteffekte auf die Signale des Regelkreises im Intervallbereich, wird eine zeitaufwändige Untersuchung aller Kombinationen der Einflussnahme auf das Regelsystem hinfällig.

6.1.1 Eigenwerte und Matrixnormen

Um die Signale eines Regelsystems im Intervallbereich bewerten zu können, müssen die relevanten Eigenschaften aus der mathematischen Beschreibung des Regelsystems (siehe Abschnitt 3.2.2) extrahiert werden. In diesem Unterabschnitt werden die dafür notwendigen mathematischen Zusammenhänge aus dem Lehrbuchwissen zusammengefasst.

Die Zustandsnorm beschreibt die euklidische Norm des Zustandsvektors und ist Bestandteil mehrerer Metriken, da sie es erlaubt den Zustand eines Systems zu bewerten und zu vergleichen. So basieren die in Abschnitt 4.3.2 vorgestellten integralen Gütekriterien auf der euklidischen Norm. Dabei wird davon ausgegangen, dass sich das Regelsystem in dem Zielzustand befindet, wenn die Norm des Zustandsvektors Null ist. Des Weiteren findet die Norm des Zustandsvektors auch bei der Optimierung von Regelsystemen und den Mechanismen zur Ereignisgenerierung in ereignisgesteuerten Regelsystemen Anwendung (Heemels et al., 2008; Lunze und Lehmann, 2010; Tabuada, 2007). Für einen Spaltenvektor a berechnet sich die euklidische Norm zu

$$|\mathbf{a}| = \sqrt{\mathbf{a}^T \cdot \mathbf{a}}$$

mit $\mathbf{a}^T$ als transponiertem Spaltenvektor.

Betrachtet man ein lineares Gleichungssystem der Form $\mathbf{b} = M \cdot \mathbf{a}$ mit quadratischer, reeller Matrix M, dann kann die Matrixmultiplikation als eine lineare Abbildung des Vektors $\mathbf{a}$ auf den Vektor $\mathbf{b}$ aufgefasst werden. In einem zweidimensionalen System ergibt sich für den Vektor $\mathbf{b}$ eine Drehung und eine Streckung respektive Stauchung gegenüber des ursprünglichen Vektors $\mathbf{a}$ wie beispielhaft in Abbildung 6.1 dargestellt. Wird ein Kreis aufgespannt, der als Mittelpunkt den Startpunkt des

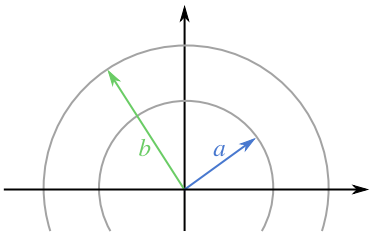


Abb. 6.1: Zweidimensionale Visualisierung der Spektralnorm.

Vektors aufweist und dessen Kreisrand auf dem Endpunkt des Vektors liegt, dann entspricht der Kreisradius der euklidischen Norm des Vektors. Für eine gegebene euklidische Norm des Vektors **a** lässt sich unter Verwendung der Spektralnorm der Matrix M der maximalen Radius, der durch den Vektor b aufgespannt wird, bestimmen.

$$|\mathbf{b}| \leqslant \|M\| \cdot |\mathbf{a}|$$

Dabei ist die Spektralnorm $\|M\|$ der Matrix M in der Numerik definiert als deren größter Singulärwert

$$\|M\| = \sigma_{max}(M) = \max_{\mathbf{a}\neq 0} \left\{ \frac{|M \cdot \mathbf{a}|}{|\mathbf{a}|} \right\} = \max_{\mathbf{a}\neq 0} \left\{ \frac{|\mathbf{b}|}{|\mathbf{a}|} \right\}. \tag{6.1}$$

Die Singulärwerte einer Matrix lassen sich anhand einer Singulärwertzerlegung bestimmen. Alternativ kann über folgenden Zusammenhang die Suche nach dem maximalen Singulärwert auf die Bestimmung des maximalen Eigenwerts zurückführt werden.

$$\sigma_{max}(M) = \sqrt{\lambda_{max}(M^T \cdot M)}$$

In den folgenden Abschnitten wird mit dem Terminus Norm die euklidische Norm für Vektoren, sowie die Spektralnorm für Matrizen bezeichnet.

6.1.2 Die Signaldichte

Ähnlich der begrenzenden Dichtefunktionen aus Kapitel 5 lässt sich aus der euklidischen Norm eines beliebigen zeitdiskreten Signalverlaufs $\mathbf{s}_k$ die Signaldichte definieren. Diese Signaldichte ergibt sich aus der Begrenzung der maximalen Summe der Vektornorm eines n-dimensionalen Signals für ein Intervall von Δ_k benachbarten Ereignisinstanzen. Ein Signal bestehend aus Skalarwerten kann dabei als eindimensionaler Vektor aufgefasst werden.

Definition 19 (Signaldichte). *Kann einer Ereignisinstanz ein Signalvektor zugewiesen werden, so ist für ein Intervall von Δ_k aufeinander folgenden Ereignissen die maximale Signaldichte ($\|\mathbf{s}(\Delta_k)\|$) eines zeitdiskreten Signalverlaufs $\mathbf{s}_k$ definiert durch*

$$\|\mathbf{s}(\Delta_k)\| \geqslant S_{i+\Delta_k} - S_i, \ \forall i \geqslant 1$$

mit

$$S_i = \sum_{k=1}^{i} |\mathbf{s}_k|$$

und

$$S_0 = 0.$$

Entsprechend zu Satz 1 ergibt sich für einen gegebenen Signalverlauf die kleinste Signaldichte zu

$$\|\mathbf{s}(\Delta_k)\| \geqslant \sup_{i \geqslant 0} \left\{ \sum_{k=1}^{\Delta_k} |\mathbf{s}_{k+i}| \right\} \; \forall \Delta_k \geqslant 1. \tag{6.2}$$

Beispiel 10. In Abbildung 6.2 sind drei diskrete Signalverläufe und deren zugehörige Signaldichte dargestellt. Für den ersten Signalverlauf ergibt sich $|\mathbf{s}_k| = 2$.
Entsprechend ergibt sich nach Gleichung (6.2) unabhängig von j die Signaldichte, welche den Signalverlauf eingrenzt $\|\mathbf{s}(\Delta_k)\| \geqslant 2 \cdot \Delta_k$. Durch die Periodizität von Signal 2 resultiert für $j = n \cdot 4$ mit $n \in \mathbb{N}_0$ die Signalfolge $|s_{k+j}| = \{0, 1, 2, 1, 0, 1, 2, \dots\}$ und weiter die Summe $\sum_{k=0}^{\Delta_k - 1} |\mathbf{s}_{k+j}| = \{0, 1, 3, 4, 4, 5, 7, \dots\}$. Entsprechend folgt für $j = n \cdot 4 + 1$ die Summe $\sum_{k=0}^{\Delta_k - 1} |\mathbf{s}_{k+j}| = \{1, 3, 4, 4, 5, 7, 8, \dots\}$. Das Supremum der daraus resultierenden vier Folgen ergibt den Verlauf der Signaldichte wie in Abbildung 6.2.

Eine derartige Begrenzung durch die Signaldichte lässt sich auf alle in einen Regelkreis involvierten Signale anwenden. Bewertet die euklidische Norm des Zustandsvektors das Regelsystem zu einem gewählten Zeitpunkt, lässt sich durch deren Folge die korrespondierende Signaldichte erstellen. Damit lässt sich nicht nur der Maximalwert des normierten Zustandsvektors spezifizieren, sondern zusätzlich dessen zeitliches Auftreten eingrenzen.
Es ist auch möglich, eine minimale Signaldichte zu definieren, welche das minimale Signalaufkommen eingrenzt. Diese ist jedoch für die Bewertung des schlimmsten Regelverhaltens nicht von Relevanz, weshalb im Folgenden die maximale Signaldichte schlichtweg als Signaldichte bezeichnet wird. Analog zu den Dichtefunktionen der Zeiteffekte ergeben sich auch für die Signaldichte Eigenschaften, die von akademischer und praktischer Wichtigkeit sind. So ist der Kurvenverlauf der Signaldichte durch die Eigenschaften der Monotonie und der Subadditivität begrenzt.

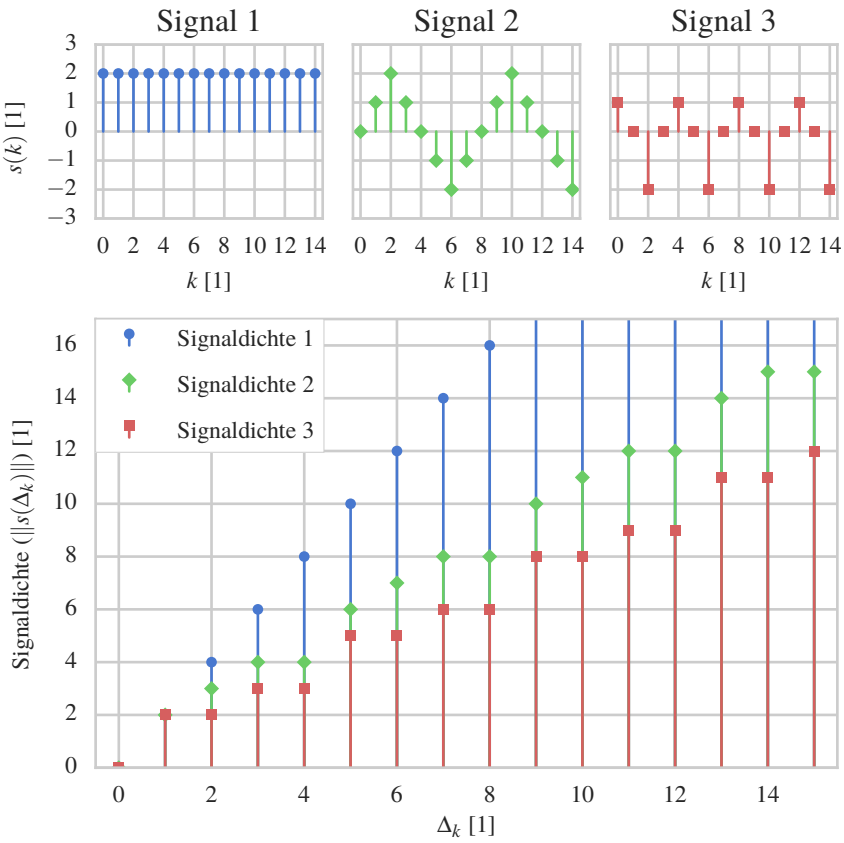


Abb. 6.2: Signaldichte für drei beispielhafte Signalverläufe.

6.1.3 Mathematische Operationen auf Signaldichten

In Wesentlichen werden in einem Regler Signale durch eine mathematische Vorschrift bearbeitet. Aber auch im Modell der Regelstrecke werden mathematische Operationen auf Signale angewendet. Werden hierbei die auf das Regelsystem einwirkenden Signale durch Signaldichten beschrieben, ergibt sich die Frage nach der resultierenden Signaldichte einer mathematischen Operation. In diesem Unterabschnitt werden deshalb die grundlegenden Rechenoperationen, welche auf Signale angewandt werden, in die korrespondierenden Algorithmen der Signaldichte transformiert.

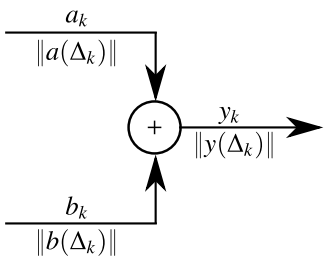


Abb. 6.3: Addition von Signalen, welche durch eine Signaldichte repräsentiert werden.

Werden Signale gleicher Dimension über eine Addition wie in Abbildung 6.3 zusammengefasst, ergibt sich bei der Betrachtung der resultierenden Signaldichte folgender Zusammenhang.

Satz 19. *Die Dichtefunktion der Summe zweier Signalfolgen* $(\mathbf{y}_k = \mathbf{a}_k + \mathbf{b}_k)$ *ist beschränkt durch*

$$\|\mathbf{y}(\Delta_k)\| \leqslant \|\mathbf{a}(\Delta_k)\| + \|\mathbf{b}(\Delta_k)\|$$

Beweis. Der Beweis des Satze ergibt sich durch Anwendung der Dreiecksungleichung auf Definition 19:

$$\begin{aligned}
\|\mathbf{y}(\Delta_k)\| &= \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} |\mathbf{a}_{k+j} + \mathbf{b}_{k+j}| \right\} \\
&\leqslant \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} |\mathbf{a}_{k+j}| + \sum_{k=0}^{\Delta_k - 1} |\mathbf{b}_{k+j}| \right\} \\
&\leqslant \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} |\mathbf{a}_{k+j}| \right\} + \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} |\mathbf{b}_{k+j}| \right\}
\end{aligned}$$

□

Durch die Dreiecksungleichungen wird ein Pessimismus eingeführt, insbesondere dann, wenn eine Korrelation der Signale existiert. So führt die obige Berechnung zu einer Überschätzung, wenn beispielsweise ausgeschlossen werden kann, dass die Maximalwerte der Signale nicht zur selben Ereignisinstanz auftreten können.
Oft ergibt sich eine zu Satz 19 umgekehrte Fragestellung. Für die gegebene Signaldichte der Summe und der eines Summanden wird die maximale Signaldichte des

zweiten Summanden gesucht, so dass die Signaldichte der resultierenden Summe nicht überschritten wird.

Satz 20. *Gegeben sei die Gleichung* $(\mathbf{y}_k = \mathbf{a}_k + \mathbf{b}_k)$ *mit den Signaldichten der Summe* $\|\mathbf{y}(\Delta_k)\|$ *und der eines Summanden* $\|\mathbf{a}(\Delta_k)\|$. *Die durch* $\|\mathbf{y}(\Delta_k)\|$ *spezifizierte Signaldichte kann dann nicht überschritten werden, wenn der weitere Summand* $\mathbf{b}_k$ *durch folgenden subadditiven Abschluss beschränkt ist*

$$\|\mathbf{b}(\Delta_k)\| \leqslant \overline{\|\mathbf{b}_e(\Delta_k)\| \otimes \|\mathbf{b}_e(\Delta_k)\|}$$

mit der endlichen Folge

$$\|\mathbf{b}_e(\Delta_k)\| = (0, \|\mathbf{b}(1)\|, \|\mathbf{b}(2)\|, \ldots, \|\mathbf{b}(\Delta - 1)\|, \gamma_{\Delta_k})$$

und

$$\gamma_{\Delta_k} = \left((\|\mathbf{y}\| - \|\mathbf{a}\|) \,\overline{\oslash}\, 0 \right) (\Delta_k) \,.$$

Beweis. Ausgangspunkt des Beweises ist die Ungleichung aus Satz 19.

$$\|\mathbf{y}(\Delta_k)\| \leqslant \|\mathbf{a}(\Delta_k)\| + \|\mathbf{b}(\Delta_k)\|$$

Eine spezifizierte Signaldichte $\|\mathbf{y}(\Delta_k)\|$ wird dann nicht durch die Signale der Summanden überschritten, wenn die Ungleichung umgekehrt wird und demnach für die gesuchte Signaldichte gilt

$$\|\mathbf{b}(\Delta_k)\| \leqslant \|\mathbf{y}(\Delta_k)\| - \|\mathbf{a}(\Delta_k)\| \,.$$

Für diese Differenz kann jedoch weder die für eine valide Dichtefunktion (vgl. Definition 15 auf Seite 64) geforderte Eigenschaft der Monotonie noch Subadditivität angenommen werden.
Durch eine Max-Plus Entfaltung (siehe Anhang B.1) mit der Nullfolge wird die Folge γ_{Δ_k} generiert, die ausgehend von der Intervalllänge Δ_k das Infimum für größere Intervalllängen bestimmt.

$$\gamma_{\Delta_k} = \inf_{i \geqslant 0} \{\|\mathbf{y}(\Delta_k + i)\| - \|\mathbf{a}(\Delta_k + i)\|\} = \left((\|\mathbf{y}\| - \|\mathbf{a}\|) \,\overline{\oslash}\, 0 \right) (\Delta_k)$$

Anschließend wird der subadditive Abschluss der Folge γ_{Δ_k} nach (Le Boudec und Thiran, 2001) gebildet. Dabei wird eine Min-Plus Faltung auf die Folge angewandt und das Ergebnis erneut gefaltet, bis sich die Folge dadurch nicht mehr ändert. □

Bei der klassischen Signalverarbeitung durch ein lineares zeitinvariantes System ergibt sich das Ausgangssignal durch die Faltung des Eingangssignals mit der Impulsantwort des Systems. Existiert für das Eingangssignal eine Begrenzung in der Form einer Signaldichte, so ergibt sich die Frage nach der Signaldichte am Systemausgang.

Satz 21. *Die ausgehende Signaldichte $\|y(\Delta_k)\|$ eines linearen zeitinvarianten Systems mit endlicher diskreter Impulsantwort g_k der Länge L und Signaldichte des Eingangssignals $\|x(\Delta_k)\|$ ergibt sich zu*

$$\|y(\Delta_k)\| \leqslant \|x(\Delta_k)\| \cdot \sum_{i=0}^{L-1} |g_i| \,.$$

Beweis. Ausgehend von der diskreten Faltungsoperation

$$y_k = (g * x)(k) = \sum_{i=0}^{L-1} g_i \cdot x_{k-i} \tag{6.3}$$

lässt sich eine Abschätzung der Signaldichte des Ausgangssignals erstellen.

$$\begin{aligned}
\|y(\Delta_k)\| &= \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} |y_{k+j}| \right\} \\
&= \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} \left| \sum_{i=0}^{L-1} g_i \cdot x_{k+j-i} \right| \right\} \\
&\leqslant \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} \sum_{i=0}^{L-1} |g_i| \cdot |x_{k+j-i}| \right\}
\end{aligned}$$

Da bei einem zeitinvarianten System der Systemausgang unabhängig von der Ereignisinstanz gleich auf ein Eingangssignal reagiert, lässt sich die Impulsantwort aus der verschachtelten Summe extrahieren, so dass gilt

$$\|y(\Delta_k)\| \leqslant \sup_j \left\{ \sum_{i=0}^{L-1} \left(|g_i| \cdot \sum_{k=0}^{\Delta_k - 1} |x_{k+j-i}| \right) \right\}.$$

Durch eine weitere Abschätzung kann die äußere Summe aufgeteilt werden zu

$$\|y(\Delta_k)\| \leqslant \sum_{i=0}^{L-1} |g_i| \cdot \sup_j \left\{ \sup_i \left\{ \sum_{k=0}^{\Delta_k - 1} |x_{k+j-i}| \right\} \right\}.$$

Der rechte Multiplikand entspricht der Signaldichte des Eingangssignals für eine Intervalllänge von Δ_k, was den obigen Satz beweist. □

Auch die resultierende Signaldichte einer Faltungsoperation induziert einen Pessimismus. In Unterabschnitt 6.1.5 wird gezeigt, wie diese Überschätzung durch weitere Kenntnisse der Eigenschaften des Signalverlaufs reduziert werden kann.

Beispiel 11. In diesem Beispiel wird ein Signal

$$x_k = \begin{cases} 0 & \text{für } k < 0 \\ \cos(k) & \text{für } k \geqslant 0 \end{cases}$$

mit der endlichen Folge $g_k = (1, 3.5, 5, 3.5, 1, -1, 0)$ gefaltet. Die Folgen, sowie der Absolutwert der diskreten Faltung $|y_k| = |(g * x)(k)|$ nach Gleichung (6.3), sind in Abbildung 6.4 dargestellt.
Ist für die Signalfolge x_k nur die Signaldichte $\|x(\Delta_k)\|$ bekannt, berechnet sich die Signaldichte des Faltungsergebnisses $\|y(\Delta_k)\|$ nach Satz 21. Im unteren Graph von Abbildung 6.4 ist die resultierende Signaldichte aus $\|x(\Delta_k)\|$ aus Satz 21, sowie das Ergebnis einer direkten Berechnung aus der Signalfolge y_k dargestellt. In dem betrachteten Fall ist nur eine geringe Überschätzung bei der Vorgehensweise nach Satz 21 zu beobachten. Dies kann jedoch nicht verallgemeinert werden. So stellt sich eine große Differenz für $|\sum g_k| \ll \sum |g_k|$ ein.

6.1.4 Anwendung der Signaldichte in Regelkreisen

Mit den hergeleiteten Gleichungen zur Beschreibung des Verhaltens eines digitalen Zustandsreglers (siehe Gl. (3.8) auf Seite 33) und dem eingeführten Modell der Signaldichte (Definition 19) lässt sich nun das dynamische Verhalten eines Regelsystems im Intervallbereich bewerten. Der Verlauf eines ungestörten Regelsystems aus einer initialen Auslenkung $(x_0 \neq 0)$ lässt sich mit Gleichung (3.8) auf Seite 33 unter der Annahme eines ungestörten Regelkreises $(w = 0)$ beschreiben. Durch wiederholte Anwendung dieser Gleichung ergibt sich eine Folge des fortschreitenden Zustandsvektors.

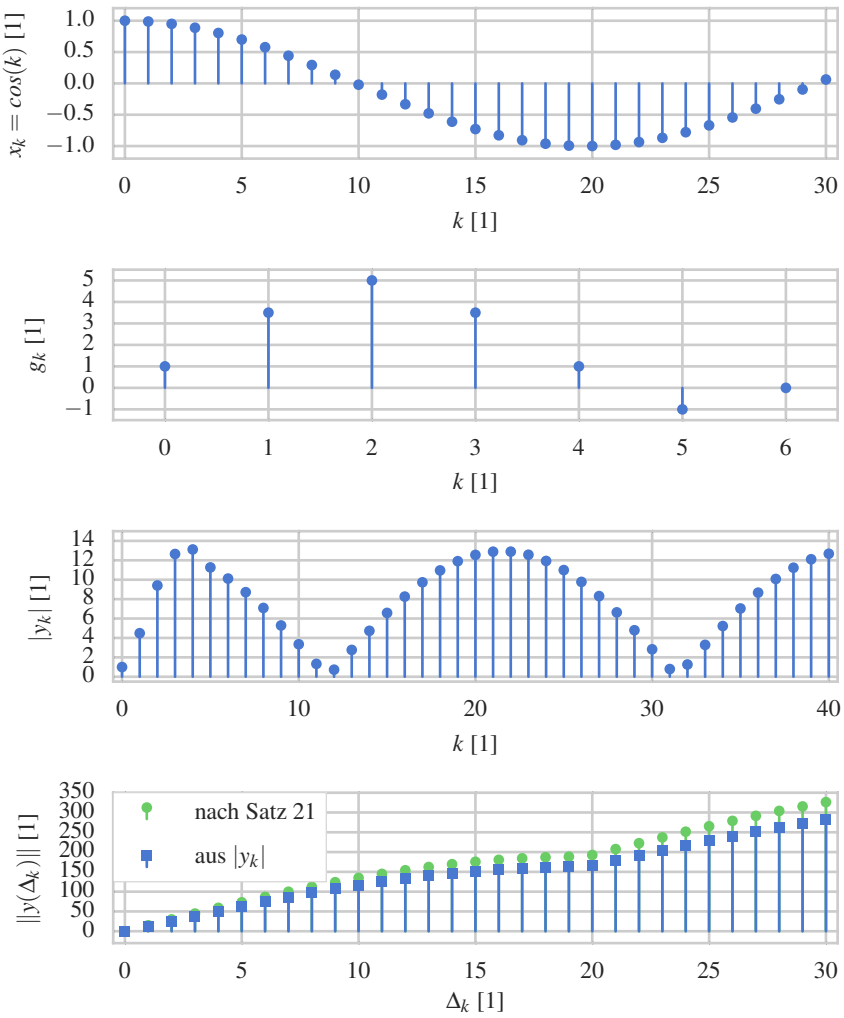


Abb. 6.4: Testsignal $x_k = \cos(k)$ mit beispielhafter Folge g_k und Faltungsergebnis $|y_k|$, sowie die resultierende Signaldichte $\|y(\Delta_k)\|$ durch direkte Berechnung nach Satz 21.

$$\mathbf{x}_1 = A_{cl} \cdot \mathbf{x}_0$$
$$\mathbf{x}_2 = A_{cl} \cdot \mathbf{x}_1$$
$$\mathbf{x}_3 = ..$$

Für ein Regelsystem mit asymptotischer Stabilität (vgl. Unterabschnitt 4.3.1) existiert eine Instanz k_ϵ, ab welcher eine Regelabweichung von ϵ nicht mehr überschritten wird.

$$|\mathbf{x}_k| < \epsilon, \ \forall k \geqslant k_\epsilon \tag{6.4}$$

Da in realen Systemen die Auflösung der zu übertragenden und zu verarbeitenden Signale begrenzt ist, kann demnach die Sequenz der Zustandsvektoren für solch ein System als endlich angenommen werden. Dies ist essentiell für die Konstruktion der Signaldichte nach obiger Definition, da das Supremum über alle Elemente der Sequenz bestimmt werden muss und eine endliche Länge Voraussetzung für deren Berechenbarkeit darstellt.

Beispiel 12. Ein beispielhaftes ungestörtes System in diskreter Zustandsraumdarstellung nach Gleichungen 3.8 auf Seite 33 weist folgende Parameter auf

$$A = \begin{bmatrix} 0,17 & 0,19 \\ -0,17 & 0,75 \end{bmatrix}, \quad B = \begin{bmatrix} 0,33 \\ -0,64 \end{bmatrix}, \quad K = \begin{bmatrix} 0,89 & 2,47 \end{bmatrix}$$

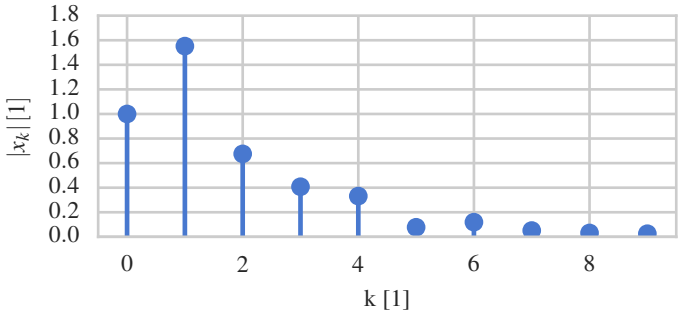
Für einen initialen Zustand von $|\mathbf{x}_0| = 1$ ergibt sich der Verlauf der Norm des Systemzustands und die dazugehörige Signaldichte wie in Abbildung 6.5 skizziert. Ausgehend von Gleichung (3.8) mit der Annahme von $w = 0$ lässt sich die Differenzengleichung (3.8) auflösen zu

$$\mathbf{x}_k = A_{cl}^k \cdot \mathbf{x}_0. \tag{6.5}$$

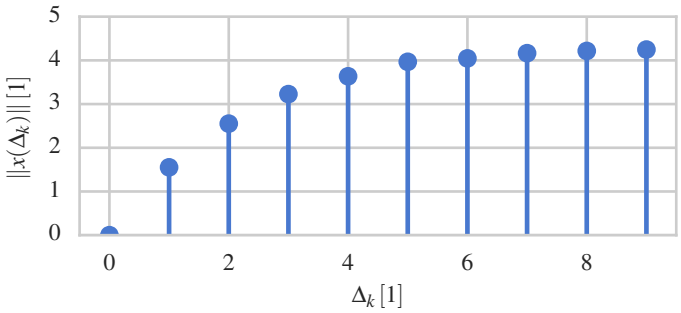
Unter Anwendung der Spektralnorm als größten Singulärwert einer Matrix lässt sich die Norm des Zustandsvektors abschätzen zu

$$|\mathbf{x}_k| \leqslant \left\|A_{cl}^k\right\| \cdot |\mathbf{x}_0| = \sigma_{max}\left(A_{cl}^k\right) \cdot |\mathbf{x}_0| \,. \tag{6.6}$$

Diese Ungleichung wird durch den Plot in Abbildung 6.5a beschrieben. Dabei ist zu beachten, dass diese Abschätzung nicht notwendigerweise eine Überschätzung ist, sondern, dass mindestens ein Initialwert mit der Eigenschaft $|\mathbf{x}_0| = 1$ existiert, für den obige Ungleichung einer Gleichung entspricht.



(a) Maximale Norm des Zustandsvektor eines beispielhaften Regelsystems für eine initiale Auslenkung von $|\mathbf{x}_0| = 1$.



(b) Signaldichte $\|\mathbf{x}(\Delta_k)\|$ für den obigen Verlauf des Zustandsvektors.

Abb. 6.5: Norm des Zustandsvektors eines ungestörten Systems mit initialer Auslenkung und korrespondierender Signaldichte.

Die Umwandlung der Elemente $|\mathbf{x}_k|$ in eine Signaldichte erfolgt wie in Definition 6.2 beschrieben. So ergibt sich für die Signaldichte eines ungestörten Regelkreises

$$\|\mathbf{x}(\Delta_k)\| \geqslant \sup_{k\in\mathbb{N}_0} \left\{ \sum_{j=0}^{\Delta_k - 1} \sigma_{max} \left(A_{cl}^{k+j} \right) \right\} \cdot |\mathbf{x}_0| = \|\alpha_0(\Delta_k)\| \cdot |\mathbf{x}_0| \,. \tag{6.7}$$

Der Term $\|\alpha_0(\Delta_k)\|$ ist dabei eine Folge, die ausschließlich von der Systemmatrix des geschlossenen Regelkreises abhängt und anhand derer aus einer Norm des initialen Zustandsvektors auf dessen Signaldichte geschlossen werden kann.
Eine Güteanforderung, spezifiziert durch eine Signaldichte, ist erfüllt, wenn diese für kein Intervall überschritten wird. Sucht man aus einer spezifizierten Signaldichte die

maximale initiale Auslenkung eines ungestörten Regelsystems, dann entspricht dies der Suche nach dem kleinsten Quotienten in der folgenden Ungleichung.

$$|\mathbf{x}_0| \leqslant \inf_{\Delta_k \in \mathbb{N}_0} \left\{ \frac{\|x(\Delta_k)\|}{\|\alpha_0(\Delta_k)\|} \right\}$$

Das erste Element der Signaldichte ($\|x(1)\|$) entspricht der größten Zustandsnorm, die im Regelsystem auftreten kann.
Anschaulich betrachtet entsprechen bei einer Sprungantwort die Werte kleiner Intervalllängen der Höhe und der Dauer des Überschwingers, während die Elemente größerer Intervalllängen zusätzlich die Dauer bis zum Erreichen des Stabilitätspunktes beinhalten. So kann die Signaldichte als ausdrucksstarkes Gütemaß für Regelsysteme aufgefasst werden, welches eine garantierte Grenze des Regelverhaltens aufzeigt und dabei mehrere Eigenschaften wie maximaler Überschwinger, Anstiegszeit und Einstellzeit vereint. Handelt es sich um einen stabilen Regelkreis, so sättigt sich die Signaldichte für zunehmende Intervalllängen in einen konstanten Wert.

6.1.5 Störeinfluss auf die Signaldichte des Zustandsvektors

Für bekannte Sequenzen der Führungsgröße $\mathbf{w}_k$ und der Störung $\mathbf{z}_k$ lässt sich anhand Gleichung (3.9) auf Seite 35 der Verlauf des Zustandsvektors im n-dimensionalen Raum deterministisch berechnen. Eine Untersuchung der Einflüsse dieser von außen auf das System einwirkenden Signale auf den Zustandsvektor ergibt zwei grundsätzliche Problemstellungen. So ist es insbesondere für Störungen realer Systeme nicht möglich a priori deren exakten Verlauf vorherzusagen. Eine veränderte zeitliche Abfolge des Störsignals resultiert in einem veränderten Verlauf des Zustandsvektors und führt im Allgemeinen zu einer veränderten Signaldichte. Um für eine spezifizierte Güteanforderung eine valide Grenze des Störeinflusses auf ein Regelsystem abzuleiten, müssen alle plausiblen Sequenzen des Störsignals untersucht werden, um den schlimmsten Fall zu erfassen. Die zweite Problemstellung ergibt sich aus der Länge der Störfolge, die untersucht werden muss, um den schlimmsten Fall für eine gewählte Intervalllänge zu konstruieren. Diese Fragestellungen der Menge und Länge an auszuwertenden Störfolgen mündet in dem 'Coverage Problem' (Donzé und Maler, 2007), wonach durch Ausführung einer Menge an numerischen Simulationen nicht sichergestellt werden kann, alle möglichen Signalverläufe abzudecken. Bei der Validierung durch Simulation wird ein System mit zufälligen Stimuli innerhalb der

spezifizierten Grenzwerten angeregt, um so eine möglichst große funktionale Abdeckung zu erhalten (Langer et al., 2007; Donzé und Maler, 2007). Dabei steigt mit dem Abdeckungsgrad der Rechenaufwand exponentiell. Bei den geläufigen Untersuchungen des Einflusses von Störungen auf ein Regelsystem finden stochastische Methoden oder eine Sensitivitätsanalyse Anwendung, die den Zusammenhang einer Störung auf das Ausgangssignal über eine Übertragungsfunktion herstellt. Diese Methoden sind jedoch nicht in der Lage den schlimmstmöglichen Fall des Regelverhaltens zu extrahieren und zu beschreiben.
Im Kontext der in den Unterabschnitten 6.1.2, 6.1.3 und 6.1.4 eingeführten Signaldichte ist es naheliegend, Signale, die das Regelsystem beeinflussen, durch deren Signaldichte zu beschränken, um daraus mit begrenztem Rechenaufwand die Signaldichte des resultierenden Zustandsvektors abzuleiten. Auch Verfahren, die umgekehrt aus einer spezifizierten Signaldichte des Zustandsvektors die maximal zulässige Einflussnahme auf das System bestimmen, werden im Weiteren vorgestellt.

Die Einwirkung einer Störung auf das Regelsystem lässt sich aufgrund des Superpositionsprinzip (Bronstein et al., 2001) eines wie in Gleichung (3.9) beschriebenen linearen Systems gesondert betrachten. Der Anteil der Störung $\mathbf{z}_k$ auf den Zustandsvektor ergibt sich demnach zu

$$\mathbf{x}_{k+1} = \mathbf{x}_{0,k+1} + \mathbf{x}_{z,k+1}$$

mit

$$\mathbf{x}_{0,k+1} = A_{cl}^{k+1} \cdot \mathbf{x}_0$$

und

$$\mathbf{x}_{z,k+1} = A_{cl} \cdot \mathbf{x}_{z,k} + E \cdot \mathbf{z}_k,\ k \in \mathbb{N}_0 \tag{6.8}$$

$$= \sum_{i=0}^{L-1} A_{cl}^i E \cdot \mathbf{z}_{k-i} \tag{6.9}$$

Analog zur Argumentation im vorherigen Unterabschnitt kann auch hier für ein asymptotisch stabiles System angenommen werden, dass der Einfluss eines Störimpulses nach einer ausreichenden Anzahl an Ereignisinstanzen vernachlässigbar ist und so der Betrachtungshorizont wie in Gleichung (6.4) beschrieben als endlich angenommen werden kann. Betrachtet man die Norm des Zustandsvektors, verursacht durch eine Folge an Störungen, lässt sich diese durch folgenden Zusammenhang abschätzen.

$$|\mathbf{x}_{z,k+1}| \leqslant \sum_{i=0}^{L-1} \zeta_i \cdot |\mathbf{z}_{k-i}| \,. \tag{6.10}$$

$$\zeta_i = \sigma_{max}(A_{cl}^i E)$$

Es ist offensichtlich, dass die Darstellung in Gleichung (6.10) einer Faltung wie in Gleichung (6.3) entspricht und sich somit eine gültige Abschätzung der Signaldichte des Zustandsvektors aus einer Signaldichte der Störung nach Satz 21 erstellen lässt. Diese Abschätzung induziert jedoch einen erheblichen Pessimismus, da von einem beliebigen Verlauf der Folge ζ_i ausgegangen wird. Betrachtet man jedoch die Folge ζ_i für ein asymptotisch stabiles System, dann ergibt sich ein ähnlicher Verlauf wie in Abbildung 6.5a dargestellt. Nach anfänglicher Auswirkung auf die Norm des Zustandsvektors reduziert sich der Einfluss eines Störimpulses mit zunehmender Anzahl an Ereignisinstanzen. Schätzt man die Folge ζ_i durch eine monoton fallende Hüllkurve ab, lässt sich die resultierende Signaldichte weniger pessimistisch abschätzen.

Satz 22. *Der Einfluss einer Störung, welche durch die Signaldichte $\|z(\Delta_k)\|$ beschränkt ist, auf die Signaldichte des Zustandsvektors ist begrenzt durch die Ungleichung*

$$\begin{aligned}\|\mathbf{x}_z(\Delta_k)\| \leqslant & \left(\|\mathbf{z}(i)\| \cdot \sum_{k=0}^{L-1} \zeta_k \right)_{i\in\mathbb{N}_0} \\ & \otimes \left(i \cdot \sum_{k=0}^{L-1} \left(\sup_{l\in\mathbb{N}_0} \{\zeta_{k+l}\} \cdot (\|\mathbf{z}(k+1)\| - \|\mathbf{z}(k)\|) \right) \right)_{i\in\mathbb{N}_0} (\Delta_k).\end{aligned}$$

Beweis. Der linke Term der Min-Plus Faltung entspricht der Abschätzung einer Signaldichte aus einer Faltung wie sie in Satz 21 beschrieben wurde. Der rechte Ausdruck begrenzt darüber hinaus den maximalen Wert den das Faltungsergebnis annehmen kann. Dabei wird die Folge ζ_i durch den Term

$$\sup_{l\in\mathbb{N}_0} \{\zeta_{k+l}\}$$

ersetzt und so durch eine monoton fallende Hüllkurve abgeschätzt. Für das theoretisch maximale Faltungsergebnis der Störung mit dieser fallenden Funktion wird aus der gegebenen Signaldichte der Störung die Störfolge mit der Eigenschaft

$$\sum_{k=0}^{\Delta_k - 1} \mathbf{z}_k = \|\mathbf{z}(\Delta_k)\|$$
$$\mathbf{z}_k = \|\mathbf{z}(k+1)\| - \|\mathbf{z}(k)\|$$

konstruiert. Bei solch einer Störfolge $(\mathbf{z}_k)$ wird für jedes Element der maximale Wert angenommen, der durch die begrenzende Signaldichte $(\|\mathbf{z}(\Delta_k)\|)$ ermöglicht wird. Das Ergebnis der Faltung dieser Folgen entspricht einer Begrenzung der resultierenden Signaldichte für ein Intervall von eins. Um nun das Ergebnis nach Satz 21 um diese Beschränkung aufzuwerten, wird eine Min-Plus Faltung mit der Treppenfunktion der Beschränkung durchgeführt. □

Kann der fallende Verlauf der Folge ζ_i nicht angenommen werden, so liefert unter Umständen der rechte Ausdruck in der obigen Min-Plus Faltung eine pessimistischere Abschätzung der resultierenden Signaldichte als die Beschränkung durch Satz 21. Durch die Min-Plus Faltung wird jedoch dadurch das Gesamtergebnis nicht verschlechtert, weshalb für Faltungen mit beliebigen Gewichtsfolgen Satz 22 immer Satz 21 vorzuziehen ist.

6.2 Verworfene Ereignisse im Regelkreis

Die Ausführungen in Abschnitt 6.1 definieren die Signaldichte und zeigen deren Eignung zur Begrenzung von Regelsignalen unter dem Einfluss von Störungen. Im Folgenden sollen diese Betrachtungen um den Einfluss von verworfenen Ereignissen in einem Regelkreis erweitert werden. So soll aus einer wie in Abschnitt 5.1 eingeführten Verlustdichte, die vergrößerte Signaldichte des Zustandsvektors bestimmt werden. Umgekehrt soll es auch möglich sein, aus einer Güteanforderung an den Regelkreis die maximale Verlustdichte abzuleiten, der das Regelsystem ausgesetzt werden kann.
In dem kommenden Abschnitt werden aus dem Einfluss der verworfenen Ereignisse und der damit einhergehenden Veränderung des Dynamikverhaltens im Zeitbereich, die Wechselbeziehung im Intervallbereich hergeleitet. Für die Auswirkung eines verworfenen Ereignisses auf das Regelverhalten ist es irrelevant, an welcher Stelle im Signalpfad des geschlossenen Regelkreises das betrachtete Ereignis verworfen wird. So wird in den folgenden Betrachtungen der Einfluss einer zusammengefassten Dichtefunktion auf das Regelsystem untersucht. Sind mehrere Tasks in dem Regelkreis involviert, denen jeweils eine Verlustdichte zugeordnet werden kann, so sind diese durch die Vorschriften aus Unterabschnitt 5.1.4.4 auf Seite 78 zusammenzufassen.

6.2.1 Auswirkungen verworfener Ereignisse im Zeitbereich

Unabhängig von der Ursache eines verworfenen Ereignisses, trifft in Folge an dem Aktorknoten kein aktualisierter Stellwert ein. Bei der gängigen Implementierung von Aktorknoten wird bei Ausbleiben einer Signalinstanz die Regelstrecke fortwährend mit dem alten Stellwert beaufschlagt. Erst wenn ein neuer Stellwert den Aktorknoten erreicht, wird dieser auf die Regelstrecke angewandt. Für das Ausbleiben des k-ten Ereignisses $(d_k = 1)$ ergibt sich somit für den Stellwert $\mathbf{u}_k = \mathbf{u}_{k-1}$, wie es auch in (Ling und Lemmon, 2002) angenommen wird. Es besteht auch die Möglichkeit zur Anwendung eines Nullsignals $(\mathbf{u}_k = 0)$ bei einem verworfenen Ereignis. Nach der Auswertung in (Schenato et al., 2007) ist jedoch keine der beiden Mechanismen der anderen überlegen. Aufwändigere Umsetzungen des Aktorknoten erlauben es, aus der Signalhistorie die verworfene Signalinstanz durch Interpolieren zu schätzen. Dieses Verfahren führt jedoch nicht immer zu einer Verbesserung des Regelverhaltens (Ling und Lemmon, 2003) trotz Zunahme des Rechenaufwandes.

In diesem Abschnitt werden die Auswirkungen ausbleibender Stellgrößen auf den Regelkreis im Zeitbereich aufgezeigt und durch eine äquivalente Störfolge ersetzt. Durch die Überführung der verworfenen Ereignisse in eine Störfolge ist die Annahme eines zeitinvarianten Systems wieder gültig und es können die konventionellen Formalismen auf das Regelsystem angewandt werden.

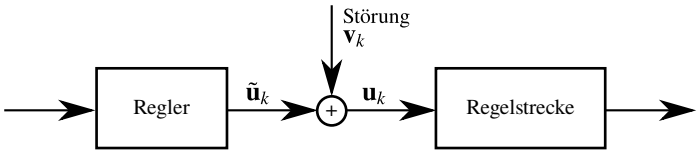


Abb. 6.6: Die durch ein verworfenes Ereignis verursachte Störung v_k auf die Stellgröße des Regelkreises.

Wird der Einfluss eines verworfenen Ereignisses wie in Abbildung 6.6 durch eine äquivalente Störung auf den Stellwert aufgefasst, dann lässt sich deren Auswirkung auf den Zustandsvektor wie in Unterabschnitt 6.1.5 bestimmen. Der Stellwert, welcher auf die Regelstrecke zur k-ten Ereignisinstanz einwirkt, wird als u_k bezeichnet, während der durch die Berechnungsvorschrift des Reglers ermittelte Stellwert mit $\tilde{u}_k$ angegeben wird. Die Auswirkungen eines verworfenen Ereignisses kann durch folgende Beziehung aufgezeigt werden.

$$\mathbf{u}_k = \begin{cases} \tilde{\mathbf{u}}_k & \text{wenn } d_k = 0 \\ \mathbf{u}_{k-1} & \text{wenn } d_k = 1 \end{cases} \tag{6.11}$$

Werden verworfene Ereignisse durch eine äquivalente Störfolge v_k abgebildet, so ergibt sich für diese

$$\mathbf{v}_k = (\mathbf{u}_{k-1} - \tilde{\mathbf{u}}_k) \cdot d_k. \tag{6.12}$$

Die Sequenz der Störsignale, ergibt sich aus der Differenz der durch den Regler berechneten Stellgrößen und der zur vorherigen Ereignisinstanz angewandten Stellgröße multipliziert mit der Sequenz an verworfenen Ereignissen. Gleichung (6.12) erschließt sich dabei aus Abbildung (6.6), wonach gilt $(\mathbf{u}_k = \tilde{\mathbf{u}}_k + \mathbf{v}_k)$. Für den Fall eines verworfenen Ereignisses zur j-ten Instanz $(d_j = 1)$ ergibt sich durch Erweiterung mit Gleichung (6.12)

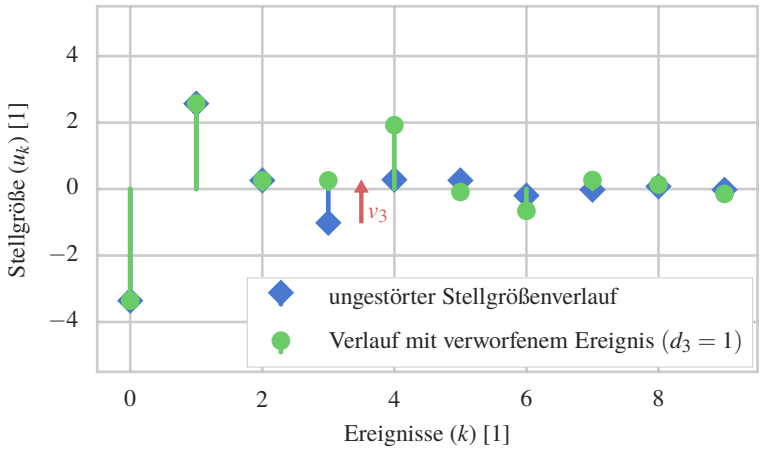


Abb. 6.7: Verlauf der Stellgröße eines asymptotisch stabilen Regelkreises mit einem verworfenen Ereignis.

$$\mathbf{u}_j = \tilde{\mathbf{u}}_j + \mathbf{u}_{j-1} - \tilde{\mathbf{u}}_j = \mathbf{u}_{j-1}.$$

Dies entspricht exakt dem Verhalten der Fallunterscheidung in Gleichung (6.11).

Beispiel 13. In Abbildung 6.7 ist der skalare Stellgrößenverlauf des Regelsystems aus Beispiel 12 mit $\mathbf{x}_0 = [1, 1]^T$ aufgetragen. Wird ein Ereignis an der Stelle $k = 3$ verworfen, so zeigt sich, dass der vorherige Stellwert für eine weitere Ereignisinstanz auf die Regelstrecke angewandt wird ($u_3 = u_2$). Auch die Größe der äquivalenten Störung v_3 auf die Stellgröße wird aus der Abbildung ersichtlich. Durch die Einwirkung der Störung v_3 verweilt die Stellgröße für eine weitere Ereignisinstanz auf dem selben Wert und divergiert dann im weiteren Verlauf von der ungestörten Stellgröße. Diese Zunahme der Dynamik vergrößert wiederum die Auswirkung etwaiger Ereignisverluste im weiteren Signalverlauf.

Wird die Berechnungsvorschrift der Stellgröße aus dem Zustandsvektor nach Gleichung (3.5) ($\mathbf{u}_k = K \cdot \mathbf{x}_k$) auf die Störfolge (Gl. (6.12)) angewandt, so ergibt sich

$$\mathbf{v}_k = \left(K \left(\mathbf{x}_{k-1} - \mathbf{x}_k\right) + \mathbf{v}_{k-1}\right) \cdot d_k. \tag{6.13}$$

Diese Gleichung macht deutlich, dass der Einfluss eines verworfenen Ereignisses direkt von der Differenz des Zustandsvektors ($\mathbf{x}_k - \mathbf{x}_{k-1}$) abhängt. Gleichung (6.13) bestätigt auch die intuitive Annahme, dass verworfene Ereignisse keinen Einfluss

auf ein Regelsystem haben, wenn dieses in einem stationären Zustand verweilt. Andererseits ist die Auswirkung verworfener Ereignisse groß, wenn sich das System in einer transienten Phase befindet.
Analog zur Gleichung (6.9) kann auch für den durch verworfene Ereignisse verursachten Einfluss auf den Zustandsvektor $\mathbf{x}_{v,k}$ auf das Regelsystem eine Differenzengleichung und deren Lösung erstellt werden.

$$\begin{aligned}\mathbf{x}_{v,k+1} &= A_{cl} \cdot \mathbf{x}_{v,k} + B \cdot \mathbf{v}_k \\ &= \sum_{j=0}^{\infty} \left(A_{cl}^{j} B \cdot \mathbf{v}_{k-j} \right)\end{aligned} \tag{6.14}$$

Die Differenz des Zustandsvektors spielt nach Gleichung (6.13) eine entscheidende Rolle bei der Bestimmung der Auswirkungen eines verworfenen Ereignisses. Für ein ungestörtes System mit initialer Auslenkung lässt sich nach dem Superpositionsprinzip die Zustandsdifferenz aufteilen in

$$\mathbf{x}'_{k+1} = \mathbf{x}_{k+1} - \mathbf{x}_k = \mathbf{x}'_{0,k+1} + \mathbf{x}'_{v,k+1} \tag{6.15}$$

$$\mathbf{x}'_{0,k+1} = \left(A_{cl}^{k+1} - A_{cl}^{k} \right) \cdot \mathbf{x}_0 \tag{6.16}$$

$$\mathbf{x}'_{v,k+1} = \left(A_{cl} - I \right) \cdot \mathbf{x}_{v,k} + B \cdot \mathbf{v}_k . \tag{6.17}$$

Diese Gleichungen ergeben sich direkt aus den Zustandsgleichungen (6.5) und (6.14). Ein zentraler Punkt bei der Evaluierung eines Regelsystems hinsichtlich dem Ereignisverlust ist die gegenseitige Abhängigkeit der Differenz des Zustandsvektors und der Auswirkung eines verworfenen Ereignisses auf den Zustandsvektor. Einerseits erhöht sich die Auswirkung eines verworfenen Ereignisses nach Gleichung (6.13) mit zunehmender Differenz des Zustandsvektors. Andererseits führen verworfene Ereignisse über Gleichungen (6.17) und (6.15) zu einer Veränderung und möglicherweise zu einer Vergrößerung der Zustandsdifferenz zukünftiger Signale. Durch diese wechselseitige Beziehung hängt der Einfluss eines verworfenen Ereignisses auf das Regelsystem von der Historie der verworfenen Ereignisse ab, wodurch sich die Evaluierung der Auswirkung verworfener Ereignisse auf ein Regelsystem im Zeitbereich aufwändig gestaltet.

6.2.2 Betrachtung im Intervallbereich

Basierend auf den Erkenntnissen aus dem vorangegangenen Unterabschnitt wird im Folgenden eine Methode eingeführt, welche die Auswirkung verworfener Ereignisse im Intervallbereich betrachtet. Zentrales Element sind dabei Folgen, die ähnlich einer Impulsantwort durch eine Faltung der vergangenen äquivalenten Störungen $(\mathbf{v}_{k-n})$ mit einer reglerspezifischen Gewichtungsfunktion (μ_n) die Auswirkungen des k-ten Ereignisverlustes bestimmen.

Satz 23. *Der betragsmäßige Störeinfluss durch den Verlust eines Ereignisses zur k-ten Ereignisinstanz ist begrenzt durch*

$$|\mathbf{v}_k| \leqslant \sum_{n=1}^{\infty} \mu_{v,n} \cdot |\mathbf{v}_{k-n}| + \theta_{v,k} \cdot |\mathbf{x}_0| \tag{6.18}$$

mit $\mathbf{v}_k = d_k = 0,\ \forall k \leqslant 0$ *und*

$$\theta_{v,k} = \left|K \cdot \left(A_{cl}^{k} - A_{cl}^{k-1}\right)\right| \tag{6.19}$$

$$\mu_{v,k} = \begin{cases} |K \cdot B - I| & \text{wenn } k = 1 \\ \left|K \cdot \left(A_{cl}^{k-1} - A_{cl}^{k-2}\right) \cdot B\right| & \text{wenn } k > 1. \end{cases} \tag{6.20}$$

Beweis. Durch das Superpositionsprinzip unterteilt sich die Zustandsdifferenz in Gleichung (6.13) für $d_k = 1$ in einen Anteil der initialen Auslenkung (Gl. (6.16)) und in einen Anteil des Störeinflusses (Gl. (6.17)).

$$\begin{aligned} \mathbf{v}_k &= -K \cdot \mathbf{x}_k^{'} + \mathbf{v}_{k-1} \\ &= -K \cdot \mathbf{x}_{v,k}^{'} + \mathbf{v}_{k-1} - K \cdot \mathbf{x}_{0,k}^{'} \end{aligned} \tag{6.21}$$

Für eine betragsmäßige Betrachtung ergibt sich daraus

$$|\mathbf{v}_k| \leqslant \left|K \cdot \mathbf{x}_{v,k}^{'} - \mathbf{v}_{k-1}\right| + \left|K \cdot \mathbf{x}_{0,k}^{'}\right|. \tag{6.22}$$

Der Einfluss einer initialen Auslenkung $(\mathbf{x}_0)$ auf die Auswirkung des Verlusts der k-ten Ereignisinstanz wird durch die Folge $\theta_{v,k}$ beschrieben. Diese Folge resultiert aus der euklidischen Norm von Gleichung (6.16), eingesetzt in die rechten Summanden von Gleichung (6.22).
Die Folge $\mu_{v,k}$ begrenzt die Auswirkung zuvor verworfener Ereignisse auf den aktuellen Ereignisverlust. Setzt man Gleichung (6.14) in Gleichung (6.17) ein, so ergibt

sich

$$\mathbf{x}'_{v,k+1} = \sum_{j=0}^{\infty} \left(A_{cl}^{j+1} - A_{cl}^{j} \right) \cdot B \cdot \mathbf{v}_{k-j-1} + B\mathbf{v}_k.$$

Daraus folgt für den linken Summanden aus Gleichung (6.22)

$$\begin{aligned} \left| K \cdot \mathbf{x}'_{v,k} - \mathbf{v}_{k-1} \right| &= \left| K \cdot \left(\sum_{j=0}^{\infty} \left(A_{cl}^{j+1} - A_{cl}^{j} \right) B\mathbf{v}_{k-j-2} + B\mathbf{v}_{k-1} \right) - \mathbf{v}_{k-1} \right| \\ &= \left| \sum_{j=0}^{\infty} K \cdot \left(A_{cl}^{j+1} - A_{cl}^{j} \right) B\mathbf{v}_{k-j-2} + (KB - I)\,\mathbf{v}_{k-1} \right| \\ &\leqslant \sum_{j=0}^{\infty} \left| K \cdot \left(A_{cl}^{j+1} - A_{cl}^{j} \right) B \right| \cdot |\mathbf{v}_{k-j-2}| + |KB - I| \cdot |\mathbf{v}_{k-1}| \end{aligned}$$

Aus diesem Zusammenhang ergibt sich direkt die Gewichtungsfunktion $\mu_{v,k}$. □

Die Gewichtungsfunktion $\mu_{v,k}$ kann als Sensitivitätsfunktion des Regelsystems hinsichtlich verworfener Ereignisse angesehen werden. Sie bestimmt wie stark sich vergangene Ereignisverluste auf das gegenwärtig verworfene Ereignis auswirken. Diese Sensitivitätsfunktion wird nach Gleichung (6.20) maßgeblich von der Matrixexponentialfunktion der Systemmatrix des geschlossenen Regelkreises A_{cl} beeinflusst.

Behauptung. Soll aus einer Verlustdichte die Degradation der Regelgüte bestimmt werden, dann darf eine kleinere maximale Dichte (größerer Abstand zwischen verworfenen Ereignissen) nicht dazu führen, dass sich die destabilisierenden Auswirkungen vergrößern. Diese Forderung bedingt, dass die Sensitivitätsfunktion $\mu_{v,k}$ folgende Eigenschaft erfüllt.

$$\mu_{v,k}^2 \leqslant \mu_{v,k+i} \cdot \mu_{v,k-i} \;\forall k, i \in \mathbb{N} \,|\, k - i > 0 \tag{6.23}$$

Ein einfaches Beispiel bestätigt die Notwendigkeit dieser Eigenschaft.

Beispiel 14. Es werden drei verworfene Ereignisse angenommen, die an den Ereignisinstanzen (k_0, k_1, k_2) auftreten. Wird die Auswirkung einer initialen Auslenkung außer Acht gelassen, lässt sich aus Gleichung (6.18) die äquivalente Störung des dritten verworfenen Ereignisses in Abhängigkeit der zuvor verworfenen Ereignisse abschätzen durch

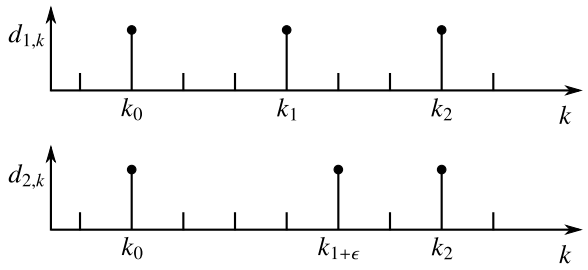


Abb. 6.8: Zwei Szenarien verworfener Ereignisse mit $iD^-(3) = 7$.

$$\begin{aligned}|\mathbf{v}_{k2}| &\leqslant \mu_{v,k2-k1} \cdot |\mathbf{v}_{k1}| + \mu_{v,k2-k0} \cdot |\mathbf{v}_{k0}| \\ &\leqslant (\mu_{v,k2-k1} \cdot \mu_{v,k1-k0} + \mu_{v,k2-k0}) \cdot |\mathbf{v}_{k0}| \,.\end{aligned}$$

Nach dieser Ungleichung resultieren Variationen in den Abständen der verworfenen Ereignisse in einem veränderten Störeinfluss ($|\mathbf{v}_{k2}|$).
Wir betrachten in diesem Beispiel zwei Szenarien, wie sie in Abbildung 6.8 skizziert sind. Im ersten Szenario sind die Abstände in denen verworfene Ereignisse auftreten gleich groß ($k_2 - k_1 = k_1 - k_0 = n$), womit gilt $\mu_{v,k2-k1} = \mu_{v,k1-k0}$. Eine veränderte Position des zweiten verworfenen Ereignisses $\left(\tilde{k}_1 = k_1 + \epsilon\right)$ im zweiten Szenario resultiert in einer höheren Dichte, da der minimale Abstand zweier Ereignisse verringert wird. Damit sich aus einer erhöhten Dichte kein verminderter Einfluss auf das Regelsystem ergibt, muss gelten

$$\left(\mu^2_{v,k_1-k_0} + \mu_{v,k_2-k_0}\right) \cdot |\mathbf{v}_{k_0}| \overset{!}{\leqslant} (\mu_{v,k_1+\epsilon} \cdot \mu_{v,k_1-\epsilon} + \mu_{v,k_2-k_0}) \cdot |\mathbf{v}_{k_0}| \,.$$

Es ist offensichtlich, dass diese Ungleichung gilt, wenn die Eigenschaft der Sensitivitätsfunktion $\mu_{v,k}$ aus Gleichung (6.23) erfüllt ist.

Um die in Gleichung (6.23) aufgeführte Eigenschaft der Sensitivitätsfunktion zu erfüllen wird diese durch ihre einhüllende Kurve mit exponentiellem Abfall

$$\overline{\mu}_{v,k} = \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot k} \geqslant \mu_{v,k} \tag{6.24}$$

mit positiven Parametern $\alpha_{v,\mu}$ und λ_v abgeschätzt.
Solch eine Funktion erfüllt die geforderte Eigenschaft da gilt

$$\alpha^2_{v,\mu} \cdot e^{-2\cdot\lambda_v \cdot k} = \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot (k+i)} \cdot \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot (k-i)}$$

Die Bestimmung des Exponenten λ für die Sensitivitätsfunktion $\mu_{v,k}$ lässt sich durch eine Begrenzung einer Matrixexponentialfunktion nach (Van Loan, 1977) ableiten. Demnach ergibt sich für eine quadratische Matrix M

$$\left|e^{M \cdot n}\right| \leqslant c \cdot e^{-\lambda \cdot n} \tag{6.25}$$

mit einem multiplikativen Faktor c und dem größten Realteil aus den Eigenvektoren λ_i der Matrix M

$$\lambda = -\left(\max_i \left\{Re(\lambda_i)\right\}\right). \tag{6.26}$$

Um diesen Zusammenhang bei der Sensitivitätsfunktion anzuwenden, wird die generische Matrix M durch den Matrixlogarithmus der Systemmatrix des geschlossenen Regelkreises ersetzt.

$$M = \ln\left(A_{cl}\right) \tag{6.27}$$

Dies begründet sich durch die Dominanz des Matrixexponential $\left(A_{cl}^k\right)$ in der Definition der Folge $\mu_{v,k}$ aus Gleichung (6.20). So gilt

$$\mu_{v,k} = \left|K \cdot A_{cl}^{k-1} \cdot \left(I - A_{cl}^{-1}\right) \cdot B\right|, \; k > 1.$$

Da auch die Folge $\theta_{v,k}$ in Gleichung (6.19) in ihrem Verlauf von dem selben Matrixexponential $\left(A_{cl}^k\right)$ dominiert wird, lässt sich auch für den Einfluss einer initialen Abweichung des Zustandsvektors auf die Auswirkungen eines verworfenen Ereignisses $\left(\theta_{v,k}\right)$ eine Abschätzung mit dem selbem exponentiellem Abfall λ erstellen.

$$\overline{\theta}_{v,k} = \alpha_{v,\theta} \cdot e^{-\lambda_v \cdot k} \geqslant \theta_{v,k} \tag{6.28}$$

Im weiteren Verlauf der Arbeit wird angenommen, dass die Folgen $\theta_{v,k}$ und $\mu_{v,k}$ den geforderten exponentiellen Abfall aufweisen und dadurch Ungleichung (6.23) erfüllen.

Beispiel 15. In Abbildung 6.9 sind die Folgen $\theta_{v,k}$ und $\mu_{v,k}$ für das Regelsystem aus den Beispielen 12 und 13 dargestellt. Der Exponent der Hüllkurve $\left(\lambda_v\right)$ berechnet sich aus den Gleichungen (6.25) bis (6.27) mit

$$A_{cl} = A - B \cdot K = \begin{bmatrix} 0,17 & 0,19 \\ -0,17 & 0,75 \end{bmatrix} - \begin{bmatrix} 0,33 \\ -0,64 \end{bmatrix} \cdot \begin{bmatrix} 0,89 & 2,47 \end{bmatrix} = \begin{bmatrix} 0,46 & 1,01 \\ -0,74 & 0,83 \end{bmatrix}$$

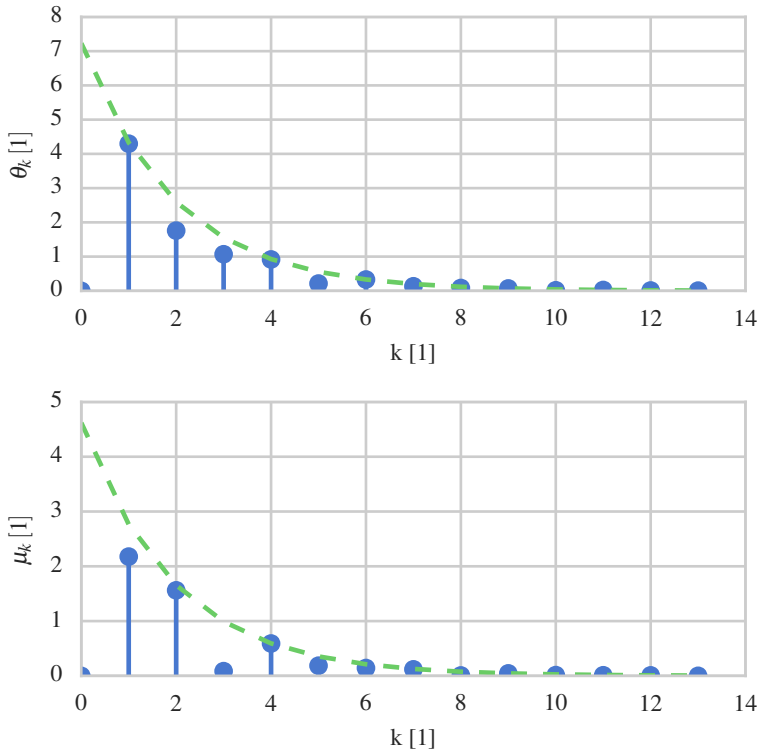


Abb. 6.9: Verlauf der Folgen $\theta_{v,k}$ und $\mu_{v,k}$, sowie deren exponentiell abfallende Hüllkurve.

$$M = \ln(A_{cl}) = \begin{bmatrix} 1,63 & 3,32 \\ -2,44 & -2,65 \end{bmatrix}$$

$$\lambda_v = -\max_i \{Re(\lambda_i)\} = -\left(\max\{-0.51,\ -0.51\}\right)$$

und ergibt für das Beispielsystem $\lambda_v = 0.51$. Die multiplikativen Faktoren der Hüllkurve ergeben sich zu $\alpha_{v,\theta} = 7,23$ und $\alpha_{v,\mu} = 4,62$.

Werden die Parameter dieser Exponentialfunktionen $(\lambda_v, \alpha_{v,\theta}, \alpha_{v,\mu})$, welche die exponentiell abfallenden Folgen beschreiben in Gleichung (6.18) eingesetzt, so ergibt sich folgender Satz:

Satz 24. *Die betragsmäßige Störung, verursacht durch den n-ten Ereignisverlust zur k-ten Ereignisinstanz ist begrenzt durch*

$$\begin{aligned}\left|\mathbf{v}_k^n\right| &\leqslant \alpha_{v,\theta} \cdot \left(\alpha_{v,\mu}+1\right)^{n-1} \cdot e^{-\lambda_v \cdot k} \cdot \left|\mathbf{x}_0\right| \\ &\leqslant \alpha_{v,\theta} \cdot e^{\ln(\alpha_{v,\mu}+1)\cdot(n-1)-\lambda_v \cdot k} \cdot \left|\mathbf{x}_0\right|\end{aligned} \tag{6.29}$$

mit

$$\sum_{i=1}^{k} d_i = n.$$

Beweis. Der Beweis dieses Satzes erfolgt über eine vollständige Induktion. In Anhang C.1 findet sich der ausformulierte Induktionsbeweis. □

Der Vorteil des obigen Satzes gegenüber Gleichung (6.18) ist die Unabhängigkeit von der Kenntnis über die konkrete Ereignisinstanz an der das Ereignis verworfen wird. So ist allein die Anzahl der verworfenen Ereignisse, die in den ersten k Ereignisinstanzen auftreten ausreichend, um die äquivalente Störung zu berechnen. Diese Beschreibungsform entspricht daher weitestgehend einer Intervallbetrachtung wie sie angestrebt wird. Ein weiterer Vorteil dieser Darstellung gegenüber Gleichung (6.18) ist der Verzicht auf die Summennotation. Die Anzahl der Summanden ergibt sich aus der Länge der Folge $\mu_{v,k}$ und würde insbesondere für Regelsysteme mit langsamen Einschwingverhalten einen erheblichen Rechenaufwand bei der Bestimmung von $|\mathbf{v}_k|$ verursachen.
Soll aus obigem Formalismus die maximale äquivalente Störung $\left(\sup_n \left\{\left|\mathbf{v}_k^n\right|\right\}\right)$ bestimmt werden, kommt die maximale Verlustdichte und derer minimalen Intervallfunktion zur Anwendung. So resultiert aus der Annahme der maximalen Anhäufung an verworfenen Ereignissen die maximalen Störung.

Satz 25. *Die maximale Signaldichte* $(\|\mathbf{v}(\Delta_k)\|)$*, der durch verworfene Ereignisse verursachten Störung ergibt sich zu*

$$\|\mathbf{v}(\Delta_k)\| \leqslant \sup_j \left\{ \sum_{k=1}^{dD^+(\Delta_k)} \left(\alpha_{v,\theta} \cdot e^{\ln(\alpha_{v,\mu}+1)\cdot(j+k-1)-\lambda_v \cdot iD^-(j+k)} \right) \right\} \cdot \left|\mathbf{x}_0\right|. \tag{6.30}$$

Beweis. Ausgehend von Gleichung (6.29) lässt sich eine Funktion $(\Upsilon(n) : \mathbb{N}_0 \to \mathbb{N}_0)$ definieren, welche die aufsummierte Störung für n verworfene Ereignisse in ähnlicher Weise wie die eingeführten Dichtefunktionen beschreibt.

$$\begin{aligned}\Upsilon(n) &= \sup_j \left\{ \sum_{i=1}^{n} \left| \mathbf{v}_{iD^-(j+i)}^{j+i} \right| \right\},\ n \in \mathbb{N} \\ &= \sup_j \left\{ \sum_{i=1}^{n} \left(\alpha_{v,\theta} \cdot e^{\ln(\alpha_{v,\mu}+1)\cdot(j+i-1)-\lambda_v \cdot iD^-(j+i)} \right) \right\} \cdot \left|\mathbf{x}_0\right|,\ n \in \mathbb{N}\end{aligned}$$

mit

$$\Upsilon(0) = 0.$$

Dabei ist das minimale Ereignisintervall in welchem diese kumulative Störung auftreten kann begrenzt durch $k = iD^-(n)$. Über die maximale Verlustdichte wird des Weiteren der Zusammenhang zwischen einem Ereignisintervall (Δ_k) und der Anzahl an verworfenen Ereignissen (n) hergestellt.

$$n = dD^+(\Delta_k)$$

Dadurch lässt sich durch eine Verkettung der Funktionen die Folge $\Upsilon(n)$ in eine Signaldichte überführen.

$$\|\mathbf{v}(\Delta_k)\| \leqslant \Upsilon(dD^+(\Delta_k))$$

Zusammengefasst ergibt sich daraus der obige Satz. □

Für die Bewertung der maximalen Störung auf das Regelsystem ergibt sich aus Ungleichung 6.30 für $\Delta_k = 1$ die vereinfachte Darstellung

$$\|\mathbf{v}(1)\| \leqslant \sup_{j>0} \left\{ a_{v,\theta} \cdot e^{\ln(\alpha_{v,\mu}+1)\cdot(j-1)-\lambda_v \cdot iD^-(j)} \right\} \cdot |\mathbf{x}_0| \,. \tag{6.31}$$

Die Auswirkungen dieser Störung auf den Zustandsvektor kann im Anschluss durch die Anwendung von Satz 22 bestimmt werden.

Beispiel 16. Entscheidend für die Auswirkungen verworfener Ereignisse auf ein Regelsystem ist die Frage, wie sich der Exponent in Gleichung (6.29) entwickelt. Dieser lässt sich aufteilen in einen destabilisierenden Anteil, welcher zur Erhöhung der betragsmäßigen Störung $|\mathbf{v}_k|$ beiträgt und einen Anteil, der diesem entgegenwirkt. Der destabilisierenden Part wird von der Anzahl an verworfenen Ereignissen und dem Faktor $\alpha_{v,\mu}$ bestimmt. Der entgegengesetzte Anteil wirkt sich mit zunehmenden Parameter λ_v stabilisierend auf das Regelsystem aus, da er zu einer Verkleinerung der äquivalenten Störung beiträgt. In Abbildung 6.10 ist der Exponent aus Gleichung (6.29), sowie der stabilisierende und destabilisierende Anteil unter der Annahme der maximalen Dichte an verworfenen Ereignissen $(k = iD^-(n) = \{0, 1, 3, 7, 11, 15\})$ für die Kurven aus Beispiel 15 dargestellt. In dem Plot stellt der destabilisierende Anteil eine Gerade dar, deren Steigung von dem Parameter α_μ abhängt. Der stabilisierende Anteil vergrößert sich mit abnehmender Dichte an verworfenen Ereignissen, da diese zu einer Erhöhung der minimalen Intervallfunktion $iD^-(n)$ führt. Ein

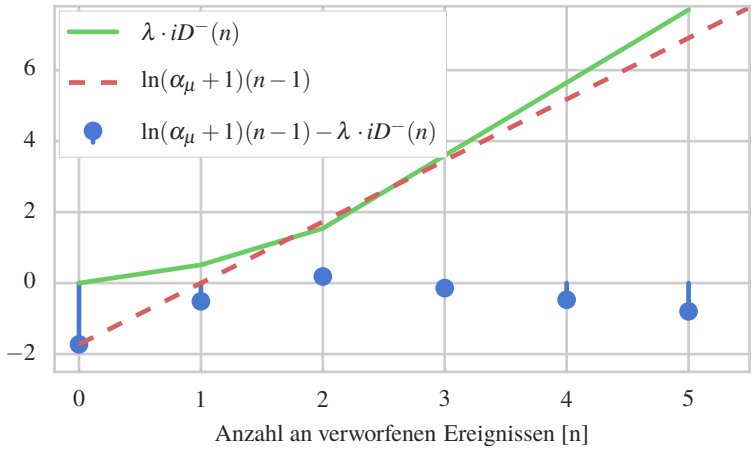


Abb. 6.10: Verlauf der stabilisierenden und destabilisierenden Anteile des Exponenten aus Gleichung (6.29).

weiterer stabilisierender Faktor ist der Parameter λ_v, der als Maß für die Robustheit der Regelung angesehen werden kann, da dieser die Rate bestimmt, mit der die Auswirkung eines verworfenen Ereignisses abnimmt. Für ein stabiles Verhalten des Regelsystems ist es zwingend notwendig, dass der stabilisierende Anteil den destabilisierende Part für eine fortschreitende Anzahl an verworfenen Ereignissen (n) dauerhaft überschreitet. Dies führt zu einer Verringerung des Exponenten in Gleichung (6.29) und damit zu einer Abnahme der durch ein verworfenes Ereignis verursachten Störung.

Um zu bewerten in welchem Umfang die vorgestellten Methoden zur Bestimmung der größtmöglichen Beeinflussung eines Regelsystems durch verworfene Ereignisse das reale Verhalten überschätzen, wurde eine Testmenge an Sequenzen verworfener Ereignisse (d_k) erzeugt. Die Testmenge besteht aus Sequenzen aller Binärkombinationen mit der Länge von 16 Ereignissen, welche die oben aufgeführte minimale Intervallfunktion in keinem Wert unterschreiten. Das erste Element der Sequenzen wird zu $d_0 = 0$ angenommen, um nach Gleichung (6.13) einen gültigen Startwert zu erhalten. Die Testmenge beinhaltet 440 Sequenzen für deren Evaluation jeweils 100 gleich verteilte Werte auf dem Einheitskreis gewählt wurden $(|\mathbf{x}_0| = 1)$.

Aus der schrittweisen Evaluation dieser 44.000 Experimente anhand der Formeln (6.13) und (6.14) kann für jedes Experiment die maximale Störung bestimmt werden, welche in Form eines Histogramms in Abbildung 6.11 dargestellt sind. Die ma-

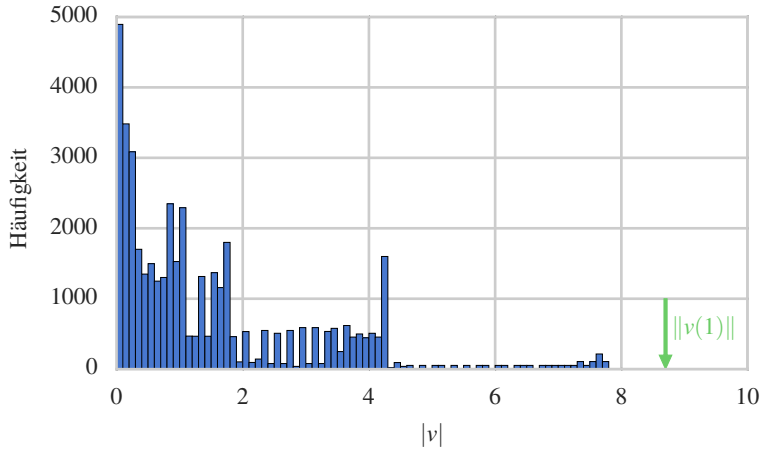


Abb. 6.11: Histogramm der maximalen Störungen durch Auswertung aller Binärkombinationen und valider Grenzwert durch Anwendung der Intervallbetrachtung.

ximal auftretende Störung aller Experimente beträgt $7,72$.
Wird hingegen der vorgestellte direkte Ansatz nach Gleichung (6.31) angewandt, resultiert für das gegebene System eine maximale Störung von $\|\mathbf{v}(1)\| = 8,71$ was einer Überschätzung von $12,8\%$ entspricht.
Vergleicht man die Auswertung der Binärkombinationen und die vorgestellte Methode zur direkten Bestimmung der Grenzwerte hinsichtlich der benötigten Rechenzeit, so verursacht das Erstellen und Auswerten der 44.000 Experimente eine Rechenzeit von $25,9$ Sekunden (Intel I7 , 8GB DDR3 RAM). Dies unterscheidet sich signifikant von den 30 Millisekunden, welche für die Berechnung nach der in diesem Abschnitt vorgestellten Methode benötigt wird. Werden durch die Auswertung von Binärkombinationen Systeme mit größerem Zustandsraum untersucht, erhöht sich der Rechenaufwand unverhältnismäßig stark, da weitere Dimensionen betrachtet werden müssen. Auch Systeme, die sich nahe der Stabilitätsgrenze befinden, können nicht effizient ausgewertet werden, da die zu betrachtenden Sequenzen sehr lang werden.
Die vorgestellte Betrachtung im Intervallbereich erlaubt jedoch, durch ihre rechenzeiteffiziente Analyse des schlimmsten Falls, Parameterstudien und ermöglicht die Entwurfsraumexploration eingebetteter Regelsysteme.
Die für dieses Beispiel gewählte Dichte an verworfenen Ereignissen erlaubt zwei verworfene Ereignisse in einem Intervall von drei aufeinander folgenden Ereignis-

instanzen. Das Regelsystem verhält sich dennoch stabil, da für größere Intervalle restriktivere Beschränkungen vorliegen. Würde man zwei verworfene Ereignisse in einem Intervall von drei aufeinander folgenden Ereignisinstanzen durch ein periodisches Modell oder eine (m,k)-Begrenzung beschreiben, würde dies in einem instabilen Regelsystem resultieren.

Satz 26. *Ein Regelsystem mit exponentieller Stabilität (vgl. Unterabschnitt 4.3.1 auf Seite 45) verhält sich unter dem Einfluss verworfener Ereignisse weiterhin global asymptotisch stabil, wenn der Quotient aus den Elementen des dominanten Punktes folgende Ungleichung erfüllt:*

$$\frac{\Delta_k^d}{n^d} \geqslant \frac{\ln(\alpha_{v,\mu}+1)}{\lambda_v} \tag{6.32}$$

Beweis. Voraussetzung dafür, dass sich ein stabiles System weiterhin stabil verhält ist, dass die einwirkende Störung begrenzt ist. So muss für die Stabilitätsuntersuchung eines Systems mit verworfenen Ereignissen bei einem ausreichend großen Betrachtungshorizont der stabilisierende Anteil den destabilisierenden Part in dem Exponenten von Ungleichung (6.30) dauerhaft überschreiten. Der destabilisierende Anteil besteht aus aus dem Term $(\ln(\alpha_{v,\mu}+1)\cdot(j+k-1))$, durch den der Exponent und damit die Störung $\left|\mathbf{v}_k^n\right|$ vergrößert wird. Stabilisierend auf das Regelsystem wirkt sich demnach der Term $(\lambda \cdot iD^-(n))$ aus. Für ein asymptotisch stabiles System muss demnach gelten:

$$\lim_{n\to\infty}\left(\frac{iD^-(n)}{n-1}\right) \geqslant \frac{\ln(\alpha_{v,\mu}+1)}{\lambda_v}.$$

Für große Intervalle kann die Intervallfunktion durch den dominanten Punkt der Verlustdichte approximiert werden. So ergibt sich aus der Abschätzung der Intervallfunktion in Gleichung (5.14) auf Seite 72

$$iD^-(m\cdot n^d+1) = m\cdot\Delta_k^d+1\,,\ \forall m\in\mathbb{N}.$$

Eingesetzt in die Ungleichung ergibt sich

$$\lim_{m\to\infty}\left(\frac{m\cdot\Delta_k^d+1}{m\cdot n^d}\right) \geqslant \frac{\ln(\alpha_{v,\mu}+1)}{\lambda_v}$$

$$\frac{\Delta_k^d}{n^d}+\lim_{m\to\infty}\left(\frac{1}{m\cdot n^d}\right) \geqslant \frac{\ln(\alpha_{v,\mu}+1)}{\lambda_v}.$$

Es zeigt sich aus Ungleichung (6.32), dass die Stabilität unabhängig von der Wahl des initialen Anfangszustands ist. Für ein (m, k)-Modell ist diese Stabilitätseigenschaft erfüllt, wenn gilt

$$\frac{m}{k} \leqslant \frac{n^d}{\Delta_k^d}.$$

Auch für eine Rate an verworfenen Ereignissen lässt sich nach Satz 8 auf Seite 72 die Stabilität des Regelsystem für den gegebenen dominanten Punkt ableiten.

Beispiel 17. Dieses Regelsystem der vorangehenden Beispiele verhält sich weiterhin stabil, wenn nach Gleichung (6.32) gilt

$$\frac{\Delta^d}{n^d} \geqslant \frac{\ln(\alpha_{v,\mu} + 1)}{\lambda_v} = \frac{\ln(4,62 + 1)}{0,51} = 3,38.$$

So kann Stabilität des Regelsystems garantiert werden, wenn nicht mehr als 3 Ereignisse in einem Intervall von 10 Ereignisinstanzen verworfen werden.

6.3 Latenzen im Regelkreis

Die Untersuchung verworfener Ereignisse auf ein Regelsystem, wie sie in Abschnitt 6.2 beschrieben wurde, erfordert von der Plattform ein zeitinvariantes Verhalten. Sowohl die Zeitpunkte, an denen die Sensorsignale erfasst werden, als auch die Zeitpunkte, an denen die Stellwerte das Regelsystem beeinflussen, werden als frei von Jitter angenommen. So impliziert die Bewertung von verworfenen Ereignissen in Abschnitt 6.2, dass in dem Regelkreis keine variablen Latenzen auftreten, oder dass diese bis zu einer vordefinierten Zeitschranke verzögert werden.
Werden hingegen Stellsignale eines ereignisbasierten verteilten Systems ohne zusätzliche Verzögerungen auf die Regelstrecke angewandt, führt dies im Allgemeinen zu variierenden Latenzen im Regelkreis. Solch eine Implementierung weist eine geringe durchschnittliche Latenz auf und kann daher als effiziente Umsetzung bezeichnet werden. Durch eine geringere durchschnittliche Latenz kann das Regelsystem schneller auf Änderungen reagieren, da der Stellwert nicht bis zu einer vordefinierten Zeitschranke zurückgehalten wird. Auch die Bereitstellung von Zwischenspeichern und Zeitmessungen (siehe (Verma et al., 1991)) ist bei der direkten Anwendung des Stellwerts hinfällig. Diese Effizienzsteigerung wird zu Kosten eines deterministischen Zeitverhaltens erreicht. Um Regelsysteme mit solch einem Zeitverhalten in einem sicherheitskritischen Umfeld einzusetzen zu können, müssen Schranken des Zeitverhaltens wie der in Abschnitt (5.2) vorgestellte Latenzdichte spezifiziert und validiert werden können. Daher wird im weiteren Verlauf des Abschnittes die Auswirkungen von Signallatenzen auf das dynamische Verhalten des Regelsystems im Zeitbereich und im Anschluss auf die Signaldichte des Zustandsvektors beschrieben.

6.3.1 Auswirkungen von Latenzen im Zeitbereich

Die Vorgehensweise bei der Bewertung von Latenzen im Regelkreis ist analog zu der in Unterabschnitt 6.2.1. Aus der Latenz eines Ereignisses wird eine äquivalente Störung konstruiert, durch die sich ein latenzfreies Regelsystem in gleicher Weise verhält, als wäre es den Pfadlatenzen ausgesetzt. Im Wesentlichen verändert sich durch Latenzen in einem periodisch abgetasteten Regelkreis die Dauer mit der Stellwerte auf die Regelstrecke einwirken. Die Dauer, in der die Stellgröße einer Ereignisinstanz auf die Regelstrecke einwirkt, ist in Abbildung 6.12 dargestellt. Zu den Aktivierungszeitpunkten des Sensortasks $t^r_{s,k}$ wird der Zustand des Systems erfasst

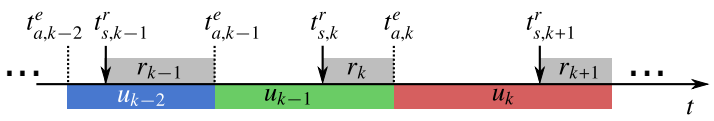


Abb. 6.12: Geltungsbereich der Stellgrößen bei variablen Latenzen.

welcher durch den Zustandsvektor $\mathbf{x}_k$ repräsentiert wird. Nach der Pfadlatenz r_k der digitalen Komponente des Regelsystems wird zum Zeitpunkt der Beendigung des Aktortasks $t^e_{a,k}$ der aktualisierte Stellwert auf die Regelstrecke angewandt. Der Stellwert $\mathbf{u}_k$ wird nach Gleichung (3.5) durch Multiplikation mit der Rückführmatrix aus dem Zustandsvektor $\mathbf{x}_k$ berechnet.
In Abhängigkeit von der Latenz r_{k-1} ergibt sich der Zustand $\mathbf{x}_k$ zum Zeitpunkt $t^r_{s,k}$ aus dem vorherigen Zustand $\mathbf{x}_{k-1}$ zum Zeitpunkt $t^r_{s,k-1}$ unter den anteiligen Einflüssen der Stellgrößen $\mathbf{u}_{k-1}$ und $\mathbf{u}_{k-2}$. Formal ändert sich unter dem Einfluss von Latenzen $r_k \leqslant T_s$ die Zustandsraumdarstellung eines ungeregelten Systems (siehe Gl. (3.6), Seite 33) nach (Åström und Wittenmark, 2011b) zu

$$\mathbf{x}_{k+1} = A \cdot \mathbf{x}_k + B_{0,k}\mathbf{u}_k + B_{-1,k}\mathbf{u}_{k-1}$$

mit

$$B_0(r_k) = \int_0^{T_s - r_k} e^{A_c \alpha} d\alpha \cdot B_c$$

und

$$B_{-1}(r_k) = \int_{T_s - r_k}^{T_s} e^{A_c \alpha} d\alpha \cdot B_c = A_c^{-1} A \left(I - e^{-A_c r_k}\right) \cdot B_c \tag{6.33}$$

Es ist trivial, dass mit Gleichung (3.7) auf Seite 33 gilt

$$B_0(r_k) + B_{-1}(r_k) = B.$$

Für ein ungestörtes Regelsystem mit Zustandsrückführung nach (3.5) auf Seite 32 ergibt sich daraus

$$\begin{aligned}\mathbf{x}_{k+1} &= A \cdot \mathbf{x}_k + B_0(r_k)K\mathbf{x}_k + B_{-1}(r_k)K\mathbf{x}_{k-1} \\ &= A \cdot \mathbf{x}_k + BK\mathbf{x}_k + B_{-1,k}K\mathbf{x}_{k-1} - B_{-1}(r_k)K\mathbf{x}_k \\ &= A_{cl} \cdot \mathbf{x}_k + B_{-1}(r_k)K\left(\mathbf{x}_{k-1} - \mathbf{x}_k\right) \end{aligned} \tag{6.34}$$

Äquivalent zu der Vorgehensweise in Abschnitt 6.2 lässt sich anhand Gleichung (6.34) der Einfluss auf den Zustandsvektor durch eine Störung $(\mathbf{w}_k)$ modellieren. So ergibt sich für den Einfluss auf den Zustandsvektor

$$\begin{aligned}\mathbf{x}_{w,k+1} &= A_{cl} \cdot \mathbf{x}_{w,k} + \mathbf{w}_k \\ &= \sum_{j=0}^{\infty}(A_{cl}^{j}\mathbf{w}_{k-j}) \end{aligned} \tag{6.35}$$

mit

$$\mathbf{w}_k = B_{-1}(r_k)K\left(\mathbf{x}_{k-1} - \mathbf{x}_k\right). \tag{6.36}$$

Es ist zu beachten, dass im Gegensatz zu der Störung durch verworfene Ereignisse $(\mathbf{v}_k)$, die äquivalente Störung durch Latenzen $(\mathbf{w}_k)$ nicht auf die Stellgröße, sondern direkt auf den Zustandsvektor einwirkt. Das Störsignal $\mathbf{w}_k$ weist jedoch erhebliche Parallelen zur Beschreibung der äquivalenten Störung verworfener Ereignisse in Gleichung (6.12) auf. So sind beide Störungen in hohem Maße abhängig von der Differenz des Zustandsvektors.

Für eine grenzwertige Betrachtung ergibt sich aus Gleichung (6.33) für ein latenzfreies System $B_{-1}(0) = 0$ und damit keine Störung. Entspricht hingegen die Latenz r_k der Abtastzeit T_s, dann folgt daraus $B_{-1}(T_s) = B$. Bei solch einer Latenz erfolgt die Beeinflussung der Regelstrecke mit dem Stellwert $\mathbf{u}_k$ zu dem Zeitpunkt der Erfassung des nächsten Systemzustandes $\left(t_{a,k}^{e} = t_{s,k+1}^{r}\right)$. Hierbei hat der Stellwert $\mathbf{u}_k$ keinen Einfluss auf den Zustandsvektor $\mathbf{x}_{k+1}$.

Die Differenz des Zustandsvektors, verursacht durch Latenzen ergibt sich aus Gleichung (6.35) zu

$$\begin{aligned}\mathbf{x}_{w,k+1}^{'} &= \mathbf{x}_{w,k+1} - \mathbf{x}_{w,k} \\ &\quad (A_{cl} - I) \cdot \mathbf{x}_{w,k} + \mathbf{w}_k. \end{aligned} \tag{6.37}$$

6.3.2 Betrachtung im Intervallbereich

Basierend auf den Auswirkungen von Latenzen auf ein Regelsystem im Zeitbereich wird im Folgenden der Zusammenhang zwischen der maximalen Latenzdichte $(dR^{+}(\Delta_k))$ und der Signaldichte der Störung $(\|w(\Delta_k)\|)$ hergestellt. Dies ermöglicht die Ableitung einer Begrenzung des Signalverhaltens eines Regelsystems aus der Beschränkung der Pfadlatenz.

Satz 27. *Der betragsmäßige Störeinfluss durch die Latenzen der k-ten Ereignisinstanz ist begrenzt durch*

$$|\mathbf{w}_k| \leqslant |B_{-1}(r_k)| \cdot \left(\sum_{n=1}^{\infty} (\mu_{w,n} \cdot |\mathbf{w}_{k-n}|) + \theta_{w,k} \cdot |\mathbf{x}_0| \right) \tag{6.38}$$

mit $\mathbf{w}_k = r_k = 0,\ \forall k \leqslant 0$ *und*

$$\theta_{w,k} = \left| K \cdot \left(A_{cl}^{k} - A_{cl}^{k-1} \right) \right| \tag{6.39}$$

$$\mu_{w,k} = \begin{cases} |K| & \textit{wenn } k = 1 \\ \left| K \cdot \left(A_{cl}^{k+1} - A_{cl}^{k} \right) \right| & \textit{wenn } k > 1. \end{cases} \tag{6.40}$$

Beweis. Der Beweis dieses Satzes ergibt sich analog zu dem Beweis von Satz 23. Aus Gleichung (6.36) folgt

$$\begin{aligned} \mathbf{w}_k &= -B_{-1}(r_k) \cdot K \cdot \mathbf{x}_{w,k}^{'} - B_{-1}(r_k) \cdot K \cdot \mathbf{x}_{0,k}^{'} \\ |\mathbf{w}_k| &\leqslant \left| B_{-1}(r_k) \cdot K \cdot \mathbf{x}_{w,k}^{'} \right| + \left| B_{-1}(r_k) \cdot K \cdot \mathbf{x}_{0,k}^{'} \right|. \end{aligned}$$

Der Einfluss einer initialen Auslenkung $(\mathbf{x}_0)$ auf die Auswirkung der verzögerten k-ten Ereignisinstanz wird durch die Folge $\theta_{w,k}$ beschrieben. So ergibt sich für den rechten Summanden mit Gleichung (6.16):

$$\begin{aligned} \left| B_{-1}(r_k) \cdot K \cdot \mathbf{x}_{0,k}^{'} \right| &\leqslant \left| B_{-1}(r_k) \cdot K \cdot \left(A_{cl}^{k} - A_{cl}^{k-1} \right) \cdot \mathbf{x}_0 \right| \\ &\leqslant |B_{-1}(r_k)| \cdot \left| K \cdot \left(A_{cl}^{k} - A_{cl}^{k-1} \right) \right| \cdot |\mathbf{x}_0| \end{aligned}$$

Der linke Summand lässt sich mit den Gleichungen (6.35) und (6.37) abschätzen zu

$$\begin{aligned}
\left|B_{-1}(r_k)\cdot K\cdot \mathbf{x}'_{w,k}\right| &= |B_{-1}(r_k)\cdot K\cdot((A_{cl}-I)\cdot \mathbf{x}_{w,k-1}+\mathbf{w}_{k-1})| \\
&= \left|B_{-1}(r_k)\cdot K\cdot\left((A_{cl}-I)\cdot\sum_{j=0}^{\infty}\left(A_{cl}^{j}\mathbf{w}_{k-j-2}\right)+\mathbf{w}_{k-1}\right)\right| \\
&= |B_{-1}(r_k)|\cdot\left|\sum_{j=0}^{\infty}\left(K\cdot\left(A_{cl}^{j+1}-A_{cl}^{j}\right)\cdot\mathbf{w}_{k-j-2}\right)+K\cdot\mathbf{w}_{k-1}\right| \\
&\leqslant |B_{-1}(r_k)|\cdot\sum_{j=0}^{\infty}\left|K\cdot\left(A_{cl}^{j+1}-A_{cl}^{j}\right)\right|\cdot|\mathbf{w}_{k-j-2}|+|K|\cdot|\mathbf{w}_{k-1}|
\end{aligned}$$

□

Die Sensitivitätsfunktion $\mu_{w,k}$ bestimmt den Einfluss zuvor verzögerter Ereignisinstanzen auf die gegenwärtige äquivalente Störung. Nähert sich die Folge $\mu_{w,k}$ in wenigen Elementen k gegen Null, ist das Regelsystem robuster gegenüber Latenzen. Verweilt die Folge jedoch lange auf hohem Niveau, weist das Regelsystem ein sensitives Verhalten gegenüber Latenzen auf.

Im Folgenden soll für Pfadlatenzen, beschränkt durch eine maximale Latenzdichte, die Signaldichte des Zustandsvektors bestimmt werden. Diese Signaldichte soll derart bestimmt werden, dass für eine Sequenz an Latenzen, welche die Latenzdichte in keinem Intervall überschreitet, die gewählte Signaldichte nicht verletzt wird. Aus diesem Anspruch resultieren zwei Einschränkungen an die Parameter in Satz 27.
Zum einen dürfen geringere Pfadlatenzen nicht größere Auswirkungen auf die äquivalente Störung haben. Dies impliziert, dass die Norm $|B_{-1}(r_k)|$ in den Gleichungen (6.39) und (6.40) für zunehmende Latenzen ein größeres Resultat ergibt.

$$r > r_0 \Rightarrow |B_{-1}(r)| \geqslant |B_{-1}(r_0)|$$

Andernfalls muss die Norm durch ein monoton steigendes Verhalten überschätzt werden.
Zum anderen darf bei gleichbleibenden Latenzen eine Entspannung des Auftretens durch eine verringerte Latenzdichte nicht in einer vergrößerten äquivalenten Störung resultieren. So muss analog zur Konsequenz der Ungleichung 6.23 auf die Sensitivitätsfunktion verworfener Ereignisse auch die Sensitivitätsfunktion des Regelsystems auf Latenzen $(\mu_{w,k})$ und $(\theta_{w,k})$ durch eine exponentiell fallende Folge abgeschätzt werden.

$$\mu_{w,k} \leqslant \alpha_{w,\mu}\cdot e^{-\lambda_w\cdot k} \tag{6.41}$$

$$\theta_{w,k} \leqslant \alpha_{w,\theta}\cdot e^{-\lambda_w\cdot k} \tag{6.42}$$

Mit dieser Abschätzung folgt eine direkte Berechnungsvorschrift für die äquivalente Störung durch eine Latenz.

Satz 28. *Die betragsmäßige Störung, verursacht durch eine Latenz r_k zur k-ten Ereignisinstanz ist begrenzt durch*

$$|\mathbf{w}_k| \leqslant |B_{-1}(r_k)| \cdot \prod_{n=1}^{k-1} \left(\alpha_{w,\mu} |B_{-1}(r_n)| + 1\right) \cdot e^{-\lambda_w \cdot k} \cdot \alpha_{w,\theta} \cdot |\mathbf{x}_0| \tag{6.43}$$

Beweis. Der Beweis dieses Satzes erfolgt über eine vollständige Induktion. In Anhang C.2 findet sich der ausformulierte Induktionsbeweis. □

Vorteilhaft an der Darstellung nach Gleichung (6.43) ist die Unabhängigkeit von der Reihenfolge in der die Latenzen auftreten. Es stellt sich nun die Frage, wie aus der Kenntnis über die maximale Latenzdichte die betragsmäßige Störung aus Gleichung (6.43) abgeschätzt werden kann.

Satz 29. *Für die maximale Signaldichte der äquivalenten Störung des Zustandsvektors, verursacht durch die maximale Latenzdichte folgt*

$$\|\mathbf{w}(\Delta_k)\| \leqslant \sup_j \left\{ \sum_{k=0}^{\Delta_k - 1} |\mathbf{w}_{k+j}| \right\}$$

mit

$$r_n = \frac{dR^+(k)}{k} \; \forall n \in [1, k]$$

für

$$\alpha_{w,\mu} \geqslant \frac{|B_{-1}(r_k)| - 1}{|B_{-1}(r_k)|}$$

Beweis. Zu Beginn wird Ungleichung (6.43) abgeschätzt durch

$$|\mathbf{w}_k| \leqslant \prod_{n=1}^{k} \left(\alpha_{w,\mu} |B_{-1}(r_n)| + 1\right) \cdot e^{-\lambda_w \cdot k} \cdot \alpha_{w,\theta} \cdot |\mathbf{x}_0| \,. \tag{6.44}$$

Dies ist dann erlaubt, wenn gilt

$$\alpha_{w,\mu} |B_{-1}(r_n)| + 1 \geqslant |B_{-1}(r_n)|$$

woraus sich für $\alpha_{w,\mu}$ die untere Schranke

$$\alpha_{w,\mu} \geqslant \frac{|B_{-1}(r_k)| - 1}{|B_{-1}(r_k)|}$$

ergibt.
Für ein zu betrachtendes Intervall an Ereignissen, wird dann das Maximum der Störung $|\mathbf{w}_k|$ erreicht, wenn die Latenzen in dem betreffenden Intervall den gleichen Wert annehmen. Dies ist dem Aufbau des Produkts in Gleichung (6.44) geschuldet, wonach gilt:

$$(\alpha_{w,\mu}\,|B_{-1}(r_n)|+1)^2 \geqslant (\alpha_{w,\mu}\,(|B_{-1}(r_n)|+\epsilon)+1)\cdot(\alpha_{w,\mu}\,(|B_{-1}(r_n)|-\epsilon)+1)$$

Die resultierende Signaldichte $\|w(\Delta_k)\|$ begrenzt dabei das maximale Vorkommen der maximalen Störung $|\mathbf{w}_k|$. □

Die so ermittelte Signaldichte der äquivalenten Störung verzögerter Ereignisse im Regelkreis erlaubt eine effiziente Abschätzung der größtmöglichen Auswirkungen von Latenzen auf ein Regelsystem.

Kapitel 7
Co-Design und Entwurfsstudie

Die in den Kapiteln 5 und 6 aufgeführten mathematischen Beziehungen bilden die Grundlage für die Validierung eines eingebetteten Regelsystems hinsichtlich der Zeiteffekte einer verteilten Rechenplattform. In diesem Kapitel wird die Anwendung dieser Formalismen anhand einer Entwurfsstudie dargestellt. Dazu werden die Wechselwirkung einer mechanischen Störeinwirkung und der zeitliche Einfluss auf ein sicherheitskritisches Regelsystem untersucht und gegeneinander abgewogen.

7.1 Modell einer Gleichstrommaschine

Die Regelstrecke des Regelsystems bildet eine permanenterregte Gleichstrommaschine, wie sie durch das Ersatzschaltbild in Abbildung 7.1 repräsentiert wird. Der

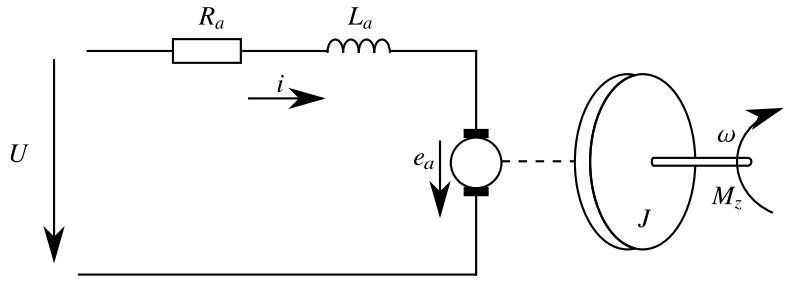


Abb. 7.1: Ersatzschaltbild des Ankers einer Gleichstrommaschine nach (Stölting und Beisse, 1987).

elektrische Teil wird durch Anwendung der Maschengleichung

$$U = R_a \cdot i + L_a \cdot \frac{di}{dt} + e_a \tag{7.1}$$

mit dem Ankerwiderstand R_a, der Induktivität der Ankerwicklung L_a und der Induktionsspannung e_a charakterisiert. Die generierte Induktionsspannung erhöht sich mit zunehmender Drehzahl und wird durch Multiplikation einer Maschinenkonstante c mit der Rotationsgeschwindigkeit ω gebildet:

$$e_a = c \cdot \omega \tag{7.2}$$

Der resultierende Ankerstrom i erwirkt ein Drehmoment, welches einem von außen aufgeprägten Drehmoment M_z entgegenwirkt, oder zur Beschleunigung der Maschine führt:

$$c \cdot i = J \cdot \frac{d\omega}{dt} + M_z \tag{7.3}$$

Anhand der Gleichungen (7.1) bis (7.3) lässt sich ein Zustandsraummodell der Gleichstrommaschine in der Form einer kontinuierlichen Matrixdifferentialgleichung nach Gleichung (3.2) auf Seite 30 erstellen.

$$\begin{aligned} \dot{\mathbf{x}}(t) &= A_c \cdot \mathbf{x}(t) + \mathbf{B_c} \cdot u(t) + \mathbf{E_c} \cdot z(t) \\ y(t) &= \mathbf{C_c} \cdot \mathbf{x}(t) \end{aligned}$$

Die drei Zustandsvaribalen und die Systemein-/ausgänge sind so gewählt, dass gilt

$$\mathbf{x} = \begin{bmatrix} \varphi \\ \omega \\ i \end{bmatrix},\ u(t) = U,\ y(t) = \omega,\ z(t) = M_z$$

mit φ als Drehwinkel des Ankers $\left(\omega = \frac{d\varphi}{dt}\right)$.
Aus dieser Konvention folgt für die Koeffizienten des Differentialgleichungssystems

$$A_c = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & \frac{c}{J} \\ 0 & -\frac{c}{L_a} & -\frac{R_a}{L_a} \end{bmatrix},\ \mathbf{B}_c = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L_a} \end{bmatrix},\ \mathbf{E}_c = \begin{bmatrix} 0 \\ -\frac{1}{J} \\ 0 \end{bmatrix},\ \mathbf{C}_c = \begin{bmatrix} 0\ 1\ 0 \end{bmatrix}.$$

Die Parameter des Motors werden zu

$$c = 0.03 \left[\frac{Nm}{A}\right] / \left[\frac{Vs}{rad}\right],\ J = 10^{-3} \left[Nm^2\right],\ R_a = 0.1\,[\Omega],\ L_a = 0.001\ [H]$$

angenommen, was einer elektrischen Kleinmaschine entspricht. Sensoren erfassen mit einer Abtastzeit von $T_s = 0.01\ s$ die Zustandsvariablen der kontinuierlichen Regelstrecke. Durch die zeitdiskrete Erfassung der Zustandsvariablen lässt sich nach Gleichung (3.7) eine äquivalente diskrete Darstellung der Regelstrecke erstellen.
Schließlich wird durch Polfestlegung ein Regler entworfen, der folgende Reglermatrix aufweist:

$$\mathbf{K} = -\begin{bmatrix} 0 & 0.614 & 0.136 \end{bmatrix}$$

Die Motorregelung erfolgt über ein verteiltes Regelsystem, bei dem Sensor, Aktor und Recheneinheit auf verschiedenen Ressourcen ausgeführt werden und dabei über ein Bussystem kommunizieren. Es wird dabei spezifiziert, dass die sicherheitskritische Aufgabe des Regelkreises erfüllt ist, wenn im Betrieb die Winkelgeschwindigkeit der Gleichstrommaschine nicht mehr als $10\ \frac{rad}{s}$ von der Solldrehzahl abweicht $\left(\|\hat{y}(1)\| = 10\ \frac{rad}{s}\right)$. Darüber hinaus soll über ein Intervall von 5 Ereignissen die Differenz der Winkelgeschwindigkeit im Mittel nicht mehr als $5\ \frac{rad}{s}$ Umdrehungen pro Sekunde betragen $\left(\|\hat{y}(5)\| = 25\ \frac{rad}{s}\right)$. Diese beiden Einschränkungen entsprechen den charakteristischen Punkten einer Signaldichte des Systemausgangs und definieren somit die Güteanforderungen an das Regelsystem.
Aus dieser spezifizierten Güteanforderung soll für das verteilte Regelsystem, unter dem im nächsten Unterabschnitt festgelegten Störeinfluss, die maximale zeitliche Einflussnahme auf der Plattformarchitektur auf die Gleichstrommaschine abgeleitet werden.

7.2 Validierung des Störverhaltens

In diesem Abschnitt wird untersucht, in wieweit sich unter Annahme einer idealen Rechenplattform ein störendes Drehmoment, spezifiziert als eine Signaldichte, auf die Drehzahländerung der geregelten Gleichstrommaschine auswirkt.
Die Signaldichte, welche das Vorkommen des störenden Drehmoments z_k begrenzt, ist in Abbildung 7.2 dargestellt. Ein mögliche Signalvolge, der diese Begrenzung maximal ausschöpft wäre $z_k = \{0,\ 0.2,\ 0,\ 0.4,\ 0,\ 0.2,\ 0,\ 0.4, \dots\}\ [Nm]$.
Die maximalen Auswirkungen solch einer Störung auf die Signaldichte des Zustandsvektors $\|x(\Delta_k)\|$ lässt sich anhand der Berechnungsvorschriften aus Satz 22

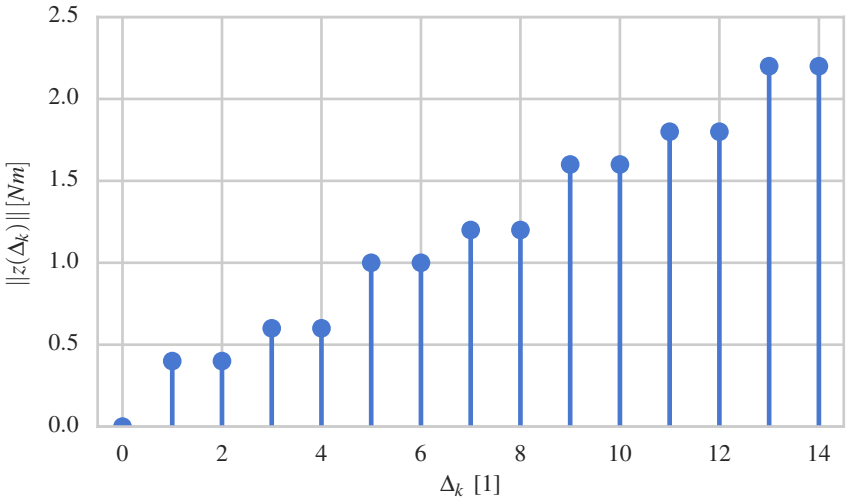


Abb. 7.2: Signaldichte des Störsignals ($\|z(\Delta_k)\|$).

auf Seite 118 bestimmen. In diesem Anwendungsfall ist jedoch die Abweichung des Zustandsvektors von untergeordneter Relevanz. Vielmehr ist die Abweichung der Winkelgeschwindigkeit, verursacht durch die Störung von Interesse, da diese im vorherigen Unterabschnitt spezifiziert wurde. Durch eine Modifikation von Gleichung (6.10)

$$\zeta_i = \sigma_{max}(A_{cl}^i \cdot \mathbf{E})$$

zu

$$\zeta_i^{yz} = \left|\mathbf{C} \cdot A_{cl}^i \cdot \mathbf{E}\right|$$

wird die Auswertung der Norm des Zustandsvektors in die betragsmäßige Abweichung des Systemausgangs $|y_{z,k}|$ überführt. Die numerische Evaluation dieses Szenarios nach Satz 22 ergibt für die Abweichung der Winkelgeschwindigkeit die Signaldichte wie in Abbildung 7.3 dargestellt. Des Weiteren ist die aus dem vorangegangen Unterabschnitt spezifizierte Begrenzung der Abweichung der Rotationsgeschwindigkeit ($\|\hat{y}(\Delta_k)\|$) in der Abbildung abgebildet. Die Spezifikation überschreitet in jedem Ereignisintervall die maximale Drehzahldifferenz, so dass für die angegebene Störeinwirkung ein ordnungsgemäßer Betrieb des Regelsystems garantiert werden kann.

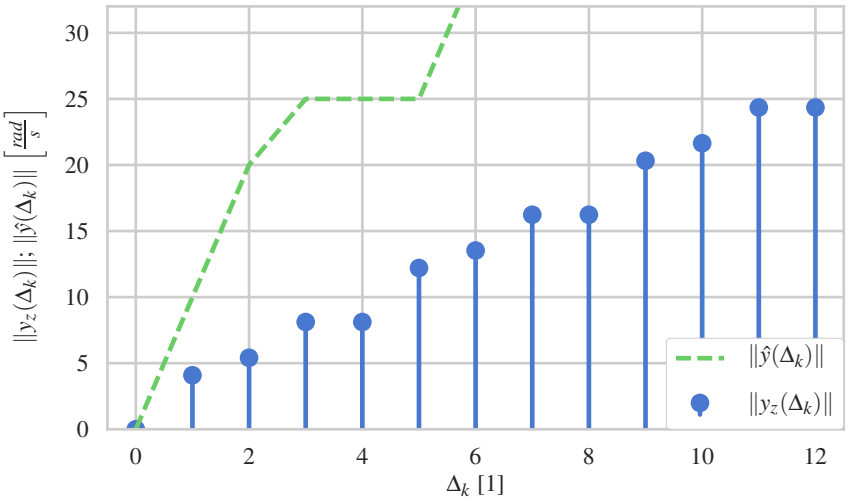


Abb. 7.3: Signaldichte der Abweichung der Winkelgeschwindigkeit $\|y_z(\Delta_k)\|$ und spezifizierte Begrenzung $\|\hat{y}(\Delta_k)\|$.

7.3 Validierung verworfener Ereignisse

Wird die durch ein Drehmoment gestörte Regelstrecke zusätzlich durch verworfene Ereignisse destruktiv beeinflusst, erhöht sich die Abweichung der Winkelgeschwindigkeit von der Solldrehzahl und damit auch die in Abbildung 7.3 dargestellte Signaldichte. Die Vorgehensweise bei der Bestimmung dieser erhöhten Regelabweichung ist dabei wie folgt:
Zu Beginn wird die Sensitivitätskurve $\mu_{v,k}$ (vgl. Unterabschnitt 6.2.2 auf Seite 124) und die damit verknüpften Parameter $\alpha_{v,\mu}$ und λ_v des geschlossenen Regelkreises bestimmt. Die Funktion $\mu_{v,k}$ und die begrenzende Exponentialfunktion

$$\mu_{v,k} \leqslant \alpha_{v,\mu} e^{-\lambda_v \cdot k} \tag{7.4}$$

mit den Parametern $\lambda_v = 1$ und $\alpha_{v,\mu} = 16.4$ sind in Abbildung 7.4 veranschaulicht. Der Sensitivitätskurve ist zu entnehmen, dass ein verworfenes Ereignis nach wenigen Abtastzeitpunkten einen verschwindenden Einfluss auf das Regelsystem hat. So ist auch der Stabilitätsuntersuchung nach Satz 26 auf Seite 133 entnehmen, dass zur Erfüllung der Stabilität der dominante Punkt verworfener Ereignisse folgende Eigenschaft aufweisen muss:

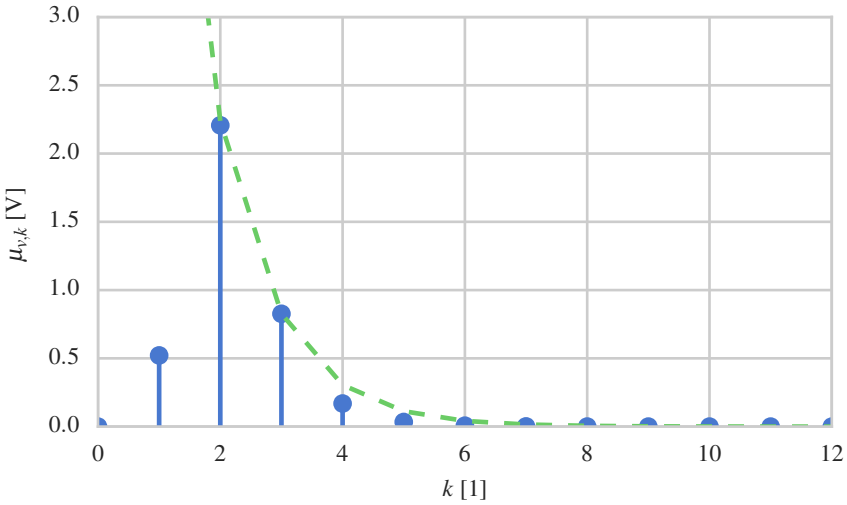


Abb. 7.4: Sensitivitätsfunktion des Regelkreises $(\mu_{v,k})$ bezüglich verworfener Ereignisse.

$$\frac{\Delta_k^d}{n^d} \geqslant \frac{\ln(\alpha_{v,\mu} + 1)}{\lambda_v} = \frac{\ln(16.4 + 1)}{1} = 2.87$$

Es sei an dieser Stelle angemerkt, dass die Sensitivitätsfunktion $\mu_{v,k}$ in der Einheit Volt angegeben wird, da die äquivalente Störung des verworfenen Ereignisses nach Unterabschnitt 6.2.1 auf Seite 120 auf den Eingang der Regelstrecke einwirkt. In dem hier beschriebenen Anwendungsfall entspricht dies der Ankerspannung der Gleichstrommaschine.

Die Annahmen des Unterabschnitts 6.2.2 zur Bestimmung der äquivalenten Störung verworfener Ereignisses lässt sich nur bedingt auf das beschriebene Szenario anwenden. So liegt hier keine initiale Auslenkung vom Stabilitätspunkt, sondern eine Beeinflussung des Regelsystems durch ein Störsignal (z_k) vor. Dabei stellt sich für Gleichung (6.18) und folgende die Frage nach der maximalen Änderung des Stellsignals $\mathbf{u}'_{k+1} = K \cdot \mathbf{x}'_{z,k+1}$ (vgl. Gl. (3.5)), welche durch das Störsignal verursacht wird, da diese Signaländerung die Auswirkungen eines verworfenen Ereignisses bestimmt.

Ausgehend von Gleichung (6.9) ergibt sich

$$
\begin{aligned}
\mathbf{u}^{'}_{z,k+1} &= \mathbf{K} \cdot \mathbf{x}^{'}_{z,k+1} \\
&= \mathbf{K} \cdot \left(\sum_{i=0}^{\infty} A_{cl}^{i} \cdot \mathbf{E} \cdot z_{k-i} - \sum_{i=1}^{\infty} A_{cl}^{i-1} \cdot \mathbf{E} \cdot z_{k-i} \right) \\
&= \mathbf{K} \cdot \left(\mathbf{E} \cdot z_k + \sum_{i=1}^{\infty} A_{cl}^{i} \cdot \mathbf{E} \cdot z_{k-i} - \sum_{i=1}^{\infty} A_{cl}^{i-1} \cdot \mathbf{E} \cdot z_{k-i} \right) \\
&= \sum_{i=1}^{\infty} \mathbf{K} \cdot (A_{cl} - I) \cdot A_{cl}^{i-1} \cdot \mathbf{E} \cdot z_{k-i} + \mathbf{KE} \cdot z_k
\end{aligned}
$$

Durch eine erneute Modifikation von Gleichung (6.10)

$$\zeta_i = \sigma_{max}(A_{cl}^{i} \cdot \mathbf{E})$$

zu

$$\zeta_i^{u'z} = \begin{cases} |\mathbf{KE}| & \text{wenn } i = 0 \\ \left|\mathbf{K} \cdot (A_{cl} - I) \cdot A_{cl}^{i-1} \cdot \mathbf{E}\right| & \text{wenn } i \geqslant 1 \end{cases}$$

resultiert über die Auswertung nach Satz 22 aus der Signaldichte der Störung ($\|z(\Delta_k)\|$) die maximalen Änderung der Stellgröße von

$$\max_k \{u_{z,k} - u_{z,k-1}\} = \left\| u_z^{'}(1) \right\| = 4.72\,V.$$

Dies ist zugleich die maximale Störung der Stellgröße, die ein einzelnes verworfenes Ereignis verursachen kann (vgl. Gl. (6.12)).

Zur Evaluation der Auswirkung einer Dichte an verworfenen Ereignissen wird angenommen, dass diese maximale Änderung der Stellgröße, verursacht durch die Störung des Drehmoments, zu jeder Ereignisinstanz auftreten kann. Der Zusammenhang im Zeitbereich, welcher durch Gleichung (6.18) auf Seite 124 beschrieben wird, ändert sich daher für diesen Anwendungsfall von

$$|\mathbf{v}_k| \leqslant \sum_{n=1}^{\infty} \mu_{v,n} \cdot |\mathbf{v}_{k-n}| + \theta_{v,k} \cdot |\mathbf{x}_0|$$

zu

$$|\mathbf{v}_k| \leqslant \sum_{n=1}^{\infty} \mu_{v,n} \cdot |\mathbf{v}_{k-n}| + \left\| u_z^{'}(1) \right\|.$$

Unter Annahme des maximalen Auftretens verworfener Ereignisse, wie durch die maximale Verlustdichte spezifiziert, und unter Anwendung von Gleichung (7.4) ergibt sich für die maximale Störung der Stellgröße

$$\begin{aligned}\|v(1)\| &= \left\|u_z'(1)\right\| + \alpha_{v,\mu} \cdot e^{\lambda_v \cdot iD^-(2)} \cdot \left\|u_z'(1)\right\| + \alpha_{v,\mu}^2 \cdot e^{\lambda_v \cdot iD^-(3)} \cdot \left\|u_z'(1)\right\| + \dots \\ &= \left(1 + \sum_{n=1}^{\infty} \alpha_{v,\mu}^n \cdot e^{\lambda_v \cdot iD^-(n+1)}\right) \cdot \left\|u_z'(1)\right\|\end{aligned}$$

Wird angenommen, dass jedes fünfte Ereignis verworfen wird $(iD^-(n) = \{0, 1, 6, 11, 16, 21, \dots\})$, dann ergibt sich in dem beschriebenen Szenario

$$\|v(1)\| = 4.72\,V + 0.2\,V + 0.02\,V + \dots$$

Die Folge der Summanden konvergiert rasch gegen Null und für die maximale Störung stellt sich $\|v(1)\| = 4.95\,V$ ein.
Um schließlich die daraus resultierende Signaldichte der Drehzahlabweichung des Gleichstrommotors zu bestimmen, werden die Störeinflüsse des Drehmoments, welches durch $\|z(\Delta_k)\|$ begrenzt ist, und die der verworfenen Ereignisse im Regelkreis, spezifiziert durch eine Dichte- oder Intervallfunktion $dD^+(\Delta_k)/iD^-(n)$, zusammengefasst. Nach dem Superpositionsprinzip und Satz 19 auf Seite 109 ergibt sich für die Signaldichte der Drehzahlabweichung

$$\|y(\Delta_k)\| = \|y_z(\Delta_k)\| + \|y_v(\Delta_k)\|$$

mit $\|y_z(\Delta_k)\|$ aus Abbildung 7.3 des vorherigen Unterabschnittes.
Die Drehzahlabweichung der Gleichstrommaschine, verursacht durch verworfene Ereignisse, wird durch Einwirkung folgender Signaldichte auf die Stellgröße erwirkt.

$$\|v(\Delta_k)\| = dD^+(\Delta_k) \cdot \|v(1)\| \tag{7.5}$$

Diese Gleichung impliziert, dass jedes verworfene Ereignis eine Stellgrößenänderung von $\|v(1)\|$ verursacht, was eine gültige obere Abschätzung darstellt. Eine weitere Variation von Gleichung (6.10) aus Satz 22 mit

$$\zeta_i^{yv} = \left|\mathbf{C} \cdot A_{cl}^i \cdot \mathbf{B}\right|$$

führt zu der Signaldichte $\|y_v(\Delta_k)\|$, welche unter Anderem in Abbildung 7.5 aufgeführt ist. Die Abbildung zeigt, dass für die gegebene Drehmomentstörung die Anforderungen an das Regelsystem nicht verletzt werden, wenn zugleich nicht mehr als jedes fünfte Ereignis verworfen wird. Eine weitere Evaluation zeigt, dass selbst dann zwei verworfene Ereignisse in einem Intervall von fünf Ereignissen zu einer

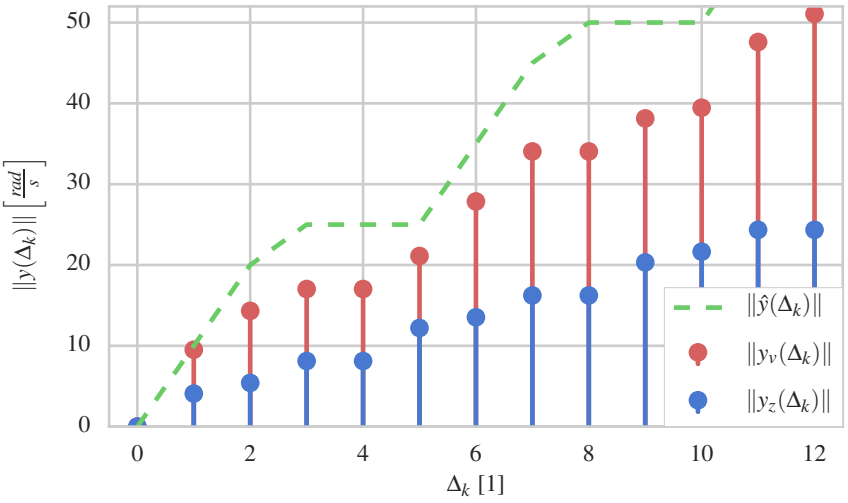


Abb. 7.5: Signaldichte der Abweichung der Winkelgeschwindigkeit $\|y(\Delta_k)\|$ bei einer Drehmomentstörung und verworfenen Ereignissen.

Überschreitung der spezifizierten Signaldichte führt, wenn diese einmalig auftreten

$$cP = \{(4, 1), (\infty, 2)\}.$$

Die Situation stellt sich wesentlich anders dar, wenn die Signaldichte der Drehmomentstörung nur 50% der in Abbildung 7.2 dargestellte Signaldichte aufweist. So ist es dadurch möglich, dass die durch verworfene Ereignisse verursachte Drehzahlabweichung $\|y_v(\Delta_k)\|$ einen größeren Anteil einnehmen kann. Ein weiterer begünstigender Faktor ist die geringere Stellgrößenänderung, was einen verminderten Einfluss eines verworfenen Ereignisses zur Folge hat. Ist diese ursprünglich $\|u_z'(1)\| = 4.72\,V$, so verringert sich diese durch die halbierte Störung zu $\|u_z'(1)\| = 2.36\,V$. Entsprechend vergrößert sich die erlaubte Anzahl an verworfenen Ereignissen, so dass durch den Verlust jedes dritten Ereignisses $cP_1 = \{(3, 1)\}$ die ursprüngliche Spezifikation nicht verletzt wird. Dies entspricht maximal zwei verworfene Ereignisse in einem Ereignisintervall von vier.

Die Ausdrucksstärke der Verlustdichte erlaubt es nun auch, dass in einem Intervall von nur drei Ereignissen, zwei Ereignisse verworfen werden können. Hierfür ist jedoch eine Einschränkung für größere Ereignisintervalle von Nöten $cP_2 = \{(2, 1), (7, 2)\}$. Solch eine Begrenzung mit mehreren charakteristischen Punk-

ten lässt sich durch bestehenden Modellen (vgl. (m,k)-Beschränkung in Unterabschnitt 5.1.4.3) nicht beschreiben.

Die maximale Verlustdichte, definiert durch die charakteristischen Punkte cP_2 ist dabei nicht zwingend dem charakteristischen Punkt cP_1 vorzuziehen. Zwar ist es durch cP_2 möglich, dass in einem Intervall von drei Ereignissen zwei verworfen werden, jedoch erlaubt cP_1 drei verworfene Ereignisse in einem Intervall von nur sieben Ereignissen, was wiederum durch cP_2 nicht möglich ist. Eine weitere maximale Verlustdichte ist durch $cP_3 = \{(2,1),(6,2),(11,3)\}$ gegeben, welche basierend auf cP_2 in einem kleineren Intervall drei verworfene Ereignisse zulässt.

Für eine verteilte Rechenplattform ist es nun ausreichend zu gewährleisten, dass nur eine der aufgeführten Verlustdichten in keinem Intervall überschritten wird, um sicherzustellen dass für die spezifizierte Drehmomentstörung die Güteanforderung an das geregelte System nicht verletzt wird. Eine weitere Kenntnis über das Systemverhalten der Gleichstrommaschine ist dabei nicht notwendig. Dadurch lässt sich isoliert von der funktionalen Sichtweise des Systems nach den Methoden aus Kapitel 5 die Zeiteffekte des verteilten Systems verifizieren.

Kapitel 8
Zusammenfassung

Seit den Anfängen der Echtzeitanalyse mit Liu und Layland (1973) vor über 40 Jahren, entwickelte sich die Bestimmung garantierter Zeitschranken von Rechenaufgaben auf Computersystemen zu einem eigenständigen Forschungsbereich. Eine Herausforderung bei der Analyse von harten Echtzeitsystemen ist es dabei, einen möglichst gering überschätzten Grenzwert des Zeitverhaltens in effizienter Weise zu bestimmen. Das Ergebnis solch einer Echtzeitanalyse wird für die Verifikation von sicherheitskritischen Funktionen eingesetzt, bei denen das Überschreiten einer Zeitspanne zu einem Systemversagen führt.

Bei der Verifikation des Verhaltens eines physikalischen Systems, welches durch ein Computersystem beeinflusst wird, kann die alleinige Kenntnis über eine obere Zeitschranke zu pessimistischen Annahmen führen. In dieser Arbeit wurde daher mit der Latenzdichte ein erweitertes Modell zur Beschreibung einer Zeitschranke eingeführt, welche das maximale respektive minimale Auftreten von Verzögerungen in Relation zu Ausführungsinstanzen von Kommunikations- und Berechnungsaufgaben setzt. Dieser Informationszugewinn wirkt sich positiv auf die Untersuchung von Computersystemen aus, welche durch wiederkehrende Verarbeitungsschritte eine Aufgabe erfüllen. Digitale Regelsysteme sind besondere Vertreter dieser Klasse, da sie durch periodische Stellwertberechnungen einen physikalischen Prozess beeinflussen, dessen Verhalten sich wiederum über einen Sensor auf zukünftige Ausführungsinstanzen des Reglers auswirkt.

Durch die in dieser Arbeit skizzierte Anknüpfung der Latenzdichte an die etablierten Methoden des Echtzeitkalküls, welches für ein breites Anwendungsfeld den Stand der Technik darstellt, lässt sich die Latenzdichte in bestehende Entwurfsabläufe verteilter Echtzeitsysteme einbinden.

Neben der Latenzdichte beschreibt diese Arbeit mit der Verlustdichte auch die ma-

ximale respektive minimale Anzahl verworfener Ereignisse in Relation zu den Ausführungsinstanzen. Sowohl die Latenzdichte als auch die Verlustdichte wurden mathematisch beschrieben und darauf aufbauend Berechnungsvorschriften für die Anwendungen in verteilten Systemen entwickelt.

Eine verteilte Plattformarchitektur entsteht durch die Vernetzung von Computersystemen, deren Vielzahl an möglichen Interferenzen einen erhöhten Aufwand bei der Zeitanalyse verursacht. Um dabei die Verlust- und Latenzdichte effizient zu speichern und zu verarbeiten, wurden diese auf deren wesentlichen charakteristischen Punkte reduziert. Über die charakteristischen Punkte ist es darüber hinaus möglich, die Dichtefunktionen zu approximieren und in weniger aussagekräftige Modelle zu überführen.

Die Bewertung eines Regelsystems gliedert sich im Wesentlichen in Stabilitätsuntersuchungen und der Beurteilung anhand eines Gütekriteriums. In dieser Arbeit wurde ein neues Gütemaß eingeführt, welches die Maximalwerte der im Regelkreis involvierten Signale begrenzt. So kann ähnlich der Latenz- und Verlustdichte die maximale Höhe und Dauer einer Regelabweichung bewertet werden. In einem anderen Anwendungsfall kann untersucht werden, ob es möglich ist, dass ein unerlaubter Systemzustand eintritt.
Ein Beitrag dieser Arbeit ist es, ein Regelsystem durch das Gütemaß der Signaldichte zu beschreiben und wie anhand einer Entwurfsstudie gezeigt, eine kombinierte Betrachtung von Signalstörungen und dem Einfluss von Zeiteffekten vorzunehmen. Die Gleichartigkeit von Signal-, Verlust- und Signaldichte durch den Bezug zu Ereignisintervallen führt zu einer inhärenten Überlagerung des grenzwertigen Verhaltens von Zeiteffekten und Signalen des Regelsystems. Mit dieser Betrachtung des schlimmsten Falls ist es daher möglich, einen Entwurfsraum hinsichtlich der Güte sicherheitskritischer Regelanwendungen rechenzeiteffizienten zu explorieren.

Die dargestellten Modelle und Methoden können in weiterführenden Forschungsarbeiten hinsichtlich ihrer Einschränkungen erweitert und auf angrenzende Forschungsgebiete angepasst werden. So wäre eine Erweiterung hinsichtlich nichtlinearer Regelsysteme und Regelungen mit einer ereignisbasierten Anregung lohnend. Besonders die Verbindung der in dieser Arbeit eingeführten Latenz- und Verlustdichte mit den reduzierten, aperiodischen Anregungen ereignisbasierter Regelsysteme könnte zu einem hocheffizienten Ressourcenverbrauch von verteilten Regelsystemen führen.

Literaturverzeichnis

J. Ackermann. Der Entwurf linearer Regelungssysteme im Zustandsraum. *at-Automatisierungstechnik*, 20(1-12), Seiten 297–300, 1972.

A. Aminifar, S. Samii, P. Eles, Z. Peng, und A. Cervin. Designing high-quality embedded control systems with guaranteed stability. In *Real-Time Systems Symposium (RTSS), 2012 IEEE 33rd*, Seiten 283–292, 2012.

A. Aminifar, E. Bini, P. Eles, und Z. Peng. Designing bandwidth-efficient stabilizing control servers. In *Real-Time Systems Symposium (RTSS), 2013 IEEE 34th*, Seiten 298–307, 2013a.

A. Aminifar, P. Eles, Z. Peng, und A. Cervin. Control-quality driven design of cyber-physical systems with robustness guarantees. In *Design, Automation Test in Europe Conference Exhibition (DATE), 2013*, Seiten 1093–1098, 2013b. doi: 10.7873/DATE.2013.230.

A. Aminifar, P. Tabuada, P. Eles, Z. Peng, Y. Zhang, J. Jiang, H. Li, M. Fujita, N. Aghaee, S. Chattopadhyay, et al. Self-Triggered Controllers and Hard Real-Time Guarantees. *Foundations and Trends in Electronic Design Automation*, 8 (3-4), Seiten 199–356, 2014.

A. Aminifar, P. Eles, und Z. Peng. Jfair: a scheduling algorithm to stabilize control applications. In *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2015 IEEE*, Seiten 63–72, 2015.

S. Anssi, K. Albers, M. Dörfel, und S. Gérard. chronVAL/chronSIM: A Tool Suite for Timing Verification of Auto-motive Applications. *Proc. Embedded Real-Time*

Software and Systems, ERTS, 2012.

A. Anta und P. Tabuada. To sample or not to sample: Self-triggered control for nonlinear systems. *Automatic Control, IEEE Transactions on*, 55(9), Seiten 2030–2042, 2010.

K.-E. Årzén, A. Cervin, J. Eker, und L. Sha. An introduction to control and scheduling co-design. In *Proceedings of the 39th IEEE Conference on Decision and Control, 2000.*, volume 5, Seiten 4865–4870, 2000.

K. J. Åström und B. Wittenmark. *Computer-controlled systems: theory and design*. Dover Publications, 3rd, 2011a.

K. J. Åström und B. Wittenmark. *Computer-controlled systems: theory and design*, chapter 2, Seiten 38–41. Dover Publications, 3rd, 2011b.

F. Baccelli, G. Cohen, G. J. Olsder, und J.-P. Quadrat. *Synchronization and Linearity: An Algebra for Discrete Event Systems*. Wiley, 1992.

R. Bhattacharya und G. J. Balas. Anytime control algorithm: Model reduction approach. *Journal of Guidance, Control, and Dynamics*, 27(5), Seiten 767–776, 2004.

E. Bini und A. Cervin. Delay-aware period assignment in control systems. In *Real-Time Systems Symposium, 2008*, Seiten 291–300, 2008.

I. Bronstein, K. Semendjajew, G. Musiol, und H. Mühlig. *Taschenbuch der Mathematik*. Verlag Harry Deutsch, Thun und Frankfurt am Main, 5. Auflage, 2001.

T. Bund und F. Slomka. A delay density model for networked control systems. In *Proceedings of the 21st International conference on Real-Time Networks and Systems*, Seiten 205–212, 2013.

T. Bund und F. Slomka. Controller/platform co-design of networked control systems based on density functions. In *Proceedings of the 4th ACM SIGBED International Workshop on Design, Modeling, and Evaluation of Cyber-Physical Systems*, Seiten 11–14, 2014.

T. Bund und F. Slomka. Worst-case performance validation of safety-critical control systems with dropped samples. In *Proceedings of the 23rd International Conference on Real Time and Networks Systems*, Seiten 319–326, 2015.

T. Bund, S. Moser, S. Kollmann, und F. Slomka. Guaranteed Bounds for the Control Performance Evaluation in Distributed System Architectures. In *Proceedings of the International Conference on Real-Time and Embedded Systems (RTES 2010)*, Singapore, Sept. 2010.

T. Bund, B. Menhorn, und F. Slomka. Event density analysis for event triggered control systems. In *Proceedings of the Conference on Design, Automation and Test in Europe*, Seiten 1111–1116, 2013.

G. Buttazzo und A. Cervin. Comparative assessment and evaluation of jitter control methods. *RTNS'07*, Seite 163, 2007.

A. Cervin. Using Jitterbug to derive control loop timing requirements. In *Proceedings of CERTS'03–Co-Design of Embedded Real-Time Systems Workshop*, 2003.

A. Cervin. Stability and worst-case performance analysis of sampled-data control systems with input and output jitter. In *American Control Conference (ACC), 2012*, Seiten 3760–3765, 2012.

A. Cervin und K.-E. Årzén. TrueTime: Simulation Tool for Performance Analysis of Real-Time Embedded Systems. In G. Nicolescu und P. J. Mosterman, editors, *Model-Based Design for Embedded Systems*. CRC Press, Nov. 2009.

A. Cervin, J. Eker, B. Bernhardsson, und K.-E. Årzén. Feedback–feedforward scheduling of control tasks. *Real-Time Systems*, 23(1-2), Seiten 25–53, 2002.

A. Cervin, B. Lincoln, J. Eker, K.-E. Årzén, und G. Buttazzo. The Jitter Margin and Its Application in the Design of Real-Time Control Systems. In *10th International Conference on Real-Time and Embedded Computing Systems and Applications (RTCSA)*, Göteborg, Sweden, Aug. 2004.

S. Chakraborty, S. Künzli, L. Thiele, A. Herkersdorf, und P. Sagmeister. Performance evaluation of network processor architectures: Combining simulation with analytical estimation. *Computer Networks*, 41(5), Seiten 641–665, 2003.

J. Cullyer. Safety-critical control systems. *Computing & Control Engineering Journal*, 2(5), Seiten 202–210, 1991.

A. Donzé und O. Maler. Systematic simulation using sensitivity analysis. In *Hybrid Systems: Computation and Control*, Seiten 174–189. Springer, 2007.

D. Fontanelli, L. Palopoli, und L. Greco. Deterministic and stochastic QoS provision for real-time control systems. In *Real-Time and Embedded Technology and Applications Symposium (RTAS), 2011 17th IEEE*, Seiten 103–112, 2011.

G. F. Franklin, J. D. Powell, und M. L. Workman. *Digital control of dynamic systems*, volume 3. Addison-wesley Menlo Park, 1998.

G. Frehse, A. Hamann, S. Quinton, und M. Woehrle. Formal analysis of timing effects on closed-loop properties of control software. In *Real-Time Systems Symposium (RTSS), 2014 IEEE*, Seiten 53–62, 2014.

A. Ghosal, T. A. Henzinger, C. M. Kirsch, und M. A. Sanvido. Event-driven programming with logical execution times. In *Hybrid Systems: Computation and Control*, Seiten 357–371. Springer, 2004.

T. E. Gibson und A. M. Annaswamy. Adaptive Control and the Definition of Exponential Stability. In *American Control Conference (ACC), 2015*, Seiten 1549–1554, 2015.

D. Goswami, R. Schneider, und S. Chakraborty. Relaxing signal delay constraints in distributed embedded controllers. *Control Systems Technology, IEEE Transactions on*, 22(6), Seiten 2337–2345, 2014.

L. Greco, D. Fontanelli, und A. Bicchi. Design and stability analysis for anytime control via stochastic scheduling. *Automatic Control, IEEE Transactions on*, 56 (3), Seiten 571–585, 2011.

K. Gresser. An Event Model for Deadline Verification of Hard Real-Time Systems. In *Real-Time Systems, 1993. Proceedings., Fifth Euromicro Workshop on*, Seiten 118–123, June 1993. doi: 10.1109/EMWRT.1993.639067.

W. Heemels, J. Sandee, und P. van den Bosch. Analysis of event-driven controllers for linear systems. *Int. J. of Control*, 81(4), Seiten 571 – 590, 2008.

D. Henriksson, A. Cervin, und K.-E. Årzén. TrueTime: Real-time Control System Simulation with MATLAB/Simulink. In *Proceedings of the Nordic MATLAB Conference*, Copenhagen, Denmark, Okt. 2003.

T. Henzinger, B. Horowitz, C. M. Kirsch, et al. Giotto: A time-triggered language for embedded programming. *Proceedings of the IEEE*, 91(1), Seiten 84–99, 2003.

T. A. Henzinger, B. Horowitz, und C. M. Kirsch. Embedded control systems development with Giotto. In *ACM SIGPLAN Notices*, volume 36, Seiten 64–72, 2001.

J. P. Hespanha, P. Naghshtabrizi, und Y. Xu. A survey of recent results in networked control systems. *PROCEEDINGS-IEEE*, 95(1), Seiten 138, 2007.

I. M. Horowitz. *Synthesis of feedback systems*. Academic, New York, 1963.

R. Isermann und M. Münchhof. *Identification of dynamic systems : an introduction with applications*. Springer, Heidelberg, New York, 2011.

C.-Y. Kao und B. Lincoln. Simple stability criteria for systems with time-varying delays. *Automatica*, 40, 2004.

H. Kopetz. Event-triggered versus time-triggered real-time systems. In A. Karshmer und J. Nehmer, editors, *Operating Systems of the 90s and Beyond*, volume 563 of *Lecture Notes in Computer Science*, Seiten 86–101. Springer Berlin Heidelberg, 1991.

H. Kopetz. *Real-time systems: design principles for distributed embedded applications*. Springer Science & Business Media, 2011.

H. Kopetz und G. Bauer. The time-triggered architecture. *Proceedings of the IEEE*, 91(1), Seiten 112–126, 2003.

H. Kopetz und G. Grünsteidl. TTP-A time-triggered protocol for fault-tolerant real-time systems. In *Fault-Tolerant Computing, 1993. FTCS-23. Digest of Papers., The Twenty-Third International Symposium on*, Seiten 524–533, 1993.

P. Kumar, D. Goswami, S. Chakraborty, A. Annaswamy, K. Lampka, und L. Thiele. A hybrid approach to cyber-physical systems verification. In *Proceedings of the 49th Annual Design Automation Conference*, Seiten 688–696, 2012.

S. Künzli, A. Hamann, R. Ernst, und L. Thiele. Combined Approach to System Level Performance Analysis of Embedded Systems. In *Proceedings of the 5th IEEE/ACM International Conference on Hardware/Software Codesign and System Synthesis*, CODES+ISSS '07, Seiten 63–68, New York, USA, 2007.

N. V. Kuznetsov. *Stability and oscillation of dynamical systems: theory and applications*. University of Jyväskylä, 2008.

J. Langer, U. Heinkel, V. Jerinic, und D. Müller. Analyse von Corner Cases und funktionaler Abdeckung auf Basis von Entscheidungsdiagrammen. 2007.

J.-Y. Le Boudec und P. Thiran. *Network calculus: a theory of deterministic queuing systems for the internet*. Springer, 2001.

E. Lee. Cyber Physical Systems: Design Challenges. In *Object Oriented Real-Time Distributed Computing (ISORC), 2008 11th IEEE International Symposium on*, Seiten 363–369, May 2008. doi: 10.1109/ISORC.2008.25.

E. A. Lee. The Future of Embedded Software, May 2006. URL http://chess.eecs.berkeley.edu/pubs/873.html.

B. Lincoln und A. Cervin. Jitterbug: A Tool for Analysis of Real-Time Control Performance. In *Proceedings of the 41st IEEE Conference on Decision and Control*, Las Vegas, NV, Dez. 2002.

Q. Ling und M. D. Lemmon. Robust performance of soft real-time networked control systems with data dropouts. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 2, Seiten 1225–1230, 2002.

Q. Ling und M. D. Lemmon. Optimal dropout compensation in networked control systems. In *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, volume 1, Seiten 670–675, 2003.

C. L. Liu und J. W. Layland. Scheduling algorithms for multiprogramming in a hard-real-time environment. *Journal of the ACM (JACM)*, 20(1), Seiten 46–61, 1973.

J. W. S. W. Liu. *Real-Time Systems*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st, 2000. ISBN 0130996513.

D. Luenberger. Observing the State of a Linear System. *IEEE Transactions on Military Electronics*, 8(2), Seiten 74–80, April 1964. ISSN 0536-1559. doi: 10.1109/TME.1964.4323124.

J. Lunze. *Ereignisdiskrete Systeme*. Oldenbourg Wissenschaftsverlag, 2006.

J. Lunze. *Regelungstechnik 1*. Number Bd. 1 in Springer-Lehrbuch. Springer Berlin Heidelberg, 2008. ISBN 9783540689096. URL https://books.google.de/books?id=fTylHoV0DNIC.

J. Lunze. *Regelungstechnik 2: Mehrgrößensysteme, Digitale Regelung*. Springer-Verlag, 2012.

J. Lunze und D. Lehmann. A state-feedback approach to event-based control. *Automatica*, 46(1), Seiten 211–215, 2010.

A. M. Lyapunov. The general problem of the stability of motion. *International Journal of Control*, 55(3), Seiten 531–534, 1992.

A. Markov. Extension of the limit theorems of probability theory to a sum of variables connected in a chain. 1971.

P. Marwedel. *Eingebettete Systeme*. Springer-Verlag, 2007.

Mathworks. SimEvents 3.1, Aug. 2010. URL http://www.mathworks.com/products/simevents/.

D. Maxim und L. Cucu-Grosjean. Response time analysis for fixed-priority tasks with multiple probabilistic parameters. In *Real-Time Systems Symposium (RTSS), 2013 IEEE 34th*, Seiten 224–235, 2013.

W. Michiels, V. Van Assche, und S.-I. Niculescu. Stabilization of time-delay systems with a Controlled time-varying delay and applications. *Automatic Control, IEEE Transactions on*, 50(4), Seiten 493–504, April 2005. ISSN 0018-9286. doi: 10.1109/TAC.2005.844723.

M. Morales, S. Rau, M. J. Palma, M. Venkatesan, F. Pulskamp, und A. Dugar. Intelligent Systems: The Next Big Opportunity, 2015. URL http://download.microsoft.com/download/9/8/7/987535CB-834C-

4652-84ED-4D0A215E7007/IDC%20-%20Intelligent%20Systems%20-%20Next%20Big%20Opportunity%20-%20Final.pdf.

J. R. Moyne und D. Tilbury. The Emergence of Industrial Control Networks for Manufacturing Control, Diagnostics, and Safety Data. *Proceedings of the IEEE*, 95 (1), Seiten 29–47, Jan 2007. ISSN 0018-9219. doi: 10.1109/JPROC.2006.887325.

T. Naumova. Abscha?tzung von Verzo?gerungsdichten aus den begrenzenden Kurven des Real-Time Calculus. Master's thesis, Universität Ulm, Institut für Eingebettete Systeme / Echtzeitsysteme, 2016.

J. Nilsson et al. *Real-time control systems with delays*. PhD thesis, Lund institute of Technology Lund, Sweden, 1998.

S. Perathoner, T. Rein, L. Thiele, K. Lampka, und J. Rox. Modeling structured event streams in system level performance analysis. In *ACM Sigplan Notices*, volume 45, Seiten 37–46, 2010.

Y. Piguet und P. Muellhaupt. Fast Step-Response Evaluation of Linear Continuous-Time Systems with Time Delay in the Feedback Loop. In *Proceedings of the 17th IFAC World Congress. Seoul (Korea)*, 2008.

J. Postel. User Datagram Protocol (RFC 768). IETF Request For Comments, Aug. 1980a.

J. Postel. Transmission Control Protocol (RFC 761). IETF Request For Comments, Jan. 1980b.

J. D. Powell und P. Katz. Sample rate selection for aircraft digital control. *AIAA Journal*, 13(8), Seiten 975–979, 1975.

S. Quinton, M. Hanke, und R. Ernst. Formal analysis of sporadic overload in real-time systems. In *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2012*, Seiten 515–520, 2012.

P. Ramanathan. Overload management in real-time control applications using (m, k)-firm guarantee. *Parallel and Distributed Systems, IEEE Transactions on*, 10 (6), Seiten 549–559, 1999.

O. Redell und M. Sanfridson. Exact best-case response time analysis of fixed priority scheduled tasks. In *Real-Time Systems, 2002. Proceedings. 14th Euromicro Conference on*, Seiten 165–172, 2002. doi: 10.1109/EMRTS.2002.1019196.

K. Richter. *Compositional scheduling analysis using standard event models: the SymTA/S approach*. PhD thesis, University of Braunschweig - Institute of Technology, 2005.

S. Samii, A. Cervin, P. Eles, und Z. Peng. Integrated scheduling and synthesis of control applications on distributed embedded systems. In *Proceedings of the conference on design, automation and test in Europe*, Seiten 57–62, 2009.

L. Schenato. Optimal estimation in networked control systems subject to random delay and packet drop. *Automatic Control, IEEE Transactions on*, 53(5), Seiten 1311–1317, 2008.

L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, und S. S. Sastry. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 95(1), Seiten 163–187, 2007.

P. Scholz. Echtzeit, Echtzeitsysteme, Echtzeitbetriebssysteme. *Softwareentwicklung eingebetteter Systeme: Grundlagen, Modellierung, Qualitätssicherung*, Seiten 39–73, 2005.

D. Seto, J. P. Lehoczky, L. Sha, und K. G. Shin. On task schedulability in real-time control systems. In *Real-Time Systems Symposium, 1996., 17th IEEE*, Seiten 13–21, 1996.

M. Simulink. The Mathworks, 1993. URL http://de.mathworks.com/products/simulink/.

F. Slomka. *Mehrkriterienoptimierung verteilter Echtzeitsysteme mit Tabu-Search*. PhD thesis, Universität Erlangen-Nürnberg, 2002.

F. Slomka. *Skript zur Vorlesung Eingebettete Systeme*. 2008.

F. Slomka, S. Kollmann, S. Moser, und K. Kempf. A Multidisciplinary Design Methodology for Cyber-physical Systems, 2011.

D. Soudbakhsh, L. T. Phan, O. Sokolsky, I. Lee, und A. Annaswamy. Co-design of control and platform with dropped signals. In *Proceedings of the ACM/IEEE 4th International Conference on Cyber-Physical Systems*, Seiten 129–140, 2013.

H.-D. Stölting und A. Beisse. *Elektrische Kleinmaschinen*. Vieweg+Teubner Verlag, 1987. doi: 10.1007/978-3-663-01104-0.

P. Tabuada. Event-triggered real-time scheduling of stabilizing control tasks. *Automatic Control, IEEE Transactions on*, 52(9), Seiten 1680–1685, 2007.

M. S. Tavazoei. Notes on integral performance indices in fractional-order control systems. *Journal of Process Control*, 20(3), Seiten 285–291, 2010.

J. Teich. *Digitale Hardware/Software-Systeme: Synthese und Optimierung*. Springer-Verlag, 2013.

L. Thiele, S. Chakraborty, und M. Naedele. Real-time calculus for scheduling hard real-time systems. In *Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on*, volume 4, Seiten 101–104, 2000.

K. Tindell und J. Clark. Holistic Schedulability Analysis for Distributed Hard Real-Time Systems. *Microprocessing and Microprogramming*, 40, Seiten 117–134, 1994.

M. Trivellato und N. Benvenuto. State control in networked control systems under packet drops and limited transmission bandwidth. *Communications, IEEE Transactions on*, 58(2), Seiten 611–622, 2010.

C. Van Loan. The sensitivity of the matrix exponential. *SIAM Journal on Numerical Analysis*, 14(6), Seiten 971–981, 1977.

M. Velasco, J. Fuertes, und P. Marti. The self triggered task model for real-time control systems. In *Work-in-Progress Session of the 24th IEEE Real-Time Systems Symposium (RTSS03)*, volume 384, 2003.

D. C. Verma, H. Zhang, und D. Ferrari. Delay jitter control for real-time communication in a packet switching network. In *Communications Software, 1991,'Com-*

munications for Distributed Applications and Systems', Proceedings of TRI-COMM'91., IEEE Conference on, Seiten 35–43, 1991.

E. Wandeler. *Modular Performance Analysis and Interface-Based Design for Embedded Real-Time Systems*. PhD thesis, ETH Zurich, September 2006.

R. Wilhelm, J. Engblom, A. Ermedahl, N. Holsti, S. Thesing, D. Whalley, G. Bernat, C. Ferdinand, R. Heckmann, T. Mitra, F. Mueller, I. Puaut, P. Puschner, J. Staschulat, und P. Stenström. The worst-case execution-time problem overview of methods and survey of tools. *ACM Trans. Embed. Comput. Syst.*, 7(3), Seiten 36:1–36:53, Mai 2008. ISSN 1539-9087. doi: 10.1145/1347375.1347389. URL http://doi.acm.org/10.1145/1347375.1347389.

J. C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *Automatic Control, IEEE Transactions on*, 16(6), Seiten 621–634, 1971.

L. Xiao, A. Hassibi, und J. P. How. Control with random communication delays via a discrete-time jump system approach. In *American Control Conference, 2000. Proceedings of the 2000*, volume 3, Seiten 2199–2204, 2000.

W. Zhang, M. S. Branicky, und S. M. Phillips. Stability of networked control systems. *Control Systems, IEEE*, 21(1), Seiten 84–99, 2001.

Anhang A
Publikationsliste

- T. Bund und F. Slomka. *Worst-Case Performance Validation of Safety-Critical Control Systems with Dropped Samples.* In Proceedings of the 23rd International Conference on Real Time and Networks Systems. 2015.
- T. Bund und F. Slomka. *Sensitivity Analysis of Dropped Samples for Performance-Oriented Controller Design.* In Proceedings of the Symposium on Real-Time Computing. 2015.
- T. Bund und F. Slomka. *Controller/Platform Co-Design of Networked Control Systems Based on Density Functions.* Fourth Workshop on Design, Modeling and Evaluation of Cyber Physical Systems. 2014.
- T. Bund und F. Slomka. *A Delay Density Model for Networked Control Systems.* In Proceedings of the 21th International Conference on Real-Time and Network Systems. 2013.
- T. Bund und F. Slomka. *A new Delay Specification for Cyber-Physical Systems Development.* Third Workshop on Design, Modeling and Evaluation of Cyber Physical Systems. 2013.
- T. Bund, B. Menhorn and F. Slomka. *Event Density Analysis for Event Triggered Control Systems.* In Proceedings of the Design Automation and Test Conference in Europe. 2013.
- T. Bund, S. Moser, S. Kollmann und F. Slomka. *Jitter Considerations for Worst-Case Performance Generation in Digital Controller Design.* Ninth GMM/ITG/GI-Workshop Cyber-Physical Systems - Enabling Multi-Nature Systems. 2012.
- T. Bund, S. Moser, S. Kollmann, und F. Slomka. *Guaranteed Bounds for the Control Performance Evaluation in Distributed System Architectures*. In Proceedings of the International Conference on Real-Time and Embedded Systems. 2010.

Anhang B
Formeln

B.1 Max-Plus-Algebra

Min-Plus Faltung

$$(f \otimes g)(k) = \inf_{0 \leqslant s \leqslant k} \{f(k-s) + g(s)\}$$

Min-Plus Entfaltung

$$(f \oslash g)(k) = \sup_{s \geqslant 0} \{f(k+s) - g(s)\}$$

Max-Plus Faltung

$$(f \overline{\otimes} g)(k) = \sup_{0 \leqslant s \leqslant k} \{f(k-s) + g(s)\}$$

Max-Plus Entfaltung

$$(f \overline{\oslash} g)(k) = \inf_{s \geqslant 0} \{f(k+s) - g(s)\}$$

B.2 Echtzeitkalkül nach (Wandeler, 2006)

Ausgehende Ereignisdichte

$$\bar{\eta}^+(\Delta t) = \inf\left\{(\eta^+ \otimes \beta^+) \oslash \beta^-, \beta^+\right\}$$
$$\bar{\eta}^-(\Delta t) = \inf\left\{(\eta^- \oslash \beta^+) \otimes \beta^-, \beta^+\right\}$$

Ausgehende Kapazitätskurve

$$\bar{\beta}^+ = (\beta^+ - \eta^-) \bar{\oslash} 0$$
$$\bar{\beta}^- = (\beta^- - \eta^+) \bar{\otimes} 0$$

Maximale Antwortzeit

$$r^+ \leqslant \sup_{\lambda > 0}\left\{\inf\left\{\tau \geqslant 0 : \eta^+(\lambda) \leqslant \beta^-(\lambda + \tau)\right\}\right\}$$

Anhang C
Beweise

C.1 Beweis zu Satz 24

Satz: *Die betragsmäßige Störung, verursacht durch den n-ten Ereignisverlust zur k-ten Ereignisinstanz ist begrenzt durch*

$$\begin{aligned}\left|v_k^n\right| &\leqslant \alpha_{v,\theta} \cdot (\alpha_{v,\mu}+1)^{n-1} \cdot e^{-\lambda_v \cdot k} \cdot |\mathbf{x}_0| \\ &\leqslant \alpha_{v,\theta} \cdot e^{\ln(\alpha_{v,\mu}+1)\cdot(n-1)-\lambda_v \cdot k} \cdot |\mathbf{x}_0|\end{aligned}$$

mit

$$\sum_{i=1}^{k} d_i = n.$$

Beweis. Der Beweis dieses Satzes erfolgt über eine vollständige Induktion.
Induktionsanfang $(n = 1)$**:**
Nach Gleichung (6.18) gilt mit der Abschätzung aus Gleichung (6.28) für den Verlust des Ereignis k_1

$$\left|\mathbf{v}_{k_1}^1\right| \leqslant \theta_{v,k_1} \cdot |\mathbf{x}_0| \leqslant \alpha_{v,\theta} \cdot e^{-\lambda_v \cdot k_1} \cdot |\mathbf{x}_0| \leqslant \alpha_{v,\theta} \cdot (\alpha_{v,\mu}+1)^{1-1} \cdot e^{-\lambda_v \cdot k_1} \cdot |\mathbf{x}_0|$$

Induktionsschritt $(n = 2)$**:**
Für den Verlust der Ereignisinstanzen $k_1$und k_2 mit $(k_1 < k_2)$ gilt entsprechend

$$\begin{aligned}\left|\mathbf{v}_{k_2}^2\right| &\leqslant \mu_{v,k_2-k_1} \cdot \left|\mathbf{v}_{k_1}^1\right| + \theta_{v,k_2} \cdot |\mathbf{x}_0| \leqslant \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot (k_2-k_1)} \cdot \left|\mathbf{v}_{k_1}^1\right| + \alpha_{v,\theta} \cdot e^{-\lambda_v \cdot k_2} \cdot |\mathbf{x}_0| \\ &\leqslant \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot (k_2-k_1)} \cdot \alpha_{v,\theta} \cdot e^{-\lambda_v \cdot k_1} \cdot |\mathbf{x}_0| + \alpha_{v,\theta} \cdot e^{-\lambda_v \cdot k_2} \cdot |\mathbf{x}_0| \\ &= \alpha_{v,\mu} \cdot e^{-\lambda \cdot k_2} \cdot \alpha_{v,\theta} \cdot |\mathbf{x}_0| + \alpha_{v,\theta} \cdot e^{-\lambda \cdot k_2} \cdot |\mathbf{x}_0| \\ &= \alpha_{v,\theta} \cdot (\alpha_{v,\mu}+1)^{2-1} \cdot e^{-\lambda_v \cdot k_2} \cdot |\mathbf{x}_0|\end{aligned}$$

In einer allgemeineren Form, lässt sich der Ausdruck als Rekursion darstellen

$$\begin{aligned}\left|\mathbf{v}_{k_n}^n\right| \leqslant \mu_{v,k_n-k_{n-1}} \cdot \left|\mathbf{v}_{k_{n-1}}^{n-1}\right| + \mu_{v,k_n-k_{n-2}} \cdot \left|\mathbf{v}_{k_{n-2}}^{n-2}\right| + \ldots \\ + \mu_{v,k_n-k_1} \cdot \left|\mathbf{v}_{k_1}^1\right| + \theta_{v,k_2} \cdot \left|\mathbf{x}_0\right|, \quad n \in \mathbb{N}\end{aligned}$$

Unter der Annahme der Induktionsbehauptung, dass obiger Satz für alle $n \in \mathbb{N}$ gilt, ergibt sich

$$\begin{aligned}\left|\mathbf{v}_{k_n}^n\right| &\leqslant \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot k_n} \cdot \alpha_{v,\theta} \cdot (\alpha_{v,\mu}+1)^{n-2} \cdot |\mathbf{x}_0| \\ &\quad + \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot k_n} \cdot \alpha_{v,\theta} \cdot (\alpha_{v,\mu}+1)^{n-3} \cdot |\mathbf{x}_0| + \ldots \\ &\quad + \alpha_{v,\mu} \cdot e^{-\lambda_v \cdot k_n} \cdot \alpha_{v,\theta} \cdot |\mathbf{x}_0| + \alpha_{v,\theta} \cdot e^{-\lambda_v \cdot k_n} \cdot |\mathbf{x}_0| \\ &= \left(\alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-2} + \alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-3} + \cdots + \alpha_{v,\mu}+1\right) \cdot \alpha_{v,\theta} \cdot e^{-\lambda_v \cdot k_n} \cdot |\mathbf{x}_0|\end{aligned}$$

Die Induktionsbehauptung ist bewiesen, da gilt

$$\begin{aligned}(\alpha_{v,\mu}+1)^{n-1} &= (\alpha_{v,\mu}+1)^{n-2} \cdot (\alpha_{v,\mu}+1) = \alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-2} + (\alpha_{v,\mu}+1)^{n-2} \\ &= \alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-2} + \alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-3} + (\alpha_{v,\mu}+1)^{n-3} \\ &= \alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-2} + \alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-3} + \alpha_{v,\mu}(\alpha_{v,\mu}+1)^{n-4} + \ldots \\ &\quad + \alpha_{v,\mu}(\alpha_{v,\mu}+1)^1 + \alpha_{v,\mu} + 1\end{aligned}$$

□

C.2 Beweis zu Satz 28

Satz: Die betragsmäßige Störung des Zustandsvektors, verursacht durch eine Latenz r_k zur k-ten Ereignisinstanz ist begrenzt durch

$$|\mathbf{w}_k| \leqslant |B_{-1}(r_k)| \cdot \prod_{n=1}^{k-1} \left(\alpha_{w,\mu} \, |B_{-1}(r_n)| + 1\right) \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot k} \cdot |\mathbf{x}_0| \, .$$

Beweis. Der Beweis dieses Satzes erfolgt über eine vollständige Induktion.
Induktionsanfang $(k = 1)$**:**
Nach Gleichung (6.38) gilt mit der Abschätzung aus Gleichung (6.42) für die Verzögerung des ersten Ereignisses

$$|\mathbf{w}_1| \leqslant |B_{-1}(r_1)| \cdot \theta_{w,1} \cdot |\mathbf{x}_0| \leqslant |B_{-1}(r_1)| \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot 1} \cdot |\mathbf{x}_0|$$

Induktionsschritt $(k = 2)$**:**
Für die äquivalente Störung des Zustandsvektors durch die Latenz des zweiten Ereignisses gilt entspreche

$$\begin{aligned}
|\mathbf{w}_2| &\leqslant |B_{-1}(r_2)| \cdot (\mu_{w,1} \cdot |\mathbf{w}_1| + \theta_{w,2} \cdot |\mathbf{x}_0|) \\
&\leqslant |B_{-1}(r_2)| \cdot \left(\alpha_{w,\mu} \cdot e^{-\lambda_w \cdot 1} \cdot |B_{-1}(r_1)| \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot 1} \cdot |\mathbf{x}_0| + \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot 2} \cdot |\mathbf{x}_0|\right) \\
&= |B_{-1}(r_1)| \cdot \left(|B_{-1}(r_2)| \cdot \alpha_{w,\theta} \cdot \alpha_{w,\mu} \cdot e^{-\lambda_w \cdot 2} \cdot |\mathbf{x}_0| + \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot 2} \cdot |\mathbf{x}_0|\right) \\
&= |B_{-1}(r_2)| \cdot (|B_{-1}(r_1)| \cdot \alpha_{w,\theta} + 1) \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot 2} \cdot |\mathbf{x}_0|
\end{aligned}$$

Für eine allgemeine Darstellung ergibt sich die Rekursion

$$\begin{aligned}
|\mathbf{w}_k| \leqslant |B_{-1}(r_k)| \cdot (&\mu_{w,1} \cdot |\mathbf{w}_{k-1}| + \mu_{w,2} \cdot |\mathbf{w}_{k-2}| + \ldots \\
&+\mu_{w,k-1} \cdot |\mathbf{w}_1| + \theta_{w,k} \cdot |\mathbf{x}_0|) \, , \quad k \in \mathbb{N}
\end{aligned}$$

Unter der Annahme der Induktionsbehauptung, dass obiger Satz für alle $k \in \mathbb{N}$ gilt, ergibt sich

$$
\begin{aligned}
|\mathbf{w}_k| \leqslant\ & |B_{-1}(r_k)| \cdot \Bigg(\alpha_{w,\mu} \cdot |B_{-1}(r_{k-1})| \prod_{n=1}^{k-2} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot k} \cdot |\mathbf{x}_0| \\
& + \alpha_{w,\mu} \cdot |B_{-1}(r_{k-2})| \prod_{n=1}^{k-3} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot k} \cdot |\mathbf{x}_0| + \dots \\
& + \alpha_{w,\mu} \cdot |B_{-1}(r_1)| \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot k} \cdot |\mathbf{x}_0| \\
& + \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot k} \cdot |\mathbf{x}_0| \Bigg) \\
=\ & |B_{-1}(r_k)| \cdot \Bigg(\alpha_{w,\mu} \cdot |B_{-1}(r_{k-1})| \prod_{n=1}^{k-2} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) \\
& + \alpha_{w,\mu} \cdot |B_{-1}(r_{k-2})| \prod_{n=1}^{k-3} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) + \dots \\
& + \alpha_{w,\mu} \cdot |B_{-1}(r_1)| \\
& + 1) \cdot \alpha_{w,\theta} \cdot e^{-\lambda_w \cdot k} \cdot |\mathbf{x}_0|
\end{aligned}
$$

Die Induktionsbehauptung ist bewiesen, da gilt

$$
\begin{aligned}
\prod_{n=1}^{k-1} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) &= \prod_{n=1}^{k-2} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) \cdot (\alpha_{w,\mu} |B_{-1}(r_{k-1})| + 1) \\
&= \alpha_{w,\mu} |B_{-1}(r_{k-1})| \cdot \prod_{n=1}^{k-2} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) \\
&\quad + \prod_{n=1}^{k-2} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) \\
&= \alpha_{w,\mu} \cdot |B_{-1}(r_{k-1})| \prod_{n=1}^{k-2} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) \\
&\quad + \alpha_{w,\mu} \cdot |B_{-1}(r_{k-2})| \prod_{n=1}^{k-3} (\alpha_{w,\mu} |B_{-1}(r_n)| + 1) + \dots \\
&\quad + \alpha_{w,\mu} \cdot |B_{-1}(r_1)| + 1
\end{aligned}
$$

□