Universität Ulm

Fakultät für Informatik



Berücksichtigung lokaler Randbedingungen bei globaler Zieloptimierung mit neuronalen Netzen am Beispiel Truck Backer-Upper

Boris Hartmann

Universität Ulm

Nr. 95-10

Ulmer Informatik-Berichte

September 1995

0
QAA 5
A 4.95,10

Berücksichtigung lokaler Randbedingungen bei globaler Zieloptimierung mit neuronalen Netzen am Beispiel Truck Backer-Upper

Boris Hartmann (kyb)

Abteilung Neuroinformatik
Universität Ulm

Vorwort

Das Problem der Berücksichtigung systemimmanenter Randbedingungen bei der Reglerentwicklung, gemäß dem neuronalen Ansatz nach B. Widrow und P. Nguyen in [4], hat sich herauskristallisiert aus der nachträglichen Bewertung des BPtt-Verfahrens in [1]. Unklarheiten ergaben sich bei der theoriegemäßen Übertragung des Verfahrens auf zeitdiskrete Regelungen, wenn bei Vorhandensein von Randbedingungen der Strecke nichtdifferenzierbare Korrekturfunktionen notwendig wurden, die zudem eine sprungartige Veränderung der Stellsignale verursachten. Diese Problematik wurde in [4] nicht angesprochen. In [1] ist sie zwar erwähnt, konnte aber dort nicht mehr vertieft werden und wurde deshalb durch Anwendung einer behelfsmäßigen Methode umgangen.
In dieser Arbeit steht dieses Thema im Zentrum der Untersuchungen. Es erfolgt eine Verallgemeinerung der Behandlung solcher systemimmanenter Randbedingungen auf die Einbeziehung zusätzlicher Optimierungskriterien in das Reglerlernen. Dies wird exemplarisch an zwingend auftretenden und zusätzlich denkbaren Randbedingungen der TBU-Regelung demonstriert.
Da es sich um eine relativ spezielle Thematik handelt, werden beim Leser bereits grundsätzliche Kenntnisse in den Gebieten der neuronalen Netze (speziell Backpropagation) und Optimierungsverfahren (speziell Gradientenverfahren) vorrausgesetzt. Als an das Thema heranführende Literatur werden vom Autor u.a. die Quellen [2], [3] und [5] vorgeschlagen.

Der thematische Aufbau gliedert sich wie folgt:
Kapitel 1 führt in die Problematik der Regleroptimierung hinsichtlich einer primären Regelungsaufgabe unter Berücksichtigung zusätzlicher Kriterien ein. Die Optimierungskriterien werden nach ihrer Bedeutung für die Streckenregelung eingeteilt und deren mit Lokalisationscharakter bezeichnete Eigenschaft herausgestellt. Gegen Ende des Kapitels wird die allgemeinere Betrachtung auf den Einsatz des Gradientenverfahrens bei BPtt hin konkretisiert.
Kapitel 2 stellt die Regelstrecke TBU, die TBU-Regelungsaufgabe und ihre begleitenden Randbedingungen vor.
Kapitel 3 betrachtet die zur Verhinderung von Randbedingungs-Verletzungen notwendigen Korrekturen des Stellsignals. Dies geschieht unter dem Aspekt ihrer Integration in das BPtt-Verfahren zur Adaption des neuronalen Reglers. Unter anderem werden hier verschiedene Propagierungsstrategien vorgestellt, die bei den Testläufen Verwendung finden.
Kapitel 4 bereitet die Untersuchungen vor. Hier wird die Zahl der zu variierenden Lernparameter eingeschränkt. Die Analysemethoden werden beschrieben und der strukturelle Aufbau der Testkonfigurationen sowie deren grundlegenden Lernparameterwerte angegeben.

In Kapitel 5 schließlich erfolgt die Präsentation der wichtigsten Testergebnisse und ihrer Analysen. Diese beziehen sich auf die vorher angesprochenen Aspekte des Reglerlernens. Abschließend werden die Ergebnisse in einem Fazit zusammengefaßt. Die experimentellen Untersuchungen mögen vielleicht etwas zu umfangreich und detailliert erscheinen. Da jedoch im Bereich der Arbeit mit neuronalen Netze zu leicht Spekulationen und anhand einzelner Ergebnisse nur unsichere Aussagen zu machen sind, ist eine möglichst umfangreiche Betrachtung zahlreicher Netzbeispiele durchaus geboten.

Realisiert wurden die hier vorgestellten Tests mit einem vom Autor selbst geschriebenen Programmsystem. Das Programmkonzept ist stark am modularen Aufbau des BPtt-Lernmodells orientiert und besitzt folgende Kennzeichen:

- Einfaches Hinzufügen neuer Funktionsmodule aufgrund einheitlicher Daten- und Funktionsschnittstellen.
- Konfigurierung der Modul-Verknüpfungen und Parametereinstellungen zusammen in einem Konfigurationsfile.
- Steuerung des Lernens und der nachfolgenden Testläufe ist interaktiv möglich. Mehrere leicht variierende Lernkonfigurationen können ausgehend von einer Grundkonfiguration nur durch Angabe ihrer Variationen definiert und als Batchjob ohne Eingreifen des Anwenders nacheinander abgearbeitet werden.
- Als Programmiersprache wurde C gewählt. Eine objektorientierte Anpassung dürfte aber aufgrund des bereits sehr modularen Ansatzes leicht durchführbar sein.
- Einfache modulspezifische grafische Aufbereitung der Moduldaten, wie z.B. Netzinformationen oder die Regelstrecken-Trajektorie mittels der unter UNIX frei verfügbaren Xview-Bibliotheksroutinen.

Die in [1] bereits gelungene Implementierung unter Zuhilfenahme der kommerziellen NeuralWorks-Simulationsumgebung[1] zur Simulation neuronaler Netze wurde nicht weiterverwendet. Die Vorteile der eigenen Entwicklung gegenüber dieser hybriden Realisierung sind folgende:

- Es konnte eine auf die spezielle Anwendung zugeschnittene, modulare Struktur verwirklicht werden. Eine Einbindung des NeuralWorks-Programms über die dort vorhandene Schnittstellenfunktionen in die angestrebte Modulstruktur ist nur mit großem Aufwand möglich, da eine Steuerung der Netzabarbeitung vom C-Schnittstellenprogramm aus nicht möglich ist.
- Die einheitliche modulare Grundstruktur gewährleistet eine gewisse Übersichtlichkeit, die erforderlich ist, wenn viele unterschiedliche Funktionen integriert werden müssen (Neuronales Netz, Streckensimulation, Wertebereichstransformationen, Fehlerfunktionen, etc.). NeuralWorks dagegen besteht aus mehreren voneinander abhängigen Programmteilen. Zur Anpassung an die spezielle Anwendung steht zudem keine einheitliche Programmiersprache zur Verfügung. Eine Realisierung des BPtt-Reglerlernens geriet letztendlich zu aufwendig und unübersichtlich, so daß auf die eigene Entwicklung umgestellt wurde.

[1] In der Version NeuralWorks Professional II / PLUS, Version 4.0 .

- Die Ausführungsgeschwindigkeit des speziell auf die Anwendung hin entwickelten Programms konnte gegenüber der allgemeiner ausgerichteten kommerziellen Simulationsumgebung erheblich gesteigert werden. Der dortige Verabeitungsoverhead sowie der Datentransfer zwischen Anwender- und Simulationsprogramm scheinen stark zu bremsen. Während ein Lerndurchgang der in [1] durchgeführten Berechnungen in etwa 8 Stunden auf einer SUN-Sparcstation II benötigten, waren mit dem hier verwendeten Programm auf einem IBM-kompatiblen 80486 DX2/50 unter LINUX für ähnliche Konfigurationen nur ca. 20 Minuten Rechenzeit nötig.

Die funktionale Korrektheit des hier eingesetzten Simulationsprogramms wurde unter anderem kontrolliert, indem die mit den beiden Implementierungen erzielten Reglerperformanzwerte für die in [1] vorgestellte Lernkonfiguration verglichen wurden. Bis auf die Simulationsroutinen des TBU-Simulationsmoduls, das hier in nur wenigen Punkten abgeändert wurde, handelt es sich um völlig unabhängig voneinander entwickelte Implementierungen. Die Ähnlichkeit der Testwerte weist daher auf eine prinzipielle funktionale Übereinstimmung der beiden Programmsysteme hin:

Implementierung	Standardabweichungen im Endpunkt der Regelung	
	Winkel	Höhe
Mit NeuralWorks, s. [1]	7.64°	$1.91m$
Eigenes Sim.programm hier	7.24°	$1.91m$

Die kleinen Abweichungen lassen sich durch eine andere Zufallszahlengenerierung in den beiden Implementierungen erklären. Dadurch erfolgt eine unterschiedliche Gewichtsinitialisierung und Auswahl der einzelnen Trainingsbeispiele.

Kontakt

Die E-Mail-Adressen des Autors lauteten zur Zeit der Erstellung dieser Diplomarbeit:
boris_hartmann@ccwn.fido.de
bhartma@ibm.net

Inhaltsverzeichnis

1 Randbedingungen in Regelsystemen 1
- 1.1 Einteilung der Regelungsanforderungen . . . 1
- 1.2 Lokalisationscharakter von Randbedingungen . . . 3
- 1.3 Integration von Randbedingungen mit BPtt . . . 5

2 Beispiel TBU 11
- 2.1 TBU, Modell und Regelung . . . 11
- 2.2 Systemimmanente Randbedingungen . . . 13
  - 2.2.1 Lenkwinkelbegrenzung . . . 13
  - 2.2.2 TBU-Selbstblockade . . . 14
- 2.3 Weitere Optimierungskriterien . . . 15
  - 2.3.1 Reduktion der Lenkwinkeldifferenz . . . 16
  - 2.3.2 Orientierung an der Horizontalen . . . 16

3 Aspekte der Lernintegration 19
- 3.1 Notwendigkeit von Korrekturmodulen . . . 21
- 3.2 Integrale/separate Korrektur . . . 21
- 3.3 Rückpropagierung durch Korrekturfunktionen . . . 23
- 3.4 Propagierungs-Strategien . . . 26
- 3.5 Bewertung von Randbedingungen . . . 27
  - 3.5.1 Sigmoidfunktion als Korrekturfunktion . . . 28
  - 3.5.2 Bewertungen des Regelstreckenzustands . . . 29
  - 3.5.3 Bewertungen der TBU-Randbedingungen . . . 29
- 3.6 Separate Gewichtungen . . . 30

4 Tests 33
- 4.1 Testvorbereitungen . . . 33
  - 4.1.1 Variable Lernparameter . . . 33
  - 4.1.2 Hohe Dimension des Lernparameterraums . . . 36
  - 4.1.3 Restriktionen . . . 37
  - 4.1.4 X-Positionsfehler . . . 38
- 4.2 Analysen . . . 39
  - 4.2.1 Mathematische Analysewerte . . . 39
  - 4.2.2 Testbilder . . . 43
- 4.3 Testkonfigurationen . . . 44
  - 4.3.1 Grundkonfiguration . . . 46

4.3.2 Übersicht 50

5 Testergebnisse **53**
5.1 Grundkonfiguration ohne lokale Bewertungen 53
5.1.1 Variation der Propagierungsstrategie 56
5.1.2 Variation der Lernparameter 58
5.1.3 Zusammenfassung 63
5.2 Lokale Bewertungen zur Blockadevermeidung 64
5.2.1 Bewertung des Stellsignals 64
5.2.2 Bewertung des Blockadewinkels 72
5.2.3 Bewertung der Blockadeposition 76
5.2.4 Lokalisationsuntersuchung 79
5.2.5 Zusammenfassung 81
5.3 Lenkwinkelbegrenzung mittels lokaler Bewertung 82
5.3.1 Probleme bei fehlender Sigmoidfunktion 82
5.3.2 Bewertung des Lenkwinkelfehlers 87
5.3.3 Bewertung von Lenkwinkelbegrenzung und Blockade 91
5.3.4 Zusammenfassung 94
5.4 Bewertung der Lenkwinkeldifferenz 96
5.4.1 Zusammenfassung 96
5.5 Fazit 99

A TBU-Modell **101**

B Fehlerskalierung und Wertebereiche **105**

C Abkürzungen **109**

Kapitel 1

Randbedingungen in Regelsystemen

Bei der Regelung der meisten technischen Systeme sind in der Praxis je nach Anwendung unterschiedliche Randbedingungen der Regelstrecke oder des Stellapparats zu beobachten. Diese müssen unter Umständen in geeigneter Weise in die Regelung miteinbezogen werden. Typische Arten von Randbedingungen in realen technischen Systemen sind Begrenzungen des Stellsignals (z.B. Ventilhub) oder vom Streckenzustand abhängige Randbedingungen (z.B. nicht zu große Querbeschleunigungen eines Automobils während einer Kurvenfahrt). Auch wenn diese systemimmanenten Randbedingungen in einer realisierten Regelung nicht immer in Erscheinung treten, ist eine Kenntnis dieser Systemeigenschaft vorab für eine fundierte Reglerentwicklung und ein sicheres Funktionieren unbedingte Vorraussetzung. Wenn ein Verletzen solcher Randbedingungen in der Folge nicht ausgeschlossen werden kann, ist weiter für geeignete Maßnahmen zu sorgen, mittels derer eine Einhaltung der Stellsignalgrenzen garantiert werden kann, um Beschädigungen an Stellapparat oder Strecke und Umgebung zu vermeiden. Während dies bei Randbedingungen allein des Stellsignals meist bereits in einfacher Form als Begrenzung durch den Stellapparat selbst gegeben ist (Anschlag bei mechanischen Stellern, Leistungsgrenzbereiche von Motoren, etc.), bedürfen Randbedingungen der Strecke einer Beobachtung und Analyse des Streckenzustands. Aufgrund eines komplizierteren Zusammenhangs zwischen Stellsignal und resultierender Zustandsveränderung muß auch eine aufwendigere Behandlung hinsichtlich einer Bewahrung vor unerwünschten Strecken-Zuständen erfolgen.

1.1 Einteilung der Regelungsanforderungen

Neben der eigentlichen Regelungsaufgabe, dem Ausgleich der Ist-Soll-Differenz, und den angesprochenen Randbedingungen des Systems Strecke/Stellglied sind weitere Anforderungen möglich, der die fertige Regelung genügen soll. Diese Optimierungskriterien sind (gegenüber systemimmanenten) als zusätzlich auferlegte Randbedingungen zu interpretieren. Auch die primäre Regelungsaufgabe kann als eine solche betrachtet werden: "Beeinflusse das zu regelnde System in einer Weise, die den Ist-Fehler des Streckenzustands minimiert". Zusätzliche Kriterien erweitern diese primäre Randbedingung der Regelung, indem sie deren Anforderung um weitere qualitative Aspekte ergänzen.
Eine Einteilung der für die Regelung in einen Sollzustand relevanten Randbedingungen

kann deshalb wie folgt aussehen:

Primäres Regelungsziel:
Die eigentliche Regelungsaufgabe, die Beseitigung der Ist-Soll-Differenz des Streckenzustands.

Randbedingungen des Systems Strecke/Stellglied:

1. Stellsignalbegrenzungen, Bsp: Lenkwinkelbegrenzungen bei Fahrzeugen, Ventilanschlag.
2. Bereiche unerlaubter Streckenzustände, die vermieden werden müssen. Das sind Bedingungen, die sich aus der Formulierung des Streckenmodells bzw. aus zwingenden, weil für eine realistische Modellierung notwendigen, Einschränkungen ergeben. Oder auch Bedingungen, die, um es vorsichtig auszudrücken, verhindern sollen, daß ein Wechsel der Systemmodellierung erfolgen muß. Bsp: Vermeiden von Selbstblockaden des TBU-Systems, Ausbrechen eines PKWs in Kurven oder Temperaturüberschreitungen in Reaktoren.

Optimalitätskriterien:
Zusätzliche Anforderungen an die Regelung, wie z.B. ein möglichst geringer Energieaufwand für den Stellvorgang, der Strecke entlang ihrer Trajektorie oder Zeitoptimalität. Diese Kriterien werden von außen vorgegeben, sind also nicht durch das Modell der Regelstrecke gegeben.

Während die Einregelung auf die Sollposition und die Optimalitätskriterien im Laufe der Reglerentwicklung prinzipiell im Regler als eine Funktion implementiert werden dürfen, deren Optimalität nicht hundertprozentig gewährleistet sein muß, stellen die systemimmanenten Randbedingungen der Strecke / des Stellglieds eine i. allg. absolute Forderung an die Einhaltung ihrer Bereichsgrenzen. Diese Randbedingungen sind demnach durch Zuordnung zu einem von zwei Typen zu unterscheiden:

- Unbedingt einzuhaltende Begrenzungs-Randbedingungen werden hier folgend als *limitierende Randbedingungen* bezeichnet.
- Randbedingungen, die nur eine tendenzielle Beeinflussung des Reglerverhaltens zum Ziel haben, werden im folgenden als *forcierende Randbedingungen* bezeichnet.

Speziell die restriktive Forderung der absoluten Einhaltung des Zulässigkeitsbereichs bei limitierenden Randbedingungen kann bei der Reglerkonstruktion zum Problem werden, wenn

- nach gelungener Integration dieser Regelungsfunktion die erforderliche Sicherheit vor Verletzungen der Randbedingungen nicht nachgewiesen werden kann,
- die Reglerentwicklung dadurch in nicht akzeptabler Weise erschwert wird,
- die Güte der dadurch beeinträchtigten Gesamtreglerfunktion nicht mehr den vorgegebenen Anforderungen entspricht.

Um in diesen Fällen dennoch zu einer Lösung zu kommen, kann zwar versucht werden, mit nach der Reglergenerierung zwischen Regler und Strecke geschalteten Modulen korrigierend einzugreifen (die Reglergenerierung erfolgt vorab, ohne die problematischen Randbedingungen miteinzubeziehen). Randbedingungen des Streckenzustandes werden dabei in zusätzliche Randbedingungen für das Stellsignal umformuliert, um solchen Zuständen präventiv entgegenzuwirken, soweit dies möglich ist[1]. Durch diesen Eingriff wird dann aber die Regelungsfunktion in entscheidender Weise verfremdet, so daß bei analytischer Herleitung der Reglerfunktion die berechneten Regelungskennwerte nicht mit den tatsächlichen übereinstimmen werden. Mathematische Aussagen über die ursprünglich generierte Regelungsfunktion, die eine unverfälschte Funktion vorraussetzen, werden damit unsicher und verlieren ihre Aussagekraft. Z.B. kann nach Entwicklung einer zeitoptimalen Regelung deren Optimalität durch nachträglich installierte Begrenzungsmodule verloren gehen. Erst ein Konstruktionsansatz, der die limitierenden Randbedingungen bereits mitberücksichtigt, vermag diese Optimierung zu garantieren.
Zudem sind durch wechselseitige Einwirkungen unabhängig von der eigentlichen Regelung arbeitender Korrekturmodule große und schnelle Bewegungen des Stellsignals denkbar. Diese führen dann bei Regelstrecke oder Stellapparat zu unerwünschten starken Beanspruchungen und verhindern so einen realen Einsatz der konzipierten Regelung. Ein geeigneter Konstruktionsansatz, der die Einhaltung der limitierenden Randbedingungen mit in die Regelungsfunktion integriert, kann derartige Wechselwirkungen möglicherweise vermeiden helfen. Wird die Regelfunktion (wie z.B. beim neuronalen Regler) als stetige Abbildung realisiert und gelingt eine Randbedingungs-Integration, dann werden aus mit externen Korrekturfunktionen produzierten Stellsignalsprüngen implizit im Regler enthaltene stetige Korrektur-Kennlinien. Dies kann eine Verminderung der Beanspruchung des Stellapparats bewirken und hilft, unerwünschte Sprungantworten des zu regelnden Systems zu vermeiden. Eine so resultierende Glättung der Korrekturkennlinie bedeutet dann allerdings auch, den zur Verfügung stehenden Stellsignalwertebereich dynamisch nicht optimal auszunutzen, was zwangsläufig zu einer gegenüber der theoretisch vorstellbaren Regelung suboptimalen Lösung der globalen Regelungsaufgabe führt.

1.2 Lokalisationscharakter von Randbedingungen

Mit der eben erwähnten *globalen* Regelungsaufgabe ist die Positionierung auf den Sollzustand charakterisiert. Global deswegen, weil das Erreichen dieses Zustandes (bzw. der dabei entstehende Fehler) aus Startpositionen eines genügend großen Zustandsbereichs heraus von den Regelungsaktionen entlang der gesamten Trajektorie abhängt. Eine Optimierung des Reglers für diese Aufgabe stellt eine Parametrisierung der Regelungsfunktion für den gesamten zu betrachtenden Arbeitsbereich hinsichtlich einer bestmöglichen Zielpositionierung dar.
Anders dagegen darf die Einstellung des Reglers hinsichtlich einer Vermeidung von Verletzungen von *lokalen* Randbedingungen die Reglerfunktion nur für diejenigen nahen Zustandsbereiche beeinflussen, in denen das Reglerverhalten diese Verletzungen entscheidend mitbeeinflußt.

[1] Anderenfalls ist für einen sofortigen Stopp der Regelung oder eine Regelung in einen sicheren Bereich zu sorgen.

Die Vermeidung von bestimmten Streckenzuständen oder die Einhaltung von Stellsignalbereichen sollte dann streng genommen nur von den lokalen Zustandseingaben in den Regler abhängig sein und nicht vom Verlauf der Trajektorie weit vorher[2]. Entsprechend dieser fehlenden Abhängigkeit des auftretenden Fehlers von weiter vorausgegangenen Regelungsvorgängen darf sich die Parametrisierung der Regelungsfunktion nur auf den so beschriebenen Zustandswertebereich der Reglereingabe beziehen, um nicht die globale Regelungsfunktion zu stark zu beeinträchtigen. Sie soll nur die Funktion der zu ersetzenden Korrekturmodule in die Regelungsfunktion miteinbauen.
Dem global wirkenden Kriterium kommt somit die Rolle einer der Regelungs-Trajektorie formgebenden Kraft zu. Dieser Grundform überlagert sorgen die durch lokale Kriterien verursachten Regeleigenschaften entlang der Trajektorie für die Einhaltung der Randbedingungen.

Einschub: Bewertung von Randbedingungen

Mit Bewertung einer Randbedingung sei im folgenden die mathematische Formulierung der Höhe der Abweichung einer zu beobachtenden Variablen von den Grenzen ihres Sollwertebereichs abgekürzt. Diese Formulierung muß in einer für das darauf anzuwendende Optimierungsverfahren geeigneten Weise erfolgen. Speziell für das in dieser Diplomarbeit mit BPtt angewandte Gradientenverfahren erfolgt diese Formulierung in Form positiv definiter Fehlerfunktionen (Kostenfunktionen), wie z.B. das Quadrat einer Sollabweichung. Der Begriff der Bewertung beinhaltet auch die aus dieser Formulierung des Fehlers abgeleitete und mit dem Optimierungsverfahren durchgeführte Parametrisierung der zu optimierenden Funktion. Die Fehlerfunktion sei im folgenden auch mit Bewertungsfunktion bezeichnet.
Für die Begrenzung auf einen zulässigen geschlossenen Wertebereich (limitierende Randbedingung) setzt man die Bewertung innerhalb des zulässigen Bereichs gleich Null und als Sollabweichung die minimale Entfernung des Werts der zu bewertenden Variablen zur Bereichsgrenze. Handelt es sich um eine forcierende Randbedingung, so stellt sich deren Bewertung wie eben beschrieben dar und wird zum Spezialfall, wenn es z.B. um eine Minimierung eines Energieaufwandes E geht. Der Zulässigkeitsbereich im obigen Sinne reduziert sich dann auf den Punkt $E = 0$.

Um eine Zuordnung eines eher lokalen oder globalen Charakters zu einer Randbedingung (ob systemimmanente Randbedingung oder Optimierungskriterium) vornehmen zu können, muß abgeschätzt werden, in welchem Maße vorangehende Regelungsvorgänge in anderen Zustandsbereichen für die Bewertung der Randbedingung von Bedeutung sind. Eine derartige Bedeutung besteht anscheinend dann, wenn die Bewertung der Randbedingung Abhängigkeiten von der Position der Strecke im Zustandsraum aufweist. Ein Beispiel:
Der Energieaufwand, der notwendig ist, ein Fahrzeug zu beschleunigen, hängt in unebenem Gelände von dessen Position ab. Bei einer Optimierung hinsichtlich eines Minimums an Energieaufwand für eine Regelung sind dann sinnvollerweise die Regelungsvorgänge mitzubeeinflussen, die ursächlich für die Bewegung in eine bestimmte Anstiegsposition des Fahrzeugs mit hohem Energieaufwand sind. Bei einer Betrachtung in der Ebene, bei der der Energieaufwand ausschließlich von der Größe des Stellsignals (z.B. Treibstoffzufuhr) abhängt, spielt das vorausgegangene Reglerverhalten keine

[2] Eine gewisse lokale Ausdehnung des Abhängigkeitsbereich von den vorherigen Regelungsaktionen kann, wie später noch genauer ausgeführt wird, unter Umständen vorhanden sein.

Rolle.
Das Stellsignal wird zum Zustandsraum hinzugezählt, wenn eine Abhängigkeit der Bewertung vom Wert des Stellsignals selber ebenfalls zu einer Erweiterung des Beeinflussungsbereichs durch benachbarte Regelvorgänge führt. Beispiel:
Bei einem Federventil hängt der Energieaufwand der Stellbewegung (Δu) von der vorherigen Position des Stellelements ab ($\Delta E = \int_{u_1}^{u_2} F(u)du$). Somit ist die vorangegangene Regleraktion mitverantwortlich für den nun zu leistenden Energieaufwand.
Aufgrund dieser Überlegungen ist auch anzunehmen, daß es zwischen der rein globalen Wirkung der Bewertung durch das primäre Regelungskriterium und einer rein lokal beeinflussenden Bewertung beliebige Grade von Abstufungen dieses *Lokalisations*-charakters (im folgenden als Oberbegriff verwendet) gibt. Diese werden je nach Strecke und Regelungsanforderungen individuell ausgeprägt sein. Ebenso werden derartige Abschätzungen in ihren Abstufungen oft nur tendenziell gelingen, wenn sie sich einer exakten Analyse aufgrund eines hohen Aufwands oder einer schlechten Praktikabilität entziehen.
Bei einer Verletzung einer begrenzenden Randbedingung wäre eine quantitative Angabe des Lokalisationsbereichs der Bewertung z.B. der mit dem Bewertungszustand zusammenhängende Bereich, innerhalb dessen regelungstechnisch gesehen keine Steuerbarkeit in einen Zulässigkeitsbereich dieser Randbedingung für $t \rightarrow \infty$ gegeben ist. Beispiel:
Um die Randbedingungen 'Kollision mit Wand vermeiden' eines auf eine Wand zufahrenden Fahrzeugs nicht zu verletzen, muß eine Wende möglich sein. Der Lokalisationsbereich in der eben beschriebenen strengen Art ist dann der Bereich vor der Wand, innerhalb dem der Wendekreis zur Vermeidung einer Kollision nicht mehr ausreicht.

1.3 Integration von Randbedingungen mit BPtt

Die Art und Weise, wie sich anhand einer konkreten Optimierungsvorgabe mit Hilfe der Methode des BPtt ein neuronaler Regler realisieren läßt, macht sie auch für die Anpassung des Reglers an mehrere Kriterien interessant. Zum einen, weil es sich um einen dynamischen Lernvorgang handelt, infolge dessen sich eine geeignete Approximation der gesuchten Funktion selbständig herausbilden sollte. Zum anderen, weil sich die zur Optimierung verwendete Fehlerfunktion in einfacher Weise mit zusätzlichen Anforderungen erweitern läßt. Dazu eine kurze Beschreibung des Prinzips von BPtt[3]:
Wie bei BP, werden bei BPtt die einzustellenden Parameter $\underline{w}$ eines Ein/Ausgabe-Systems durch Anwendung des Gradientenverfahrens optimiert. Die Fehlerfunktion des Gesamt-Ausgabefehlers des Systems, im Raum dieser Parameter dargestellt, wird so iterativ zu minimieren versucht.
Die Parameter werden in einem iterativen Verfahren gemäß dem Gradienten der Fehlerfunktion E bezüglich dieser Parameter verändert (*Update*). Für einen Parameter w_i der Abbildungsfunktion ist dies

$$\Delta w_i = -\eta \frac{\partial E}{\partial w_i}$$

Der Faktor η ist der sogenannte Lernkoeffizient, mit dem die Stärke der Veränderungen gesteuert werden kann.
Gemäß der verallgemeinerten Kettenregel der Differentialrechnung setzt sich die Ableitung

[3] Eine detailliertere Herleitung kann z.B. [4], [5], [1] entnommen werden

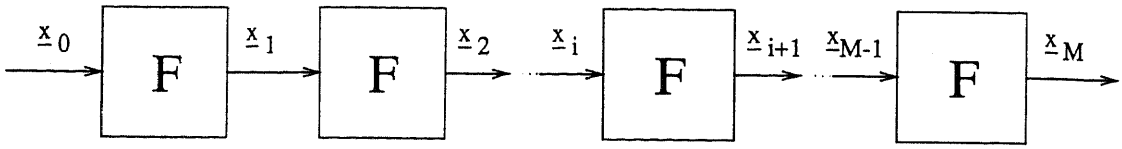


Abbildung 1.1: Über die Zeit abgerollte Funktionsiteration

der Fehlerfunktion nach dem Parameter w_i aus der direkten partiellen und den indirekten, über andere abhängige Parameter resultierenden Ableitungsketten zusammen. Mit $E(\underline{x})$ und $\underline{x}$, dem die Parameter $\underline{w}$ und die anderen Variablen $\underline{v}$ des Systems vereinenden Variablen-Vektor, kann dies geschrieben werden als

$$\frac{\partial E}{\partial w_i} = \sum_{j=1}^{N} \frac{\partial E}{\partial x_j} \frac{\partial x_j}{\partial w_i}$$

Die Berechnung des Updates muß nicht über den Gradienten der Gesamt-Fehlerfunktionswert erfolgen (Offline-Verfahren). Es kann auch ein Update aus der Bewertung jedes einzelnen Ein/Ausgabe-Beispiels berechnet werden (Online-Verfahren). Zudem ermöglicht das Prinzip der verallgemeinerten Kettenregel eine schrittweise Multiplikation und Summierung der Ableitungen der in der Abbildungsfunktion enthaltenen einzelnen Funktionen gemäß ihrer Anordnung in der Kette der Funktionen von w_i bis zu E (*Rückpropagierung*). BPtt ist nichts weiter als die Erweiterung dieses Verfahrens auf rückgekoppelte zeitdiskrete Funktionen. Die nach M Zeitschritten bewertete Funktionsausgabe setzt sich aus den über der Zeitachse abgerollten und aufeinanderfolgenden einzelnen Funktionsberechnungen zusammen (s. Abbildung 1.1). Die Menge der N Systemvariablen vergrößert sich so auf $M * N$, da durch die zeitliche Erweiterung der Berechnung jede Variable x_j nun durch M Instanzen x_{jt} in den unterschiedlichen Zeitschritten repräsentiert wird. Der resultierende Update eines Parameters ergibt sich dann als Summe der in den einzelnen Zeitschritten berechneten Updates dieses Parameters. Der Update darf also erst erfolgen, nachdem die Rückpropagierung bis zur letzten zu berücksichtigenden Parameterinstanz dieser Funktionskaskade durchgeführt ist.
Die obige Erwähnung der Bewertung der Funktionsausgabe nach dem letzten Zeitschritt (die Berechnung der Fehlerfunktion) suggeriert, daß nur die Ausgaben der Funktion und auch nur die des Zeitindexes t_M bewertet werden. Dies ist in der Tat auch der Ansatz, der aus der Herleitung von BP folgt. Es spricht jedoch nichts dagegen, die Fehlerfunktion in der Art zu erweitern, daß sie explizit interne Variablen der Abbildungsfunktion mitbewertet. Dies können dann auch Variablenwerte eines niedrigeren Zeitindexes sein. Man führt sozusagen diese in diesem Fall als eigene Variablen aufzufassenden Variableninstanzen zur Bewertung nach außen. Die Bewertung kann separat erfolgen oder in eine vorhandene mitintegriert werden, da mehrere Fehlerfunktionen sich durch einfache Addition zu einer zusammenfassen lassen.
Die schrittweise Fehlerpropagierung der Bewertung einer internen Variable x_{jt} beginnt erst in dem ihrem Zeitindex entsprechenden Abschnitt innerhalb der Funktionskaskade. Für die rechnerbasierte Realisierung des Verfahrens ist daher interessant, daß die Rückpropagierung einer separaten Bewertung durch einfache Addition der jeweiligen Gradienten an einer bestimmten Position der Funktionskette mit einer anderen Bewertung kombiniert werden kann. Dies kann aber nur dann geschehen, wenn die Rückpropagierung der beiden

Bewertungen an der gleichen Stelle der Kaskade gestoppt werden soll. Eine nachträgliche Extraktion des Gradienten einer der Bewertungen kann nämlich nach weiterer Rückpropagierung des kombinierten Wertes nicht mehr auf einfache Weise erfolgen. Unter diesem Aspekt ist eine Konsequenz, daß Bewertungen mit unterschiedlichen Rückpropagierungs-Endpunkten separat behandelt werden müssen und ihre Fehlerfunktionen aufgrund der anderen Bestimmung des Gradienten nicht zu einer zusammengefaßt werden dürfen. Anders ausgedrückt:

Zwei Fehlerfunktion können zu einer kombinierten Funktion $L3$

$$L3 = L1 + L2$$

addiert werden, wenn für alle w_i aus $\underline{w}$ gilt

$$\Delta w_i(L3) \equiv \Delta w_i(L1) + \Delta w_i(L2)$$

Mit

$$\Delta w_i(L_k) = \sum_{\check{t}_k}^{M} \Delta w_{it} \quad , \quad 1 \leq \check{t}_k \leq M$$

Dies ist unter der Vorraussetzung einer nicht möglichen nachträglichen Separierbarkeit des Gradienten zweier kombinierter Bewertungen L_1 und L_2 bei Rückpropagierungen mit unterschiedlichen Endzeitschritt-Indizes $\check{t}_1$, $\check{t}_2$ nicht gegeben.

Ein vorzeitiger Abbruch der Rückpropagierung von Randbedingungs-Bewertungen kommt dann ins Spiel, wenn versucht wird, die Lokalisationseigenschaft von Bewertungen in BPtt zu berücksichtigen. Die Multiplikation von Neuronenausgangs- und Gewichts-Wert $o_i w_{ik}$ der Verbindung zum nachfolgenden Neuron k hat zur Folge, daß der Gradient der Fehlerfunktion zum Gewicht proportional zum anliegenden Neuronenoutput ist:

$$\begin{aligned} I_k &= \sum_{i=1}^{n_k} w_{ik} o_i \\ \Delta w_{ik} &= -\eta \frac{\partial L}{\partial w_{ik}} = -\eta \frac{\partial L}{\partial I_k} \frac{\partial I_k}{\partial w_{ik}} = -\eta \frac{\partial L}{\partial I_k} o_i \end{aligned}$$

Dies bedeutet für einen bestimmten Eingabezustand des neuronalen Reglers (der Meßvektor, oft gleich dem Streckenzustand), daß die Gewichte, deren Source-Neuronen stärker aktiviert sind, eine um so stärkere Veränderung erfahren werden. Dies kann somit als eine bezüglich der Eingabe bereichsspezifische Parametrisierung interpretiert werden.
Um nun lokal wirkende Parameteranpassungen mit BPtt vornehmen zu können, sollte sich im Idealfall die Rückpropagierung nur über den dem Lokalisationscharakter der jeweiligen Bewertung entsprechenden Trajektorienbereich erstrecken. Da, wie oben schon erwähnt, eine anschließende Separierung des Rückpropagierungs-Fehlers der lokal begrenzten Bewertung vom kombinierten Fehlersignal nicht möglich ist, muß diese Bewertung dann in einem separaten Durchgang erfolgen.
Tatsächlich ist es aber so, daß eine Gewichtsänderung sich nicht nur auf das Reglerverhalten im abgegrenzt gewünschten Zustandsbereich auswirkt, sondern, aufgrund der funktionalen Verknüpfung aller beteiligten Netzparameter, auch Nebeneffekte im restlichen

Bereich verursacht. Letztendlich führt dies für alle an der Darstellung der Reglungsfunktion beteiligten Komponenten zu einem Kompromiß zwischen den unterschiedlichen Optimierungsanforderungen. Einbußen in der Güte der Gesamtregelungsfunktion durch Hinzufügen neuer Randbedingungen könnte man versuchen entgegenzuwirken, indem eine größere Zahl von Neuronen (=Funktionen) und gewichteter Verbindungen (=Parametern) bereitgestellt wird, um damit die erhöhte Komplexität der Gesamtregelung in Teilbereichen realisieren zu können. Damit kann vermieden werden, daß Ressourcen zur Optimierung der anderen Kriterien abgezogen werden müssen.
Der Versuch, die Regelung einer Randbedingung auf einen möglichst kleinen Zustandsbereich zu begrenzen, z.B. auf den durch die Grenze der Steuerbarkeit gegebenen, wie auf Seite 5 beschrieben, käme jedoch einer Implementierung von Schaltvorgängen zwischen grundsätzlich verschiedenen Funktionen des Reglers gleich. Die Approximation der dabei auftretenden Unstetigkeiten ist mit neuronalen Strukturen, wie in BP verwendet, mit erheblichem Aufwand verbunden und nur angenähert zu realisieren.
Anders dagegen bewirkt eine bewußte Aufweitung der Bewertungs-Lokalisation dann zwar auch eine stärkere Veränderung des globaleren Regelungsverhaltens, kann jedoch auch dazu führen, daß durch nur kleine Veränderungen in früheren Regelungsschritten die aktuelle Randbedingungs-Verletzung ganz vermieden wird. Für das auf Seite 5 verwendete Beispiel hieße das: Anstatt starke Abweichungen von der globalen Regelungsfunktion in Wandnähe in die Reglerfunktion einzubauen, könnte schon eine leichte Veränderung des globalen Reglerverhaltens in weiter entfernten Punkten eine gefährliche Annäherung an die Wand ganz verhindern. Dann wäre auch eine Implementierung der exakten Korrekturfunktion für die Randbedingungen in den neuronalen Regler unnötig und somit der Realisierungsaufwand insgesamt kleiner zu halten.
Bei der Wahl der Rückpropagierungsstrategie für eine Randbedingungs-Bewertung sollte daher neben ihrem eigenen Lokalisationscharakter auch berücksichtigt werden, inwieweit dies für das Lernen der Gesamtfunktion (primäres Regelungsziel und andere Randbedingungen) von Nachteil sein wird, sei es durch zu hohe Inanspruchnahme von Ressourcen (Parameterfunktionen) oder durch Seiteneffekte auf die anderen parallel auszuführenden Regelungsaufgaben. Das Resultat wird stets ein Kompromiß sein, der aber nach Vorüberlegungen über die Randbedingungs-Eigenschaften den unterschiedlichen Anforderungen an die Regelung möglicherweise besser angepaßt werden kann.
Im Interesse der Untersuchungen liegen deshalb folgende Problemstellungen:

- Ist ein paralleles Lernen mehrerer und unterschiedlicher Regelungsanforderungen in zufriedenstellender Weise möglich ?
- Inwieweit wird die primäre Zielregelung durch die zusätzlichen Regelungsaufgaben beeinträchtigt ?
- Kann dann eine Verschlechterung der Zielregelung durch bestimmte Gegenmaßnahmen ausgeglichen werden ?
- Ist die Lokalisationseigenschaft von Randbedingungen nachweisbar ?
- Welche Schwierigkeiten prinzipieller/spezieller Art treten bei einer Behandlung von lokalen Randbedingung auf ?

Einige weitere Eigenschaften allgemeinerer Art müssen bei der Realisierung eines neuronalen Reglers mittels BPtt zusätzlich beachtet werden und seien hier der Vollständigkeit halber angeführt:

- In der Regel wird man sich mit einem suboptimalen Ergebnis des Lernvorgangs begnügen müssen.
- Wegen der Verwendung von nur stetigen Funktionen im neuronalen Netz sind keine Unstetigkeiten in dessen Ausgabe möglich. Deshalb lassen sich zum Beispiel keine Randbedingungen des Stellsignals in die Abbildungsfunktion integrieren, die voneinander getrennte Zulässigkeitsbereiche beschreiben.
- Die Bewertungsfunktionen müssen differenzierbar sein.
- Es kann keine positive Bewertung des Regelergebnisses erfolgen. Die Bewertung wirkt immer negativ. Das hat zur Folge, daß eine Ausrichtung der Reglerparameter auf das Optimierungsziel nur durch Bewertung der Fehler und nicht durch ein direktes Bestärken eines erfolgreichen Verhaltens erfolgen kann.
- Das Ergebnis des Lernvorgangs ist nicht hundertprozentig verifizierbar. Eine Garantie kann für die Einhaltung von Randbedingungen nur für die in nachfolgenden Tests überprüften Regelungs-Situationen gegeben werden. Ein Verzicht auf die Korrekturmodule zur Vermeidung von Randbedingungen wird deshalb in der Regel nicht möglich sein. Allerdings können starke Stellbewegungen aufgrund spontan eingreifender Korrekturen durch die Integration der Randbedingungen in die Reglerfunktion gedämpft werden, mit Vorteilen durch
  - eine geringere Belastung der Stellapparats sowie der Regelstrecke, mit Auswirkungen auf die konstruktionelle Auslegung dieser Systeme,
  - eine Vermeidung unerwünschter Sprungantworten der Regelstrecke,
  - eine Verminderung des Energieaufwands des Stellapparats, der sich sonst bei schnelleren Regelkorrekturen typischerweise erhöhen würde.

Kapitel 2

Beispiel TBU

Um das Vorgehen für die Einbeziehung von Randbedingungen ins Reglerlernen an einem konkreten Beispiel zu zeigen, soll hier zuerst das aus [4] bekannte und hier leicht modifizierte TBU-Modell[1] in kurzer Form vorgestellt werden. Eine Beschreibung der TBU-eigenen Systemrandbedingungen, möglicher Zusatzkriterien und ihrer besonderen Eigenschaften schließt sich an.

2.1 TBU, Modell und Regelung

Das zu regelnde System stellt eine 2-achsige Zugmaschine mit 1-achsigem Anhänger (feste Achse) dar, s. Abbildung 2.1. Der Antrieb befindet sich an der auf ein zentrales Rad reduzierten Vorderachse. Je nach Auflagepunkt der Anhängerdeichsel vor oder hinter der festen Hinterachse der Zugmaschine wird das Fahrverhalten eines PKWs mit Anhänger oder LKW-Aufliegers eingestellt. Zur Vereinfachung der Regelungsfunktion (1 Ausgabesignal) wird eine konstante Antriebsgeschwindigkeit angenommen. Der Regler hat somit nur den Wert der Lenkradstellung zu liefern.
Die Regelungsaufgabe besteht nun darin, das Fahrzeug aus einer Startposition eines vor-

[1] für eine ausführliche mathematische Systembeschreibung s. [1].

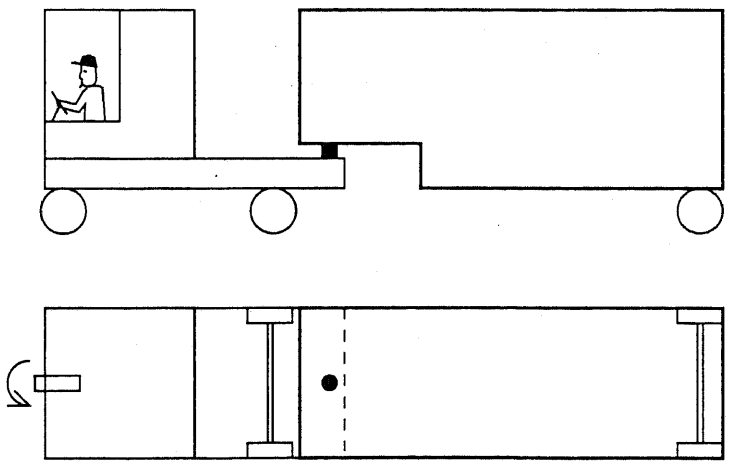


Abbildung 2.1: Schematischer Aufbau des Truck Backer-Upper-Fahrzeugs

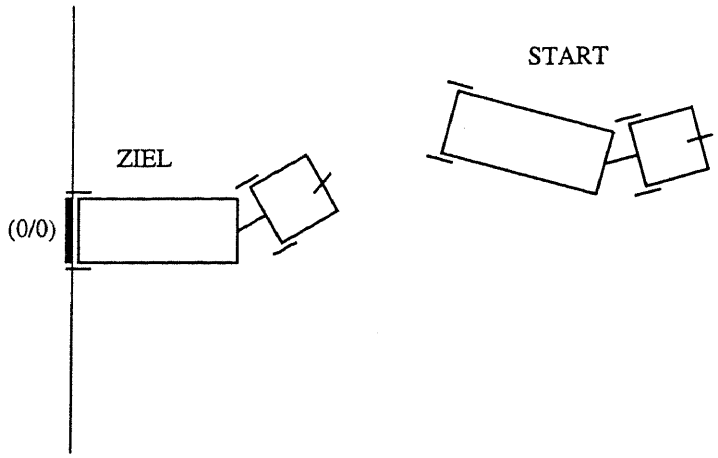


Abbildung 2.2: Regelung des TBU aus Startposition in Zielzustand

gegebenen Startbereichs in eine Zielposition zu bringen (Abbildung 2.2), die durch eine Andockposition der Anhängerrückwand an eine feststehende Wand im Punkt (0/0) gegeben ist. Deren Koordinatenwahl ist ohne Beschränkung der Allgemeinheit festgelegt, da jede andere Andockstellung durch Koordinatentransformation in den Nullpunkt überführt werden kann.
Durch die Wahl einer negativen Geschwindigkeit im Antriebspunkt bekommen die Systemgleichungen einen stark nichtlinearen Charakter. Die erfolgreiche Regelung dieser Rückwärtsfahrt in [4] und [1] zeigt daher besonders gut die Eignung neuronaler Netze zur nichtlinearen Regelung. Im Anhang A sind die Systemgleichungen, wie sie hier verwendet wurden, nochmals angeführt.
Wichtig für die erfolgreiche Reglerentwicklung mittels BPtt ist der Aufbau des sogenannten Lernplans, über den die Auswahl der Lernbeispiele aus dem Startbereich erfolgt. Zu Beginn des Lernens, wenn das Reglernetz noch untrainiert ist, müssen sehr einfache, nah bei der Zielposition gelegene Startzustände vorgegeben werden. Durch ein noch unentwickeltes Reglerverhalten würde sonst die Bandbreite der Fehler in den Endpunkten der Regelungen uneinschätzbar hoch und ein erfolgreiches Einschwingen des neuronalen Netzes in Richtung eines vernünftigen Regelverhaltens erschweren (der Abbruch eines Regelungsdurchgangs erfolgt, wenn das Fahrzeug über die begrenzende Wand gelangt oder die Zahl der Regelungsschritte ein Maximum überschreitet). Durch etappenweises oder kontinuierliches Erweitern des zu lernenden Startbereichs kann die Reglerfunktion schrittweise entwickelt werden.
Als Eingabe bekommt der Regler[2] den aktuellen Zustand der Strecke $\underline{x}$. An die Reglerausgabe schließt sich dann die Berechnung des aus Lenkwinkel θ und $\underline{x}$ resultierenden neuen Zustands $\underline{x}'$ an. Infolge der Zeitdiskretisierung der Regelung wird das aus dem alten Streckenzustand vom Regler berechnete Stellsignal während der folgenden Strecken-Simulationsphase konstant gehalten. Die in Abbildung 2.3 dargestellte Regelschleife läßt sich so entsprechend der in Kapitel 1.1 auf Seite 6 formulierten Weise als Funktionskaskade zur Verwendung mit BPtt auffassen.
Die Bewertung der primär wichtigen Regelung in den Sollzustand setzt an dem durch das Auftreffen auf die Wandbegrenzung gegebenen Endzustand der Regelstrecke an. Die dort

[2] In der Grundkonfiguration gemäß der in [1] durchgeführten Untersuchung

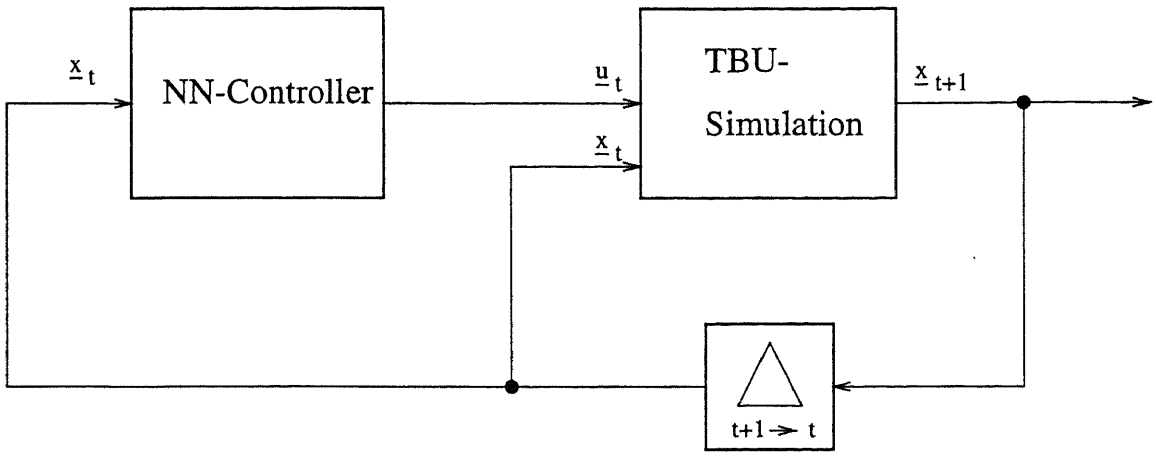


Abbildung 2.3: Regelungsschleife mit NN und TBU-Simulation

durch die Verletzung der Randbedingung

$$||x_{Soll} - x_{Ist}|| \stackrel{!}{=} Min.$$

vorgenommene Bewertung des Positionierungsfehlers wird als globale Bewertung aufgefaßt und durch die gesamte Regelungstrajektorie zurückpropagiert[3]. Zwischen die Module Regler und Strecke lassen sich andere Module einfügen, wie z.B. Koordinatentransformationen und Stellsignal-Korrekturfunktionen. Zur Bewertung von Systemzuständen kann man Variablen aus der Funktionskette abgreifen und daran spezielle Bewertungsmodule anfügen. Dies ist auf jeden Fall am Endpunkt der Kaskade der Fall, wo der dort auftretende Istfehler beurteilt wird. Genauso aber können in jedem Abschnitt Bewertungen einer oder mehrerer Variablen der Funktionenkette ansetzen, um sie als separate Bewertung im BPtt-Verfahren der Fehlerrückpropagierung zu behandeln. Um die Möglichkeiten der zusätzlichen Bewertungen zu betrachten, werden nun zuerst die im System TBU auftretenden Randbedingungen und einige denkbare Zusatzkriterien für die Regelung genauer beschrieben.

2.2 Systemimmanente Randbedingungen

2.2.1 Lenkwinkelbegrenzung

Die durch die Lenkmechanik bedingte Begrenzung des Lenkwinkels θ des Vorderrads ist in der Regel als symmetrisch um den Winkel 0 anzunehmen. Allgemein kann man für die Einhaltung des erlaubten Lenkwinkelbereichs jedoch ansetzen

$$\check{\theta}_G \quad <= \quad \theta \quad <= \quad \hat{\theta}_G$$

Falls, wie hier vorausgesetzt, die Grenzen $\check{\theta}_G$ und $\hat{\theta}_G$ konstant sind (unabhängig vom Systemzustand und vorausgegangenen Stellsignalwerten), handelt es sich um eine generell zu

[3]Man kann sich überlegen, daß diese Charakterisierung nicht immer zutreffend sein muß. Bei sehr langen Regelungen wird es sicher weniger bedeutend für die Zielregelung sein, wie das Regelverhalten in der Startposition durch Zielfehler beeinflußt wird. Und umgekehrt werden lokale Regelungsanforderungen für die frühe Regelungsphase bedeutender sein als die grobe Ausrichtung des dortigen Regelverhaltens auf das Ziel.

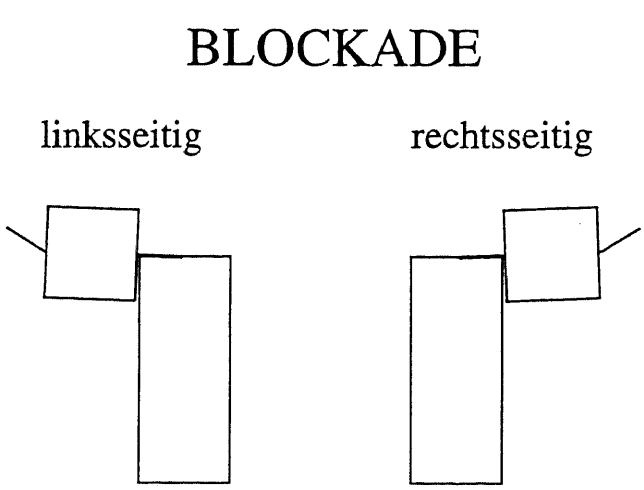


Abbildung 2.4: Blockadestellungen des TBU-Systems

erlernende Regelungsaufgabe, die unabhängig von anderen Variablen in die Reglerfunktion miteingebaut werden muß. Diese Randbedingung hat somit einen absolut lokal begrenzten Einfluß auf die Reglerparametrierung, da benachbarte Regelungsvorgänge nicht ursächlich für eine auftretende Randbedingungs-Verletzung sind.
Eine Verletzung dieser Randbedingung kann durch eine nachfolgende Anpassung der Reglerausgabe an den erlaubten Bereich vermieden werden. Ein Abbruch der Regelung bei Bereichsüberschreitung ist deshalb leicht zu vermeiden. Die Anpassung erfolgt im realen System eines Fahrzeugs bereits durch den mechanischen Anschlag an den Lenkwinkelgrenzen. Da dies aber zu einer starken mechanischen Beanspruchung des Stellapparats führen kann, ist es sinnvoller, die erforderliche Korrektur bereits am numerischen Reglersignal vorzunehmen. Dies mag, um eine differenzierbare Kennlinie zu verwenden, bereits in Form einer Sigmoidfunktion am Ausgabeneuron des Reglernetzes erfolgen. Näheres hierzu in Kapitel 3.5.1.

2.2.2 TBU-Selbstblockade

In der Nähe eines bestimmten Grenzwinkels Φ zwischen Zugmaschine und Anhänger und bei einem diese Differenz noch tendenziell vergrößernden Lenkwinkel kommt es zwangsläufig zur Blockade der Fahrbewegung (s. Abbildung 2.4). Diese Blockade kann beidseitig auftreten und je nach Fahrzeuggeometrie mit Winkeln aus bestimmten Lenkwinkelbereichen wieder gelöst werden. Als Randbedingung für Φ formuliert gilt ähnlich wie bei der Lenkwinkelbegrenzung folgendes:

$$\check{\Phi}_G \quad <= \quad \Phi \quad <= \quad \hat{\Phi}_G$$

Dabei sind $\check{\Phi}_G$ und $\hat{\Phi}_G$ die Grenzwinkel der beiden Blockadefälle, bei denen eine Blockade gerade noch nicht auftritt (Berührung). Um eine Blockaderegelung aus einem blockadenahen Zustand mittels eines Korrekturmoduls zu vermeiden, muß aber präventiv über das Stellsignal eingegriffen werden, so daß die Randbedingung für Φ in eine des Lenkwinkels transformiert werden muß. Für weite Bereiche von Φ gibt es keine Beschränkung für θ, da aus diesen Zuständen keine Blockade angefahren werden kann. Daher muß ein Weg

gefunden werden, die nur für eine unmittelbare Blockaderegelung gefährlichen Bereiche von Φ zu ermitteln und für diese den jeweiligen Lenkwinkelbereich zu bestimmen, mit dem die drohende Blockade umfahren wird.
Für den Spezialfall, daß der Blockadegrenzwinkel $\Phi_G \in \{\check{\Phi}_G, \hat{\Phi}_G\}$ gerade konstant bleibt, wurde in [1] bereits eine implizite Lösung für den dafür benötigten Lenkwinkel θ_G gefunden. Diese kann bei gegebenem Φ_G entweder numerisch gelöst werden oder für den Spezialfall $\Phi_G \in \{-\pi/2, +\pi/2\}$ sogar exakt angegeben werden. Um für blockadenahe Bereiche den Lenkwinkel θ_g zu berechnen, der gerade noch nicht in eine Blockade lenkt, müßte dieser jedesmal mittels eines Iterationsverfahrens aus den Systemgleichungen der Strecke numerisch ermittelt werden[4]. Um diesen Berechnungsaufwand zu vermeiden, kann man auf zwei alternative Näherungsmethoden ausweichen:

1. Man schätzt den nachfolgenden Zustand und setzt bei Detektion einer Blockade den vorher errechneten Winkel θ_G des Grenzfalls $\Phi = \Phi_G$ als sicheren Ausweichwert ein.

2. Man wählt einen Winkel Φ_B nahe Φ_G, aus dem heraus aber noch nicht in die Blockade gelenkt werden kann und setzt für alle Winkel Φ zwischen Φ_B und Φ_G den Korrekturwert θ_G anstelle des Reglersignals.

In der Diplomarbeit fand aus Genauigkeitsgründen die erste Methode Verwendung, indem als Schätzfunktion die Simulationsgleichungen der Strecke verwendet wurden.
Die Randbedingung für das Stellsignal im in Abbildung 2.4 dargestellten Fall bei Rückwärtsfahrt lautet damit also

$$\begin{array}{lcccccl} \Phi'(\theta, \underline{x}) > \hat{\Phi}_G & \Rightarrow & \theta_G(\hat{\Phi}_G) + \pi & <= & \theta & <= & \theta_G(\hat{\Phi}_G) \\ \Phi'(\theta, \underline{x}) < \check{\Phi}_G & \Rightarrow & \theta_G(\check{\Phi}_G) & <= & \theta & <= & \theta_G(\check{\Phi}_G) - \pi \end{array}$$

mit $\Phi'(\theta, \underline{x})$ der Winkeldifferenz Φ im vorrausgeschätzten nächsten Zustand.
Das Fahrzeug ist somit in der gezeigten Geometrie immer aus dem Gefahrenbereich der Blockade steuerbar. Die Annäherung an die Blockadegrenze Φ_G wird jedoch durch vorherige Reglerausgaben verursacht. Was die Beurteilung des Lokalisationscharakters der Randbedingung betrifft, kann man in diesem Fall von einem weiter ausgedehnten Einflußbereich ausgehen, was sich letztendlich auch auf die zu wählende Länge der Bewertungsrückpropagierung auswirken mag. Denn ein Lernen der genauen Blockade-Korrekturfunktion würde die Regelungsfunktion in diesem Bereich stark verändern. Eine Vermeidung der Blockadesituation durch frühzeitiger ansetzende und deshalb kleinere Modifikationen an der Regelungsfunktion wäre dann im Hinblick auf eine weniger komplexe Regelungsfunktion eventuell von Vorteil.

2.3 Weitere Optimierungskriterien

Zur besseren Anpassung des Reglers an ein gewünschtes Verhalten oder bezüglich bestimmter Optimalitätsanforderungen kann versucht werden, mittels zusätzlicher Kriterien in Form von Bewertungsfunktionen (Straffunktionen) das Lernen geeignet zu beeinflussen. Es folgt nun eine Auswahl von für das TBU-System in Erwägung zu ziehenden Formulierungen.

[4] Jedoch ist auch eine existierende explizite Funktion vorstellbar, die die Beziehung $\theta_g = f(\theta, \Phi)$ exakt wiedergibt oder diese auf der sicheren Seite in guter Näherung wiedergibt. Jedoch konnte dieser Weg aus Gründen des Aufwandes nicht weiterverfolgt werden.

2.3.1 Reduktion der Lenkwinkeldifferenz

In einem mit konstanter Geschwindigkeit fahrenden Fahrzeug führen zu schnelle Lenkbewegungen zur Instabilität der Bewegung. Deshalb scheint es sinnvoll in dem hier vorhandenen zeitdiskreten Regelsystem die Differenz aufeinanderfolgender Lenkwinkel (=Reglerausgaben) nicht zu groß werden zu lassen, um eine möglichst realistische Simulation zu ermöglichen.
Die Absolut-Differenz $\Delta\theta = |\theta_{i+1} - \theta_i|$ sollte daher auf einen bestimmten Bereich begrenzt werden. Die Randbedingung lautet so

$$0^\circ \quad <= \quad \Delta\theta \quad <= \quad (\Delta\hat{\theta})_G$$

bzw. für die Reglerausgabe θ_{i+1}

$$\theta_i - (\Delta\hat{\theta})_G \quad <= \quad \theta_{i+1} \quad <= \quad \theta_i + (\Delta\hat{\theta})_G$$

und ist somit abhängig von dem vorherigen Stellsignal.
Damit der Regler diese Randbedingung überhaupt erlernen kann, muß ihm daher der vorherige Stellsignalwert als Eingabe präsentiert werden. Dabei stellt sich die Frage, ob dies im Fall einer vorangegangenen Stellsignal-Korrektur der Originalwert oder der korrigierte Wert sein soll. Eines der Ziele der Reglergenerierung ist die Vermeidung von nachfolgenden Korrekturen. Daher macht es wenig Sinn, den Regler bezüglich dieser während der Lernphase auftretenden Korrekturen, die zudem noch in wechselnder Form auftreten werden, auf eine Verminderung der Lenkwinkeldifferenz einzustellen. Es empfiehlt sich daher, die Abhängigkeit der Reglerausgabe von der vorherigen Originalausgabe bereits früh als Eigenschaft der Regelung zu erlernen, was zu weniger Störungen in der Entwicklung der Gesamtregelfunktion führen sollte. Verletzungen dieser Randbedingung in der Lernphase sind entweder zu tolerieren oder mittels einer zusätzlichen Korrekturfunktion zu verhindern.

2.3.2 Orientierung an der Horizontalen

Für die Zielpositionierung des TBU ist es von Vorteil, wenn das Fahrzeug bereits frühzeitig in die Horizontale der Andockposition einlenkt. Bei schrägerer Anfahrt wird die Schluß-Regelung wesentlicher aufwendiger. Ohne erst eine Reihe möglicher, aber komplizierter Formulierungen von Randbedingungen für diese Anforderung anzugeben, sei auf eine bei der Herleitung der Ausgangs-Lernkonfiguration gefundene Formulierung verwiesen (vgl. Kapitel 4.1.4). Ein bei der Bewertung der Endposition vorgetäuschter Positionierungsfehler in x-Richtung dergestalt, daß das Anhängerende sich anscheinend hinter der Wand befindet, führt bei globaler Rückpropagierung durch die Regelungs-Trajektorie zu einer vorteilhaften Korrektur der Netzparameter. Diese wird derart sein, daß in Zukunft der Anhänger weiter von der Wand weg positioniert werden wird um den angeblichen Endpunktfehler auszugleichen. Da die Abbruchbedingung der Ziel-Regelung aber gerade die Kollision mit der Wand ist (die dann auch weiterhin eintreten wird), wird der Versuch des Reglers einer Positionierung vor die Wand nur dazu führen, die Einlenkphase in die Zielposition weiter von der Wand weg zu verschieben. Somit kann dann mittels dieser globalen Bewertung das gewünschte Verhalten erreicht werden. Die Formulierung dieser Randbedingung ist

$$\tilde{x}_3 \stackrel{!}{=} 0$$

mit $\tilde{x}_3$ als dem mit einem festen Fehlerbetrag modifizierten Wert der x-Position x_3 des Anhängers. Die daraus berechnete Bewertung der Fehlpositionierung wird wie die globale Endpunktbewertung behandelt, da sie sich ebenfalls auf die gesamte Trajektorie auswirken soll. Der auftretende Fehler läßt sich dann auch zum Endpositionierfehler hinzuaddieren und so als Fehler eines kombinierten Kriteriums rückpropagieren.

Kapitel 3

Aspekte der Lernintegration lokaler Randbedingungen

In Abbildung 3.1 ist eine vereinfachte Modul-Grundstruktur für die verwendeten Methoden zum parallelen Erlernen mehrerer Randbedingungen dargestellt. Folgende Fakten sind nicht mit in der Abbildung berücksichtigt:

- Es ist nur der Informationsfluß der Vorwärtspropagierungs-Phase dargestellt (der eigentlichen Regelung). Die Pfade für die Rückpropagierungs-Phase ergeben sich durch einfache Rückverfolgung der über die Zeitachse abgewickelten Funktionenkette.

- Prinzipiell können jedem Modul alle im Gesamtsystem vorhandenen Variablen zugeführt werden. Abgebildet sind nur die Eingänge, die speziell für das betreffende Modul von Bedeutung sind.

- Die im Gesamt-Korrekturmodul beinhalteten einzelnen Korrekturen sind in der Regel solche von limitierenden Randbedingungen, deren Einhaltung unbedingt gewährleistet sein muß. Es können selbstverständlich auch weitere Korrekturen definiert werden.

- Jedes Bewertungsmodul ist separat zu behandeln und mit eigenen Eingängen und Funktionen versehen. Die Gruppierung in Bewertungen der limitierenden und forcierenden Randbedingungen dient in der Abbildung nur der Übersicht. In Wahrheit gliedern sich diese in die einzelnen Bewertungen aller Randbedingungen auf.

- Korrekturmodule und die sich darauf beziehenden Bewertungsmodule sind hier getrennt dargestellt. In Wahrheit sind Randbedingungs-Korrektur und -Bewertung eng miteinander verknüpft, was eine kompaktere Form in der Realisierung nahelegt.

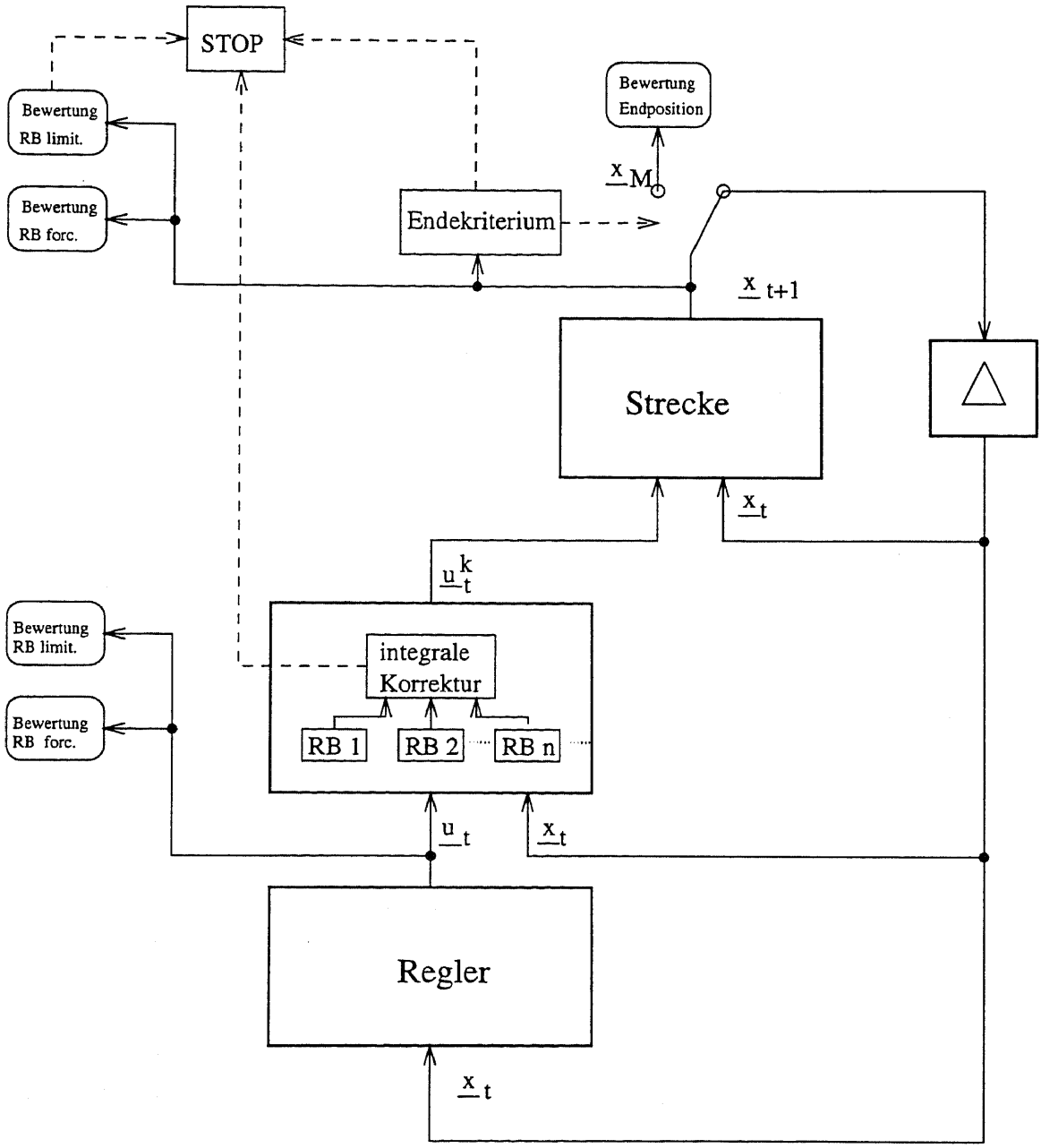


Abbildung 3.1: Modulare Struktur der RB-Bewertung

- Das Ende-Kriterium bestimmt den Abbruch-Zeitpunkt der Regelung, um aus diesem Endzustand dann die Bewertung der Zielpositionierung einzuleiten. Das Stoppmodul bricht ebenfalls die Regelung ab, jedoch nicht aus dem Grund der Zielbewertung, sondern wenn eine nicht zu korrigierende Verletzung einer limitierenden Randbedingung auftrat. Dies ist bei vom Streckenzustand abhängigen limitierenden Randbedingungen stets bei Verletzung der Fall; bei rein vom Stellsignal abhängigen Randbedingungen jedoch nur, wenn sich die Korrekturen der separaten Korrekturmodule nicht miteinander vereinbaren lassen.
- Für den Beginn eines Regelungsdurchgangs wird ein definierter Startzustand aller Variablen mit dem Zeitindex 0 angenommen.

3.1 Notwendigkeit von Korrekturmodulen

Wie schon erwähnt kann für den trainierten neuronalen Regler keine Garantie für eine absolut korrekte Regelungsfunktion gegeben werden. Daneben existieren aber auch lerntechnische Gründe, die die Notwendigkeit der Bereitstellung von Korrekturmodulen begründen. Bei einem Verzicht auf die Korrekturmodule für das Lenksignal wäre ein Stopp notwendig, sobald eine Verletzung einer limitierenden Randbedingung auftritt, um keine unrealistischen Regelungen zu simulieren. Bei Vorhandensein mehrerer Randbedingungen (wozu ja auch die Minimierung des Zielfehlers gehört) ist ein Stopp nach dem ersten Auftreten einer solchen Randbedingungs-Verletzung jedoch wenig sinnvoll. Dies hätte bei regelmäßigen Verletzungen viele unvollständige Lernvorgänge zur Folge mit folgenden Konsequenzen:

- Das primäre Regelungsziel wird nur ungenügend erlernt. Das BPtt-Lernen setzt eine gleichmäßige, mit der Verbesserung der Regelungseigenschaften schritthaltende Steigerung der Regelungsproblematik voraus. Die Fehler, die bei der Bewertung der Endposition auftreten, sollen dadurch in einem abschätzbaren, begrenzten Rahmen gehalten werden, um die Einstellung der übrigen Lernparameter hinsichtlich einer guten Performanz des trainierten Reglers zu erleichtern. Durch die Bewertung der Istposition in einem zu weit von der Sollposition entfernten Punkt würde ein für die Optimierung der Gewichte zu großer Fehler rückpropagiert werden.
- Der Lernvorgang für sämtliche zu integrierende Regelungsfunktionen verlängert sich erheblich, da immer nur eine Bewertung einer Randbedingung stattfindet.
- Das Lernen scheitert unter Umständen, wenn ein bestimmtes Regelungsverhalten nicht antrainiert werden kann, das zur Einhaltung einer Randbedingung notwendig wäre und deshalb diese fortwährend verletzt würde.

Der Einbau von Korrekturmodulen in die Funktionskette ist deshalb empfehlenswert, um

a) während einer Regelung in einen Endzustand mehrere Bewertungen verschiedener Randbedingungen durchführen zu können,

b) unabhängig von einem Mißlingen des Erlernens einzelner Randbedingungen zu sein.

3.2 Integrale/separate Korrektur

Betrachtet man jede limitierende Randbedingung für sich, so besitzt sie einen eigenen Zulässigkeitsbereich, auf den ein inkorrektes Stellsignal des Reglers korrigiert werden muß. Diese Korrektur läßt sich im einfachsten Fall als Begrenzungskennlinie, wie in Abbildung 3.2 dargestellt, formulieren[1].
Bei mehreren aktiven Randbedingungen treten aufgrund unterschiedlicher Zulässigkeitsbereiche Konflikte auf, wenn der Korrekturwert einer Randbedingung im verbotenen Bereich einer anderen liegt. Durch eine UND-Verknüpfung der Zulässigkeitsbereiche vorab wird ein gemeinsamer Bereich definiert, mit dem, unter Berücksichtigung aller Randbedingungen, ein konfliktfreier, integraler Korrekturwert ermittelt werden kann (Abbildung 3.3).

[1] Es wird hier von einzelnen, beidseitig begrenzten Zulässigkeitsbereichen ausgegangen. Bei einem unterteilten Bereich würden dazwischenliegende Werte auf die jeweils näheste Segmentgrenze begrenzt. Ebenso bei zwei Halbebenen bildenden, beidseitig begrenzten Ausschlußbereichen.

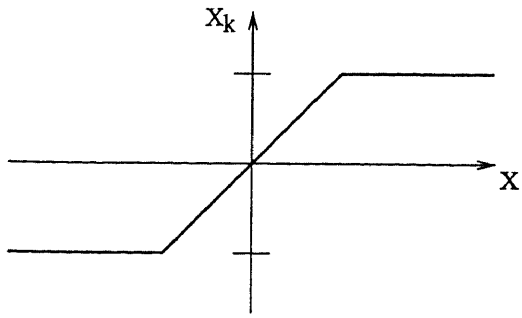


Abbildung 3.2: Begrenzungsfunktion

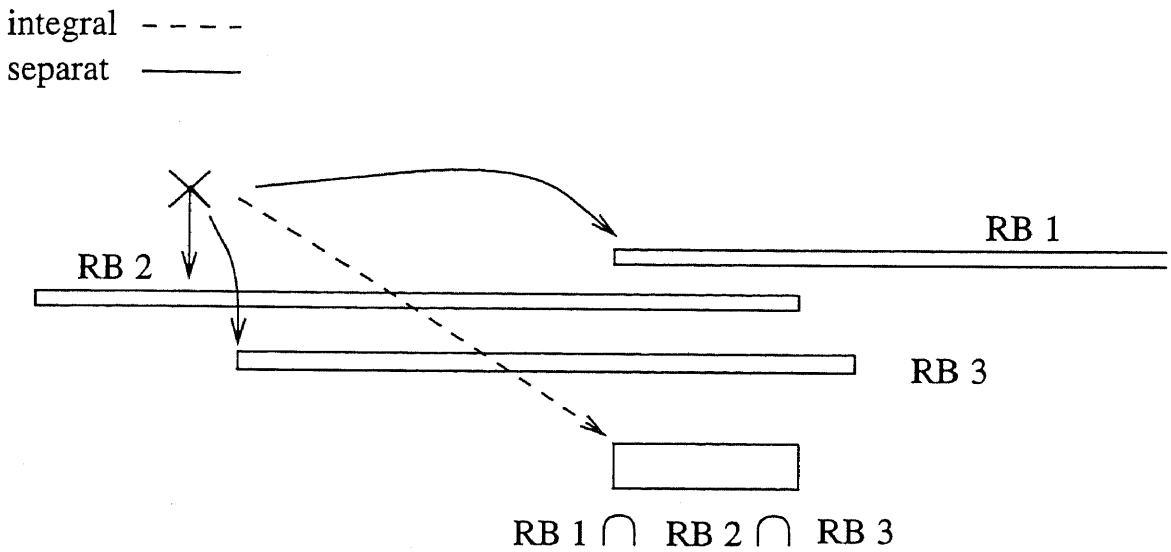


Abbildung 3.3: Unterschiedliche Korrekturwerte bei separater und integraler Korrektur

Resultiert als Schnittmenge die leere Menge, so besteht ein nicht zu behebender Widerspruch zwischen den Randbedingungen und die Regelung muß an dieser Stelle abgebrochen werden.
Bei einer notwendigen integralen Korrektur aufgrund mehrerer Verletzungen von Randbedingungen stellt sich die Frage, wie diese bewertet werden soll. Es bestehen zwei Alternativen:

Integrale Bewertung: Die Bewertung des integralen Korrekturwertes mit einer auf dessen Zulässigkeitsbereich abgestimmten Bewertungsfunktion.

Separate Bewertung: Die Bewertung aller durch die Reglerausgabe direkt verletzten Randbedingungen mit separaten Bewertungsfunktionen.

Der Vorteil einer integralen gegenüber einer separaten Bewertung ist darin zu sehen, daß die integrale Korrektur eine eindeutige Abbildungsfunktion $\theta_k = k(\theta, \underline{x})$ darstellt, die prinzipiell mit neuronalen Netzen approximierbar ist. Gegensätzliche Fehlerwerte, wie sie beispielsweise bei separater Bewertung zweier unterschiedlich korrigierender Module auftreten können, werden hierbei vermieden.
Als Nachteil gegenüber einer separaten Bewertung erkauft man sich aber, daß dann keine

Berücksichtigung der unterschiedlichen Bewertungseigenschaften der Randbedingungen mehr möglich ist (vgl. Kapitel 3.6 Seite 30). Mit der separaten Bewertung kann dagegen

- eine differenzierte Gewichtung der Bewertungen erfolgen und so die Stärke der Integration der verschiedenen Randbedingungen in die Regelungsfunktion gesteuert werden,
- der Lokalisationscharakter der Randbedingungen durch unterschiedliche Längen der Rückpropagierungen berücksichtigt werden,
- je Randbedingung die Bewertungscharakteristik unabhängig von ihrer Korrektur gestaltet werden, indem etwa für die Berechnung des Fehlers andere Sollwerte als die Bereichsgrenzen definiert werden. Bei der integralen Bewertung wird der Bereich der Bewertung abhängig von der Schnittmengenbildung aller erlaubten Bereiche und kann nicht mehr auf eine bestimmte Randbedingung hin abgestimmt werden.

Je nach Übereinstimmung dieser Randbedingungs-Eigenschaften wird man sich für eines der beiden Konzepte entscheiden. Wenn, wie in Abbildung 3.3, die Randbedingungen einen nur sehr kleinen gemeinsamen Gültigkeitsbereich aufweisen, werden deren separate Bewertungen aufgrund der häufig nicht harmonisierenden separaten Korrekturwerte zu keiner zufriedenstellenden Approximation der Randbedingungen führen. Genauso wird auch mit der integralen Bewertung von Randbedingungen keine optimale Integration der Randbedingungen in die Regelungsfunktion gelingen, wenn sie in puncto Lokalisationscharakter oder Fehlerskalierungs-Bereich zu verschieden sind.

3.3 Rückpropagierung durch Korrekturfunktionen

Die Rückpropagierung von Fehlerwerten durch eine Funktionenkette gemäß dem BP-Verfahren setzt voraus, daß alle beteiligten Funktionen differenzierbar sind. Nun gibt es aber speziell bei Korrekturfunktionen, wie den Begrenzungsfunktionen, Differenzierbarkeitslücken. Bei den stetigen Begrenzungsfunktionen kann diese Lücke durch eine der beiden seitlichen Ableitungen oder z.B. ihre Mittelwert ersetzt werden. Man kann sich die Begrenzungsfunktion ja auch als im Knickbereich gerundet vorstellen, wie in Abbildung 3.4 dargestellt, und so eine prinzipielle Differenzierbarkeit annehmen, die aufgrund der infinitesimalen Ausdehnung aber überhaupt nicht in die Berechnungen miteinfließt.

Sprungfunktionen, wie die in Abbildung 3.5 gezeigte, besitzen aufgrund ihrer Unstetigkeit eine entscheidende problematische Eigenschaft. Die Sprungstelllen lassen sich nicht als Grenzfall einer stetig differenzierbaren Funktion zur Nutzung mit BP betrachten, was an einem beispielhaften Vergleich gezeigt werden soll. Dazu aber vorher eine Überlegung, was für eine Bedeutung ein zu einer bestimmten Stelle innerhalb der Regelungskaskade zurückpropagierter Fehler für die dortige Teilfunktion hat:

Der zu einer inmitten der Regelungskaskade gelegenen Variable zurückpropagierte Fehler der Zielpositionierung zeigt an, wie der Variablenwert in Richtung auf eine Verbesserung der Zielregelung zu verändern ist[2]. Der vorliegende Fehler kann wie ein aus einer imaginären Bewertung bzgl. der Variablen abgeleiteter Fehler interpretiert werden (s. Abbildung 3.6) und separat durch

[2]Diese Veränderung findet letztendlich durch die Parametrierung der am Ende der Ableitungskette gelegenen Gewichtsvariablen statt

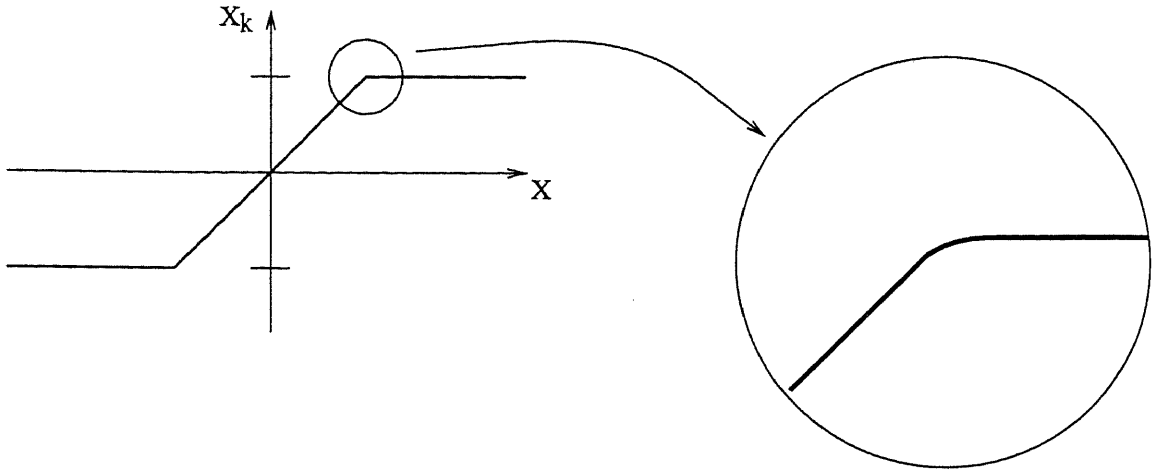


Abbildung 3.4: Differenzierbarer Knickbereich einer Begrenzungsfunktion

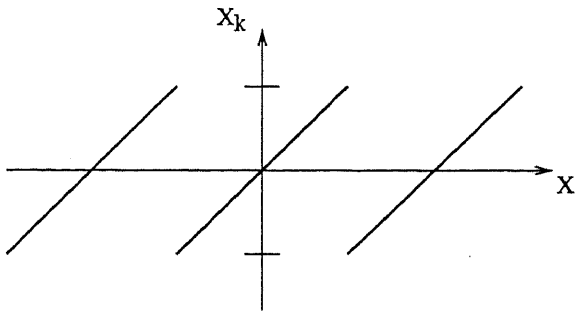


Abbildung 3.5: Sprungfunktion

die restliche Funktionenkette weiterpropagiert werden. Diese imaginäre Bewertung ist dergestalt, daß sie die Wirkung des Variablenwerts auf die Zielregelung in gleicher Weise berücksichtigt, wie dies durch die Rückpropagierung der Endpositions-Bewertung der Fall ist. Sie wird durch die komplette, von der betrachteten Variablen bis zur Endpositionsbewertung nachfolgenden Funktionenkette repräsentiert.

Wir betrachten nun einmal eine stetige, im Punkt $u = 0$ nicht differenzierbare Funktion $y_s(u)$ und eine sägezahnartige im Punkt $u = 0$ unstetige Funktion $y_{us}(u)$ (Abbildung 3.7). Beide Funktionen sollen eine Korrektur für die Reglerausgabe u beschreiben. Der zu y_s und y_{us} zurückpropagierte Fehler e_y gibt dann an, in welche Richtung der Funktionswert geändert werden müßte, um das bewertete Kriterium besser zu erfüllen. Nun nehmen wir für den gesamten Bereich von u einen positiven resultierenden Fehler e_y an, der ausdrückt, daß der Funktionswert unter einem (unbekannten) einzustellenden Sollwert y_{soll} ist[3]. Dann bewirkt die Rückpropagierung dieses Fehlers zu den zu verändernden Funktionsparametern (den Gewichten des neuronalen Netzes) eine tendenzielle Veränderung der Parameterwerte dahin, daß die Funktionen y_s und y_{us} größerer Ausgabewerte liefern werden. Dies bedeutet für die stetige Funktion ein Einpendeln um den Punkt 0, sofern die Gewichtsänderungen nicht zu groß ausfallen, was aber durch eine geeignete Wahl der Lernkoeffizienten vermieden sein sollte. Für die unstetige Funktion jedoch be-

[3] Was nicht unrealistisch ist, wenn y proportional zu einem Energieaufwand ist, der gegenüber dem optimalen Wert möglicherweise stets zu klein sein wird

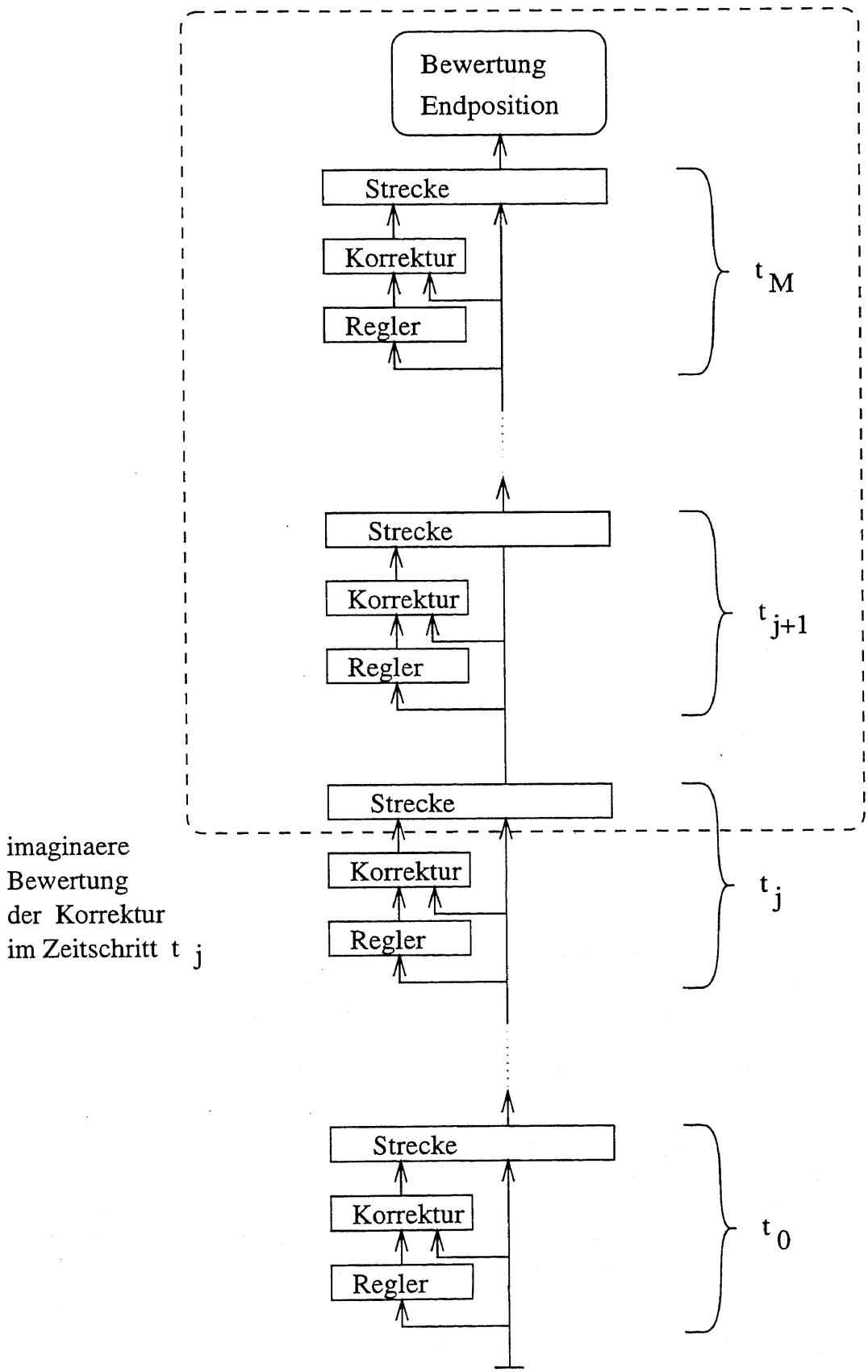


Abbildung 3.6: Imaginäre Bewertung in Zeitschritt t_j

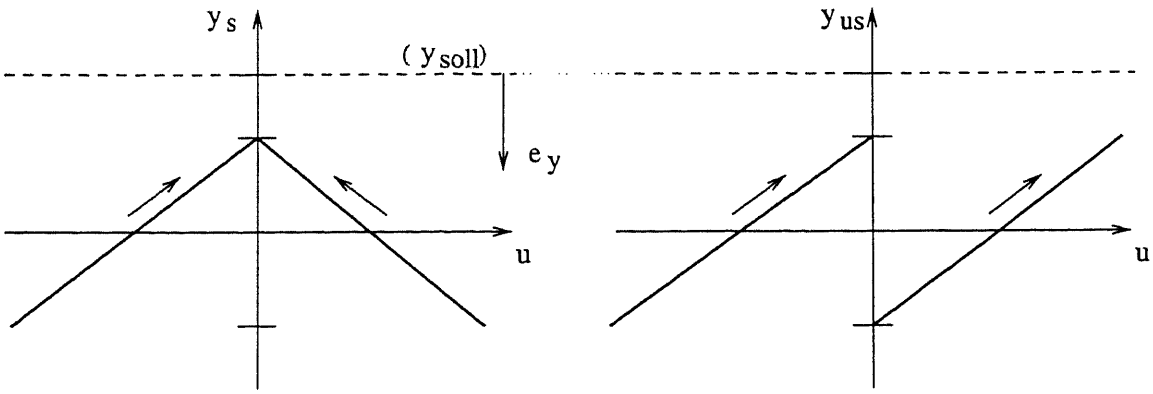


Abbildung 3.7: Wirkung des BP-Fehlers auf Reglerausgabe

deutet dies, aufgrund der vorzeichengleichen Ableitung und bei wiederholtem Lernen, eine ungewollte fortwährende Verschiebung des Arbeitspunktes in positiver u-Richtung. Die Folge ist, daß die Gewichte sich auf kein Optimum mehr einstellen, das Lernen divergiert. Denn wachsende u-Werte können nur durch ständig extremere Gewichtswerte des neuronalen Netzes erzielt werden, die dann zusätzlich noch ein Ansteigen der Fehlerwerte in der weiteren Propagierung fördern.

3.4 Propagierungs-Strategien

Die Frage ist nun, wie man trotz nicht vollständig differenzierbarer (im folgende als ND-Funktion bezeichnet) Korrekturfunktionen das Verfahren der neuronalen Regelung mittels BPtt durchführen kann, so daß alle zu berücksichtigenden Optimierungskriterien möglichst gut vom neuronalen Regler erfüllt werden. In betracht kommen u.a. folgende Vorgehensweisen:

FP-STOPP: Der Abbruch der Regelung eines Lernbeispiels, sobald in der FP-Phase eine Korrektur einer Reglerausgabe notwendig wird. Damit bei fortwährender Regelung in die gleiche Korrektur-Situation der Lernprozeß nicht stoppt, sollte an der Abbruchstelle eine Bewertung dieser Begrenzungs-Randbedingung erfolgen, die versucht, die Reglerausgabe auf deren Zulässigkeitsbereich zu beschränken. Andere Bewertungen von Systemzuständen, die bis zum Abbruch durchgeführt wurden, können ebenfalls rückpropagiert werden.
Da in diesen Fällen keine Endpositionsbewertung stattfindet, wird die Zielregelung mit dieser Methode in geringerem Umfang erlernt werden.

BP-STOPP: Vollständige Regelung bis Regelungsende. Dann Rückpropagierung aller separaten Bewertungen gemäß ihrer jeweiligen Rückpropagierungslänge (Anzahl zu durchlaufender Zeitschritte). An einer Korrekturstelle (im Begrenzungsbereich der Korrekturfunktion) wird die gerade aktive Rückpropagierung vorzeitig abgebrochen. Dies bedeutet eine Reduktion der für die Bewertung angesetzten lokalen Beeinflussung der Reglerparameter.

BP-NULLUNG: Ähnlich wie beim BP-Abbruch wird hier die Rückpropagierung durch eine Korrektur abgebrochen, indem der rückpropagierte Fehler an der Korrektur-

stelle zu Null gesetzt wird. Im Unterschied zum BP-Abbruch wird aber in parallelen Funktionszweigen, die nicht über die Korrekturfunktion führen, die Rückpropagierung fortgesetzt. Dies ist zulässig, da eine separate Propagierung von Fehlern in verschiedenen Punkten der Funktionenkette prinzipiell möglich ist. Außerdem kann dann wenigstens noch eine Parametrisierung für die noch zu durchlaufenden Regelungsbereiche erfolgen, die eine Optimierung der anderen Eingangsvariablen der imaginären Bewertung (vgl. Seite 23 und Abbildung 3.6) verspricht.

BP-PROP: Mit der Argumentation aus Kapitel 3.3 auf Seite 24 kann versucht werden, den Fehler auch durch die ND-Funktionen der Korrekturen zu propagieren. Dabei sind die möglichen Auswirkungen durch die Form mancher, zugegebenermaßen exotischer, Funktionen auf die Einstellung der Reglerparameter zu bedenken. Bei nicht möglicher analytischer Herleitung der Ableitungsfunktion kann diese numerisch angenähert werden. Um dabei zu große Werte an Sprungstellen zu vermeiden, sollte diese Näherung in beide Richtungen ausgeführt werden und der absolut kleinere Wert gewählt werden. Für gewöhnliche Begrenzungsfunktionen wird die Ableitung im Begrenzungsbereich 0. Dann stimmt diese Vorgehensweise genau mit der der BP-NULLUNG überein.

BP-BYPASS: Ein etwas riskant erscheinender Ansatz ist, die Kennlinie der Korrekturfunktion in der BP-Phase zu ignorieren und als Ableitung den Wert eins zu verwenden, so daß der Fehler am Ausgang der Funktion direkt zur Eingabe durchgegeben wird. Für eine Begrenzungsfunktion bedeutete dies, eine ähnliche Wirkung auf die Veränderung der Gewichte zu haben, wie im obigen Beispiel der Sägezahn-Kennlinie bereits beschrieben. Ein positiver Fehler für einen im oberen Bereich der Funktion begrenztes Stellsignal des Reglers würde dann noch zu einer Verstärkung des Reglerverhaltens in diesem Zustandsbereich führen. Dadurch werden aber auch aufgrund der funktionalen Auswirkung der Gewichtsänderungen auf benachbarte Zustandsbereiche dort extremere Regelungseingriffe impliziert. Bleiben diese Gewichtsänderungen in einem für das Lernen insgesamt nicht schädlichen Rahmen, dann kann, wie an einigen Testkonfigurationen innerhalb dieser Arbeit ersichtlich, die Performanz der Zielregelung verbessert werden. Allerdings wirkt dies einem Erlernen der Randbedingungs-Korrekturen entgegen und sollte daher nicht gleichzeitig mit diesen kombiniert werden.

3.5 Bewertung von Randbedingungen

Es sei u die auf einen Bereich $[\check{u}_G, \hat{u}_G]$ zu begrenzende Variable. Weiter seien $\check{u}_B$ und $\hat{u}_B$ die jeweils untere und obere Bewertungsgrenze, die von den unbedingt einzuhaltenden Korrekturgrenzen $\check{u}_G$ und $\hat{u}_G$ einer limitierenden Randbedingung abweichen können. Sinnvollerweise liegen die Bewertungsgrenzen dann innerhalb des Zulässigkeitsbereichs der limitierenden Randbedingung, um einen gewissen Sicherheitsabstand von der Randbedingungs-Verletzungsgrenze zu erhalten. Bei forcierenden Randbedingungen gibt es keinen unbedingt einzuhaltenden Zulässigkeitsbereich, so daß die Werte für $\check{u}_B$ und $\hat{u}_B$ relativ frei gewählt werden dürfen. Oft fallen beide dann auf den Nullpunkt (z.B. Energieminimierung).
Die Bewertungsfunktion beurteilt den Fehler $|u_B - u_{ist}|$ zwischen dem Wert der zu begren-

zenden Variablen und der nähesten Bewertungsgrenze $u_B \in \{\check{u}_B, \hat{u}_B\}$, falls der Wert der Variablen außerhalb des Bereichs $[\check{u}_B, \hat{u}_B]$ liegt. Statt einer Variablen kann auch ein Variablenvektor bewertet werden, dessen Werte auf einen geschlossenen Bereich eingegrenzt werden sollen.
Eine übliche Bewertungsfunktion, wie auch in dieser Arbeit verwendet, ist der quadratische Fehler der zu bewertenden Variablen u_j (j=1,2,..,N)

$$E = \frac{1}{2}\sum_{j}^{N}(u_{j,B} - u_{j,ist})^2$$

3.5.1 Sigmoidfunktion als Korrekturfunktion

Die Korrektur für eine Begrenzungs-Randbedingung des Stellsignals kann prinzipiell durch eine Sigmoidfunktion erfolgen. Die Formulierung ist für eine nur vom Stellsignal selbst abhängige Randbedingung relativ einfach. Der Vorteil dieser differenzierbaren Korrekturfunktion ist, daß sie als Bestandteil der Reglerausgabe des neuronalen Netzes im Ausgabeneuron realisiert werden kann und deshalb keine Adaptierung der Reglerfunktion auf diese Randbedingung erfolgen muß.
Die sigmoide Beschreibung wird aber z.B. für die Blockade-Randbedingung des Stellsignals schon schwieriger, da

1. deren Korrekturfunktion abhängig vom Stellsignal θ des Reglers und dem TBU-Winkel Φ ist und daher 2-dimensional zu formulieren ist,

2. die exakten Korrekturwerte in der Nähe des Blockadegrenzwinkels nicht analytisch bestimmbar sind.

Außerdem ist eine einfache funktionale Hintereinanderschaltung mehrerer Korrekturfunktionen zur Begrenzung nach verschiedenen Randbedingungen generell nicht möglich. Eine Korrektur der letzten in der Reihe der Korrekturfunktionen würde die vorherigen Korrekturen anderer Randbedingungen ignorieren. Beispiel:
Realisiert man die Lenkwinkelbegrenzung auf den Bereich $[-10°, +10°]$ als sigmoide Korrekturfunktion, dann würde eine nachfolgende Blockadekorrektur, gemäß der linksseitigen Blockade in Abbildung 2.4 Seite 14, bei einem aktuellen Lenkwinkel $\theta = 0°$ diesen auf einen Wert größer $+10°$ korrigieren, um die Blockade noch umfahren zu können.
Folglich ist die Lenkwinkelbegrenzung nur dann in das neuronale Netz als z.B. sigmoide Outputfunktion des Ausgabeneurons integrierbar, wenn man davon ausgehen kann, daß nachfolgende Korrekturen die Lenkwinkel-Randbedingung nicht erneut verletzen. Deshalb wäre eine Implementierung aller Korrekturen als Sigmoidfunktionen nur sinnvoll, wenn dies ausschließlich auf den integralen Korrekturbereich angewandt würde. Eine geeignete Sigmoidfunktion für die Schnittmenge aller Zulässigkeitsbereiche zu finden, im mehrdimensionalen Raum der Korrketureingänge, ist jedoch nicht einfach. Für die Blockadekorrektur wurde daher durchweg bei unerlaubten Stellsignalwerten die harte Korrektur auf die Bereichsgrenzen der Randbedingung durchgeführt.
Ein Grund, die sigmoide Funktion für die Stellsignalbegrenzung nicht zu verwenden, ist, die Unterschiede in der Regelung zu untersuchen, wenn die Reglerfunktion mittels einer Bewertung des Stellsignals selbst eine Begrenzung erreicht. Für Ausgabewerte im Sättigungsbereich der sigmoiden Kennlinie findet außerdem eine nur sehr langsame Anpassung

der Reglergewichte statt (Ableitung geht gegen 0). Bei Wegfall dieses Sättigungsbereichs könnte sich also ein anderes Reglerverhalten in diesen Ausgabebereichen ergeben.

3.5.2 Bewertungen des Regelstreckenzustands

Bei vom Streckenzustand abhängigen Randbedingungen, wie der TBU-Blockade, kann es zwei Alternativen für deren Formulierung und somit für den Ansatzpunkt ihrer Bewertung geben (vgl. Kapitel 2.2.2 Seite 14, Blockade-Randbedingung):

1. **Bewertung des Regelstreckenzustands:** Eine Bewertung der Verletzung einer limitierenden Streckenzustands-Randbedingung setzt am Ausgang der Streckensimulation an. Dies bleibt eine *einmalige* Bewertung, da die Regelung hier abgebrochen werden muß. Wenn aber die Bewertungsgrenzen in den Zulässigkeitsbereich hineinverschobenen werden, wird ein Sicherheitsabstand definiert, innerhalb dem bei weiterlaufender Regelung eine Bewertung der Regelung in diesen Gefahrenbereich vorgenommen werden kann. Dabei kann man auch eine präventive Korrektur am Stellsignal mitverwenden, damit durch eine Verletzung dieser Randbedingung kein Abbruch nötig wird. Man erhält so eine *begleitende* Bewertung der limitierenden Randbedingung.

2. **Transformation auf Stellsignal-Randbedingung:** Eine Formulierung der Randbedingung des Streckenzustands als Stellsignal-Randbedingung macht eine Bewertung des Stellsignalfehlers bzgl. der Randbedingungs-Einhaltung möglich. Verbunden mit der dann auch vornehmbaren Korrektur ergibt dies eine begleitende Bewertung entlang der Regelungstrajektorie.

3.5.3 Bewertungen der TBU-Randbedingungen

Hier nun eine Liste der dem Autor sinnvoll erscheinenden Behandlungsmöglichkeiten von Randbedingungen des TBU-Modells, wie sie hier vorgestellt werden.

Lenkwinkelbegrenzung

Charakterisierung: limitierende Stellsignal-Randbedingung, abhängig von Stellsignal des Reglers.
Bewertungsarten:

Sigmoidkorrektur: Korrektur als reglerinterne sigmoide Funktion, wenn die Randbedingung zusätzlich in der nachfolgenden integralen Korrektur berücksichtigt wird. Es ist in diesem Fall keine Bewertung zur Regleradaptierung auf Randbedingung notwendig, da diese bereits als Bestandteil der Reglerfunktion angesehen werden kann.

Bewertung des Stellsignals: Wird die Lenkwinkelbegrenzung als zu bewertende externe Korrekturfunktion realisiert, dann wird der Regler selbst auf eine Einhaltung der Randbedingung hin trainiert. Dies ergibt möglicherweise ein anderes Reglerverhalten, als ein Lernen mittels der sigmoiden Begrenzung.

TBU-Blockade

Charakterisierung: limitierende Streckenzustands-Randbedingung, abhängig vom Winkel Φ_{t+1} zwischen Zugmaschine und Anhänger. Eine Transformation der Randbedingung auf eine limitierende Stellsignal-Randbedingung ist möglich. Diese ist dann abhängig von Φ_t und der Reglerausgabe θ.
Bewertungsarten:

Bewertung des TBU-Winkels: Bewertung des Winkels Φ am Ausgang des Simulationsmoduls, wie in Kapitel 3.5.2 beschrieben.

Bewertung des Stellsignals: Die Bewertung des Stellsignalfehlers aufgrund der Korrektur des Stellsignals zur Vermeidung der Blockade im nächsten Simulationsschritt. Die Erkennung einer drohenden Blockade im nächsten Zustand erfolgt über dessen Schätzung (vgl. Kapitel 2.2.2 Seite 15).

Bewertung der Blockadeposition: Wie in [1] bereits geschildert, kann eine Endpositionsbewertung an der Blockadestelle ebenfalls zu einer grundsätzlichen Vermeidung dieser Regelung führen. Dies wird damit begründet, daß die Ausrichtung an der Blockadestelle auf die Sollposition am Andockpunkt ein der Blockade entgegenwirkendes Regelverhalten verlangt. Diese Bewertung findet ebenfalls am Simulationsausgang statt.

Begrenzung der Lenkwinkeldifferenz

Charakterisierung: (hier) forcierende Stellsignal-Randbedingung, abhängig von aktuellem und vorherigem Stellsignal. Sie kann auch zur limitierenden Randbedingung werden, wenn die Forderung nach einer Begrenzung der Lenkwinkeldifferenz aus modelltechnischen Gründen erforderlich ist. Beim TBU-Modell wird davon ausgegangen, daß große Lenkbewegungen zwischen zwei Regelungsschritten möglich sind (etwa durch kurzes Anhalten des Fahrzeugs).
Bewertungsarten:

Bewertung des Stellsignals: Da hier keine Korrekturfunktion vorhanden ist (forcierende Randbedingung), wird das Stellsignal bzgl. den Grenzen eines festzulegenden Bewertungsbereichs bewertet.

3.6 Separate Gewichtungen

Zur Bewertung von Randbedingungen sind bisher folgende Variationsmöglichkeiten vorgestellt worden:

- Korrekturbewertung integral/separat,
- Ablauf der Vorwärts- und Rückwärtspropagierung bei Auftreten von Korrekturen (s. Kapitel 3.4 Seite 26),
- Bewertungsalternativen für Stellsignal- und Streckenzustands-Randbedingungen.

Die separate Behandlung von Randbedingungen macht es aber auch möglich, ganz gezielt Gewichtungen zwischen den Bewertungen vorzunehmen. Diese Gewichtungen ergeben für die einzelnen zu bewertenden Variablen unterschiedliche *effektive Lernkoeffizienten*. Die Umschreibung *effektiv* steht für die Zusammenfassung aller das Maß der Gewichtsveränderungen beeinflussender Faktoren, mit Ausnahme der Teilfunktionen der Propagierungskette.
Eine Gewichtung mag z.B. erforderlich sein, wenn es darum geht, weniger gut zu erlernende Randbedingungen mit größeren Fehlerwerten zu beaufschlagen, oder wenn starke Reaktionen der Netzparametrierung auf seltene hohe Fehlerwerte zu einer Zerstörung des bereits Gelernten führen.
Es folgt eine Auflistung der am effektiven Lernkoeffizienten beteiligten Faktoren. Manche dieser Faktoren lassen sich nicht oder nur ungenau quantitativ angeben, weshalb die Angabe eines effektiven Lernkoeffizienten auch nur eine qualitative Aussage sein kann. Einige lassen sich gezielt auf die Veränderung der Gesamtbewertungs-Charakteristik anwenden, andere sind eher als unbeabsichtigter Nebeneffekt zu betrachten.

Netz-Lernkoeffizient Der Lernkoeffizient, der bei der Berechnung des Gewichtsupdates gemäß der Delta-Lernregel mit dem rückpropagierten Fehler multipliziert wird. Dieser ist gewichtsspezifisch festgelegt und kann über den Verlauf des Lernens nach dem sogenannten Lernplan variiert werden. Es ist der vom herkömmlichen BP-Verfahren übernommene, grundlegende Skalierungsfaktor.

Fehlerfunktion Die Fehlerfunktion, hier auch als Bewertungsfunktion bezeichnet, ist zwar Teil der Funktionskette, durch die propagiert wird, kann aber auf die Beeinflussung des Regelverhaltens hinzielend verändert werden, um beispielsweise zu große Sollabweichungen nicht zu stark zu bewerten.

Bewertungsskalierung Zur einfacheren Gewichtung der unterschiedlichen Bewertungen empfiehlt es sich, einen Skalierungswert für die Ausgabe jeder Fehlerfunktion vorzusehen.

Fehlerskalierung Nach Propagierung durch die Fehlerfunktion können die an den bewerteten Variablen anstehenden Fehlerwerte mit einem variablenspezifischen Faktor skaliert werden. Im Prinzip ist das nichts anderes als eine aus der Fehlerfunktion herausgeführte Modifikation derselben.
Beispiel:
Für die Fehlerbewertung zweier Variablen x_1 und x_2 wurde die quadratische Fehlerfunktion

$$E = \sum_{j=1}^{2} (x_{j,B} - x_j)^2$$

gewählt. Der zu x_1 rückpropagierte Fehlerwert soll mit dem Faktor 2 gewichtet werden. Dies läßt sich auch durch eine Modifikation der Fehlerfunktion mittels

$$\tilde{E} = 2(x_{1,B} - x_1)^2 + (x_{2,B} - x_2)^2$$

erreichen, ist in der Handhabung aber einfacher an den Variablen selber vorzunehmen.
Mittels dieser Methode kann ein gleichmäßigeres Erlernen einer Randbedingung erreicht werden. Bei der Zielpositionierung des TBU ist z.B. aufgrund der unterschiedlichen Einheiten des Anhängerwinkels β [rad] und der y-Position [m] je nach

Geometrie des Fahrzeugs ein Ungleichgewicht im Lernerfolg feststellbar. Durch eine Skalierung der Variablenfehler an der Bewertungsposition kann dieser Unterschied ausgeglichen werden. Eine Justierung mittels Fehlerskalierung kann auch durch die für die Bewertung gewählten Wertebereichs-Darstellungen der Variablen nötig werden. Siehe dazu die Anmerkungen im Anhang B.

BP-Länge Den Lokalisationscharakter einer Randbedingung kann man durch die Vorgabe der Anzahl zu durchlaufender Regelungsschleifen (Zeitschritte) berücksichtigen. Durch eine unterschiedliche Wahl dieser BP-Länge unterliegen die einzelnen Gewichte dann auch unterschiedlich großen Veränderungen ihres Wertes[4].

Lernregel Eine Variation der Delta-Lernregel kann dem eben angesprochenen akkumulierenden Effekt einer unterschiedlichen BP-Länge entgegenwirken. Die normierende Delta-Lernregel (*Norm-Delta-Regel*) multipliziert den über die Rückpropagierung summierten Update eines Gewichts mit der Quadratwurzel der BP-Länge. Um eine gleichmäßige Wirkung für kleine und größere BP-Längen zu erzielen, müßte der BP-Längen-Effekt gesichert bekannt sein. Deshalb wurde auf eine Anwendung dieser Lernregel oder einer ihrer Variationen verzichtet.

Bewertungshäufigkeit Entlang einer Trajektorie werden die einzelnen Bewertungen unterschiedlich häufig Verwendung finden. Während dies bei der Endpositionierung nur am Ende der Regelung stattfindet, kann die Blockade bei eingeschalteter Korrektur mehrmals und sogar aufeinanderfolgend bewertet werden. Dies wirkt sich ebenfalls auf den effektiven Lernkoeffizienten in annähernd proportionaler Weise aus, wenn es sich um benachbarte Regelungsschritte handelt. Eine Vervielfachung des Gewichtupdates ist auch in diesem Fall zu beobachten. Eine Möglichkeit diesem zu begegnen ist, bei einer Folge von Bewertungen stets nur eine davon zur Bewertung heranzuziehen. Jedoch stellt sich dann die Frage der Repräsentativität der ausgewählten Bewertung.

Bewertungstyp Im Fall der Behandlung der Blockade-Randbedingung mittels Bewertung der Endposition in der Blockadesituation treten bei weit von der Sollposition entfernten Blockaden sehr viel größere Positionsfehler auf als bei der durch das Endekriterium in Nähe der Sollposition aktivierten Bewertung. Dies sollte bei der Skalierung der Fehlerfunktion mit berücksichtigt werden.

[4] Es ist möglich, daß zu ähnlichen Regelungssituationen (also in benachbarten Regelungsbereichen) in etwa ähnliche Fehler rückpropagiert werden, was eine Akkumulation des Updates eines Gewichts hervorrufen würde. Jedoch kann diese Wirkung bei längeren Regelungstrajektorien durch gegenteilige Gewichtsveränderungen, aufgrund einer für die dortige Regelungssituation anderen Parametrisierung, wieder aufgehoben werden. Dieser Effekt konnte in dieser Arbeit nicht mehr näher untersucht werden und bleibt daher eine Hypothese.

Kapitel 4

Tests

4.1 Testvorbereitungen

Die hier vorgestellten Versuche der Einbindung von lokalen Randbedingungen in das BPtt-Reglerlernen sollen eine möglichst gute Vergleichs- und Analysemöglichkeit besitzen. Dazu ist es notwendig, vorab eine möglichst stabile Grundkonfiguration aller Lernparameter zu ermitteln, aus der dann die Variationen für die Untersuchungen abgeleitet werden. Bei der Veränderung von Lernparametern müssen so gut wie möglich unbeabsichtigte Seiteneffekte ausgeschlossen werden, die eine Interpretation der Wirkung der Parameteränderung auf den Lernerfolg verfälschen würden. Soweit nicht zur Einbindung der Randbedingungen unbedingt nötig, sollte daher auf Variationen von Parametern verzichtet werden, deren Einfluß auf das Lernen nicht klar oder nur spekulativ einschätzbar ist.

4.1.1 Variable Lernparameter

In Tabelle 4.1 sind die nicht Randbedingungs-spezifischen Lernparameter geordnet zusammengefaßt. Dazu einige Anmerkungen:
Eine strukturelle Erweiterung des Reglernetzes über die **Zahl der Neuronen und Layer** wirkt sich wenig auf die Lernrate der einzelnen Neuronenverbindungen aus. Vielmehr kann dadurch eine Aussage über eine Kompensation des Performanzverlusts bei erhöhten Optimierungsanfoderungen erhalten werden. Beobachtet wurde jedoch zu Beginn des Lernens eine stärkere Reaktion auf frühe Bewertungen lokaler Randbedingungen. Sind die Gewichte noch ähnlich dem Initialisierungszustand[1] eingestellt, so werden die zum Eingang rückpropagierten Fehler eines Fehlers am Ausgabelayer bei einer größeren Zahl von Neuronen im Hidden Layer größer sein, da im Eingabelayer mehr Fehleradditionen stattfinden. Die einlaufenden Fehler zweier mit verschiedener Neuronenzahl ausgestatteter Netze unterscheiden sich für den Initialzustand nicht in ihrer Verteilung. Durch die iterative Rückpropagierung entlang der Funktionskaskade akkumuliert sich deshalb dieser Fehler auch an den Gewichten entsprechend der Neuronenzahl stärker und beeinflußt auf diese Weise letztendlich den effektiven Lernkoeffizienten. Dies kann bei empfindlich auf Parameterveränderungen reagierenden Lernkonfigurationen ein erfolgreiches Einschwingen auf die Zielfunktion des Reglers verhindern.

[1] Als Initialisierung wird hier eine gleichmäßig zufällige Verteilung aus [-0.1,+0.1] vorgenommen.

Lernparameter	Variation
NN-Regler:	
Anzahl Neuronen	x
Anzahl Layer	x
Neuronenfunktionen	x[a]
Lernkoeffizienten	-
Lernregel	-
Eingangsvariablen	x
Normierungswertebereiche	-
Lernplan:	
Anzahl Lernbeispiele	-
Aufbau Startbereiche	-
Variation der Lernkoeffizienten	-
Rasterung der Startwerte	-
Bewertung:	
Bewertungsgewichtung	x
Fehlerskalierung	x
Fehlerfunktion	-
BP-Länge	x
Variablen	-
Regelungsschrittweite	-
TBU-Modell	-
Propagierungsstrategie	x

Tabelle 4.1: Lernbeeinflussende Parameter und ihre Verwendung in den Versuchen

[a]Nur Transferfunktion des Ausgabelayers

Die **Lernkoeffizienten** der Gewichte sollten, um eine Feinjustage in Richtung Optimalwerte gegen Ende des Lerndurchgangs durchzuführen, während des Lernens langsam erniedrigt werden. Da in dieser BPtt-Anwendung jedoch von einer Konstanz der Lernproblematik ausgegangen wird (stetige Erweiterung des Startbereichs), kann eine Feinjustierung allenfalls erst nach Abarbeitung des Endstartbereichs erfolgen. Weiter zu beachten ist, daß Bewertungen verschiedener Randbedingungen zu unterschiedlichen Phasen des Lerndurchgangs unterschiedlich stark ausfallen werden. Die Blockade wird z.B. meist ganz zu Beginn vom ungelernten Netz und dann erst wieder etwa zur Hälfte des Lerndurchgangs angefahren, wenn die Start-Ziel-Regelung entsprechend kurvige Regelungstrajektorien verlangt. Die Lenkwinkel-Randbedingung jedoch wird, wie das sich langsam auf den Schwierigkeitsgrad einstellende Netz, durch das stetig anwachsende Stellsignal in stetig wachsendem Maß verletzt werden. Um einen Einfluß eines variierenden Lernkoeffizienten auf das Erlernen der Randbedingungen in verschiedenen Lernphasen auszuschließen, wird deshalb dieser layerspezifisch über das gesamte Training konstant gehalten. Randbedingungsspezifische Lernkoeffizienten können dann über die **Bewertungsgewichtung** und die **Fehlerskalierung** eingestellt werden.

Weiter konstant gehalten werden die **Neuronenfunktionen**. Ausgenommen davon ist die Transferfunktion des Ausgabeneurons, die zur in das Netz integrierten Begrenzung des

Stellsignalwerts verwendet werden kann.

Die Abhängigkeit des effektiven Lernkoeffizienten von der BP-Länge einer Bewertung ist unklar und daher nicht durch eine alternative **Lernregel** (Norm-Delta-Lernregel) sicher zu kompensieren, zumal die Form der Abhängigkeit, nach den Überlegungen in Kapitel 3.6 Seite 32, wieder von der BP-Länge selber beeinflußt sein wird. Ein Batchlearning über mehrere Lernbeispiele hinweg kann eventuelle Lernstörungen durch Fehlerausreißer vermeiden helfen. Da Batchlearning aber in der Regel eine Erhöhung der Zahl der Lernbeispiele erfordert, wurde hier stets die einfache Delta-Lernregel verwendet.

Bei der Variation der **BP-Länge** einer Bewertung muß die Veränderung des effektiven Lernkoeffizienten in der Interpretation berücksichtigt werden. Zur Vereinfachung wurden daher viele Tests mit der Blockadebewertung nur mit der BP-Länge 1 durchgeführt.

Die Abbildung Streckenzustand $\rightarrow$ Reglersignal ist in ihrer Komplexität auch von der Wahl der **Eingangsvariablen** abhängig. Mit bestimmten Eingangsvariablen wird die gesuchte Regelungsausgabe einfacher als neuronale Funktion zu realisieren sein. Um auch hier so wenig wie möglich Abhängigkeiten einzubauen, ist die Zahl der Eingangskoordinaten auf den Minimalwert 4 reduziert. Zusätzlich wird dann die Wirkung einer zusätzlichen Hilfsgröße (TBU-Winkel Φ) auf das Reglerlernen untersucht.

Durch eine andere Wahl der **zu bewertenden Variablen** wird auch der Fehlercharakter verändert, da die hierbei angewandte Variablentransformation der Fehlerfunktion zuzuordnen ist (vgl. [1], Kapitel 2.3.6. Seite 13). Dies ist durch Verwendung der stets gleichen Bewertungsvariablen im Interesse der Vergleichbarkeit zu vermeiden.

Eine andere **Fehlerfunktion** kann etwa sinnvoll sein, wenn die auftretenden Fehlerwerte eine zu große Bandbreite aufweisen. Zur Bewertung des Sollfehlers wurde hier aber stets die quadratische Summe eingesetzt.

Die Wertebereichstransformationen der Variablenwerte zwischen der Darstellung im neuronalen Netz und dem Real-World-Bereich besteht je Variable aus einer affinen Abbildung mittels Skalierung und Offset. Der zu spezifizierende Arbeitsbereich einer Eingangs-Variable wird dabei auf einen zur neuronalen Verarbeitung günstigen Bereich normiert (z.B. $[-1,1]$). Am Ausgang des Reglers findet die umgekehrte Transformation für das Stellsignal statt. Eine Veränderung dieser **Normierungs-Transformation** kann prinzipiell durch andere Skalierungen und Bias-Offsets im Netzinnern mit dem Netzlernen kompensiert werden. Jedoch wurden bei Vorabtests Abhängigkeiten von der Wahl der Normierung registriert, so daß hier nur bei nicht verwandten Lernkonfigurationen unterschiedliche Wertebereichvorgaben zugelassen wurden.

Der Aufbau des **Lernplans** ist von entscheidender Bedeutung für das Lernen, da in ihm die Lerndauer (**Anzahl Lernbeispiele**) sowie der Grad und die Konstanz der Lernschwierigkeit (**Aufbau der Regelungs-Startbereiche**) definiert werden. Insbesondere was die Auswahl der Startbeispiele aus den Startbereichen betrifft, gibt es folgende Problematiken zu bedenken:

- Bei den Anfangsbeispielen zum Training des noch ungelernten Netzes kann es zu häufigen Randbedingungs-Verletzungen (Blockade) kommen. Ist diese Randbedingung nicht Teil des Lernvorgangs, so wird bei bestimmten Propagierungs-Strategien kein Lernen stattfinden und bei wiederholter Anfahrt der Blockade der gesamte Lernprozeß unvollständig bleiben.

- Es sollte im Hinblick auf eine möglichst konstante Erschwerung des Regelungsproblems verhindert werden, daß Beispiele ausgewählt werden, die entweder nicht ausregelbar sind (TBU an Wand und nicht in Andocknähe) oder bei denen zwischen benachbarten Beispielen funktional grundsätzlich verschiedene Regelungen erforderlich sind. Solche Funktionsunstetigkeiten sind mit neuronalen Netzen nur ungenügend zu approximieren, weil sie nicht in der Art einer stetigen Funktionsadaptierung generierbar sind.

- Problematisch können auch Regelungen aus Randzonen des bis zum aktuellen Zeitpunkt regelbaren Bereichs in noch unbekanntes Terrain sein. Wenn die Startbereiche zu schnell vergrößert werden, werden die Regelungen in diesen neuen Bereichen immer ungenauer und bereits zu Beginn der Regelung ein nicht-zielorientiertes 'Ausreißen' des Fahrzeugs verursachen.

- Erfolgt die Erweiterung des Startbereichs in mehreren Etappen, so wird bei sprungartiger Veränderung der Startbereichsgrenzen auch der Regelungsfehler Sprünge aufweisen. Um Lernstörungen dadurch bei empfindlichen Lernkonfigurationen zu vermeiden, wird bei den folgenden Tests stets zwischen den Stützpunkten der Bereichserweiterung interpoliert, so daß eine gleichmäßige Erweiterung erfolgt.

Auf eine Variation von anderen Lernparametern über den Lernplan wird wie auch schon bei den Lernkoeffizienten des neuronalen Netzes verzichtet.
Ebenso wurde auf eine **Rasterung der Startwerte** zur Verkleinerung der Lernbeispielmenge verzichtet. Zum einen, um einen weiteren in der Wirkung nicht genau zu quantifizierenden Lernparameter zu beseitigen. Zum anderen, weil es bei der Regelung entlang der Trajektorie automatisch zu zwischen den Rasterpunkten liegenden Systemzuständen kommt.

Die **Regelungsschrittweite** darf nicht zu groß sein, um nicht zu träge zu reagieren (mit der Folge einer überschwingenden oder sich sogar aufschaukelnden Regelung). Da bei einer Vergrößerung (Verkleinerung) der Regelungsabstände entsprechend weniger oft (öfter) durch den Regler hindurchpropagiert würde und damit weniger (mehr) Gewichtsupdates kumuliert würden, muß auch diese Beeinflussung des effektiven Lernkoeffizienten, ähnlich der der BP-Länge, unterbleiben.

Auch ein verändertes **TBU-Modell** wird den Lernerfolg beeinflussen (es gibt einfacher und aufwendiger zu kontrollierende Anhänger-/Zugmaschinen-Geometrien). Daher werden auch diese Werte über alle Versuche konstant gehalten.

4.1.2 Hohe Dimension des Lernparameterraums

Die zu variierenden Lernparameter lassen sich, was die Versuche zur Einbindung lokalen Regelverhaltens betrifft, auf die folgenden reduzieren:

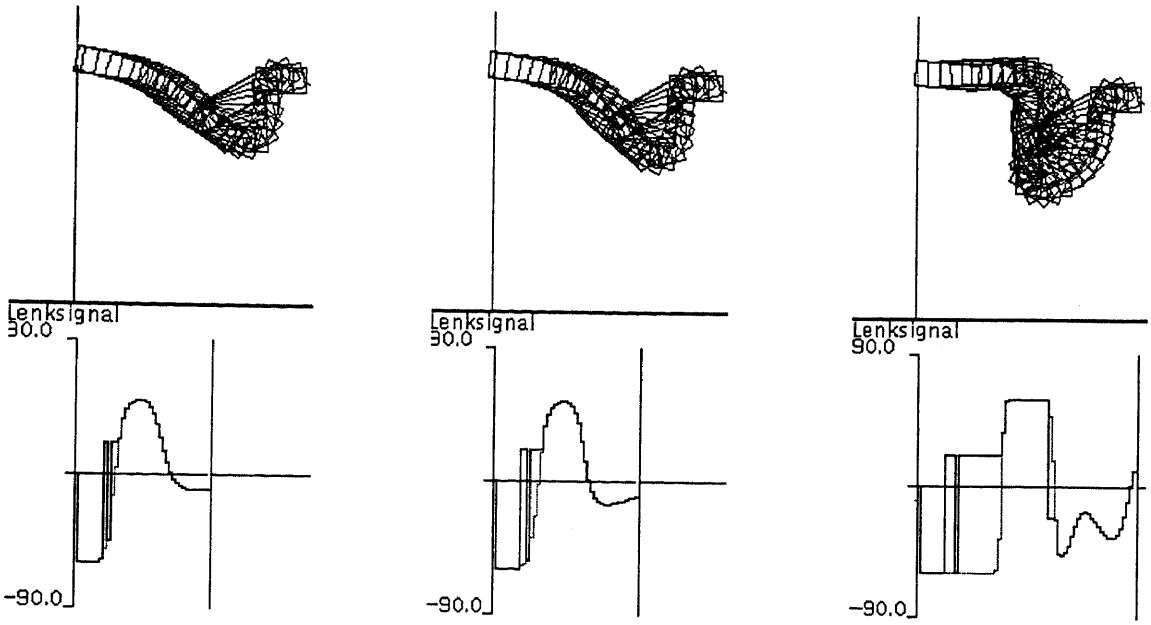
1. Anzahl Neuronen in hidden Layern
2. Anzahl Layer
3. Zusätzliche Reglereingänge
4. Bewertungsgewichtung
5. Fehlerskalierung
6. BP-Länge

Kombiniert mit den in Kapitel 3.4 beschriebenen Propagierungs-Strategien und der Auswahl zu berücksichtigender Randbedingung ergibt sich trotzdem eine hohe Dimension des Lernparameterraums. Um daher Aussagen zur Wirkung bestimmter Parametereinstellungen auf die Bewältigung der Regelungsaufgabe machen zu können, müßten streng genommen mehrdimensionale Parameterbereiche untersucht werden, was aus Zeitgründen hier nicht durchführbar war. Letztlich wurde nur stichprobenweise jeweils ein Parameter variiert und das Ergebnis unter Berücksichtigung der möglichen Seiteneffekte zu interpretieren versucht. Es muß deshalb eigentlich davon ausgegangen werden, daß die Einstellung der nicht variierten Lernparameter für eine variierte Konfiguration in gleicher Weise optimal eingestellt sind wie für die Ausgangskonfiguration. D.h. die Veränderung der Reglereigenschaft darf nur auf der Veränderung des variierten Parameters beruhen. Eine derartig partielle Abhängigkeit der Reglerfunktion von den Lernparametern ist im realen Fall selbstverständlich nicht gegeben. Um so vorsichtiger sind daher die resultierenden Ergebnisse zu interpretieren.

4.1.3 Restriktionen

Die vorab vorhandenen Einschränkungen für einen Vergleich verschiedener Konfigurationen sind:

- Der Einfluß der BP-Länge auf den effektiven Lernkoeffizienten, wenn der Lokalisationscharakter einer Randbedingung berücksichtigt werden soll.
- Die nicht genau bekannte Verteilung eines Randbedingungs-Fehlers in Betrag und Auftreten innerhalb eines Lerndurchgangs. Zwei verschiedene Bewertungen der gleichen Variablen werden bei sonst gleichen Bewertungsparametern eine unterschiedliche derartige Fehlerverteilung aufweisen und sich daher ein unterschiedlicher Lernerfolg aufgrund ihrer separaten Berücksichtigung in der Regelungsfunktion einstellen.
- Verschiedene Bewertungsverfahren der gleichen Randbedingung können in ihrem Fehlercharakter ebenfalls unterschiedlich sein. Beispiel: die mit unterschiedlichen Methoden zu bewertende Blockade. Ein Vergleich bei sonst gleichen Bewertungsparametern (Skalierung, Fehlerfunktion) wird durch die unterschiedlichen Orte der Bewertung (Stellsignal / Simulationsausgang) und die anderen Bewertungsvariablen (Stellsignal + TBU-Winkel / TBU-Winkel / Sollzustandsvariablen) schwerer möglich.



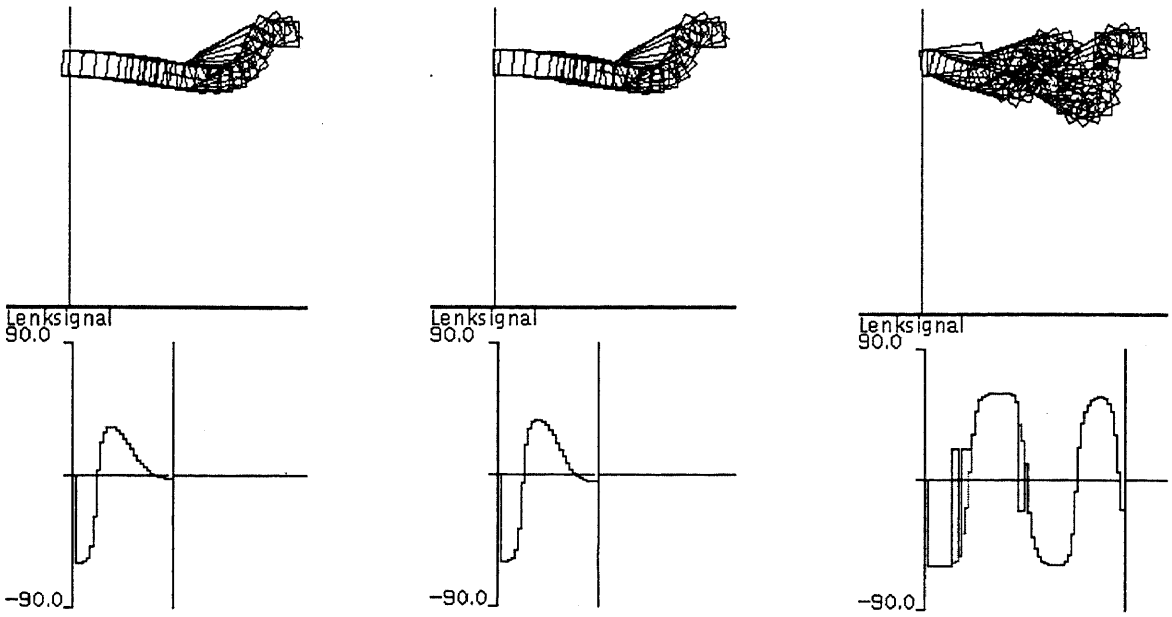
$x_s = 0.05m$ $x_s = 0.15m$ $x_s = 0.25m$

Abbildung 4.1: Vorteilhaftes frühes Einlenken durch statischen x-Positionsfehler

4.1.4 X-Positionsfehler

Aufgrund der zeitdiskreten Simulation wird das TBU-Fahrzeug durch das Endekriterium der Regelung erst nach Überfahren der Wandbegrenzung gestoppt. Wird der dabei auftretende Fehler in x-Richtung mitbewertet, wie in [4] angegeben, erhält man folgenden verdeckten Effekt:
Die Bewertung des Endzustands bewertet einen im Mittel $h/2$ großen positiven x-Fehler (h=Regelungsschrittweite). Wird dieser separat betrachtet, dann bewirkt seine Rückpropagierung durch die gesamte Trajektorie eine Veränderung der Regelungsfunktion dergestalt, daß der Regler zukünftig versucht, das Fahrzeug weiter in positive x-Richtung (vor die Wand) zu positionieren. Das Fahrzeug stoppt aber auch zukünftig erst bei $x < 0$. Der Versuch der Vorpositionierung verändert deshalb nur die Form der Anfahrtsbewegung, und zwar ist ein früheres Einschwenken in die Horizontale mit $y = 0$ erkennbar. Dieses Verhalten ist für die Regelung prinzipiell von Vorteil, da dadurch aufwendige Manöver vor der Andockposition bei schräger Anfahrt vermieden werden.
Um den Effekt der Vorpositionierung genauer zu untersuchen, wurde anstelle des x-Fehlers der Endposition ein über den Lerndurchgang statischer positiver x-Fehler x_s angesetzt. Dies wurde mit unterschiedlichen Werten für x_s durchgeführt. Es zeigte sich, daß ab einem ungefähren Wert $x_s = 0.15m$ des statischen x-Fehlers[2] das vorteilhafte frühe Einschwenken für bestimmte Startbereiche übergeht in eine sich entlang der Horizontalen $y = 0$ schlängelnde, auf $x = 0$ zu bewegende Regelung[3].
Die Bilder 4.1 und 4.2 zeigen für zwei unterschiedliche Startpositionen die Regelungen für unterschiedliche x_s-Werte, die die Wirkung eines sich erhöhenden x-Positionierungsfehlers

[2] mit der in den Tests verwendeten TBU-Geometrie, s. Kapitel 4.3.1 S. 46
[3] Salopp gesagt drückt der Regler sich immer mehr davor, in Richtung Wand zu lenken.



$x_s = 0.05m$ $x_s = 0.15m$ $x_s = 0.25m$

Abbildung 4.2: Pendelung um x-Achse in infolge eines zu hohen statischen x-Positionsfehlers

verdeutlichen.
Für die hier durchgeführten Versuche wurde ein statischer x-Fehler nur zusätzlich angesetzt, um in Kombination mit dem lokalen Bewertungslernen eine weitere Optimierung des Regelverhaltens zu erzielen. Grundsätzlich wurde ein negativer x-Fehler jedoch zu Null gesetzt.

4.2 Analysen

Zur vergleichenden Auswertung der erhaltenen Reglerfunktionen werden

1. die Regelungen jeweils für $N_T = 1000$ zufällig aus dem letzten Lern-Startbereich ausgewählten Testbeispielen mathematisch analysiert. Die Lernbereiche sind für alle Testkonfigurationen gleich; durch gleiche Initialisierung der Zufallsfunktion wird immer der gleiche Testdatensatz in gleicher Reihenfolge präsentiert,
2. die Regelungseigenschaften anhand 6 spezieller Startpositionen demonstriert,
3. zusätzlich Beobachtungen des Regelverhaltens durch den Autor auf etwaige Besonderheiten vorgenommen und dann in beispielhaften Trajektorienbildern veranschaulicht.

4.2.1 Mathematische Analysewerte

In Tabelle 4.2 sind die hier verwendeten Analysewerte zusammengefaßt, die nun noch genauer beschrieben werden.

Analysewerte							
Endposition		Trajektorie		Stellsignal			
σ_y	σ_β	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
Randbedingungen							
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$

Tabelle 4.2: Zuordnung der Analysewerte

Endposition Zur Beurteilung der Endpositionierung wird die Standardabweichung der Endpunktfehler errechnet. Da der tatsächlich auftretende x-Positionsfehler stets als Wert 0 in die Bewertung eingeht und sich ungefähr als der Mittelwert der Simulationsschrittweite ergibt (vgl. Kapitel 4.1.4 Seite 38), ist er nicht weiter von Bedeutung.

$$\sigma_y = \sqrt{\frac{1}{N_T}\sum_{i=1}^{N_T}(y_{Soll} - y_{i,Ende})^2}$$

$$\sigma_\beta = \sqrt{\frac{1}{N_T}\sum_{i=1}^{N_T}(\beta_{Soll} - \beta_{i,Ende})^2}$$

Interpretation: Güte der Zielregelung.

Trajektorie Die genaue Form kann natürlich nicht mit singulären Zahlenwerten beschrieben werden. Um dennoch ein numerisches Indiz für einen globalen Regelungscharakter zu erhalten, wird die Länge einer Regelungstrajektorie mit einer als ideal angesetzten verglichen. Aus der durchschnittlichen Länge der Trajektorie kann auf ein einfacheres (Geradeausfahrt), komplexeres (enges Einlenken in Horizontale) oder zu extremes Regelungsverhalten (aufwendige Kurvenfahrten, Pendeln) geschlossen werden. Als 'ideale' Länge wird der Abstand des Startpunktes zum Sollpunkt genommen, unter Berücksichtigung der Variablen x_3, y_3 (Anhängerheck), und β (Anhängerwinkel). Der Winkel α der Zugmaschine ist als Soll undefiniert und daher nicht berücksichtigt. Um ein Längenmaß für den Winkel β zu erhalten, wird daraus die Bogenlänge b_β errechnet. Diese ergibt sich aus der Betrachtung einer reinen Drehung des Anhängers um das Heck für die Anhängerfront durch Multiplikation des Winkels mit der Länge des Anhängers: $b_\beta(\Delta\beta) = \Delta\beta l_3$. Die 'ideale' Distanz ist dann

$$d^{id} = \sqrt{\left[(x_{Soll} - x_{Start})^2 + (y_{Soll} - y_{Start})^2 + b_\beta^2(\beta_{Soll} - \beta_{Start})\right]}$$

Die Trajektorienlänge s_i wird im Antriebspunkt der Zugmaschine (am Vorderrad) abgelesen und ist daher bei konstanter Antriebsgeschwindigkeit v_1 und konstanter Regelungsschrittweite h proportional zu M_i, der jeweiligen Zahl der Regelungsschritte. Als Analysewerte stehen dann zur Verfügung die durchschnittliche Länge s und das durchschnittliche Idealverhältnis s_d. Der Wert für s_d wird für unausgeformte Zielregelungen nahe bei 1.0 liegen. Für komplexeres Regelungsverhalten bis zu einige Zehntel darüber, so daß die prozentuale Abweichung von $s_d = 1.0$ anschaulicher

erscheint.

$$\overline{s} = \frac{1}{N_T}\sum_{i=1}^{N_T} s_i = \frac{1}{N_T}\sum_{i=1}^{N_T}\left(M_i\frac{|v_1|}{h}\right)$$

$$\overline{s_d} = \left[\left(\frac{1}{N_T}\sum_{i=1}^{N_T}\frac{s_i}{d_i^{id}}\right) - 1.0\right] 100.0$$

Interpretation: Regelungs- oder Trajektorienkomplexität.

Stellsignal Zur Beurteilung der Reglerdynamik wird einmal der Durchschnitt der absoluten Stellsignalwerte und der absoluten Stellsignaländerungen aufeinanderfolgender Regelungsausgaben berechnet. Die Stellsignaländerungen werden dabei auf die Länge eines Regelungsschritts bezogen, um einen allgemeineren, von den Simulationsparametern v_1 und h unabhängigen Wert zu verwenden. Das korrigierte Stellsignal θ_k wird separat betrachtet.

$$\overline{\theta} = \frac{1}{N_T}\sum_{i=1}^{N_T}\frac{1}{M_i}\sum_{j=1}^{M_i}|\theta_{ij}|$$

$$\overline{\Delta\theta} = \frac{1}{N_T}\sum_{i=1}^{N_T}\frac{1}{M_i}\sum_{j=2}^{M_i}\frac{|\theta_{ij}-\theta_{i(j-1)}|}{h|v_1|}$$

$$\overline{\theta_k} = \frac{1}{N_T}\sum_{i=1}^{N_T}\frac{1}{n_i}\sum_{j=1}^{M_i}|\theta_{k,ij}|$$

$$\overline{\Delta\theta_k} = \frac{1}{N_T}\frac{1}{M_i}\sum_{i=1}^{N_T}\sum_{j=2}^{M_i}\frac{|\theta_{ij}-\theta_{i(j-1)}|}{h|v_1|}$$

Interpretation: Reglerdynamik.

Randbedingungen Zur Beurteilung der Verletzung von limitierenden Randbedingungen bzw. deren durch die Bewertung stattgefundenen Vermeidung, werden die separaten Stellsignalkorrekturen ausgewertet. Je Randbedingung werden die Zahl der Trajektorien mit Randbedingungs-Verletzungen, deren Häufigkeit insgesamt bzw. innerhalb der nur betroffenen Trajektorien, sowie die Soll-Ist-Differenz des Stellsignals bezüglich der jeweiligen Bereichsgrenze registriert. Es ist dabei zu beachten, daß in den Testläufen die Regelung stets ausgefahren wird. D.h. es wird bei jeder zu korrigierenden Verletzung einer Randbedingung diese Korrektur auch ausgeführt (im Gegensatz zu mancher Propagierungs-Strategie in der Lernphase). Es ergeben sich die folgenden Analysekennwerte ($< rb >$ steht stellvertretend für die spezifizierte Randbedingung, 'b' für die der Blockade, 'l' für die der Lenkwinkelbegrenzung):

$s^{<rb>}$ Die Länge der von der Verletzung der Randbedingung betroffenen Teile aller Regelungstrajektorien in Prozent der Gesamttrajektorienlänge (Summe aller Trajektorienlängen).
Interpretation: Häufigkeit von Randbedingungs-Verletzungen.

$\#^{<rb>}$ Die Zahl der von einer Verletzung einer Randbedingung betroffenen Trajektorien in Prozent der Gesamtzahl der Testbeispiele. Gibt es eine Abhängigkeit

des Auftretens von Randbedingungs-Verletzungen dergestalt, daß bei einfacher zu regelnden Startbeispielen nur wenige Korrekturen notwendig sind, so wirken sich bereits kleine Anpassungen der Reglerfunktion auf viele Regelungstrajektorien positiv aus. Ein stärkere Verringerung dieses Werts gegenüber $s^{<rb>}$ deutet dann auf eine derartige Abhängigkeit hin.
Interpretation: Startwerteabängigkeit, Häufigkeit von Verletzungen der Randbedingung.

$\tilde{s}^{<rb>}$ Die durchschnittliche Gesamtlänge der von der Verletzung der Randbedingung betroffenen Abschnitte einer Trajektorie. Hierbei werden nur die von einer Verletzung betroffenen Trajektorien betrachtet. Dies gibt einen Eindruck von der Länge der meist geschlossen aufeinanderfolgenden Korrekturvorgänge.
Interpretation: Korrekturhäufigkeit, -dauer.

$\pm\theta^{<rb>}$ Die durchschnittliche absolute Abweichung des Stellsignals vom Zulässigkeitsbereich der jeweiligen Randbedingung, wenn diese verletzt wurde.
Interpretation: Bei der Interpretation dieses Werts ist mitzubeachten, daß hier hohe Werte auch bedeuten können, daß kleinere Abweichungen gut vom Regler vermieden werden. Dann treten möglicherweise nur noch in schwierigen Situationen, bei denen kein Kompromiß mit anderen Regelungsanforderungen erlernt werden konnte, Abweichungen auf, die dann aber auch relativ groß sind.

Das Erlernen der forcierenden Randbedingung der Lenkwinkeldifferenz-Minimierung ist anhand der oben bereits vorgestellten Stellsignal-Kennwerten $\overline{\theta}$ und $\overline{\Delta\theta}$ einzuschätzen.

Wo ein Vergleich zwischen den Regelungseigenschaften zweier Konfigurationen hilfreich ist, wird zusätzlich zu den Analysewerten eine qualitative Beurteilung der tendenziellen Veränderung bezogen auf die Analysewerte einer anderen Konfiguration angegeben. Zur symbolischen Darstellung der Vergleichsresultate werden die Symbole $--$, $-$, $=-$, $\approx$, $=+$, $+$, $++$ und $\pm$ verwendet. Diese werden gemäß dem eingeschätzten Grad der Veränderung vergeben. Dabei ist bewußt keine genaue numerische Bandbreite vorgegeben. Die tendenziellen Angaben sind relativ zu den Bezugswerten und deren zu erwartender Schwankungsbandbreite zu sehen. Ungefähr kann gesagt werden, daß im Größenmaßstab einer Reduktion auf die Hälfte oder eine Vergrößerung auf das Doppelte des Bezugwerts ein $--$ bzw. $++$ vergeben wird. Geringe tendenzielle Veränderung um wenige Prozentpunkte erhalten ein $=-$ oder $=+$. Die qualitative Beurteilung erfolgt in den Kategorien

Zielregelungs-Performanz (Ziel) als Zusammenfassung der Standardabweichungen des Endpositionsfehlers. Bei Bedarf auch auch nach den Fehlervariablen β und y getrennt. Eine positive Kennzeichnung bedeutet eine Verbesserung in der Zielpositionierung, also eine Verringerung der Fehler.

Randbedingungs-Performanz (limit. RB) als Zusammenfassung der Randbedingungskennwerte für Blockade und Lenkwinkelbegrenzung. Bei Bedarf nach den einzelnen Randbedingungen getrennt. Dann ist ein b für die Blockade- und ein l für die Lenkwinkel-Randbedingung vorangestellt. Eine positive Kennzeichnung bedeutet einen geringeren Grad der Randbedingungsverletzung.

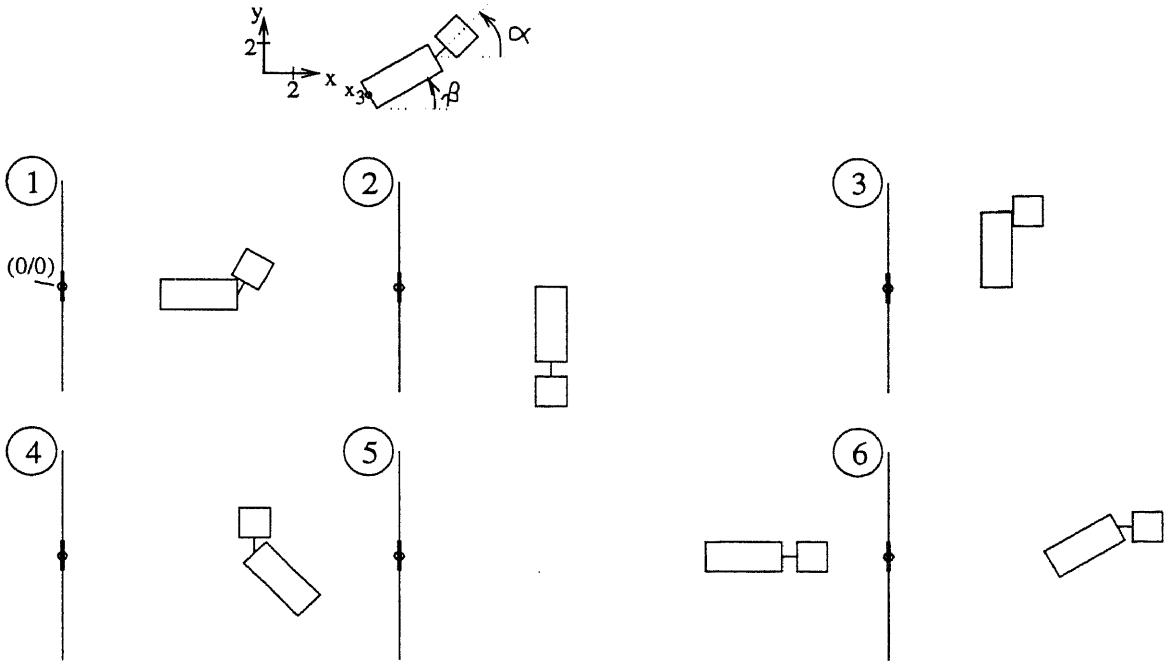


Abbildung 4.3: 6 ausgewählte Testbeispiele

Stellsignaldynamik (Dynamik) als Beurteilung für die Extremität der Stellsignalausgabe. Eine positive Kennzeichnung bedeutet eine Vergrößerung der Stellwinkelbeträge bzw. ihrer Differenzen [4].

Trajektoriencharakteristik (Trajektorie) als Beurteilung der Trajektorienlänge. Eine positive Kennzeichnung bedeutet eine Vergrößerung der Trajektorienlänge. Dies kann ein besseres, weil komplexeres Rangieren bedeuten, aber auch auf Störungen wie Pendeln um die Horizontale zum Andockpunkt hindeuten.

4.2.2 Testbilder

In Abbildung 4.3 sind die sechs zur Veranschaulichung des Regelverhaltens ausgewählten Startpositionen abgebildet. Die Positionen 1,3 und 4 entsprechen in etwa dreien bereits in [4] präsentierten Beispielen. Position 5 soll die möglichst unbeeinflußte Geradeausfahrt nachweisen. Mit Position 6 kann ein gewisser Pendelungseffekt aufgrund störender Lerneinflüsse erkannt werden (vgl. X-Positionsfehler in Kapitel 4.1.4). Die Beispiele 4 und 5 demonstrieren zudem die Fähigkeit des Reglers, aus nicht im Lern-Startbereich gelegenen Positionen heraus ins Ziel zu lenken.
Den nachfolgenden Auswertungsbildern ist zur besseren Orientierung ein Gitter mit einem $5m$-Raster unterlegt.
In den in Kapitel 4.1.4 hinzugefügten Stellsignal-Plots sind die Lenkwinkel der Regelung

[4] Die symbolische Beurteilung ist hier und bei der Trajektoriencharakteristik eine rein objektive Aussage über die Wertveränderungen. Meist bedeutet eine Vergrößerung der Stellsignal-Analysewerte auch eine komplexer ausgeformte Regelungsfunktion. In bestimmten Untersuchungen, wie beim Versuch der Begrenzung der Lenkwinkeldifferenzen, ist dagegen eine Reduzierung der θ-Analysewerte erwünscht. In diesen Fällen wird die symbolische Beurteilung ergänzend kommentiert.

über der Zeitachse abgetragen. Die starke Linie stellt das korrigierte, die schwächer ausgebildete das vom Regler ausgegebene Original-Stellsignal dar. Dieses ist nur getrennt sichtbar, wenn auch eine Korrektur stattfand. Ansonsten sind beide übereinander gezeichnet.
Außerdem sind im Stellsignal-Plot alle Regelungsschritte dokumentiert, während bei den Simulationsdarstellungen der Übersichtlichkeit wegen nur jeder zweite Regelungsschritt abgebildet ist.
Die Koordinaten für die Startpositionen sind:

1. $\alpha = 60°, \beta = 0°, x_3 = 6.3m, y_3 = -0.5m.$
2. $\alpha = -135°, \beta = -135°, x_3 = 13.9m, y_3 = 0m.$
3. $\alpha = 0°, \beta = 90°, x_3 = 7m, y_3 = 0m.$
4. $\alpha = 90°, \beta = 135°, x_3 = 16.12m, y_3 = -3.55m.$
5. $\alpha = 0°, \beta = 0°, x_3 = 20m, y_3 = 0m.$
6. $\alpha = 0°, \beta = 30°, x_3 = 10.67m, y_3 = -0.5m.$

4.3 Testkonfigurationen

In Abbildung 4.4 auf Seite 45 ist der Gesamtaufbau der für die Testkonfigurationen verwendeten Module dargestellt. Gegenüber der schematischeren Abbildung 3.1 auf Seite 20 ist hier anzumerken:

- Es sind die speziellen Bewertungen in ihrer Position und Einordnung (limitierend / forcierend), sowie die separaten Korrekturmodule angegeben.
- Zusätzlich ist die Bewertung der integralen Korrektur angeführt; diese kann aber nur alternativ zur separaten Bewertung limitierender Randbedingungen eingesetzt werden und ist hier nur der Vollständigkeit halber eingetragen. Die integrale Bewertung wird in dieser Diplomarbeit nicht untersucht.
- Die Korrekturfunktionen für die Blockade und die Lenkwinkelbegrenzung sind immer aktiviert. Die Lenkwinkelbegrenzung muß generell zur Bestimmung eines integralen Korrekturwerts trotz einer sigmoiden Begrenzung explizit nochmal vorhanden sein.
- Ein Widerspruch bei der Bestimmung des integralen Korrekturwerts für Blockade- und Lenkwinkel kann mit den hier eingestellten Randbedingungs-Parametern nicht auftreten. Daher entfällt eine Stopp-Auslösung durch das Korrekturmodul.
- Die Wertebereichstransformationen um das neuronale Netz sind miteingezeichnet. Über die Transformation an der Netzausgabe wird bei vorhandener sigmoiden Transferfunktion des Ausgabeneurons der reale Begrenzungsbereich des Lenkwinkels definiert.
- Die zur neuronalen Begrenzung des Stellsignals dienende Sigmoidfunktion des Ausgabeneurons ist separat eingezeichnet. Auf diese wird dann später beim expliziten Erlernen der Lenkwinkelbegrenzung verzichtet. Die Transferfunktion des Ausgabeneurons ist dann linear.

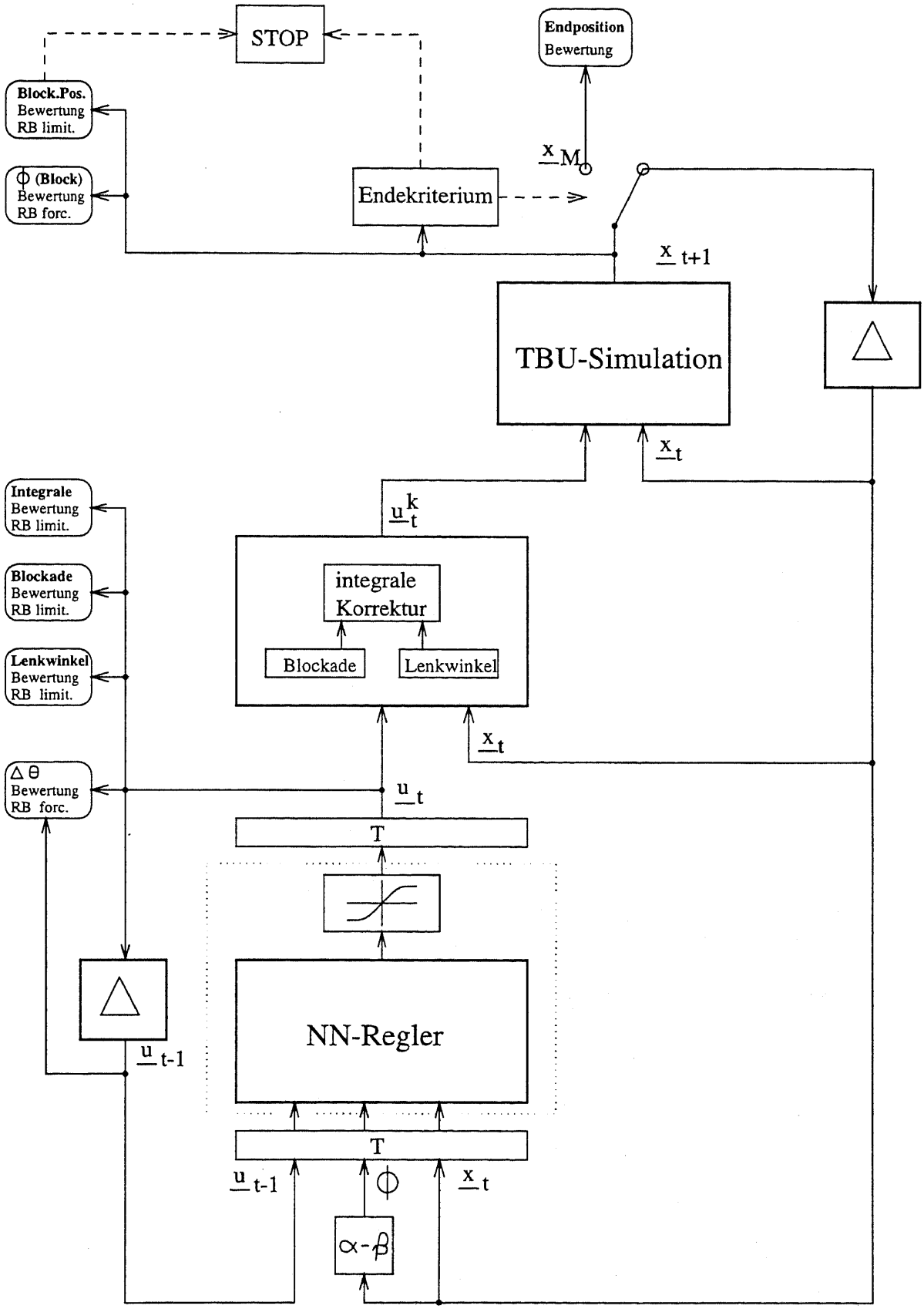


Abbildung 4.4: Überblick über alle Konfigurationsmodule

- Das bei der Lenkwinkeldifferenz zur Bereitstellung des vorigen Stellsignals dienende Verzögerungsglied ist hinzugefügt.
- Das Endekriterium prüft ein Überfahren der Begrenzungswand durch den Anhänger (Front oder Heck). Um bei einer Startpositionierung nahe der Wand und nachfolgender Kurvenfahrt nicht vorzeitig abzustoppen und so außergewöhnlich große Fehler zu produzieren, wird ein Überfahren der Wandgrenze mit der Zugmaschine toleriert.

Der Übersichtlichkeit wegen wird hier wieder von der Darstellung der engen Verknüpfung zwischen Korrektur- und Bewertungsmodulen, zusätzlicher Eingangsvariablen und der Rückpropagierungswege abgesehen. Aus einer Grundkonfiguration ohne Bewertungen zusätzlicher Randbedingungen werden die anderen Testkonfigurationen abgeleitet. Daher werden zuerst einmal die Lernparameterwerte dieser Grundkonfiguration im Detail beschrieben, um für weitere Konfigurationen nur noch Änderungen angeben zu müssen.

4.3.1 Grundkonfiguration

Der schematische Aufbau der Grundkonfiguration in Abbildung 4.5 entspricht dem in Abbildung 4.4 auf Seite 45 mit folgenden Vereinfachungen :

- Es sind bis auf die Endpositionsbewertungen keine Bewertungsmodule aktiv.
- Der Regler bekommt zudem nur den Streckenzustand x_t zugeführt.
- Der Lenkwinkel wird über die Sigmoidfunktion im Ausgabeneuron und die nachfolgende Transformation in den Realbereich auf $[-60°, +60°]$ begrenzt.

Wo nicht anders erwähnt, gelten für die Module die folgenden Einstellungswerte:

Neuronales Reglernetz

Layerzahl	3 + 1 Biasneuron		
	Layer 1	Layer 2	Layer 3
# Neuronen	4 (Eingabe)	25	1 (Ausgabe)
Lernkoeffizienten	-	2.0	1.5
Transferfunktion	linear	sigmoid	sigmoid[a]
Eingabevariablen	α β x_M y_M		
Ausgabevariable	θ		
Lernregel	Delta-Lernregel (BPtt)		

[a]Nur bei neuronaler Lenkwinkelbegrenzung, sonst linear

Im Fall der Erweiterung des Netzes um einen Hidden-Layer gelten die folgenden Grundwerte:

Layerzahl	4 + 1 Biasneuron			
	Layer 1	Layer 2	Layer 3	Layer 4
# Neuronen[a]	4 (Eingabe)	15/20/25	10/15/20	1 (Ausgabe)
Lernkoeffizienten	-	2.0	1.5	1.0
Transferfunktion	linear	sigmoid	sigmoid	sigmoid

[a]Es werden teilweise unterschiedliche Neuronenzahlen verwendet in den Kombinationen: 15+10, 20+15, 25+25.

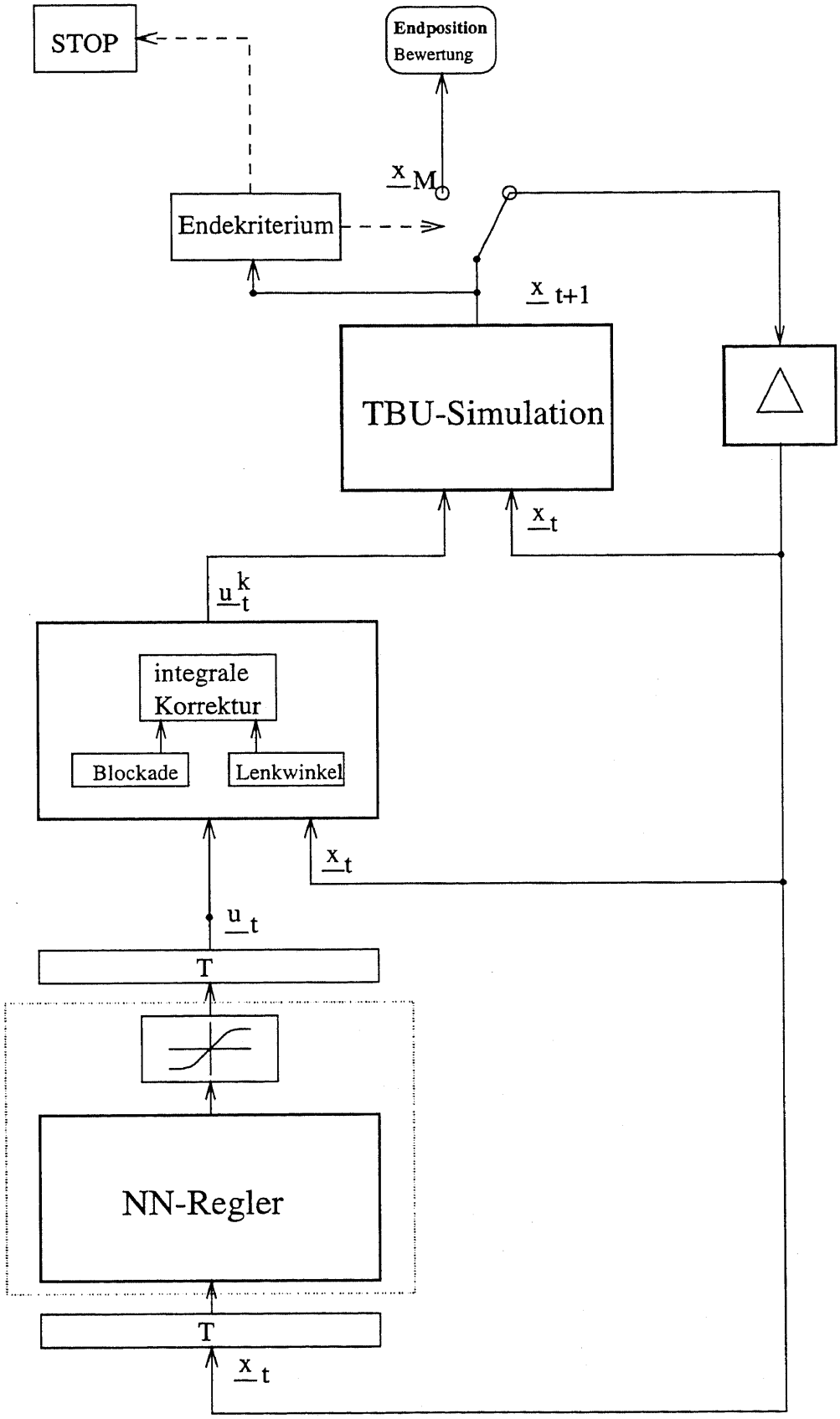


Abbildung 4.5: Grundkonfiguration ohne lokale Bewertungen

Wertebereichstransformationen

	NN-Eingang			NN-Ausgang
	θ^a, α, β	x_M	y_M	θ
Realbereich	$[-\frac{\pi}{2}, +\frac{\pi}{2}]$	$[0m, 10m]$	$[-5m, +5m]$	$[-\frac{\pi}{2}, +\frac{\pi}{2}]$
Neuronalbereich	$[-1, +1]$	$[-1, +1]$	$[-1, +1]$	$[0, 1]$

[a]Nur bei Bewertung der Lenkwinkeldifferenz

Korrekturmodule

Randbedingung	Grenzwinkel
Blockadevermeidung	$\Phi_G = \pm 90°$
Lenkwinkelbegrenzung	$\theta_G = \pm 60°$

TBU-Simulation

TBU-Modell	$l_1 = 2m$, $l_2 = 1m$, $l_3 = 5m$, $v_1 = -1\frac{m}{s}$
Regelungsschrittweite	$h = 0.5s$
Numerische Simulation	Euler-Vorwärts-Integration

Endpositionsbewertung

Bewertete Variablen	β	y_3	x_3[a]
Bewertungsgrenzen	$\beta_B = \pm 0°$	$y_{3,B} = \pm 0m$	$x_{3,B} = \pm 0m$
Fehlerskalierung[b] f_E^e	0.2	0.01	0.01
Bewertungsskalierung f_B^e	1.0		
Fehlerfunktion	quadratischer Fehler		
Anzahl BP-Zeitschritte $\#BP^e$	bis Startposition		

[a]Generell keine Bewertung negativer x-Werte. Nur wenn statischer X-Positionsfehler angesetzt wird, dann konstante negative Soll-Abweichung $\Delta x_s = -0.15m$

[b]Eine Erklärung der ungleichen Gewichtung befindet sich in Anhang B

Auswahl der Lernbeispiele (Lernplan)

Die Startposition des TBU-Fahrzeugs ist mit den aus Abbildung 4.6 auf Seite 49 ersichtlichen Koordinaten formuliert. β^* und α^* sind die bezüglich φ und $(\varphi + \beta^*)$ relativen TBU-Winkel.

Der Startbereich wird für die ersten 1000 Lernbeispiele konstant gehalten, um das 'Einschwingen' des Netzlernens zu erleichtern. Zwischen den Stützwerten wird interpoliert, um Fehlersprünge in der Zielregelung zu vermeiden. Eine Rasterung der Startwerte findet nicht statt.

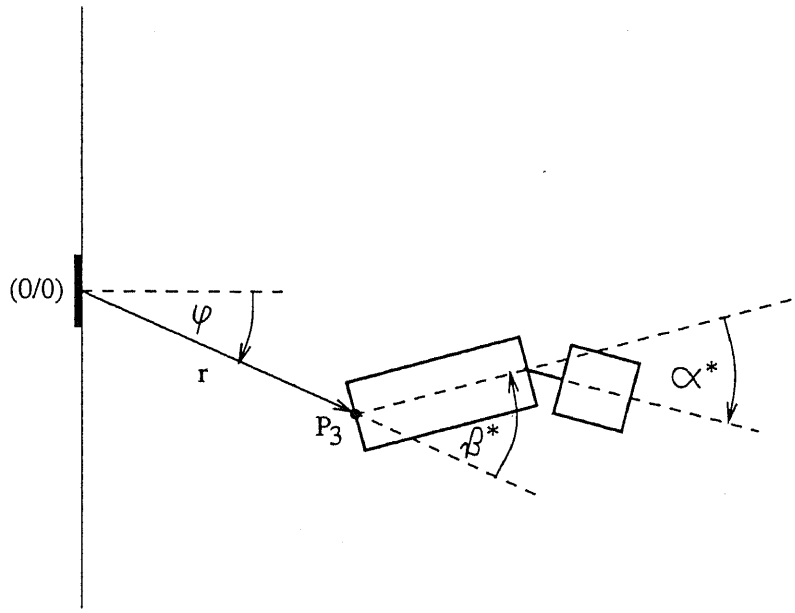


Abbildung 4.6: Koordinaten für Auswahl der Startposition

Beispielnummern (Stützstellen)	$\|\varphi\|_{max}$ [°]	r [m]..[m]	$\|\beta^*\|_{max}$ [°]	$\|\alpha^*\|_{max}$ [°]
1-1000	2.5	2.0..3.0	1	30
2000	5.0	2.2..3.5	6	34
3000	7.5	2.4..4.0	12	38
4000	10.0	2.6..4.5	18	42
5000	12.5	2.8..5.0	24	46
6000	15.0	3.0..5.5	30	50
7000	17.5	3.2..6.0	36	54
8000	20.0	3.4..6.5	42	58
9000	22.5	3.6..7.0	48	62
10000	25.0	3.8..7.5	54	66
11000	27.5	4.0..8.0	60	70
12500	30.0	4.2..8.5	66	74
14000	32.5	4.4..9.0	72	78
16000	35.0	4.6..9.5	78	82
18000	37.5	4.8..10.0	84	86
20000-22000	40.0	5.0..10.5	90	90

Folgende Abkürzungen finden für die zu variierenden Lernparameter Verwendung:

Fehlerskalierung	$f_E^{<RB>}$
Bewertungsskalierung	$f_B^{<RB>}$
Zahl der Neuronen im Hidden Layer i	$\#N_{Hi}$
Zahl der Layer im Reglernetz	$\#L$
Zahl der Rückpropagierungsschritte einer Bewertung $< RB >$	$\#BP^{<RB>}$
Statischer X-Positionsfehler	Δx_{st}

Dabei steht für den Index $< RB >$ der bewerteten Randbedingung ein **'e'** für die Endpo-

sitionsbewertung der primären Regelungsaufgabe sowie 'b', 'l' und '$\Delta\theta$' für die Blockade-, Lenkwinkelbegrenzung und Lenkwinkeldifferenz-Randbedingungen.

Für die verschiedenen Propagierungsstrategien werden die in Kapitel 3.4 auf Seite 26 eingeführten Abkürzungen verwendet, die sich auf die Behandlung der Propagierung bei Korrekturen beziehen:

FP-STOPP Stopp der Regelung bei Auftreten der ersten Korrektur. Rückpropagierung dieser Korrektur-Bewertung und aller davor ansetzenden forcierenden Bewertungen.

BP-STOPP Regelung bis Endekriterium der Zielregelung erfüllt ist. Rückpropagierung aller Bewertungen bis zum ersten Auftreten einer Korrektur auf dem Propagierungsweg.

BP-NULLUNG Regelung bis Endekriterium der Zielregelung erfüllt ist. Rückpropagierung aller Bewertungen. Bei Propagierung durch den Begrenzungsbereich einer Korrektur wird der Rückpropagierungsfehler in diesem Zweig auf Null gesetzt.

BP-PROP Berechnung der numerischen Ableitung zur Propagierung durch die Korrekturfunktion. Bei den hier verwendeten Korrekturen ist dies identisch mit der Methode der BP-NULLUNG, da die Begrenzungsfunktionen im Begrenzungsbereich die Ableitung 0 besitzen.

BP-BYPASS Statt der Ableitung oder Fehler-Nullung wird hier der zum Ausgang der Korrekturfunktion propagierte Fehler ohne Veränderung zum Eingangs-Stellsignal durchgeschleift.

4.3.2 Übersicht

Zunächst wird das Reglerlernen für die Grundkonfiguration mit Endpositionsbewertung, aber ohne zusätzliche Randbedingungs-Bewertungen betrachtet. Dabei werden sowohl die Abhängigkeit von Lernparametervariationen als auch der Einfluß der verschiedenen Propagierungsstrategien untersucht. Die Analysedaten und Regelungsbilder für die Grundkonfiguration werden als Richtlinie zur Beurteilung der anderen Tests genommen.
Es schließen sich die Behandlung der unterschiedlichen Randbedingungen an. Dabei wird auch wieder der Einfluß der Lernparametervariationen auf die Performanz des Reglers hinsichtlich der Zielpositionierung (*Ziel-Performanz*), aber auch hinsichtlich der Einhaltung der Randbedingung durch den Regler (im weiteren als *RB-Performanz* bezeichnet) ausgewertet. Die Beeinflussung der Regelungsfunktion durch den statischen x-Positionsfehler wird nur an der Grundkonfiguration analysiert, da dies eine nur auf das spezielle Regelungsproblem zugeschnittene Möglichkeit bildet.
Für die Integration der Blockade-Randbedingung in die Reglerfunktion werden die drei prinzipiellen Möglichkeiten der Bewertung (vgl. Kapitel 3.5.3 Seite 30) miteinander verglichen. An der Blockade-Randbedingung soll ebenfalls einem möglichen Effekt des in die Propagierung übertragenen Lokalisationscharakters nachgegangen werden. Letztendlich werden auch noch etwaige positive oder negative Auswirkungen einer zusätzlichen Hilfsvariable Φ des Reglers aufgezeigt.
Als nächstes wird die in der Grundkonfiguration durch die Sigmoidfunktion übernommene Lenkwinkelbegrenzung explizit versucht zu erlernen. Die dabei vorzunehmenden

Änderungen an der Konfiguration und weitere Realisierungs-Schwierigkeiten werden besonders herausgestellt. Um den Nachweis zu erbringen, daß auch mehrere lokale Optimierungskriterien vom neuronalen Regler zu berücksichtigen sind, wird die Bewertung der Blockade-Korrektur noch abschließend hinzugefügt.
Die forcierende Bewertung der Lenkwinkeldifferenz schließt die Untersuchung der zusätzlichen lokalen Bewertungen ab.

Kapitel 5

Testergebnisse

5.1 Grundkonfiguration ohne lokale Bewertungen

Ohne lokale Bewertungen der Korrekturen ist ein FP-STOPP-Lernen nicht sinnvoll, da bei Abbruch in der Korrekturposition keine Zielpositionierung erlernt wird. Wiederholt sich dies in zunehmendem Maße, wie das bei fortschreitendem Lernen geschieht, so bleibt eine große Zahl der Lernbeispiele aufgrund der unvollständigen Regelungen unberücksichtigt, was die Entwicklung der Zielregelung beeinträchtigt. Aus diesem Grund wird diese Methode hier nicht betrachtet.
Die im folgenden präsentierten Beispiele basieren auf der BP-STOPP-Methode. Die Grundeinstellung der Endpositionsbewertung ist gemäß Kapitel 4.3.1:

Konfiguration O1 Keine Lokalen Bewertungen			
Propagierungsstrategie	BP-STOPP		
Bewertete Variablen	β	y_3	x_3
Bewertungsgrenzen	$\beta_B = \pm 0°$	$y_{3,B} = \pm 0m$	$x_{3,B} = \pm 0m$
Fehlerskalierung f_E^e	0.2	0.01	0.01
Bewertungsskalierung f_B^e	1.0		
Anzahl BP-Zeitschritte $\#BP^e$	bis Startposition[a]		

[a]Dies ist die von der Wahl der Propagierungsmethode unabhängige Voreinstellung.

Der damit gelernte Regler weist die in Tabelle 5.1 eingetragenen Analysewerte auf. Lenkwinkelverletzungen treten wegen der sigmoiden Begrenzung im Reglernetz nicht auf. Erkennbar sind aber zahlreiche Blockadekorrekturen, die durch die fehlende Berücksichtigung im Reglerlernen auftreten. Abbildung 5.1 zeigt die zugehörigen Testbeispiel-Regelungen.
Im Vergleich mit den von Widrow/Nguyen in [4] vorgestellten Regelungen fällt das hier erzielte Ergebnis schlechter in der Zielperformanz aus. Dies mag zum einen am hier ausgeschlossenen Einfluß des negativen x-Positionsfehlers liegen, der die Zielregelung positiv beeinflußt. Zum anderen ist der Aufbau des Lernplans leicht verschieden und die in [4] gewählten Modell- und Lernparametereinstellungen nicht bekannt. Der Vergleich oder eine bestmögliche Optimierung des Reglers soll hier jedoch keine primäre Beachtung finden.

Tabelle 5.1: Analyse

O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

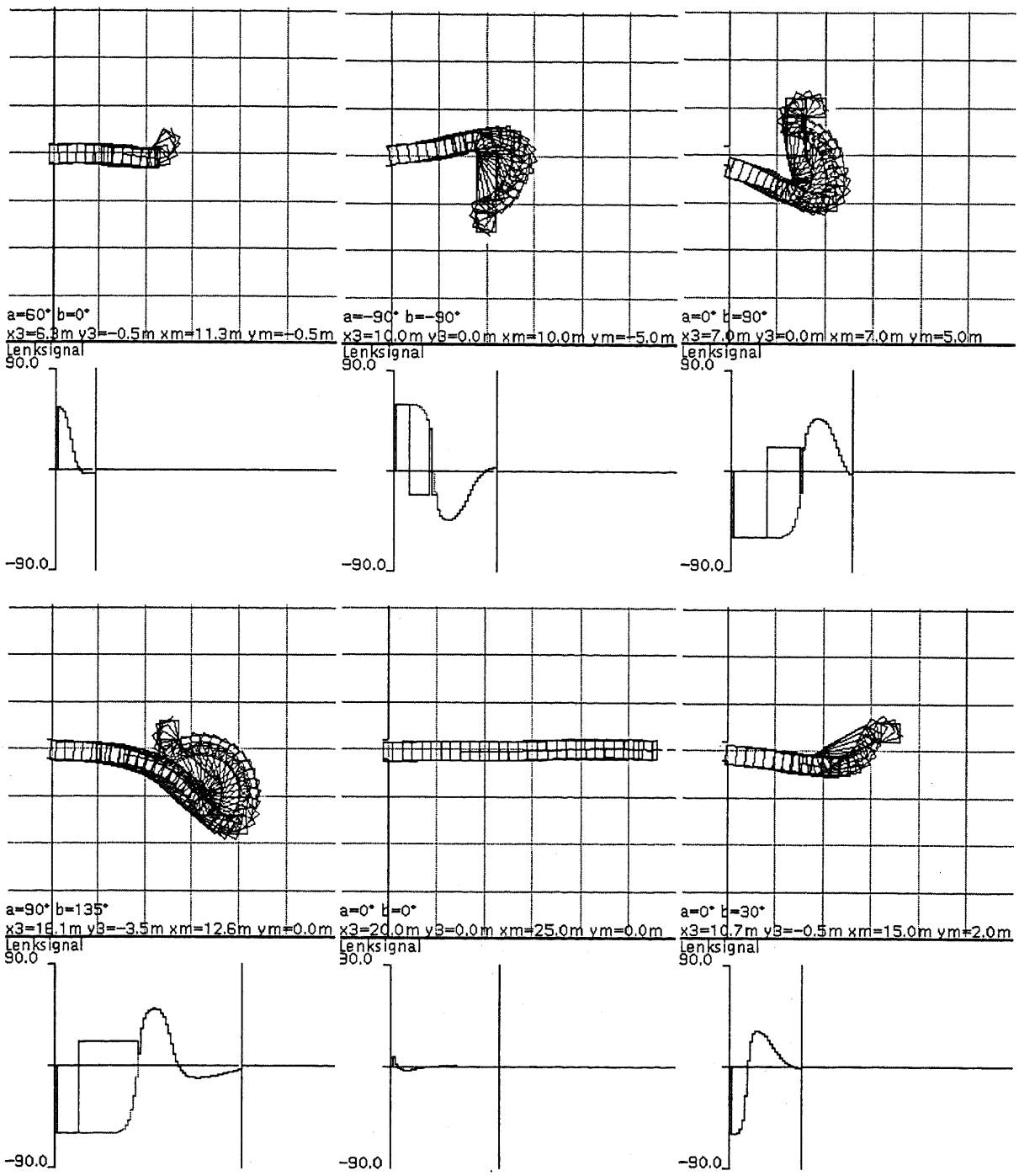


Abbildung 5.1: Beispielregelungen Konfiguration O1

Tabelle 5.2: Analyse

O1 mit BP-PROP							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.80°	$0.88m$	$12.69m$	8.18%	29.96°	$9.38\frac{°}{m}$	28.48°	$10.83\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.07%	47.00%	$2.72m$	54.69°	0%	0%	$0m$	0°

Tendenz bzgl. O1			
Ziel	limit. RB	Dynamik	Trajektorie
$\approx$	$\approx$	$\approx$	$= -$

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

5.1.1 Variation der Propagierungsstrategie

Die Wahl von BP-PROP oder BP-BYPASS als Propagierungsstrategien verändert die Resultate nur wenig, s. Tabellen 5.2 und 5.3.
Die unwesentlichen Veränderungen bei Anwendung der BP-PROP sind wohl darauf zurückzuführen, daß sich die nach einer Korrekturstelle fortgesetzte Propagierung solange nicht auf die Gewichtsveränderungen auswirkt, wie Korrekturen durchlaufen werden (da im Sättigungsbereich der Korrekturfunktionen die Fehler zu Null gesetzt werden). Betrachtet man die typischen Lenksignalverläufe (Abbildung 5.10 Seite 55), dann erkennt man, daß die Korrekturen meist zu Beginn der Regelung auftreten. Eine Rückpropagierung der Endpositionsbewertung hat dann, bei auftretenden Korrekturen, nur noch wenig Einfluß auf diese frühen Regelungsbereiche. Zudem liegen die ersten Reglerausgaben oft im Sättigungsbereich der Sigmoidfunktion des Ausgabeneurons. Dadurch wird der zu propagierende Fehler nochmals stark vermindert und bewirkt nur geringe Gewichtsänderungen im Reglernetz.
Die deutlicher ausfallende Verbesserung der Endpositionierung im Fall der BP-BYPASS-Methode belegt die in Kapitel 3.4 auf Seite 27 geäußerte Vermutung einer zu extremeren Reglerausgaben führenden Wirkung des Fehlerdurchschleifens durch die Korrekturfunktion. Die Vergrößerung der Trajektorienlänge deutet ebenfalls auf eine bessere Ausformung der TBU-Bahn hin. Allerdings ist diese Verbesserung in den Beispielplots kaum sichtbar. Ein Beispiel,in dem wenigstens die Veränderung in der Blockade-Vermeidung sichtbar wird, ist in Abbildung 5.2 zu sehen.

Tabelle 5.3: Analyse

O1 mit BP-BYPASS							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.49°	$0.86m$	$12.81m$	9.14%	30.87°	$9.79\frac{°}{m}$	28.90°	$11.29\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
11.28%	49.60%	$2.91m$	59.23°	0%	0%	$0m$	0°

Tendenz bzgl. O1			
Ziel	limit. RB	Dynamik	Trajektorie
$=+$	$-$	$=+$	$=+$

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

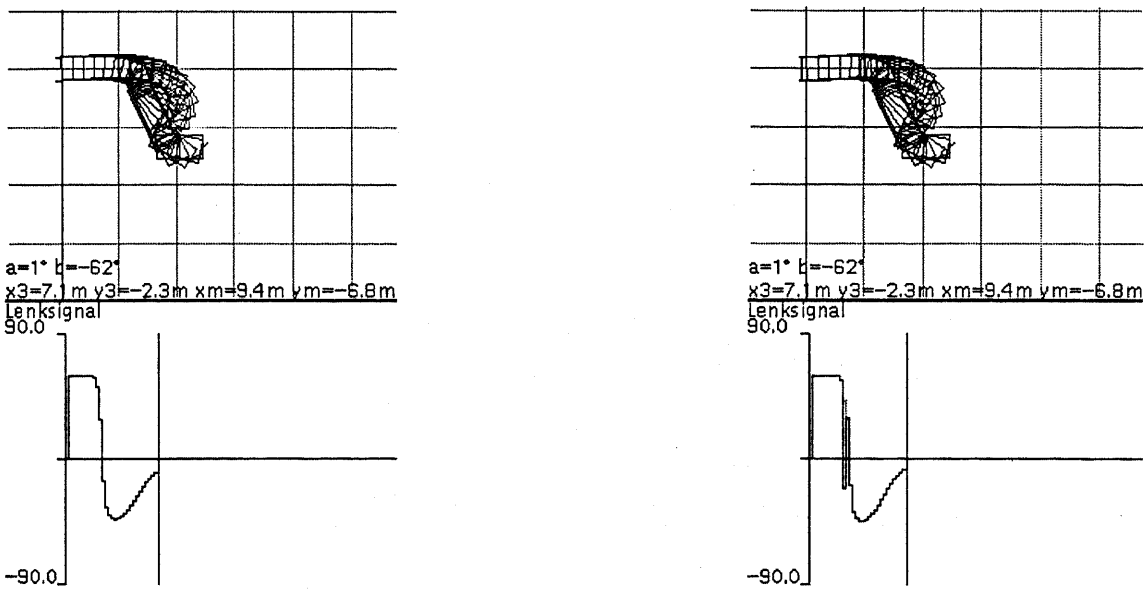


Abbildung 5.2: Effekt der BP-BYPASS-Methode (rechts) für O1 im Vergleich mit BP-STOPP (links)

5.1.2 Variation der Lernparameter

Unabhängig von der Wahl der angesprochenen Propagierungsstrategien sind die tendenziellen Auswirkungen der Lernparameter-Veränderungen für die Konfiguration O1.

Bewertungsskalierung und Netzstruktur

Das Reglerlernen zeigt sich relativ unbeeinflußt von Variationen der Bewertungsskalierung und der Zahl der Neuronen im Hidden Layer. Die Analysedaten für zusätzliche Bewertungsskalierungen mit den Werten $f_B^e = 0.5$ und $f_B^e = 2.0$ (Tabelle 5.4) und Konfigurationen mit $\#N_{H1} = 30$ bzw. $\#N_{H1} = 35$ Neuronen (Tabelle 5.5) entsprechen in etwa denen der Konfiguration O1 (auf eine bildliche Darstellung von Beispielregelungen wird verzichtet).

Der Zusatz eines Hidden-Layers zeigt eine Verbesserungsmöglichkeit in der Zielperformanz. Bei einer Konfiguration mit $\#N_{H1} = 20$ und $\#N_{H2} = 15$ Neuronen bekommt man das in Tabelle 5.6 zusammengefaßte Ergebnis.

Tabelle 5.4: Analyse

O1 mit $f_B^e = 2.0$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.78°	$0.88m$	$12.70m$	8.32%	30.33°	$9.56\frac{°}{m}$	28.62°	$11.01\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.36%	47.70%	$2.76m$	57.34°	0%	0%	$0m$	0°
Tendenz bzgl. O1							
Ziel		limit. RB		Dynamik		Trajektorie	
$\approx$		$=+$		$\approx$		$\approx$	

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

Tabelle 5.5: Analyse

O1 mit $\#N_{H1} = 35$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.81°	$0.83m$	$12.73m$	8.59%	30.07°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.16%	46.80%	$2.76m$	55.26°	0%	0%	$0m$	0°
Tendenz bzgl. O1							
Ziel		limit. RB		Dynamik		Trajektorie	
$\approx$		$\approx$		$\approx$		$\approx$	

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

Tabelle 5.6: Analyse

O1 mit BP-PROP, $\#N_{H1} = 20$, $\#N_{H2} = 15$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
10.09°	$0.85m$	$12.90m$	10.10%	29.91°	$9.24\frac{°}{m}$	28.76°	$10.52\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
9.46%	44.50%	$2.74m$	51.33°	0%	0%	$0m$	0°
Tendenz bzgl. O1							
Ziel		limit. RB		Dynamik		Trajektorie	
+		$\approx$		$\approx$		+	
Besonders der β-Fehler verbessert sich, während der y-Fehler nahezu konstant bleibt. Die Trajektorie wird insgesamt länger. Die Geradeausfahrt (Abbildung 5.3) weist jedoch starke Mängel auf. Die längeren Trajektorien sind auf Störungen in diesem Bereich zurückzuführen.							

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

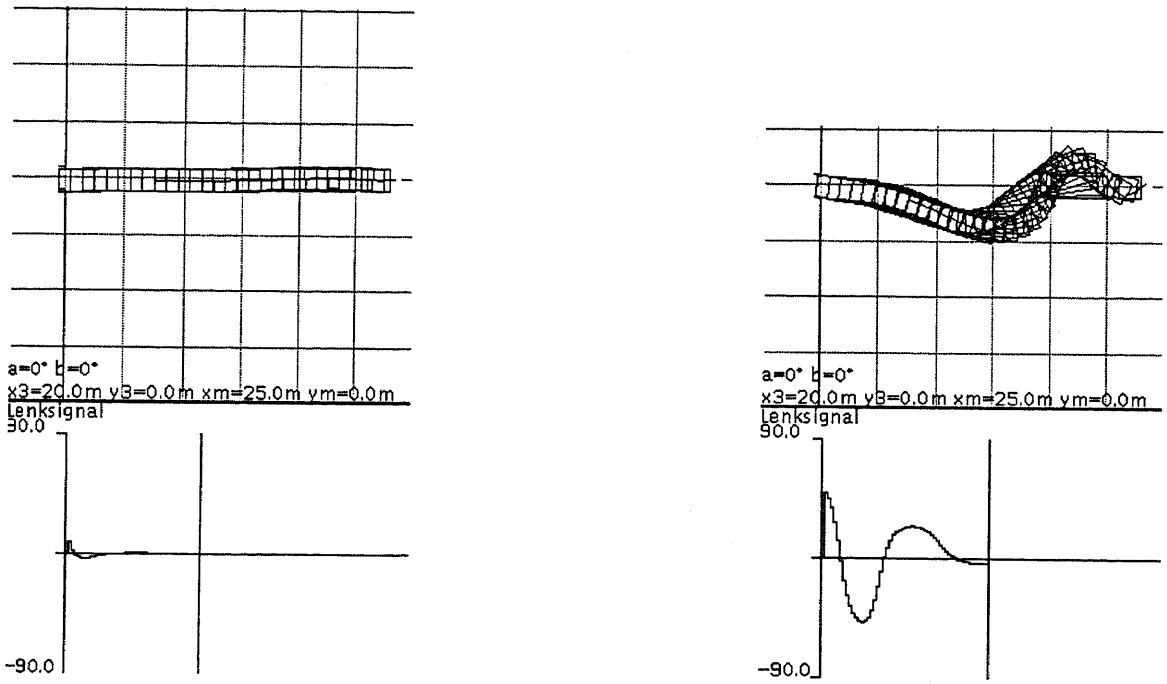


Abbildung 5.3: Störungen in der Geradeausfahrt mit 2. Hidden Layer (rechts)

Statischer X-Positionsfehler

Eine geringe Verbesserung der Zielregelung wird durch einen festen x-Positionsfehler $\Delta x_{st} = -0.15m$ erreicht. Durch die etwas extremeren Lenkbewegungen (engere Kurvenfahrten) treten in der Folge auch vermehrt Korrekturen auf (Tabelle 5.7).

Tabelle 5.7: Analyse

O1 mit $\Delta x_{st} = -0.15m$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.34°	$0.70m$	$13.05m$	11.46%	31.07°	$9.73\frac{°}{m}$	29.15°	$11.21\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
11.68%	50.10%	$3.04m$	57.26°	0%	0%	$0m$	0°
Tendenz bzgl. O1							
Ziel		limit. RB		Dynamik		Trajektorie	
+		−		= +		+	
Das gering stärkere Einschwenken in die Horizontale (vgl. Abbildung 5.4) ist begleitet von einer etwas größeren Lenkdynamik. Die dadurch verbesserte Zielpositionierung wird mit einer Verschlechterung der Einhaltung der Randbedingungen erkauft.							

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

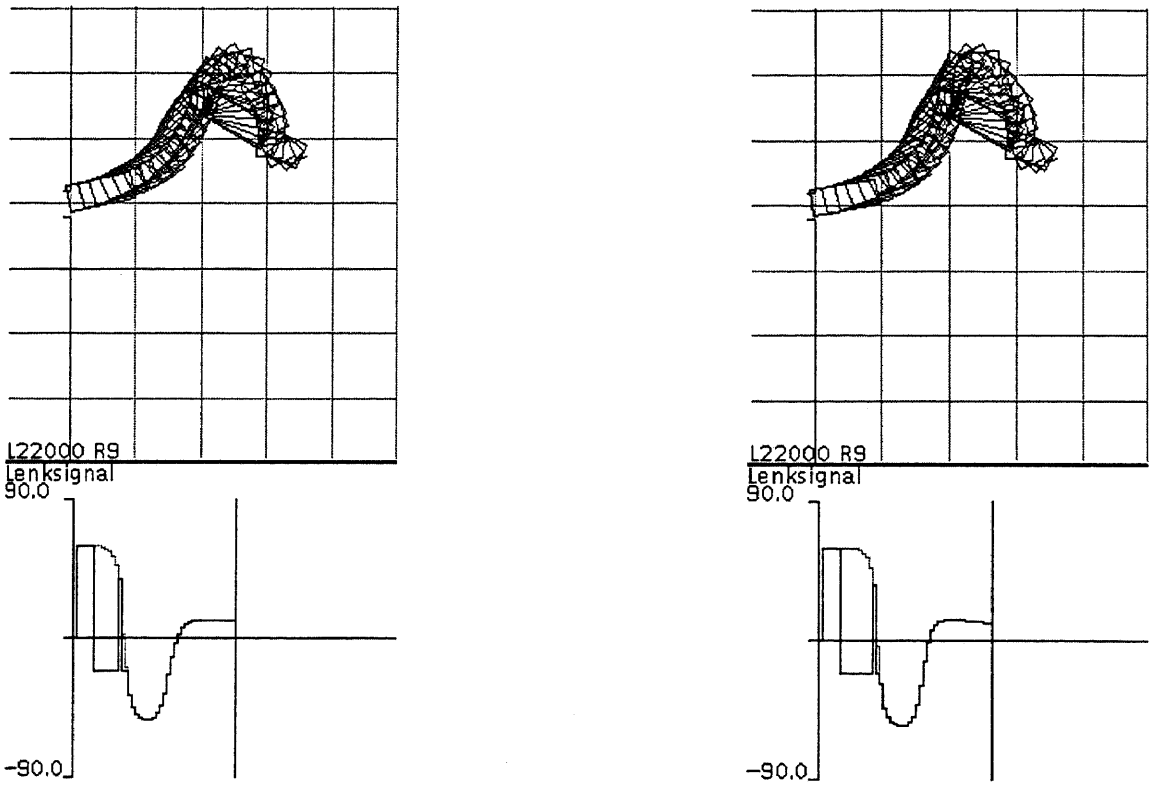


Abbildung 5.4: Effekt des statischen x-Positionsfehlers für O1. Rechts mit $\Delta x = -0.15$

5.1.3 Zusammenfassung

Die mit der Konfiguration O1 erzielte Regelung scheint sich in bezug auf die vorgenommenen Lernparametervariationen in einem relativ stabilen Zustand zu befinden. Auch durch einen zusätzlichen Hiddenlayer wird die Regelung aufgrund der dann schlechteren Bahnstabilität nicht qualitativ verbessert.
Die Abhängigkeit der Reglerentwicklung von der gewählten Propagierungsstrategie scheint den Annahmen in Kapitel 3.4 zu entsprechen, wenn auch nur sehr geringe Auswirkungen erkennbar sind.

Im weiteren gilt es nun, die durch die verschiedenen Bewertungen zusätzlicher lokaler Randbedingungen erwarteten Zielperformanz-Einbußen und den Grad der Verbesserung RB-Performanz größenmäßig zu erfassen. Danach soll der Einfluß der Lernparameter auf eine Bewahrung der ursprünglichen Zielperformanz bei gleichzeitiger Einhaltung der lokalen Randbedingungen untersucht werden.

5.2 Lokale Bewertungen zur Blockadevermeidung

Als Signal für eine Blockade gilt hier für alle betrachteten Fälle eine betragsmäßig größer als $\Phi_G = 90°$ werdende Winkeldifferenz $\Phi = \alpha - \beta$ zwischen Anhänger und Zugmaschine.

5.2.1 Bewertung des Stellsignals

Für die an der Reglerausgabe ansetzende Bewertung des Stellsignals hinsichtlich einer Vermeidung der Blockade wird für die Ausgangskonfiguration die Konfiguration O1 (Seite 53) mit folgendem hinzugefügtem Bewertungsmodul angesetzt:

Konfiguration B1 Blockade-Bewertung des Stellsignals	
Propagierungsstrategie	BP-STOPP
Bewertete Variablen	θ
Bewertungsgrenzen[a]	$\theta_B \in \{\check{\theta}_G, \hat{\theta}_G\}$
Fehlerskalierung f_E^b	1.0
Bewertungsskalierung f_B^b	0.1
Anzahl BP-Zeitschritte $\#BP^b$	1
Endpositionsbewertung wie	O1

[a] vgl. Kapitel 2.2.2 Seite 15.

Die Bewertungsgrenzen stimmen also mit den Grenzen des Zulässigkeitsbereichs der zugehörigen Blockadekorrektur überein. D.h. es findet erst bei einer aktiven Korrektur eine Bewertung der Reglerausgabe statt.
Der damit gelernte Regler weist die in Tabelle 5.8 zusammengefaßten Analysewerte auf.

Tabelle 5.8: Analyse

B1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.73°	$1.24m$	$12.27m$	4.48%	26.50°	$7.54\frac{°}{m}$	26.55°	$7.61\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
0.70%	5.50%	$1.57m$	14.67°	0%	0%	$0m$	0°

Tendenz bzgl. O1			
Ziel	limit. RB	Dynamik	Trajektorie
$\beta :\approx \; y : -$	$++$	$-$	$-$

Die Blockade wird sehr gut vermieden. Die Häufigkeit ist gegenüber O1 um etwa ein Zehntel vermindert. Dies wird jedoch mit einer schlechteren y-Positionierung erkauft. Insgesamt sind die Lenkwinkelfolgen glatter, die Lenkwinkel nicht so extrem. Die kürzeren Trajektorien weisen auf weniger ausgeprägte Kurvenfahrten hin. In Abbildung 5.5 sind die Veränderungen gegenüber O1 besonders an den Beispielen 2,3 und 4 erkennbar.

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

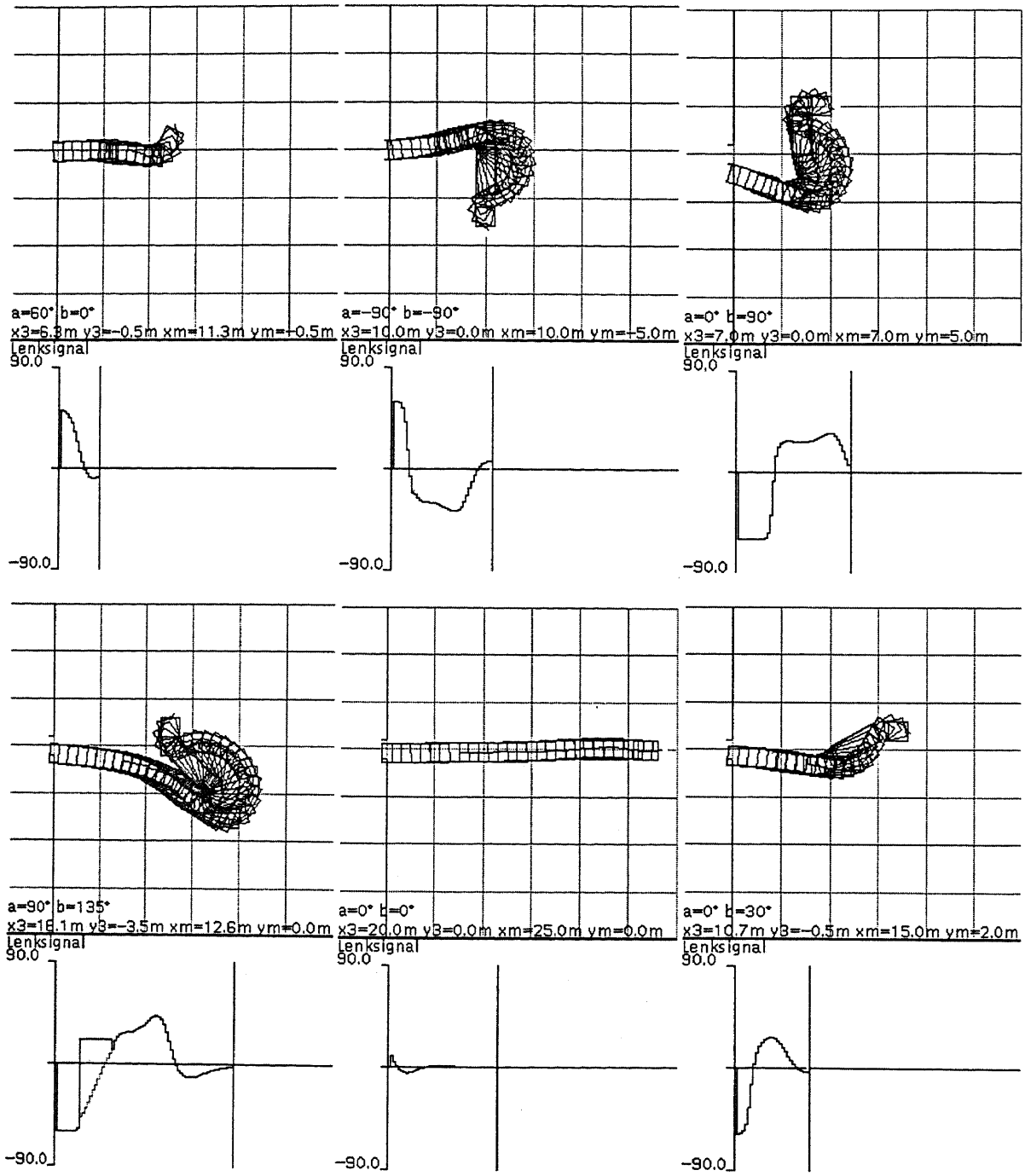


Abbildung 5.5: Beispielregelungen Konfiguration B1

5.2.1.1 Variation der Propagierungsstrategie

Ein Unterschied mit den schon bei der Konfiguration O1 getesteten Propagierungsstrategien BP-PROP und BP-BYPASS in den Analysewerten ist auch hier bei Variation der Konfiguration B1 kaum zu beobachten. Eine Extremisierung der Reglerausgabe mit der BP-BYPASS-Strategie macht sich allerdings erst bei höheren Bewertungsskalierungen bemerkbar.
Interessant ist, daß mit der FP-STOPP-Methode, bei der ja pro Lernbeispiel jeweils nur eine Bewertung einer Blockadekorrektur stattfindet, mit den sonst gleichen Konfigurationswerten ein in etwa ähnliches Ergebnis erzielt wird, wie mit der Konfiguration B1 (s. Tabelle 5.9).

Tabelle 5.9: Analyse

B1 mit FP-STOPP							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.52°	1.14m	12.37m	5.46%	26.42°	7.64$\frac{°}{m}$	26.48°	7.72$\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
0.81%	6.60%	1.52m	15.11°	0%	0%	0m	0°

Tendenz bzgl. B1			
Ziel	limit. RB	Dynamik	Trajektorie
$=+$	$=-$	$\approx$	$+$
Die etwas bessere Zielpositionierung und etwas schlechtere RB-Performanz haben ihre Ursache in der geringeren Zahl von Bewertungspropagierungen.			

Referenz: B1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.73°	1.24m	12.27m	4.48%	26.50°	7.54$\frac{°}{m}$	26.55°	7.61$\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
0.70%	5.50%	1.57m	14.67°	0%	0%	0m	0°

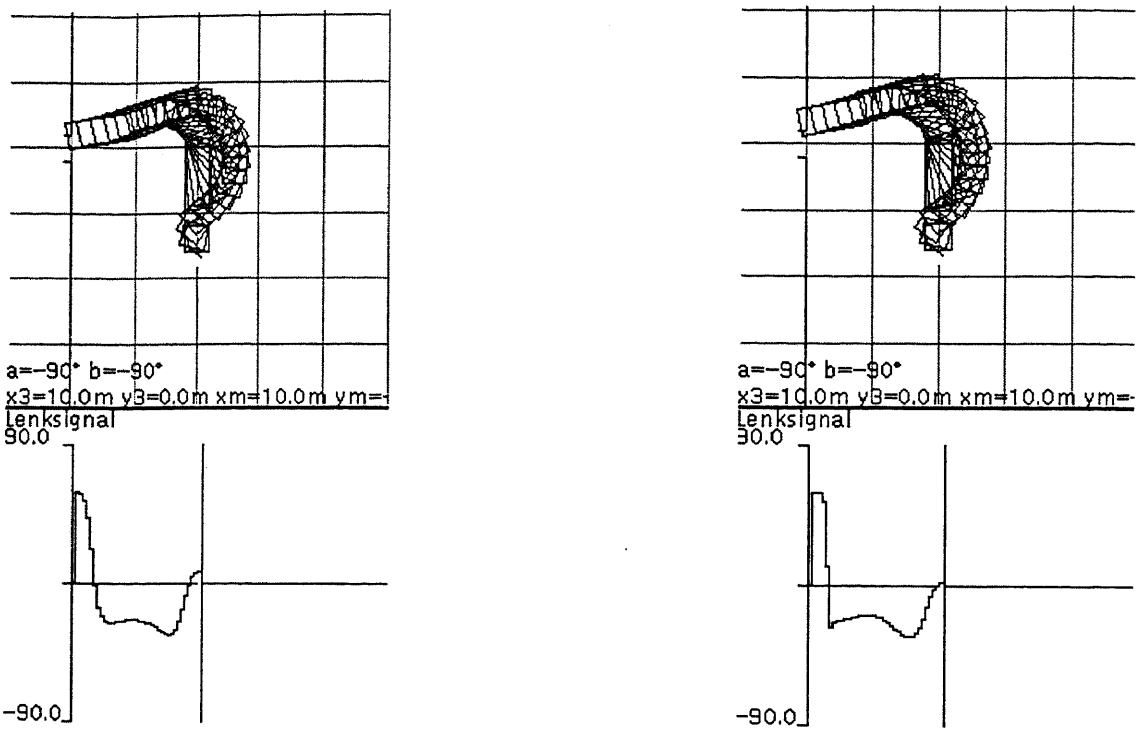


Abbildung 5.6: Testbeispiel 2 nach vergrößerter Bewertungsskalierung. Links B1 mit BP-PROP, rechts mit BP-BYPASS

5.2.1.2 Variation der Lernparameter

Eine Vergrößerung des Skalierungwerts der Blockadebewertung um die Faktoren 2 und 4 wirkt generell auf eine Verbesserung der RB-Performanz und eine weitere Verschlechterung der Zielperformanz hin. Jedoch gibt es zwischen den Propagierungsstrategien im Grad der Veränderung der Werte kleine Unterschiede.
Mit der BP-PROP-Strategie gelang sogar eine auf den Testdatensatz bezogene völlige Blockadevermeidung bei einer Bewertungsskalierung von $f_B^b = 0.4$. Trotz einer wesentlich gesteigerten Stellsignaldynamik wurde dieses annähernd auch mittels BP-BYPASS erreicht, was bedeutet, daß diese Strategie nicht unbedingt zu stärkeren Verletzungen von Stellsignal-Randbedingungen führen muß, wie bei der Konfiguration O1 angedeutet. Daß durch die starke Betonung der Blockadevermeidung die Zielpositionierung sich negativ verändert, zeigt beispielsweise das in Abbildung 5.6 vorgestellte Testbeispiel Nr. 2 für BP-PROP und BP-BYPASS.
Erwartungsgemäß fallen die Performanzänderungen, aufgrund der nur einmaligen Bewertung pro Beispiel, bei der FP-STOPP-Strategie geringer aus.

5.2.1.3 Variation der Netzstruktur

Der Versuch, mittels einer Netzstrukturerweiterung die Einbußen in der Zielpositionierung unter Beibehaltung der RB-Performanz zu kompensieren, gelingt nur zum Teil. Als Beispiel seien hier die Variationen der mittels BP-PROP trainierten Netze mit $\#N_{H1} = 30$ Neuronen und des 2-Hidden-Layer-Netzes mit $\#N_{H1} = 20$, $\#N_{H2} = 15$ vorgestellt. Diese

Tabelle 5.10: Analyse

B1 mit BP-PROP und $\#N_{H1} = 30$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.18°	$1.14m$	$12.53m$	6.70%	27.43°	$8.28\frac{°}{m}$	27.50°	$8.35\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
0.93%	8.00%	$1.46m$	11.96°	0%	0%	$0m$	0°

Tendenz bzgl. B1			
Ziel	limit. RB	Dynamik	Trajektorie
+	= −	+	+

Referenz: B1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.73°	$1.24m$	$12.27m$	4.48%	26.50°	$7.54\frac{°}{m}$	26.55°	$7.61\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
0.70%	5.50%	$1.57m$	14.67°	0%	0%	$0m$	0°

lieferten innerhalb der untersuchten Konfigurationen zur Netzstruktur-Abhängigkeit die besten Ergebnisse (Tabellen 5.10 und 5.11).

Tabelle 5.11: Analyse

B1 mit BP-PROP und $\#N_{H1} = 20$, $\#N_{H2} = 15$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
10.35°	$1.18m$	$12.52m$	6.76%	26.73°	$7.93\frac{°}{m}$	26.81°	$8.00\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
1.12%	8.50%	$1.65m$	12.03°	0%	0%	$0m$	0°
Tendenz bzgl. B1							
Ziel		limit. RB		Dynamik		Trajektorie	
+		−		+		+	
Die Verbesserung der β-Performanz fällt hier noch stärker aus. Es interessiert aber vor allem die Kompensation des Verlusts in der y-Performanz bzgl. O1, was hier mit 2 Hidden Layern ebenfalls nicht erreicht wird.							

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

5.2.1.4 Zusätzlicher Reglereingang Φ

Da die Funktion zur Vermeidung der Blockade auf einer Abhängigkeit von θ und Φ beruht, liegt es nahe, dem neuronalen Regler zusätzlich zum TBU-Zustandsvektor $(\alpha, \beta, x_M, y_M)$ noch den Winkel Φ zu präsentieren. Dabei wird ein vereinfachter Prozeß der Regleradaption erhofft, da der indirekte Zusammenhang über die Differenz von α und β entfällt.
Für die Grundkonfiguration B1 mit dem zusätzlichen Eingang Φ ergibt sich auch eine Verbesserung in der y-Performanz. Allerdings verschlechtert sich die Blockadevermeidung um 50%, was aber angesichts des Unterschieds zur Konfiguration O1 ohne Blockadebewertung noch akzeptabel sein dürfte (Tabelle 5.12).

Tendenziell verhält sich diese Konfiguration bei Variation der anderen Parameter ähnlich wie B1 mit folgenden Unterschieden:

- Die Vergrößerung der Bewertungsskalierung f_B^b wirkt sich etwas stärker auf die RB-Performanz aus, so daß ebenfalls beinahe die Blockade für den Testdatensatz ganz vermieden wird (Wert: $s^b = 0.09\%$).
- Eine Vergrößerung der Neuronenzahl hat hier sowohl bei der Ziel- als auch bei der RB-Performanz geringe positive Wirkung.

Der erhoffte Qualitätssprung wurde aber trotz dieser Unterschiede nicht erreicht.

Tabelle 5.12: Analyse

B1 mit zusätzlichem NN-Eingang Φ							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.75°	$1.13m$	$12.39m$	5.58%	26.43°	$7.85\frac{°}{m}$	26.51°	$7.94\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
1.15%	8.50%	$1.67m$	12.35°	0%	0%	$0m$	0°
Tendenz bzgl. B1							
Ziel		limit. RB		Dynamik		Trajektorie	
$\beta :\approx y : +$		$-$		$= +$		$= +$	

Referenz: B1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.73°	$1.24m$	$12.27m$	4.48%	26.50°	$7.54\frac{°}{m}$	26.55°	$7.61\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
0.70%	5.50%	$1.57m$	14.67°	0%	0%	$0m$	0°

5.2.2 Bewertung des Blockadewinkels

Statt der Bewertung des Reglerausgangs bei Auftreten einer Blockade-Korrektur wird nun die Bewertung am Ausgang TBU-Simulationsmodul angesetzt. Dort wird der aus α_{t+1} und β_{t+1} berechnete Winkel Φ_{t+1} einer forcierenden Bewertung unterzogen. Die Korrektur der Blockade ist weiterhin aktiv.
Die Bewertungsgrenze wird bei $\Phi_B = \pm 80°$ gewählt, so daß innerhalb des Sicherheitsbereichs $[-90°, -80°]$ und $[80°, 90°]$ der Blockaderegelung über die Bewertung entgegengewirkt wird.

Konfiguration B2 Blockade-Bewertung des Winkels Φ	
Propagierungsstrategie	BP-STOPP
Bewertete Variablen	Φ
Bewertungsgrenzen	$\Phi_B = \pm 80°$
Fehlerskalierung $f_E^{b,\Phi}$	1.0
Bewertungsskalierung $f_B^{b,\Phi}$	0.1
Anzahl BP-Zeitschritte $\#BP^{b,\Phi}$	1
Endpositionsbewertung wie	O1

Aufgrund der Unterschiedlichkeit der Bewertung kann nicht von vornherein davon ausgegangen werden, daß der der Konfiguration B1 entnommene Skalierungswert hier ähnlich geeignet ist. Verglichen mit dem Ergebnis von B1 ist das von B2 in der RB-Performanz auch nicht überraschend ungenügend (s. Tabelle 5.13).
Durch eine Vervierfachung der Bewertungsskalierung auf $f_B^{b,\Phi} = 0.4$ ändert sich dies jedoch bedeutend, allerdings auch auf Kosten der Zielpositionierung in y, wie aus den in Tabelle 5.14 zusammengefaßten Analysewerten und den Testbildern in Abbildung 5.7 ersichtlich.
Eine Verbesserung der RB-Performanz durch Einsatz eines Reglernetzes mit mehr Neuronen hatte keinen Erfolg. Für $\#N_{H1} = 30$ und $\#N_{H1} = 35$ ergaben sich nur unwesentliche Differenzen, und für $\#N_{H1} = 50$ scheiterte das Lernen.

Tabelle 5.13: Analyse

B2							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
12.64°	$1.10m$	$12.27m$	4.44%	27.87°	$8.49\frac{°}{m}$	27.23°	$9.57\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
6.68%	35.60%	$2.30m$	48.23°	0%	0%	$0m$	0°
Tendenz bzgl. O1							
Ziel		limit. RB		Dynamik		Trajektorie	
−		+		−		−	

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

Tabelle 5.14: Analyse

B2 mit $f_B^{b,\Phi} = 0.4$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
10.62°	$1.43m$	$12.07m$	2.65%	25.63°	$7.27\frac{°}{m}$	25.66°	$7.36\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
0.64%	5.40%	$1.44m$	21.86°	0%	0%	$0m$	0°

Tendenz bzgl. O1			
Ziel	limit. RB	Dynamik	Trajektorie
$\beta: + \; y: --$	$++$	$-$	$--$
Die Verbesserung in der β-Positionierung ist wohl in erster Linie ein Effekt der schlechteren y-Positionierung verbunden mit einer in Kurven weniger aufwendigen Regelungsbahn. Dadurch wird eine frontalere Ausrichtung in Punkten $y \neq 0$ einfacher und so eine bessere β-Performanz erreicht.			

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

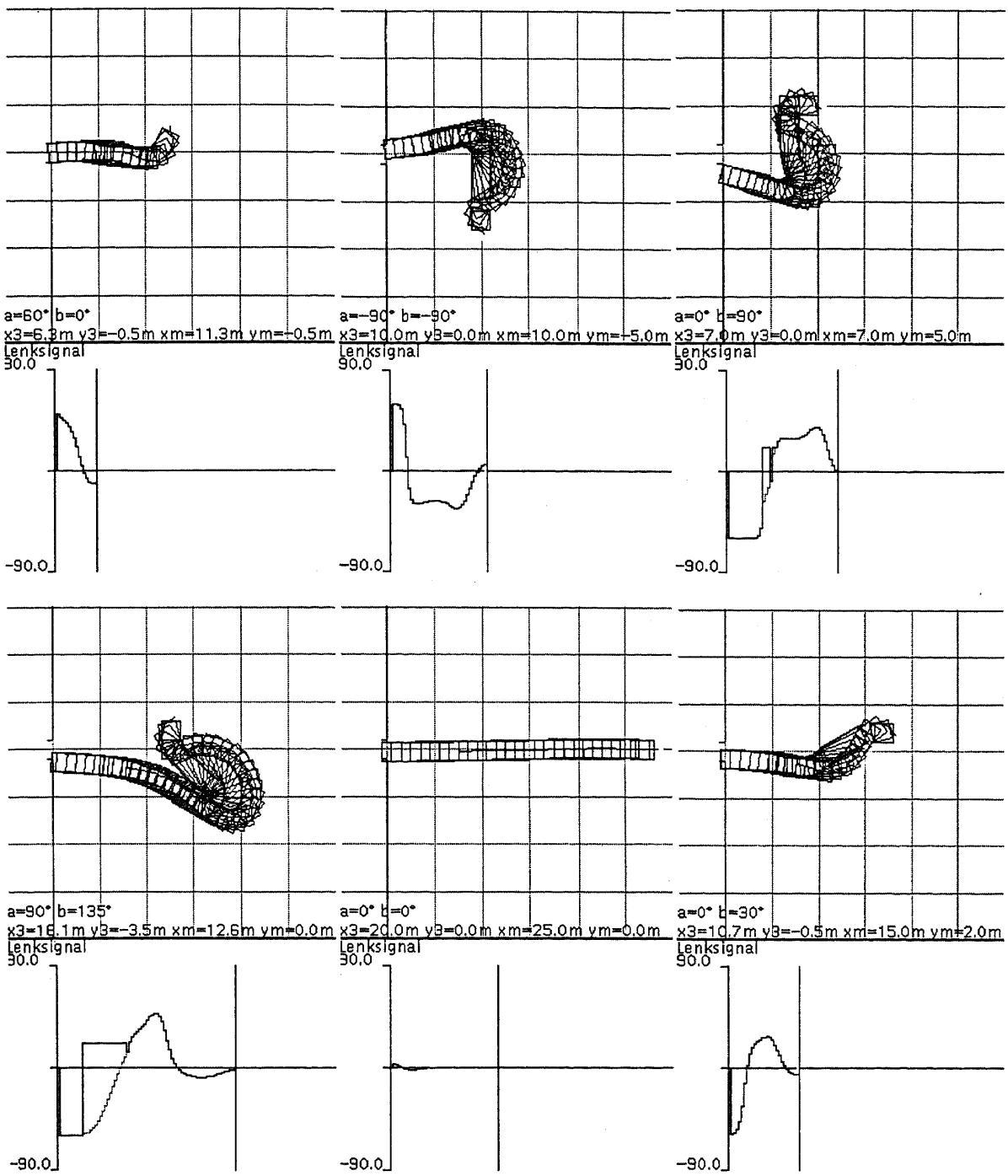


Abbildung 5.7: Beispielregelungen Konfiguration B2 mit $f_B^{b,\Phi} = 0.4$

5.2.3 Bewertung der Blockadeposition

Als letzte der Bewertungsmöglichkeiten zur Blockadevermeidung wurde noch die Bewertung der Blockadeposition implementiert. Dabei wird die Blockadeposition als Endposition aufgefaßt und der sich ergebende Fehler gemäß der Endpositionsbewertung rückpropagiert. Die Zahl der Rückpropagierungsschritte stimmt aber mit der der anderen Blockadebewertungen überein.

Konfiguration B3 Bewertung der Blockadeposition als Endposition			
Propagierungsstrategie	FP-STOPP		
Bewertete Variablen	β	y_3	x_3
Bewertungsgrenzen	$\beta_B = \pm 0°$	$y_{3,B} = \pm 0m$	$x_{3,B} = \pm 0m$
Fehlerskalierung $f_E^{b,EP}$	0.2	0.01	0.01
Bewertungsskalierung $f_B^{b,EP}$		1.0	
Anzahl BP-Zeitschritte[a] $\#BP^{b,EP}$		1	

[a]Für die eigentliche Endpositionsbewertung im Fall, daß keine Blockade auftritt und das Endekriterium erfüllt wird, gilt weiter die in der Grundkonfiguration festgelegte Propagierung bis zur Startposition.

Da als Propagierungsstrategie FP-STOPP gewählt wurde, wird die Blockade maximal einmal pro Trajektorie bewertet.
Diese Bewertung ist keine direkte Bewertung einer die Blockade beschreibenden Variablen. Unter diesem Gesichtspunkt ist das mit diesen Konfigurationswerten erzielte Ergebnis bereits akzeptabel und muß einen Vergleich mit den anderen Grundkonfigurationen B1 und B2 nicht scheuen (Tabelle 5.15).

Im Gegensatz zu den B1- und B2-Konfigurationen bleibt hier bei Erhöhung der Bewertungsskalierung die y-Performanz ungefähr konstant, während die β-Performanz leicht nachläßt. Ansonsten verändern sich die Analysewerte bei Variierung der Bewertungsskalierung tendenziell gleich.
Zu beachten ist eventuell noch die gegenüber B1 und B2 aufgetretene Verlängerung der Trajektorien, die Hand in Hand mit engeren Kurvenradien geht und deshalb wahrscheinlich die relativ gute Ziel-Positionierung ermöglicht.

Tabelle 5.15: Analyse

B3							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
10.29°	$1.04m$	$12.51m$	6.81%	26.57°	$8.03\frac{°}{m}$	26.61°	$8.38\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
2.93%	17.30%	$2.12m$	29.22°	0%	0%	$0m$	0°

Tendenz bzgl. O1			
Ziel	limit. RB	Dynamik	Trajektorie
$\beta: + \; y: -$	$++$	$= -$	$= -$

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

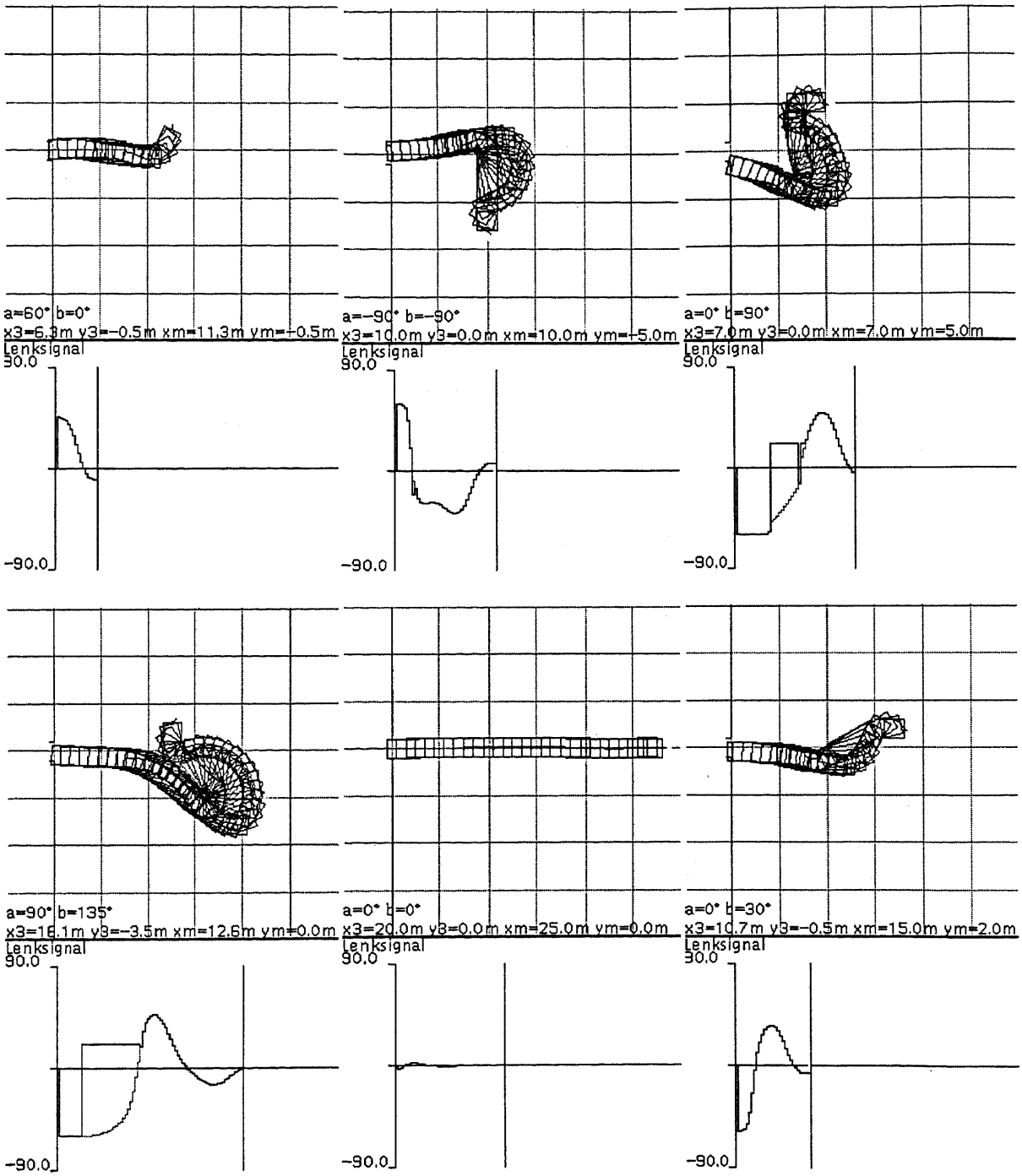


Abbildung 5.8: Beispielregelungen Konfiguration B3

5.2.4 Lokalisationsuntersuchung

Bisher wurde für die Bewertungspropagierung der Blockade stets eine Zeitschritt-Länge von $\#BP^b = 1$ verwendet. Eine Ausdehnung der Parameterbeeinflussung auf vom Bewertungspunkt weiter entferntere Wertebereiche der Reglereingangsvariablen, gemäß der vermuteten Lokalisationseigenschaft der Randbedingung (vgl. Kapitel 2.2.2 Seite 15), könnte durch eine Vergrößerung der Propagierungslänge der Blockadebewertung erreicht werden. Dabei gilt es aber folgendes zu bedenken:

- Durch die Art des Gewichtupdates und der funktionalen Abhängigkeiten innerhalb des neuronalen Netzes findet schon durch die einmalige Propagierung eine 'Verschmierung' der Regleradaption über den Eingabezustandsraum statt. Es wird also keine punktuelle lokale Veränderung der Reglerfunktion vorgenommen. Ein möglicher Lokalisationseffekt der Bewertung wird daher eventuell nicht sehr stark ausgeprägt sein.

- Die Wirkung einer iterativen Propagierung durch mehrere Zeitschritte auf den effektiven Lernkoeffizienten (vgl. Kapitel 3.6 Seite 32) ist nicht bekannt. Eine mögliche Akkumulation der Gewichtupdates durch eine längere Propagierung und somit eine Überdeckung des zu untersuchenden Lokalisationseffekts kann nicht ausgeschlossen werden.

Um trotzdem einen vorhandenen Effekt der erweiterten Lokalisation zu erkennen, wurde folgendermaßen vorgegangen:
Ausgehend von einer Konfiguration mit Blockadebewertung des Stellsignals und FP-STOPP-Propagierungsstrategie wurden a) die Bewertungsskalierung und b) die Propagierungslänge der Blockadebewertung erhöht und die sich ergebenden Performanz-Analysewerte dieses Rasters in eine Tabelle eingetragen. Durch Vergleich der Werte bei partieller Veränderung der Skalierung oder der Propagierungslänge kann dann der Einfluß der jeweiligen Lernparametervariation interpretiert werden.
Um für die Zielpositionierung einen zusammenfassenden Wert zu erhalten, wird der σ_β-Fehler in ein Längenmaß umgerechnet und anschließend mit dem σ_y-Fehler zum Zielfehlerindex z addiert. Dies geschieht durch Multiplikation des Radiant-Werts des Winkels mit der Anhängerlänge. Man erhält die Bogenlänge b der Drehung des Anhängers um sein Heck, die noch nötig ist, um den Winkelfehler auszugleichen. Im Beispiel der nun gewählten Ausgangskonfiguration errechnet sich b aus der Standardabweichung $\sigma_\beta = 11.52°$ zu $\sigma_b = 1.005m$. Mit dem Fehler $\sigma_y = 1.14m$ addiert sich das (gerundet) zum Zielfehlerindex $z = 2.15m$.
Als Ausgangskonfiguration wird die Bewertung des Stellsignals als B1-Konfiguration mit FP-STOPP gewählt (s. Seite 64). Die Werte der Propagierungslänge und Bewertungsskalierung sind $\#BP^b = 1$ und $f_B^b = 0.1$. Die Variationswerte der Propagierungslänge sind $\#BP^b = 1, 2, 4, 6$ und die der Bewertungsskalierung $f_B^b = 0.1, 0.2, 0.4, 0.6$.
Für die 16 gelernten Reglernetze erhält man dann folgende Performanzwerte:

Performanzwerte der Lokalisationsuntersuchung									
		Zielfehlerindex z [m]				Blockadehäufigkeit s^b [%]			
f_B^b	$\#BP$:	1	2	4	6	1	2	4	6
0.1		2.15	2.10	2.14	2.14	0.81	0.83	0.91	0.90
0.2		2.20	2.14	2.09	2.10	0.78	0.76	0.64	0.78
0.4		2.31	2.19	2.18	2.17	0.66	0.51	0.49	0.49
0.6		2.39	2.26	2.25	2.21	0.49	0.41	0.35	0.39

Man erkennt, daß bei Erhöhung der Bewertungsskalierung und konstantem $\#BP^b$ (Spalten) sich zwar die RB-Performanz stark erhöht, dies aber auf Kosten einer sich verschlechternden Zielperformanz geschieht. Wenn nun die Propagierungslänge einen direkten Einfluß auf den effektiven Lernkoeffizienten hätte, wie die Bewertungsskalierung, dann müßte auch bei deren Erhöhung eine ähnliche Veränderung der Performanzwerte erfolgen. Bei konstanter Bewertungsskalierung (Reihen) ist jedoch eine generell positive Veränderung beider Performanzwerte erkennbar[1]. Folglich muß die Erweiterung des Lokalisationsbereichs der Bewertung, zumindest für die hier gewählten Werte, die Parameter in einer zur Bewertungsskalierung unterschiedlichen Weise beeinflussen. Als Nachweis einer prinzipiellen Existenz dieser Lokalisationseigenschaft der Randbedingung reicht diese Stichprobe selbstverständlich nicht aus; es kann nur die Nonkonformität der Parameterbeeinflussung durch die Lernparameter $\#BP^b$ und f_B^b belegen.

[1] Für eine Skalierung $f_B^b = 0.1$ bleibt die Zielperformanz zumindest ungefähr konstant.

5.2.5 Zusammenfassung

Mit den hier realisierten Bewertungen sind ausgesprochen gute Ergebnisse in der Blockadevermeidung durch den Regler erzielt worden. Die RB-Performanz kann durch eine Erhöhung der Bewertungsskalierung noch weiter verbessert werden. Die beste RB-Performanz (Null Blockaden im Test) wurde mit der Stellsignalbewertung der BP-PROP-Strategie bei erhöhter Blockadebewertung erhalten.
Die daneben verschlechterte Zielperformanz konnte nur zum Teil über Netzstrukturveränderungen ausgeglichen werden. Ebenso ist eine leichte Verbesserung der Regelungen auf eine gleichzeitig gute Ziel- und RB-Performanz außer über die Netzstrukuränderung auch durch Vergrößerung des Lokalisationsbereichs der Bewertung möglich. An der Stellsignalbewertung konnte dies beispielhaft gezeigt werden. Aufgrund des zur Erhöhung der Skalierung unterschiedlichen Einflusses auf die Performanz ist dies ein Indiz für einen vorteilhaft nutzbaren Lokalisationscharakter der Randbedingung.
Geringe Unterschiede wurden allerdings beim Einsatz der verschiedenen Propagierungsstrategien festgestellt.
Eine qualitative Verbesserung durch Hinzufügen der Hilfsvariablen Φ am Reglereingang konnte nicht registriert werden.
Die Bewertung der Blockade über den Blockadewinkel Φ am Streckenmodul-Ausgang erbrachte etwas schlechtere Resultate als die direkte Bewertung des Stellsignals. Möglicherweise sind aber auch die hierfür gewählten Konfigurationseinstellungen nicht in gleicher Weise optimal, so daß dies keine generell gültige Aussage sein kann. Die Charakteristik der so erzielten Regelung ist denn auch denen der anderen Blockadebewertungen ähnlich.
Die Bewertung der Blockadeposition erbrachte eine leicht verbesserte Zielpositionierung bei etwas geringerer RB-Performanz. Ihre Eignung zur Randbedingungsintegration in die Reglerfunktion kann aber ebenso attestiert werden.

5.3 Lenkwinkelbegrenzung mittels lokaler Bewertung

Mit der Sigmoidfunktion ist bereits eine zufriedenstellende und mit dem BPtt-Verfahren vereinbare reglerinterne Begrenzung des Lenkwinkels vorhanden. Trotzdem soll aus zwei Gründen nun versucht werden, diese Begrenzung über eine Bewertung des Lenkwinkelfehlers in die Regelfunktion zu integrieren.

1. Es interessiert, ob sich durch das explizite Erlernen der Begrenzung eine andere Regelungscharakteristik ergibt.

2. Es soll in Kombination mit der Blockadebewertung beispielhaft untersucht werden, mit welchen Einschränkungen der globalen Zieloptimierung sich mehrere zusätzliche Optimierungskriterien mit in die Reglerfunktion einarbeiten lassen.

5.3.1 Probleme bei fehlender Sigmoidfunktion

Bei Fehlen einer sigmoiden Ausgabefunktion im neuronalen Regler-Netz muß mit folgenden neuen Problematiken umgegangen werden:

Extreme Reglerausgaben und BP-Fehler

Bei Wegfall der Begrenzung durch die Sigmoidfunktion werden prinzipiell große Stellsignalwerte des neuronalen Reglers möglich, die dann bei Bewertung einer der nachfolgenden Randbedingungs-Korrekturen zu größeren Fehlerwerten führen. Die Lernkoeffizienten sind dementsprechend für eine größere Fehlerbandbreite auszulegen, d.h. sie dürfen auch bei großen auftretenden Fehlerwerten keine zu drastischen Gewichtsveränderungen bewirken.
Durch den fehlenden Sättigungscharakter der Sigmoidfunktion werden auch die BP-Fehler, die aus hohen Reglerausgaben resultierten, nicht mehr durch die sehr kleinen Ableitungen im Sättigungsbereich gedämpft. Die Folge sind wesentlich größere in das Netz hineinpropagierte Fehler für bereits extreme Reglerausgaben.
Durch zu starke Änderungen in den Gewichten bereits in der frühen Lernphase kann es aufgrund einer noch nicht vernünftig arbeitenden Reglerfunktion zu andauernd auftretenden Blockaden kommen. Werden diese nicht von einer separaten Blockadebewertung wieder in der Häufigkeit reduziert, dann kann unter Umständen das Reglerlernen bereits hier scheitern, da durch die Korrekturfunktion der Blockade stets ein Nullfehler propagiert wird[2] und resultierend kein ausreichender Gewichtsupdate für die Endpositionsbewertung existiert.

Winkeldarstellung

Die Winkeleigenschaft des Lenksignals darf nicht bereits bei der Reglerausgabe angenommen werden. Würde die Reglerausgabe als Winkel betrachtet, dann müßten z.B. die (in den Realbereich transformierten) Ausgabewerte $\theta > 180°$ wie ihre im Bereich $[-180°, +180°]$ angesiedelten Entsprechungen behandelt werden. Bei einer Korrektur des Lenksignals auf $\theta_G = \pm 60°$ mittels der einfachen Begrenzungsfunktion wird ein Ausgabewinkel $\theta_1 = +190°$ zuerst transformiert auf $\theta'_1 - 170°$, dann begrenzt auf $\theta'_{1,k} - 60°$ und vom Bewertungsmodul ein Fehler von $e_\theta = (-60°) - (-170°) = +110°$ zurückpropagiert (vgl. Abbildung 5.9). Dieser Fehler wirkt dann auf eine Justierung

[2] Wenn die BP-BYPASS-Strategie verwendet wird, werden die Reglerausgaben noch extremer.

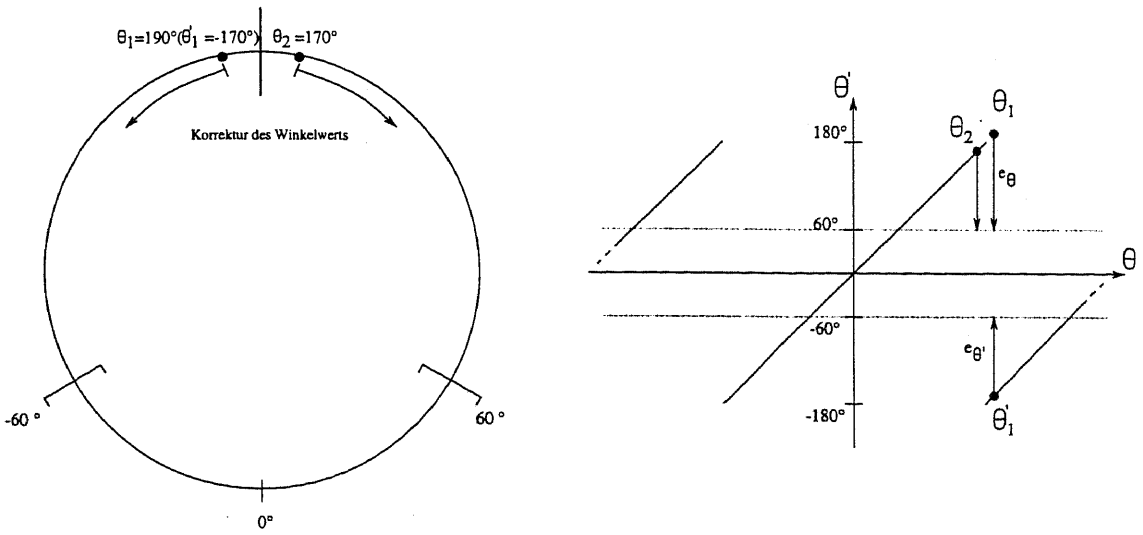


Abbildung 5.9: Unterschiedliche Korrektur und Bewertungsfehler benachbarter, als Winkel betrachteter Reglerausgaben $\theta_1 = 190°$ und $\theta_2 = 170°$

des Reglernetzes auf eine noch größere Ausgabe in Richtung $[+300°, +420°]$ hin, anstatt auf eine Reduzierung in Richtung $[-60°, +60°]$, wie dies z.B. bei einem benachbarten Ausgabewinkel von $\theta_2 = +170°$ der Fall wäre. Dies führt letztendlich zu einer Regelfunktion, die mehrere nicht zusammenhängende Zulässigkeitsbereiche aus $[360° * i - 60°, 360° * i + 60°]$, $i \in \mathbf{Z}$ der Netzausgabe berücksichtigen soll, was wegen den dazwischenliegenden 'verbotenen' Zonen mit einer stetigen Abbildungsfunktion nicht zu realisieren ist. Die Regelfunktion wird somit durch den Versuch der Regelung in nicht verbundenen Sektoren 'zerrissen', was letztendlich in einem Mißerfolg des Lernens endet.
Man muß folglich die Reglerausgabe als Nicht-Winkelwert betrachten und für die Bewertung so belassen, auch wenn der damit dargestellte Winkel die Randbedingung erfüllen würde (für das obige Beispiel z.B. ein Ausgabewert entsprechend $+360°$). Nur so kann sich die Netzfunktion der Randbedingung mit einem Zulässigkeitsbereich $[-60°, +60°]$ anpassen.

5.3.1.1 Ermittlung einer Grundkonfiguration

Die Konfigurationsdaten der O1-Konfiguration sind nicht ohne eine Anpassung der Skalierungswerte in die Konfiguration ohne sigmoide Netzausgabe übernehmbar. Daher muß zuerst einmal eine Grundkonfiguration ohne lokale Bewertungen gefunden werden, die in etwa der O1-Konfiguration ähnliche Performanzwerte aufweist. Es zeigt sich, daß, wie eben bereits vorüberlegt, die Skalierungswerte der Endpositionsbewertung erheblich geringer ausfallen müssen. Die Fehlerskalierungswerte werden um eine Zehnerpotenz gegenüber den O1-Werten verkleinert, um für die Bewertungsskalierung f_B^e mit in etwa gleichen Werten arbeiten zu können. Als Ausgangskonfiguration für die nachfolgenden Variationen eignet sich daher folgende:

Konfiguration L1 Keine NN-Sigmoidausgabe, leine Lokalen Bewertungen			
Propagierungsstrategie	BP-PROP		
Bewertete Variablen	β	y_3	x_3
Bewertungsgrenzen	$\beta_G = \pm 0°$	$y_{3,G} = \pm 0m$	$x_{3,G} = \pm 0m$
Fehlerskalierung f_E^e	0.02	0.001	0.001
Bewertungsskalierung f_B^e	2.0		
Anzahl BP-Zeitschritte $\#BP^e$	bis Startposition		

Die Wertebereichstransformation am Netzausgang wurde leicht modifiziert, da der Ausgangswertebereich nun nicht mehr auf $[0, 1]$ beschränkt ist:

	NN-Ausgang θ
Realbereich	$[-\frac{\pi}{3}, +\frac{\pi}{3}]$
Neuronalbereich	$[-1, 1]$

Die Wahl der BP-PROP-Strategie statt BP-STOPP sollte keinen bedeutenden Unterschied machen. Im Zweifelsfall sind bei Vergleichen mit der O1-Konfiguration die Analysewerte von O1 mit BP-PROP in Tabelle 5.2 auf Seite 56 heranzuziehen. Der Vergleich der Analysewerte mit O1 zeigt nur unwesentliche Unterschiede in der Ziel- und RB-Performanz und der Trajektorienlänge, dafür aber die erwartete hohe Stellsignaldynamik (Tabelle 5.16).

In den Testbildern in Abbildung 5.10 sind auch erstmals die Kombination beider Korrekturen zu erkennen. Z.B. wird in Testbild 4 zuerst der Lenkwinkel auf $-60°$ beschränkt. Dann setzt zusätzlich die Blockadevermeidung mit ein, so daß als nächstes durch die integrale Korrektur auf den betragmäßig kleineren Blockadekorrekturwinkel korrigiert wird. Wenn dann nach etwa einem Drittel der Regelung die Reglerausgabe weniger extrem wird, ist nur noch die Blockadekorrektur aktiv.
Vergleicht man noch die Lenkwinkel-Analysewerte von L1 und O1, so stellt man fest, daß trotz der größeren Dynamik der Reglerausgabe bei L1 das korrigierte Signal ähnliche Werte aufweist. Auch in den Testbildern stimmt das korrigierte Lenksignal (kräftige Linie) weitgehend überein (man beachte: für die L-Konfigurationen ist bei der Darstellung des Lenksignals ein anderer Maßstab gewählt!).

5.3.1.2 Variation der Lernparameter

Die L1-Konfiguration zeigt sich sehr empfindlich gegenüber einer Erhöhung der Bewertungsskalierung und der Neuronenzahl. Mit einer leichten Erhöhung auf $f_B^e = 3.0$ bzw. einer Strukturänderung auf $\#H_{L1} = 50$ konnte kein Regler erfolgreich trainiert werden.

Tabelle 5.16: Analyse

L1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
12.00°	$0.86m$	$12.68m$	8.07%	45.05°	$15.61\frac{°}{m}$	28.28°	$10.63\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
8.89%	45.20%	$2.49m$	55.79°	19.44%	78.30%	$3.15m$	76.01°

Tendenz bzgl. O1			
Ziel	limit. RB	Dynamik	Trajektorie
$\approx$	$b := +\ l : (-)$	$++$	$= -$

Der Vergleich der RB-Performanz der Lenkwinkelbegrenzung ist natürlich ohne Bedeutung, da diese bei O1 durch die Sigmoidfunktion stets optimal war.
Eine Verletzung der Lenkwinkel-Randbedingung tritt in fast 80% der Trajektorien auf. Die geringere relative Häufigkeit über die Trajektorienstrecke betrachtet und der hohe durchschnittliche Lenkwinkelfehler deuten bereits die in den Testbildern in Abbildung 5.10 sichtbaren Stellsignalspitzen an.

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

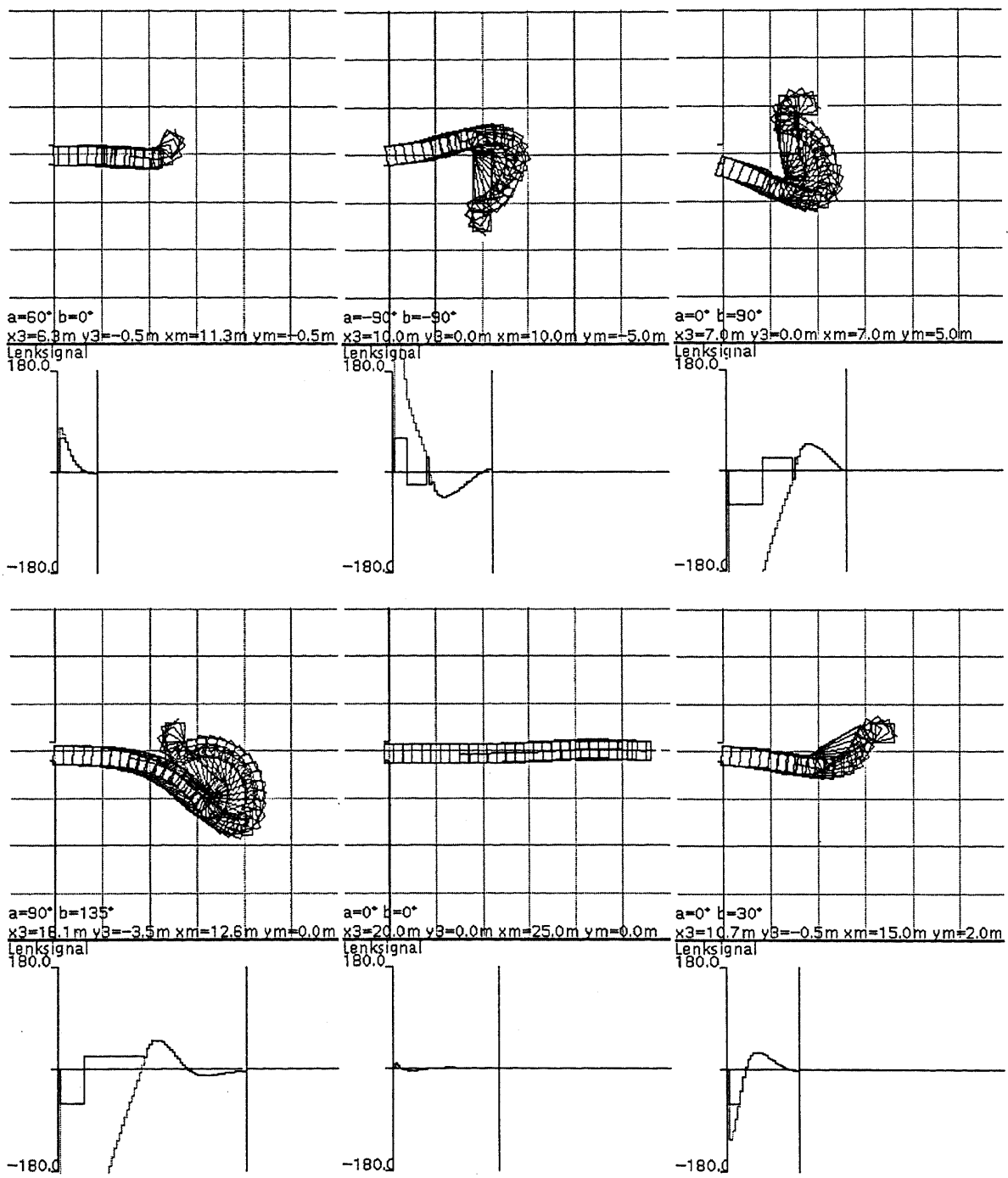


Abbildung 5.10: Beispielregelungen Konfiguration L1

5.3.2 Bewertung des Lenkwinkelfehlers

Zur Begrenzung der Lenkwinkelfehler soll nun die Bewertung des zu korrigierenden Stellsignals verwendet werden. Da eine gute Einstellung der Bewertungsskalierung vorab nicht genau eingeschätzt werden konnte, wurde innerhalb eines weiten Wertebereichs getestet. Im folgenden sind sowohl für die Lenkwinkelbewertung, als auch für die später hinzugefügte Blockadebewertung Analysen für die Skalierungswerte $f_B = 0.001$ und $f_B = 0.01$ angegeben.

Konfiguration L2 Bewertung des Lenkwinkelfehlers	
Propagierungsstrategie	BP-PROP
Bewertete Variablen	θ
Bewertungsgrenzen	$\theta_B = \pm 60°$
Fehlerskalierung f_E^l	1.0
Bewertungsskalierung f_B^l	0.001 / 0.01
Anzahl BP-Zeitschritte $\#BP^l$	1
Endpositionsbewertung wie	L1

Das mit $f_B^l = 0.001$ erzielte Ergebnis läßt in der Analyse (Tabelle 5.17) und in Abbildung 5.11 auf Seite 88 bereits eine deutliche Begrenzung des Lenkwinkels erkennen; extreme Lenkwinkelspitzen werden sicher vermieden. Gleichzeitig ist die Dynamik des Stellsignals sowie die Zahl der Blockaden reduziert worden (s. z.B. Testbild 3), ein direkter Einfluß der Lenkwinkelbegrenzung.
Dagegen sticht als (nicht erklärbarer) störender Einfluß ein negativer y-Offset ins Auge, der

Tabelle 5.17: Analyse

L2 mit $f_B^l = 0.001$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.66°	1.58m	12.88m	9.75%	27.51°	8.78$\frac{°}{m}$	25.57°	8.10$\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
3.36%	19.00%	2.28m	22.66°	7.55%	41.90%	2.32m	25.79°
Tendenz bzgl. L1							
Ziel		limit. RB		Dynamik		Trajektorie	
$\beta :\approx$ $y : --$		$b : ++$ $l : ++$		$--$		$+$	

Referenz: L1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
12.00°	0.86m	12.68m	8.07%	45.05°	15.61$\frac{°}{m}$	28.28°	10.63$\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
8.89%	45.20%	2.49m	55.79°	19.44%	78.30%	3.15m	76.01°

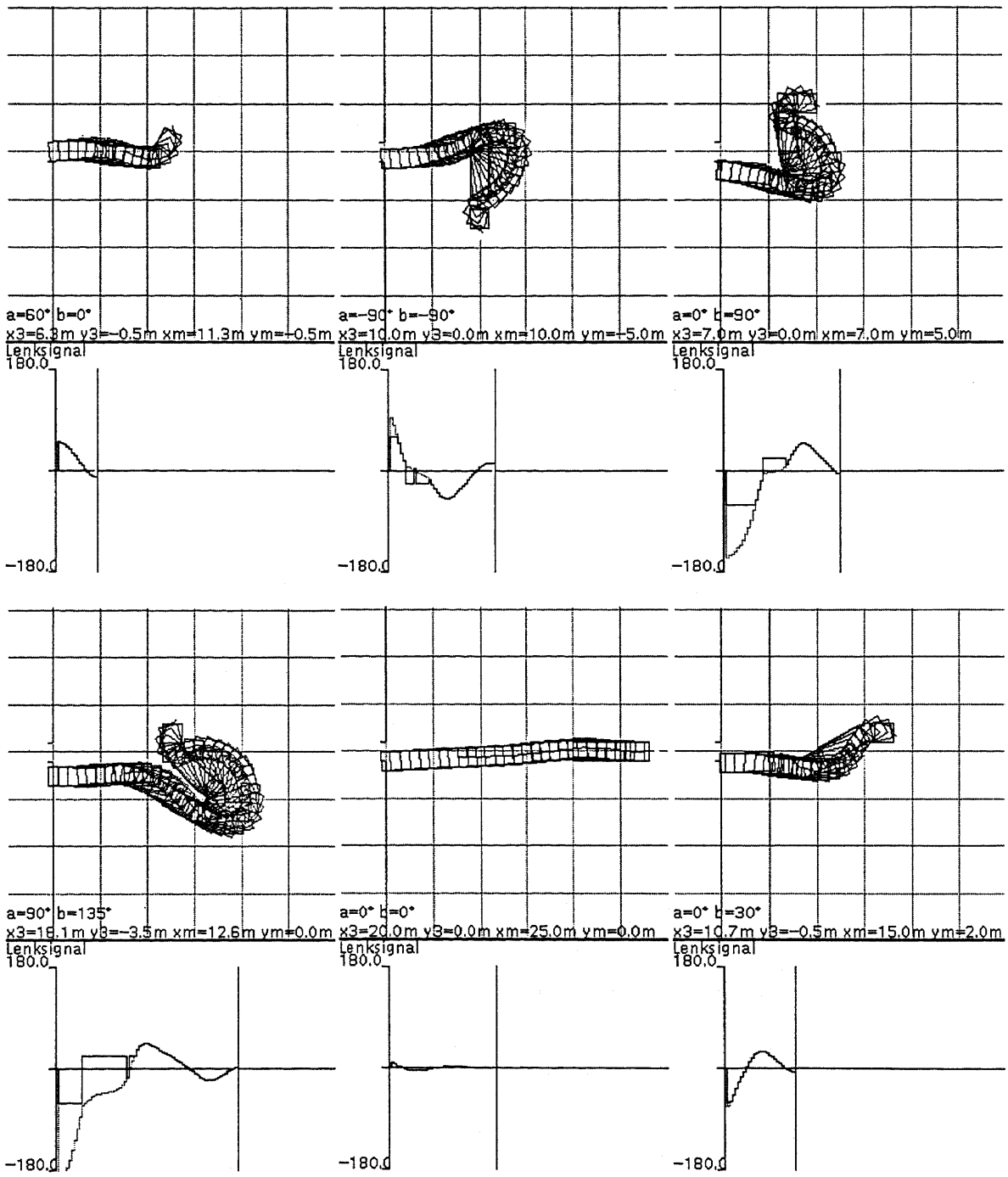


Abbildung 5.11: Beispielregelungen Konfiguration L2

vermutlich mitverantwortlich für die relativ schlechte Zielperformanz in y ist, angesichts einer optisch beurteilt sich sonst nicht gegenüber L1 verschlechterten Regelungseigenschaft. Die β-Positionierung ist sogar noch etwas verbessert.

Tabelle 5.18: Analyse

L2 mit $f_B^l = 0.01$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
14.62°	$1.64m$	$12.63m$	7.08%	25.04°	$7.11\frac{°}{m}$	24.98°	$7.42\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\overline{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\overline{s}^l$	$\pm\theta^l$
5.12%	29.90%	$2.16m$	21.42°	4.61%	28.30%	$2.06m$	10.29°

Tendenz bzgl. L2 mit $f_B^l = 0.001$			
Ziel	limit. RB	Dynamik	Trajektorie
$\beta : - \; y := -$	$b : - \; l : +$	$-$	$-$
Auch der Einfluß auf die Blockadevermeidung ist jetzt nicht mehr so stark wie mit $f_B^l = 0.001$. Das bedeutet, daß für eine Verbesserung der Blockadevermeidung durch den Regler eine gezielte Blockade-Bewertung notwendig ist.			

Referenz: L2 mit $f_B^l = 0.001$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.66°	$1.58m$	$12.88m$	9.75%	27.51°	$8.78\frac{°}{m}$	25.57°	$8.10\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\overline{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\overline{s}^l$	$\pm\theta^l$
3.36%	19.00%	$2.28m$	22.66°	7.55%	41.90%	$2.32m$	25.79°

Eine Erhöhung der Bewertungsskalierung auf den Wert $f_B^l = 0.01$ zeigt (Tabelle 5.18), daß eine noch bessere Unterdrückung des Lenkwinkelfehlers nur auf Kosten der Positioniergenauigkeit erreicht wird (s. auch die zugehörigen Testbilder 2,3 und 4 in Abbildung 5.12).

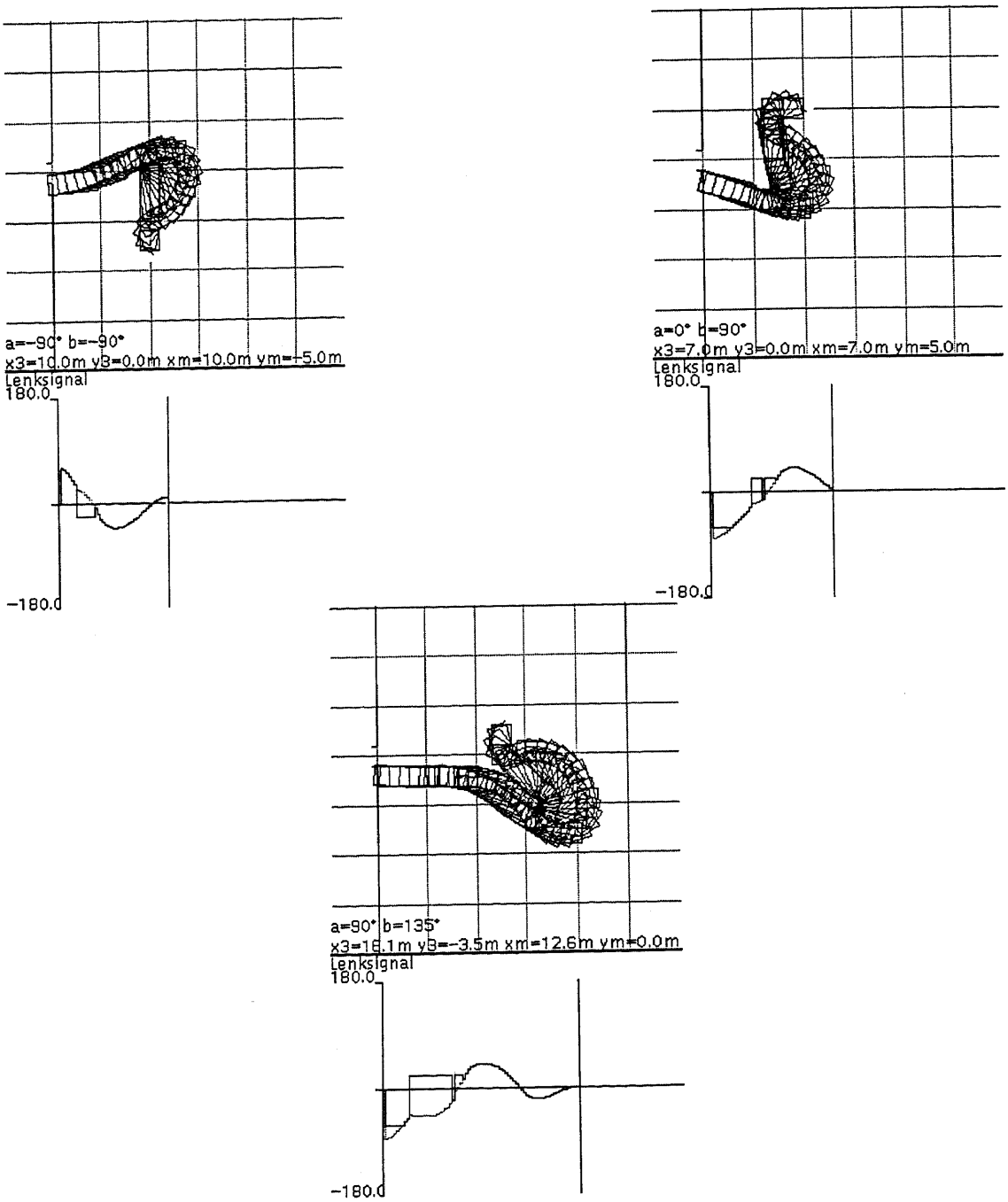


Abbildung 5.12: Konfiguration L2 mit $f_B^l = 0.01$, Beispiele 2-4

5.3.3 Bewertung von Lenkwinkelbegrenzung und Blockade

Eine zusätzliche Integration der Blockade-Randbedingung in das Reglerlernen sollte, nach den bisherigen Ergebnissen, zu einer weiteren Verschlechterung der globalen Optimierung (der Zielpositionierung) führen. Es soll nun das Ausmaß dieser Verschlechterung, gleichzeitig aber auch die dabei erzielte RB-Performanz beurteilt werden. Zusätzlich von Interesse ist der Zusammenhang zwischen den durch die separaten Bewertungen erzielten Lernerfolgen. Wie schon bei der Konfiguration L2 beobachtet, kann die Bewertung der einen Randbedingungs-Verletzung als Nebeneffekt auch eine Berücksichtigung der anderen nicht bewerteten Randbedingung durch den Regler haben.
Die Einstellung der Bewertung der Lenkwinkelbegrenzung wird von L2 übernommen, die Blockade-Bewertung ist bis auf die Skalierungsfaktoren gleich der der Konfiguration B1:

Konfiguration L3 Blockade-Bewertung des Stellsignals	
Propagierungsstrategie	BP-PROP
Bewertete Variablen	θ
Bewertungsgrenzen[a]	$\theta_B \in \{\theta_G(\hat{\Phi}_G), \theta_G(\check{\Phi}_G)\}$
Fehlerskalierung f_E^b	1.0
Bewertungsskalierung f_B^b	0.001 / 0.01
Anzahl BP-Zeitschritte $\#BP^b$	1
Endpositionsbewertung wie	L1

[a] vgl. Kapitel 2.2.2 Seite 15.

Die Wirkung der zusätzlichen Blockadebewertung kann beispielhaft an einer Konfiguration mit den Skalierungswerten $f_B^l = 0.001$ und $f_B^b = 0.01$ demonstriert werden (Tabelle 5.19).

Tabelle 5.19: Analyse

L3 mit $f_B^l = 0.001, f_B^b = 0.01$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
12.40°	$1.71m$	$12.39m$	5.27%	26.19°	$7.90\frac{°}{m}$	24.29°	$6.99\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
1.61%	10.30%	$1.94m$	17.58°	7.41%	41.50%	$2.21m$	24.92°

Tendenz bzgl. L2 mit $f_B^l = 0.001$			
Ziel	limit. RB	Dynamik	Trajektorie
$\beta := -\ y := -$	$b : +\ l := +$	$-$	$--$

Die Verbesserung der Blockade-RB-Performanz bei etwa konstanter Lenkwinkel-RB-Performanz geht mit einer Verschlechterung der Zielperformanz einher und wird vor allem durch eine niedrigere Stellsignal-Dynamik bewirkt. Dies hat auch einfachere (kürzere) Trajektorien zur Folge. Siehe die Testbeispiele in Abbildung 5.13 auf Seite 93 im Vergleich zu denen der Konfiguration L2 auf Seite 88. Man erkennt deutlich die im Lenksignal in geringerer Zahl vorkommenden Blockadekorrekturen. Läßt man den hier wieder auftretenden y-Offset außen vor, dann ist die Regelung aber durchaus als gelungen anzusehen.

Referenz: L2 mit $f_B^l = 0.001$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.66°	$1.58m$	$12.88m$	9.75%	27.51°	$8.78\frac{°}{m}$	25.57°	$8.10\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
3.36%	19.00%	$2.28m$	22.66°	7.55%	41.90%	$2.32m$	25.79°

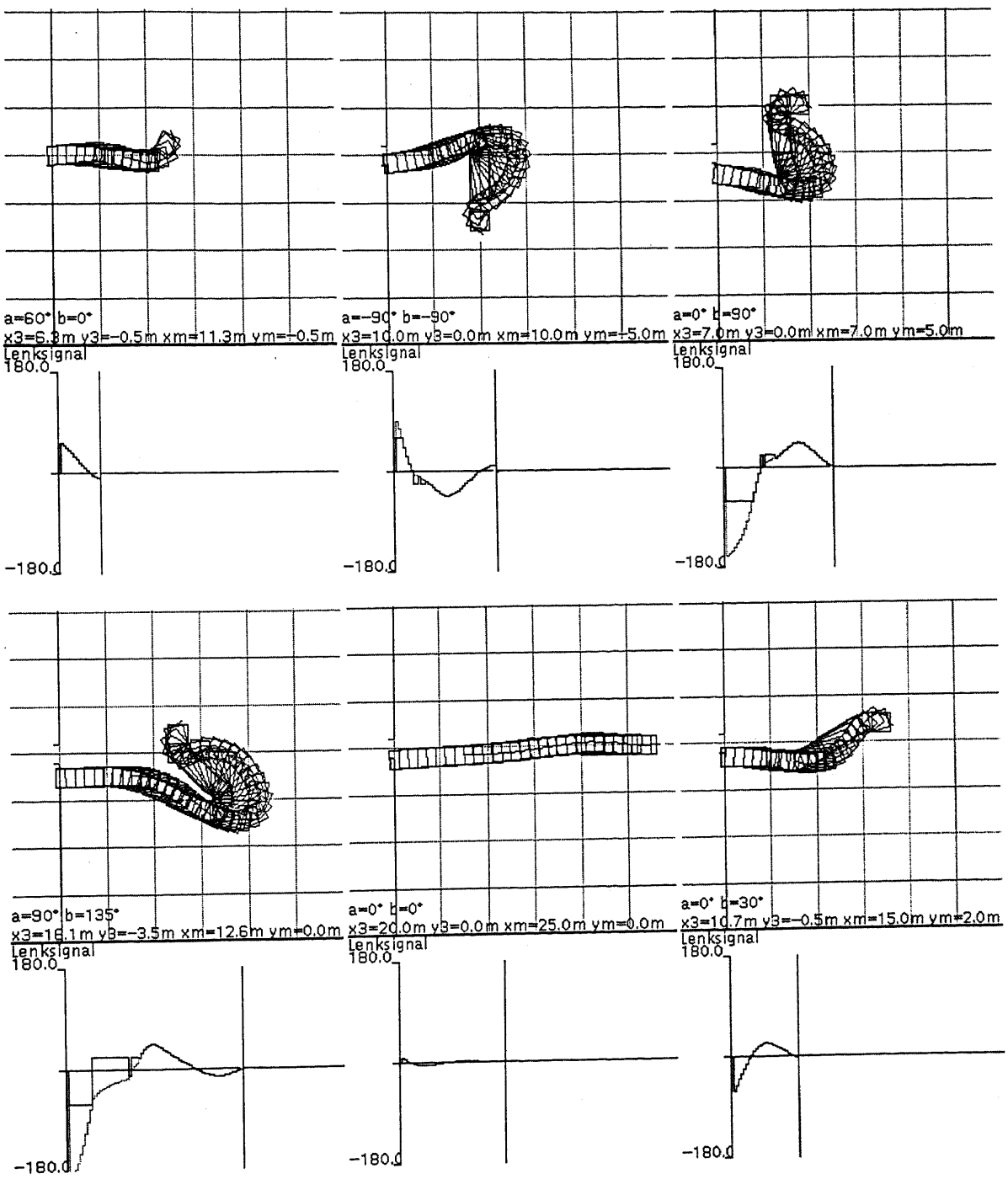


Abbildung 5.13: Beispielregelungen Konfiguration L3

5.3.3.1 Unabhängigkeit des Randbedingungs-Lernens

Betrachtet man die durch eine Variation der jeweiligen Skalierungswerte sich einstellenden RB-Performanzen, dann ist diesbezüglich eine Unabhängigkeit der Skalierungsvariation für die zwei lokalen Bewertungen feststellbar. Die Performanzen lassen sich gezielt über ihre zugehörigen Bewertungen verbessern. Zwar wird durch die einzelne Lenkwinkelbegrenzung die Blockade-Randbedingung bereits besser eingehalten (s. Analysetabelle 5.17 der Konfiguration L1 auf Seite 87). Jedoch läßt sich dann durch partielle Veränderung der Skalierung einer RB-Bewertung hier eine Performanz-Änderung erzielen, während die jeweils andere RB-Performanz relativ unbeeinflußt bleibt. In folgender Tabelle sind die Performanzergebnisse bei partieller Skalierungswerte-Änderung ablesbar. Eine Bewertungsskalierung mit dem Wert 0 bedeutet, daß die Bewertung nicht aktiviert war, wie in den Konfigurationen L1 (Blockade und Lenkwinkel) bzw. L2 (Blockade). Ein Netzlernen für die Kombination $f_B^l = 0.01$, $f_B^b = 0.001$ konnte nicht erfolgreich abgeschlossen werden. Dieses Ergebnis ist positiv zu werten, da es zeigt, daß die Reduktion der Zahl der Blockadesituationen nicht nur durch eine pauschale Begrenzung der Reglerausgaben erreicht werden kann, was der Lenkwinkel-Randbedingung genügen würde, sondern durch gezielte Veränderung an der Regelfunktion bewirkt wird, die wiederum kaum die Einhaltung der Lenkwinkelbegrenzung beeinflussen. Der Regler lernt also durchaus zwei verschiedene Regelprinzipien hinzu, was nicht von vorneherein zu erwarten war.

RB-Performanz bei partieller Variierung der Skalierungswerte						
f_B^l	f_B^b					
	0		0.001		0.01	
	s_b[%]	s_l[%]	s_b[%]	s_l[%]	s_b[%]	s_l[%]
0	8.89	19.44				
0.001	3.36	7.55	3.0	7.56	1.61	7.41
0.01	5.12	4.61	-		2.0	3.54

5.3.3.2 Variation der Netzstruktur

Der Vollständigkeit wegen sei hier noch der Einfluß der Netzstrukturerweiterung durch eine Erhöhung der Neuronenzahl auf $\#N_{H1} = 50$ angegeben. Wie schon bei anderen Konfigurationen kann weder eine bedeutende Verbesserung der RB-Performanz noch eine völlige Kompensation der Zielperformanz-Einbußen erreicht werden. Die Analysedaten für die Konfiguration L3 mit $f_B^l = 0.001$, $f_B^b = 0.01$ sind in Tabelle 5.20 angegeben.

5.3.4 Zusammenfassung

Mit der Bewertung allein der Lenksignalkorrektur ist bereits eine gute Lenksignalbegrenzung erzielbar. Die nebenbei erfolgende Blockadevermeidung ist auf die drastische Reduktion der hohen Lenksignalwerte zurückzuführen.
Als besonderes Ergebnis kann hervorgehoben werden, daß bei hinzugeschalteter Blockadebewertung eine gezielte Verbesserung der einzelnen RB-Performanzwerte durch Erhöhung des jeweiligen Bewertungsskalierungswerts möglich ist.
Die erreichten RB-Performanzen bei Bewertung beider Randbedingungen sind als gut einzustufen.

Tabelle 5.20: Analyse

L3 mit $f_B^l = 0.001, f_B^b = 0.01, \#N_{H1} = 50$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.13°	$1.75m$	$12.33m$	4.76%	26.14°	$7.68\frac{°}{m}$	24.32°	$6.82\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
1.73%	10.40%	$2.05m$	18.79°	7.23%	40.60%	$2.20m$	24.53°

Tendenz bzgl. L3 mit $f_B^l = 0.001, f_B^b = 0.01$			
Ziel	limit. RB	Dynamik	Trajektorie
$\beta : + \; y : \approx$	$b := - \; l := +$	$= -$	$= -$

Referenz: L3 mit $f_B^l = 0.001, f_B^b = 0.01$							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
12.40°	$1.71m$	$12.39m$	5.27%	26.19°	$7.90\frac{°}{m}$	24.29°	$6.99\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\tilde{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\tilde{s}^l$	$\pm\theta^l$
1.61%	10.30%	$1.94m$	17.58°	7.41%	41.50%	$2.21m$	24.92°

Allerdings ist ein störender, nicht zu beseitigender Offset in der y-Positionierung vorhanden, der einem realen Einsatz der trainierten Regler im Weg stehen dürfte.
Prinzipiell ist eine sehr gute Regelungstrajekorie erkennbar, die bei Abzug des Offsets auch eine gute Zielpositionierung aufweisen würde.
Die Wirkung der Netzstrukturveränderung auf das Netzlernen kann hier ebenfalls nur als gering eingestuft werden.

5.4 Bewertung der Lenkwinkeldifferenz

Als Beispiel einer forcierenden Randbedingung soll hier die Minimierung der Lenkwinkeldifferenz zweier aufeinanderfolgender Regelungsschritte behandelt werden. Dazu wird gemäß Abbildung 4.4 auf Seite 45 der vom Regler ausgegebene Lenkwinkel für den nächsten Regelungsschritt über ein Zeitverzögerungsglied zum Reglereingang zurückgegeben. Die Bewertung der Winkeldifferenz $\Delta\theta = \theta_t - \theta_{t-1}$ startet erst mit dem zweiten Regelungsschritt $t = 2$, da ein z.B. mit $\theta_0 = 0°$ angenommener Startlenkwinkel in der späteren Lernphase für die Startregelung zu extrem hohen Fehler führen würde; die Reglerausgaben sind zu Beginn einer Regelung betragsmäßig meist sehr groß (s. alle vorigen Regelungsbeispiele).
Als Bewertungsgrenze der Lenkwinkeldifferenz wird $\Delta\theta_B = \pm 10°$ gewählt, so daß innerhalb $[-10°, +10°]$ der Fehler zu Null gesetzt wird und keine Bewertung stattfindet.
Die Grundkonfiguration T1 des Bewertungsmoduls der Lenkwinkeldifferenz ist:

Konfiguration T1 Blockade-Bewertung der Lenkwinkeldifferenz	
Propagierungsstrategie	BP-STOPP
Bewertete Variablen	$\Delta\theta = \theta_t - \theta_{t-1}$
Bewertungsgrenzen	$\Delta\theta_B \in \{\theta_G(\hat{\Phi}_G), \theta_G(\check{\Phi}_G)\}$
Fehlerskalierung f_E	1.0
Bewertungsskalierung $f_B^{\Delta\theta}$	0.01
Anzahl BP-Zeitschritte $\#BP^{\Delta\theta}$	1
Endpositionsbewertung wie	O1

Die Propagierungslänge ist deshalb mit $\#BP = 1$ angesetzt, weil keine Ursächlichkeit der vorherigen Regelungen für das lokale Regelungsverhalten bzgl. der Verletzung der $\Delta\theta$-Begrenzung angenommen werden kann. Für das Analyse-Ergebnis der damit erzielten Regelung siehe Tabelle 5.21. In den Beispielen in Abbildung 5.14 auf Seite 98 ist eine Verbesserung der Zielregelung zu erkennen. Störungen, wie z.B. Pendelung für entferntere Startpositionen treten nicht auf. Der Effekt ist neben einer sehr guten Zielpositionierung aber auch eine gegenteilig erhoffte, erhebliche Verschlechterung der Begrenzung der Lenkwinkeldifferenz. Man hat es hier also mit einer nicht vorhersehbaren und, im Gegensatz zur Wirkung des statischen x-Positionsfehlers, schwer erklärbaren Beeinflussung der Regelung zu tun. Die Orientierung des Fahrzeugs an der Andock-Horizontalen ist deutlich zu erkennen. Für die Untersuchung der Einbeziehung von Randbedingungen ist diese Regelung trotz der besten Zielperformanz aller Konfigurationen nutzlos, da es sich um einen ungewollten Nebeneffekt einer Randbedingungs-Bewertung handelt.

5.4.1 Zusammenfassung

Das Ziel der Verringerung der Lenkwinkeldifferenz durch den Regler konnte nicht erreicht werden. Es ist eher Vergrößerung dieser Differenz erkennbar. Daher ist bereits der Bewertungsansatz in Frage zu stellen.
Die Verbesserung der Zielregelung (bei Erhöhung der eigentlich zu verringernden Stellsignaldynamik) ist ein zwar interessanter, aber nicht erklärbarer und daher für die Untersuchung nutzloser Effekt.

Tabelle 5.21: Analyse

T1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
10.17°	$0.62m$	$13.76m$	17.73%	33.31°	$11.23\frac{°}{m}$	31.01°	$12.97\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\bar{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\bar{s}^l$	$\pm\theta^l$
14.49%	58.40%	$3.42m$	56.05°	0%	0%	$0m$	0°

Tendenz bzgl. O1			
Ziel	limit. RB	Dynamik	Trajektorie
$\beta: + \; y: +$	$-$	$+$	$++$

Die zu beurteilende RB-Performanz ist hier in der Dynamik des Stellsignals enthalten (Vorsicht: ein '-' kennzeichnet hier eine Dynamikverringerung, also eine Verbesserung der Lenkwinkeldifferenz-Minimierung). Man sieht, daß sich diese bzgl. dem Wert $9.41\frac{°}{m}$ der Konfiguration O1 verschlechtert hat. Die Zielperformanz zeigt sich verbessert, was im Vergleich mit den anderen Randbedingungs-Bewertungen positiv zu werten ist. Obwohl die Trajekorienlänge stark zunimmt, wird die Regelung nicht pendelnd.

Referenz: O1							
Endposition		Trajektorie		Stellsignal			
σ_β	σ_y	$\overline{s}$	$\overline{s_d}$	$\overline{\theta}$	$\overline{\Delta\theta}$	$\overline{\theta_k}$	$\overline{\Delta\theta_k}$
11.79°	$0.83m$	$12.74m$	8.56%	30.06°	$9.41\frac{°}{m}$	28.56°	$10.86\frac{°}{m}$
Blockade				Lenkwinkel			
s^b	$\#^b$	$\bar{s}^b$	$\pm\theta^b$	s^l	$\#^l$	$\bar{s}^l$	$\pm\theta^l$
10.15%	46.90%	$2.76m$	55.22°	0%	0%	$0m$	0°

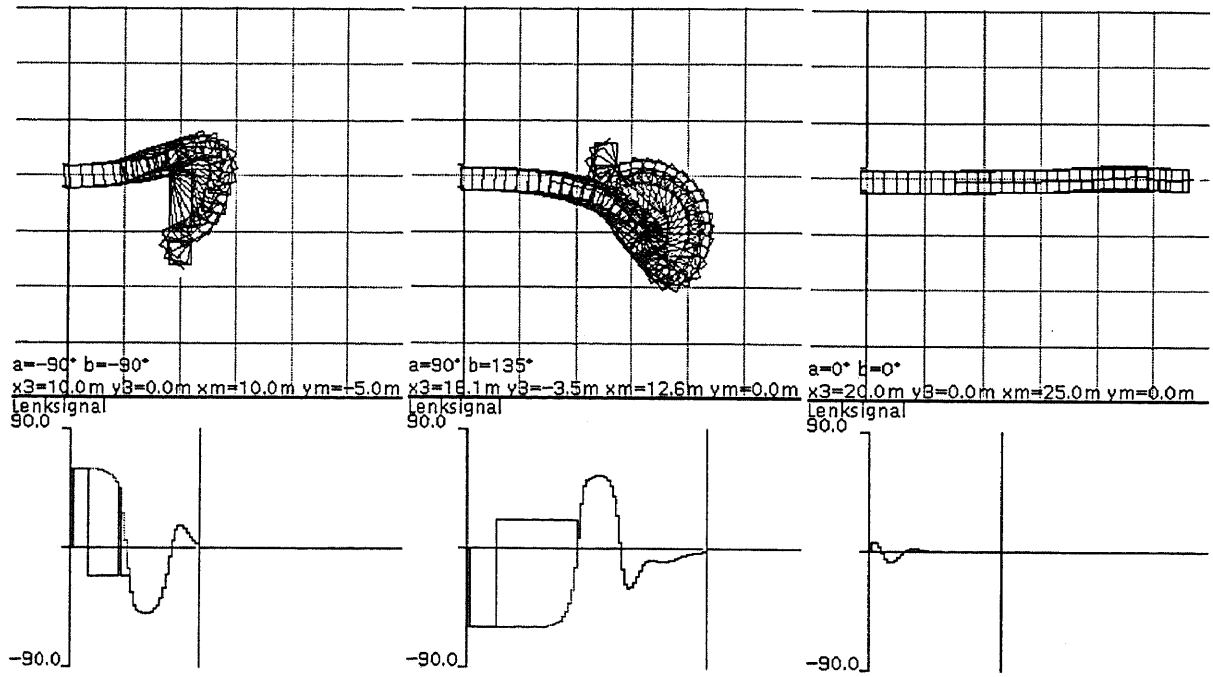


Abbildung 5.14: Ungewollte verbesserte Zielperformanz bei Bewertung der Lenkwinkeldifferenzen

5.5 Fazit

Die Einhaltung von Randbedingungen konnte mit den meisten der hier vorgestellten Bewertungsansätze vom Regler in zufriedenstellender Weise erlernt werden. Teilweise wurde eine fast optimale Vermeidung von nachfolgenden Blockade-Korrekturen erreicht.
Bei Berücksichtigung mehrerer zusätzlicher Optimierungskriterien (Blockade und Lenkwinkelbegrenzung) waren deren Randbedingungs-Performanzen nicht wie erwartet schlechter. Die Einhaltung der beiden Randbedingungen ließ sich weitgehend unabhängig voneinander über die Bewertungsskalierung beeinflussen. Was die Verschlechterung der Zielperformanz betrifft, konnte diese nicht vollständig durch gezielte Gegenmaßnahmen vermieden werden. Eine unterschiedliche Gewichtung der Bewertungsskalierungen hat praktisch nie eine globale Verbesserung aller Performanzwerte zur Folge. Der Effekt der Variation einer bestimmten Bewertungsskalierung ist für das Erlernen der primären Zielregelung meist gegenteilig ausgeprägt. Dagegen kann über eine Erweiterung der Netzstruktur in der Zahl der Neuronen oder Hidden-Layer zwar eine leichte Verbesserung der Zielregelung bei Beibehaltung der Randbedingungs-Performanzen erreicht werden. Die erhoffte Wirkung einer völligen Kompensation dieses Performanzverlusts konnte mit den getesteten Konfigurationen aber nicht festgestellt werden. Außerdem erfordert eine Veränderung der Netzstruktur wahrscheinlich eine Anpassung der anderen Lernparameter. Da dies hier nicht in vollem Umfang ausgetestet werden konnte, waren viele Trainingsdurchgänge mit einer größeren Zahl von Neuronen oder Layern nicht erfolgreich. Bei den zum Vergleich herangezogenen erfolgreich trainierten Netzen ist daher eine weitere Optimierung der Lernparametrierung denkbar. Die tendenzielle Beurteilung bei veränderten Netzstrukturen sind als dementsprechend vorsichtige Aussagen zu betrachten.
Andere Versuche, wie ein zusätzlicher Hilfseingang am Regler (Blockadewinkel Φ) oder unterschiedliche Bewertungsmethoden für die gleiche Randbedingung (Blockade), erbrachten keine grundlegenden Verbesserungen.
Genauso scheint die Wahl der Propagierungsstrategie nicht von entscheidender Bedeutung zu sein. Trotzdem sind hier spezifische Auswirkungen erkennbar, wie z.B. die leichte Extremisierung der Stellsignale bei Anwendung der BP-BYPASS-Strategie.
Als gelungene Beispiele können genannt werden:

1. Die absolute Vermeidung von Blockadesituationen für den Testdatensatz im Fall der Konfiguration B1 mit BP-PROP (Seite 68).

2. Die gleichzeitige Integration der Lenkwinkelbegrenzung und der Blockadevermeidung bei guter Zielregelung mit der Konfiguration L3 (Tabelle 5.19 Seite 92).

Bei ersterer stört aber eine gewisse Verschlechterung der Zielperformanz. Bei der zweiten ist diese eigentlich recht gut, wird aber durch einen störenden y-Offset überlagert.
Dagegen wurde die Begrenzung der Lenkwinkeldifferenz nicht zufriedenstellend bewältigt. Dies mag an einer schlecht in das BPtt-Verfahren zu integrierenden Formulierung dieser Randbedingung liegen, oder es zeigt eventuell auch die Grenzen des Gradientenverfahrens zur parallelen Optimierung hinsichtlich stark unterschiedlicher Kriterien auf. Im Gegensatz zu den halbwegs gelungenen Versuchen mit der Blockade- und Lenkwinkelbewertung wird hier zumindest die generelle Eignung des Verfahrens, beliebige Kostenfunktionen minimieren zu können, in Frage gestellt.
Im Vergleich mit der Ausgangskonfiguration ohne zusätzliche lokale Bewertungen war in

den Regelungen der anderen Konfigurationen keine grundlegend andere Regelungscharakteristik zu erkennen. Es ist zwar bei einigen Reglern, wie denen mit Lenkwinkelbegrenzung, eine verstärkte, der Zielregelung förderliche Orientierung an der Horizontalen zur Andockposition erkennbar. Dies kann jedoch auch mit anderen Mehoden der Lernbeeinflussung erzielt werden und ist deshalb nicht nur auf eine Verbesserung durch die lokalen Bewertungen beschränkt. Beispielsweise wirkt ein statischer x-Positionsfehler in diese Richtung. Auch der glückliche Nebeneffekt bei der $\Delta\theta$-Bewertung (Seite 96) hat eine derartig positive Wirkung.
Ein weiteres herauszustellendes Resultat der Untersuchungen ist die Abgrenzung der Wirkung einer vergrößerten Rückpropagierungslänge der Blockadebewertung gegenüber der Wirkung eines erhöhten effektiven Lernkoeffizienten. Die in Kapitel 5.2.4 auf Seite 79 beschriebenen Performanzverbesserungen sowohl der Zielpositionierung als auch der Korrekturvermeidung sind nicht mit einer Erhöhung des Lernkoeffizienten zu erklären, der in der Zielperformanz gegenteilige Wirkung hat. Dieser eigene Effekt kann somit als Indiz für den in Kapitel 1.2 auf Seite 3 vermuteten Lokalisationscharakter der Randbedingung betrachtet werden.
Um grundsätzliche Aussagen treffen zu können, sind die hier durchgeführten Tests in der Zahl und Repräsentativität sicher nicht ausreichend. Dennoch zeigen die positiven Ergebnisse, daß bei einer nicht zu strengen Anforderung an den Optimierungsgrad verschiedene Optimierungskriterien durch das Netzlernen mittels BPtt miteinander vereinbar sind. Deren gemeinsame Berücksichtigung in Form einer stetigen, differenzierbaren Regelungsfunktion konnte trotz unterschiedlicher Charkteristik der einzelnen Randbedingungen beispielhaft in zufriedenstellender Weise nachgewiesen werden.
Abschließend stellt sich trotzdem die Frage nach der Notwendigkeit der Einbeziehung von Randbedingungen in das Reglerlernen. Mit der Ausgangskonfiguration konnte die Reglerentwicklung ja bereits in einfacher Weise unter Verwendung nachgeschalteter Korrekturmodule bzw. sigmoider Netzausgaben demostriert werden. In der Tat scheint die Berücksichtigung zusätzlicher Optimierungskriterien das Lernen insgesamt zu erschweren und zu einem größeren Implementierungsaufwand zu führen. Dies ist dann noch verbunden mit einer geringeren Performanz in der primären Zielregelung. Allerdings gilt dieser Einwand nur soweit, wie unvermeidbare Unstetigkeiten oder Sprünge in der Bewegung des Stellsignals tolerierbar sind. Durch eine unter Umständen anders geartete neuronale Formulierung dieser Korrekturen kann die Beanspruchung des Stellapparats erheblich reduziert werden.
Ebenso wird erst mit Hilfe von Bewertungen zusätzlich zu definierender Randbedingungen eine gezielte Beeinflussung des Reglerverhaltens möglich, die bestimmte Zielvorstellungen des Anwenders mitberücksichtigt. Über die Bewertungsformulierungen lassen sich so die gewünschten (lokalen) Eigenschaften der Regelung und die damit verbundenen Auswirkungen auf die Regelstreckenzustände gezielt steuern.
Letztendlich kann nicht ausgeschlossen werden, daß das Training eines neuronalen Reglers mit dem BPtt-Verfahren ohne Einbeziehung der lokalen Randbedingungen bei einem anderen Regelungsproblem mit kritischeren Randbedingungen undurchführbar ist. Wenn mehrere Randbedingungen vorhanden und zahlreiche Korrekturvorgänge entlang der Trajektorie notwendig sind, werden erheblich weniger Bewertungen der globalen Regelungsaufgabe zu keiner ausreichenden Anpassung der Netzparameter führen (vgl. Propagierungsstrategien Kapitel 3.4). Spätestens dann wird das parallele Erlernen der Einhaltung zusätzlicher Randbedingungen unbedingt notwendig.

Anhang A

TBU-Modell

Bei konstanter Geschwindigkeit v_1 in P_1 und positivem Längenwert l_1 ergeben sich gemäß Abbildung A.1 folgende kinematischen Gleichungen für die Bewegung des Fahrzeugs:

Die Hilfsgrößen d_1, d_2, d_M (Kurvenradien der Zugmaschine) und γ (Kurvenwinkel zwischen P_M und P_2) berechnen sich zu

$$\begin{aligned} d_1 &= \frac{l_1}{\sin(\theta)} \\ d_2 &= \frac{l_1}{\tan(\theta)} \\ d_M &= \frac{l_2}{\sin(\gamma)} \\ \gamma &= \arctan\left(\frac{l_2}{l_1}\tan(\theta)\right) = \operatorname{arctan2}\left(\frac{l_2}{l_1}\sin(\theta)\,,\ \cos(\theta)\right) \end{aligned}$$

Die gleichförmige Kreisbewegung um Q ergibt für die Berechnung des Winkels $\alpha(h)$ nach der Simulationszeit h und für die Winkelgeschwindigkeit ω in Q

$$\begin{aligned} \alpha(h) &= \alpha(0) + \omega h \\ \omega &= \frac{v_1}{d_1} \end{aligned}$$

Die Koordinaten des Momentanpols Q sind

$$\begin{aligned} x_Q &= x_M - \sin(\epsilon)\, d_M \\ y_Q &= y_M + \cos(\epsilon)\, d_M \end{aligned}$$

Daraus ergeben sich die neuen Koordinaten $x_M(h)$ und $y_M(h)$ zu

$$\begin{aligned} x_M(h) &= x_Q + \sin[\epsilon(h)]\, d_M \\ y_M(h) &= y_Q - \cos[\epsilon(h)]\, d_M \end{aligned}$$

mit dem Winkel des Vektors $\overline{QP_M}$ zur Senkrechten

$$\begin{aligned} \epsilon &= \alpha - \gamma \qquad \text{bzw.} \\ \epsilon(h) &= \alpha(h) - \gamma \end{aligned}$$

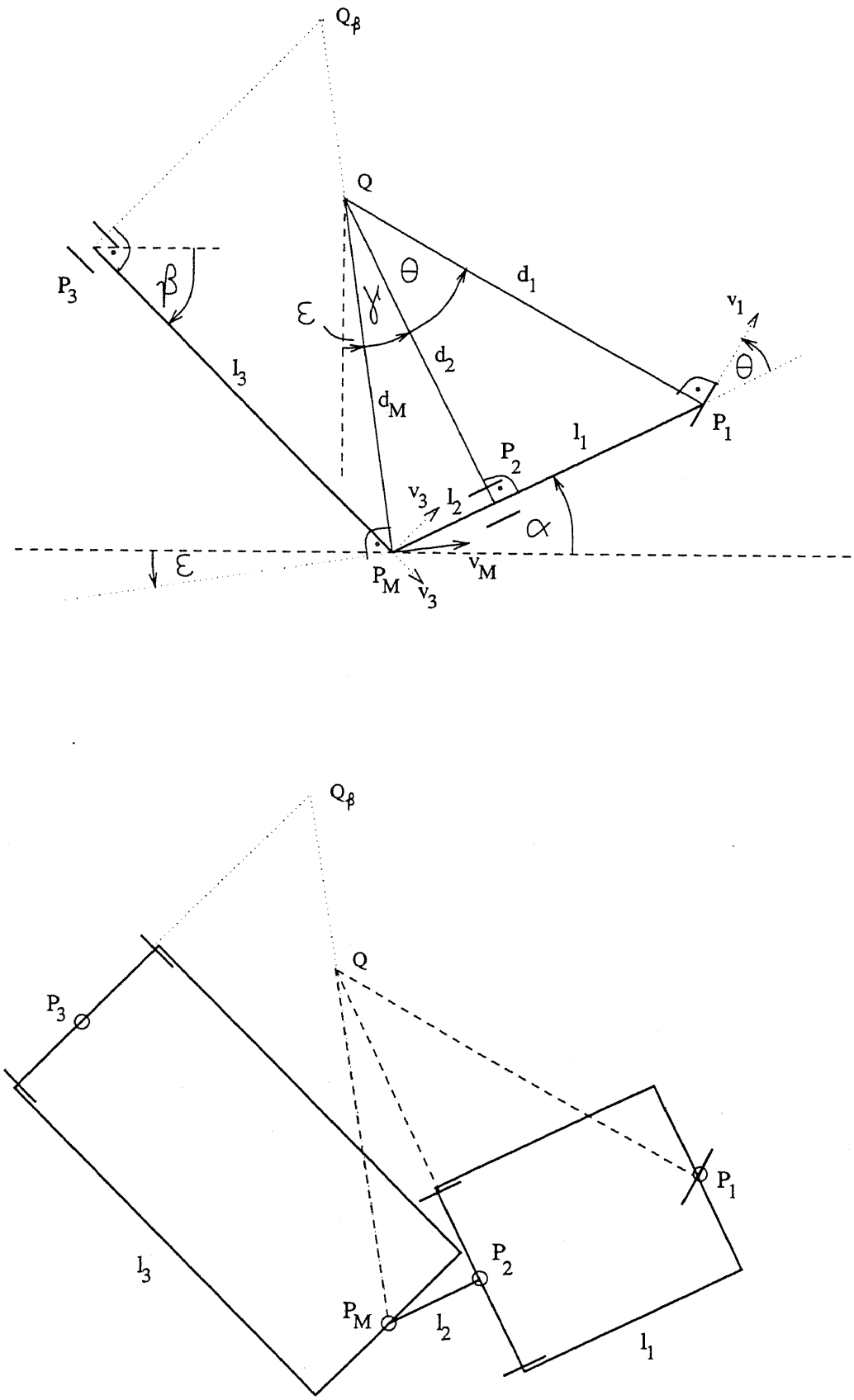


Abbildung A.1: Schematisiertes TBU-Modell

Die Winkelgeschwindigkeit um den während eines Simulationsschritts sich bewegenden Momentanpol des Anhängers ist gleichbedeutend mit der zeitlichen Ableitung des Winkels β, so daß gilt

$$\dot{\beta}(t) = \omega_\beta = \frac{v_M}{l_3} \sin[\epsilon(t) - \beta(t)]$$

mit der Geschwindigkeit v_M in Punkt P_M

$$v_M = \frac{d_M}{d_1} v_1$$

Mittels der Euler-Vorwärts-Integration ergeben sich die Simulationsgleichungen zu

$$\begin{pmatrix} \alpha' \\ x'_M \\ y'_M \end{pmatrix} = \begin{pmatrix} \alpha \\ x_M \\ y_M \end{pmatrix} + \begin{pmatrix} \omega h \\ [\sin(\epsilon') - \sin(\epsilon)] d_M \\ [\cos(\epsilon) - \cos(\epsilon')] d_M \end{pmatrix}$$

$$\beta' = \beta + \frac{v_M}{l_3} \sin(\epsilon - \beta)$$

Sonderfälle

$\underline{l_2 \neq 0}$:

1. $\sin(\theta) = 0$

$$\begin{aligned} d_1,\, d_2,\, d_M &\rightarrow \infty \\ \omega &= 0 \\ v_M &= \text{sign}[\cos(\theta)]\, v_1 \end{aligned}$$

2. $\cos(\theta) = 0$

$$\begin{aligned} d_1 &= \text{sign}[\sin(\theta)]\, l_1 \\ d_2 &= 0 \\ \gamma &= 0 \\ d_M &= 0 \\ v_M &= 0 \\ x_Q &= x_2 \\ y_Q &= y_2 \end{aligned}$$

$\underline{l_2 = 0}$:

1. $\sin(\theta) = 0$

$$\begin{aligned} d_1,\, d_2,\, d_M &\rightarrow \infty \\ \omega &= 0 \\ v_M &= \text{sign}[\cos(\theta)]\, v_1 \end{aligned}$$

2. $\cos(\theta) = 0$

$$
\begin{aligned}
d_1 &= \text{sign}[\sin(\theta)]\, l_1 \\
d_2 &= 0 \\
\gamma &= \text{arctan2}\left(\frac{l_2}{l_1}\sin(\theta)\ ,\ 0\right) \\
d_M &= \text{sign}[\sin(\gamma)]\, l_2 \\
v_M &= \omega d_M \\
x_Q &= x_2 \\
y_Q &= y_2
\end{aligned}
$$

Anhang B

Fehlerskalierung und Wertebereiche

Die Wahl der Fehlerskalierung zu bewertender Variablen hängt entscheidend von der Wertebereichsdarstellung der Variablen ab. Die unterschiedlichen Skalierungen für die Fehler der Positionsvariablen bei der Endpositionsbewertung der Grundkonfiguration (Kapitel 4.3 Seite 44) ist dadurch erklärbar, daß deren Bewertung nicht in einer neuronalen Wertebereichsdarstellung der Variablen erfolgt, sondern in der Realbereichsdarstellung. Dies wird durch folgendes Gedankenmodell deutlich:
Wir gehen von folgender einheitlich neuronalen Wertebereichsdarstellung der Zustandsvariablen in allen Modulen aus:

Variablen	θ	α	β	x	y
Wertebereich neuronal			$[0,1]$		
Entsprechung im Realen	$[-\frac{\pi}{2},+\frac{\pi}{2}]$	$[-\frac{\pi}{2},+\frac{\pi}{2}]$	$[-\frac{\pi}{2},+\frac{\pi}{2}]$	$[-5,5]$	$[-5,5]$
Trafo-Skalierungsfaktor	π	π	π	10	10

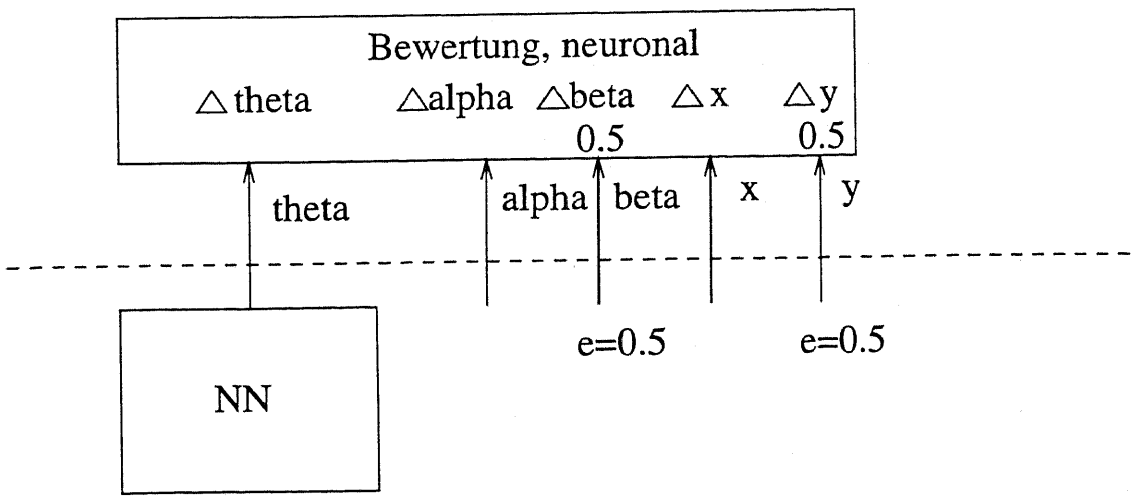


Abbildung B.1: Rückpropagierte Fehler der Bewertung auf neuronaler Ebene

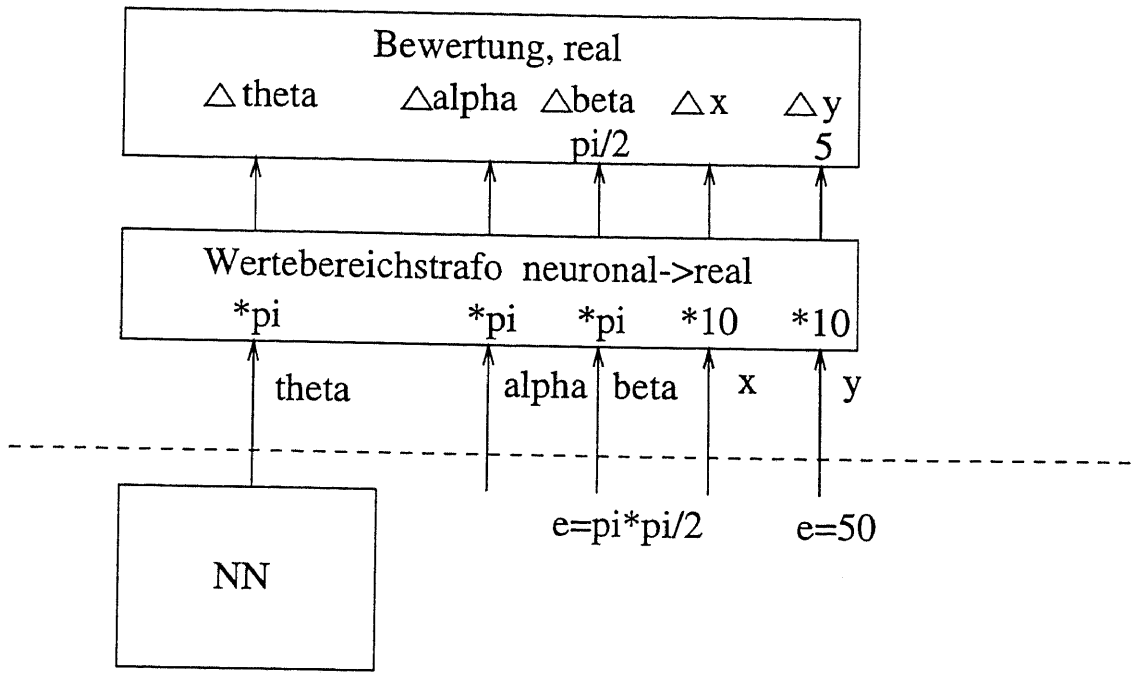


Abbildung B.2: Rückpropagierte Fehler der Bewertung auf realer Ebene

Zum β-Eingang des neuronalen Bewertungsmoduls (Bild B.1) wird dann bei neuronalen Sollabweichungen $\Delta\beta = 0.5$ (real: $\frac{\pi}{2}$) in der neuronalen Darstellung ein Fehler von $e^n_\beta = 0.5$ zurückpropagiert. Für eine Sollabweichung $\Delta y = 0.5$ (real: 5) liegt am y-Eingang der Bewertung in der neuronalen Darstellung ein Fehler $e^n_y = 0.5$ an.
Will man nun die Bewertung in der realen Bereichsdarstellung so durchführen (Bild B.2), daß auf neuronaler Ebene gleich große Rückpropagierungs-Fehler auftreten, so muß eine doppelte Skalierung jedes Variablenfehlers vorgenommen werden:

1. Die Transformation vom Real- in den Neuronalbereich skaliert bei der Rückpropagierung die Fehler mit den in obiger Tabelle angegebenen Werten.

2. Da der rückpropagierte Fehler bei Verwendung der quadratischen Fehlerfunktion proportional zur Sollabweichung ist, tritt hier nochmals die Skalierung der neuronal→real-Wertebereichstransformation auf.

Im Beispiel: Zum Eingang des realen Bewertungsmoduls werden die realen Sollabweichungen $\Delta\beta = \frac{\pi}{2}$ und $\Delta y = 5$ als Fehler $\tilde{e}^r_\beta = \frac{\pi}{2}$ und $\tilde{e}^r_y = 5$ rückpropagiert[1]. Dies bedeutet hier gegenüber dem Fehler am Eingang der neuronalen Bewertung eine Skalierung mit den Faktoren $s_\beta = \pi$ und $s_y = 10$. Die Rückpropagierung durch die Wertebereichstransformation zur neuronalen Darstellung der Zustandskoordinaten bewirkt eine nochmalige Multiplikation mit diesen Skalierungsfaktoren, so daß die dort am Eingang anliegenden Fehler nun allgemein $\tilde{e}^n_\beta = e^n_\beta * s^2_\beta$ bzw. $\tilde{e}^n_y = e_y * s^2_y$ betragen. Um an diesem Punkt also mittels der realen Bewertung die gleichen Propagierungsfehlerwerte zu erhalten, wie mittels der neuronalen Bewertung, müssen der β-Fehler und der y-Fehler mit den Kehrwerten von s^2_β bzw. s^2_y skaliert werden. Entsprechendes gilt für die Fehlerskalierungen der anderen

[1] bei Verwendung der quadratischen Fehlerfunktion $E(\underline{x}) = \frac{1}{2}\sum_i (x_{i,Soll} - x_{i,Ist})^2$.

Eingangs-Variablen des Bewertungsmoduls.
Geht man also von einer in etwa gleichgewichteten Fehlerskalierung für die Eingangs-Variablen der Bewertung auf neuronaler Ebene aus[2], so kann man mittels Kenntnis der Skalierungsfaktoren der Wertebereichstransformationen die Fehlerskalierung einer Bewertung mit einer alternativen Darstellung der Variablenwerte berechnen. Dies wurde so auch für die in dieser Diplomarbeit im Realen stattfindenden Bewertungen durchgeführt, bei Kenntnis der sich eignenden Skalierungsfaktoren für Bewertungen im Neuronalen aus [1].

[2] was durchaus sinnvoll erscheint, da die Adaption der Gewichte ja ebenfalls auf neuronaler Ebene erfolgt.

Anhang C

Abkürzungen

BP	Backpropagation.
BPtt	Backpropagation through time.
FP	Forwardpropagation.
NN	Neuronales Netz.
RB	Randbedingung.
TBU	Truck Backer-Upper. Die Bezeichnung der Regelstrecke, einem rückwärtsfahrenden 2-achsigen Fahrzeug mit 1-achsigem Anhänger.

Propagierungsstrategien

FP-STOPP	Stopp der Regelung bei Auftreten der ersten Korrektur. Rückpropagierung dieser Korrektur-Bewertung und aller davor ansetzenden forcierenden Bewertungen.
BP-STOPP	Regelung bis Endekriterium der Zielregelung erfüllt ist. Rückpropagierung aller Bewertungen bis zum ersten Auftreten einer Korrektur auf dem Propagierungsweg.
BP-NULLUNG	Regelung bis Endekriterium der Zielregelung erfüllt ist. Rückpropagierung aller Bewertungen. Bei Propagierung durch den Begrenzungsbereich einer Korrektur wird der Rückpropagierungsfehler in diesem Zweig auf Null gesetzt.
BP-PROP	Berechnung der numerischen Ableitung zur Propagierung durch die Korrekturfunktion. Bei den hier verwendeten Korrekturen ist dies identisch mit der Methode der BP-NULLUNG, da die Begrenzungsfunktionen im Begrenzungsbereich die Ableitung 0 besitzen.
BP-BYPASS	Statt der Ableitung oder Fehler-Nullung wird hier der zum Ausgang der Korrekturfunktion propagierte Fehler ohne Veränderung zum Eingangs-Stellsignal durchgeschleift.

Literaturverzeichnis

[1] Hartmann, B.: *Regelung mit neuronalen Netzen am Beispiel Truck Backer-Upper*, FAW Ulm, Technical Report FAW-TR-94008, 1994

[2] Hecht-Nielsen,R.: *Neurocomputing*, Addison Wesley, 1989

[3] Rumelhart,D. E., Hinton,G. E., Williams,R. J.:"Learnin Internal Representations by Error Propagation" in: Rumelhart,D. E., McClelland,J. L., the PDP Research Group: *Parallel Distributed Processing, Volume 1: Foundations*, 318-362, MIT-Press, 1986

[4] Widrow,B., Nguyen,D.: "The Truck Backer-Upper: An Example of Self-Learning in Neural Networks" in: Miller,W. Th., Sutton,R. S., Werbos,P. J.: *Neural Networks for Control*, 287-299, MIT Press, 1990. [Mit zusätzlichen Angaben von Trainingsdaten und modifizierter Problembeschreibung im dortigen Anhang A.5]

[5] Zell,A.: *Simulation Neuronaler Netze*, Addison Wesley, 1994.

Liste der bisher erschienenen Ulmer Informatik-Berichte

Einige davon sind per FTP von `ftp.informatik.uni-ulm.de` erhältlich
Die mit * markierten Berichte sind vergriffen

List of technical reports published by the University of Ulm

Some of them are available by FTP from `ftp.informatik.uni-ulm.de`
Reports marked with * are out of print

91-01 *Ker-I Ko, P. Orponen, U. Schöning, O. Watanabe*
Instance Complexity

91-02* *K. Gladitz, H. Fassbender, H. Vogler*
Compiler-Based Implementation of Syntax-Directed Functional Programming

91-03* *Alfons Geser*
Relative Termination

91-04* *J. Köbler, U. Schöning, J. Toran*
Graph Isomorphism is low for PP

91-05 *Johannes Köbler, Thomas Thierauf*
Complexity Restricted Advice Functions

91-06* *Uwe Schöning*
Recent Highlights in Structural Complexity Theory

91-07* *F. Green, J. Köbler, J. Toran*
The Power of Middle Bit

91-08* *V.Arvind, Y. Han, L. Hamachandra, J. Köbler, A. Lozano, M. Mundhenk, A. Ogiwara, U. Schöning, R. Silvestri, T. Thierauf*
Reductions for Sets of Low Information Content

92-01* *Vikraman Arvind, Johannes Köbler, Martin Mundhenk*
On Bounded Truth-Table and Conjunctive Reductions to Sparse and Tally Sets

92-02* *Thomas Noll, Heiko Vogler*
Top-down Parsing with Simulataneous Evaluationof Noncircular Attribute Grammars

92-03 *Fakultät für Informatik*
17. Workshop über Komplexitätstheorie, effiziente Algorithmen und Datenstrukturen

92-04* *V. Arvind, J. Köbler, M. Mundhenk*
Lowness and the Complexity of Sparse and Tally Descriptions

92-05* *Johannes Köbler*
Locating P/poly Optimally in the Extended Low Hierarchy

92-06* *Armin Kühnemann, Heiko Vogler*
Synthesized and inherited functions -a new computational model for syntax-directed semantics

92-07* *Heinz Fassbender, Heiko Vogler*
A Universal Unification Algorithm Based on Unification-Driven Leftmost Outermost Narrowing

92-08* *Uwe Schöning*
On Random Reductions from Sparse Sets to Tally Sets

92-09* *Hermann von Hasseln, Laura Martignon*
Consistency in Stochastic Network

92-10 *Michael Schmitt*
A Slightly Improved Upper Bound on the Size of Weights Sufficient to Represent Any Linearly Separable Boolean Function

92-11 *Johannes Köbler, Seinosuke Toda*
On the Power of Generalized MOD-Classes

92-12 *V. Arvind, J. Köbler, M. Mundhenk*
Reliable Reductions, High Sets and Low Sets

92-13 *Alfons Geser*
On a monotonic semantic path ordering

92-14* *Joost Engelfriet, Heiko Vogler*
The Translation Power of Top-Down Tree-To-Graph Transducers

93-01 *Alfred Lupper, Konrad Froitzheim*
AppleTalk Link Access Protocol basierend auf dem Abstract Personal Communications Manager

93-02 *M.H. Scholl, C. Laasch, C. Rich, H.-J. Schek, M. Tresch*
The COCOON Object Model

93-03 *Thomas Thierauf, Seinosuke Toda, Osamu Watanabe*
On Sets Bounded Truth-Table Reducible to P-selective Sets

93-04 *Jin-Yi Cai, Frederic Green, Thomas Thierauf*
On the Correlation of Symmetric Functions

93-05 *K.Kuhn, M.Reichert, M. Nathe, T. Beuter, C. Heinlein, P. Dadam*
A Conceptual Approach to an Open Hospital Information System

93-06 *Klaus Gaßner*
Rechnerunterstützung für die konzeptuelle Modellierung

93-07 *Ullrich Keßler, Peter Dadam*
Towards Customizable, Flexible Storage Structures for Complex Objects

94-01 *Michael Schmitt*
On the Complexity of Consistency Problems for Neurons with Binary Weights

94-02 *Armin Kühnemann, Heiko Vogler*
A Pumping Lemma for Output Languages of Attributed Tree Transducers

94-03 *Harry Buhrman, Jim Kadin, Thomas Thierauf*
On Functions Computable with Nonadaptive Queries to NP

94-04 *Heinz Faßbender, Heiko Vogler, Andrea Wedel*
Implementation of a Deterministic
Partial E-Unification Algorithm for Macro Tree Transducers

94-05 *V. Arvind, J. Köbler, R. Schuler*
On Helping and Interactive Proof Systems

94-06 *Christian Kalus, Peter Dadam*
Incorporating record subtyping into a relational data model

94-07 *Markus Tresch, Marc H. Scholl*
A Classification of Multi-Database Languages

94-08 *Friedrich von Henke, Harald Rueß*
Arbeitstreffen Typtheorie: Zusammenfassung der Beiträge

94-09 *F.W. von Henke, A. Dold, H. Rueß, D. Schwier, M. Strecker*
Construction and Deduction Methods for the Formal Development of Software

94-10 *Axel Dold*
Formalisierung schematischer Algorithmen

94-11 *Johannes Köbler, Osamu Watanabe*
New Collapse Consequences of NP Having Small Circuits

94-12 *Rainer Schuler*
On Average Polynomial Time

94-13 *Rainer Schuler, Osamu Watanabe*
Towards Average-Case Complexity Analysis of NP Optimization Problems

94-14 *Wolfram Schulte, Ton Vullinghs*
Linking Reactive Software to the X-Window System

94-15 *Alfred Lupper*
Namensverwaltung und Adressierung in Distributed Shared Memory-Systemen

94-16 *Robert Regn*
Verteilte Unix-Betriebssysteme

94-17 *Helmuth Partsch*
Again on Recognition and Parsing of Context-Free Grammars: Two Exercises in Transformational Programming

94-18 *Helmuth Partsch*
Transformational Development of Data-Parallel Algorithms: an Example

95-01 *Oleg Verbitsky*
On the Largest Common Subgraph Problem

95-02 *Uwe Schöning*
Complexity of Presburger Arithmetic with Fixed Quantifier Dimension

95-03 *Harry Buhrman, Thomas Thierauf*
The Complexity of Generating and Checking Proofs of Membership

95-04 *Rainer Schuler, Tomoyuki Yamakami*
Structural Average Case Complexity

95-05 *Klaus Achatz, Wolfram Schulte*
Architecture Indepentent Massive Parallelization of Divide-And-Conquer Algorithms

95-06 *Christoph Karg, Rainer Schuler*
Structure in Average Case Complexity

95-07 *P. Dadam, K. Kuhn, M. Reichert, T. Beuter, M. Nathe*
ADEPT: Ein integrierender Ansatz zur Entwicklung flexibler, zuverlässiger kooperierender Assistenzsysteme in klinischen Anwendungsumgebungen

95-08 *Jürgen Kehrer, Peter Schulthess*
Aufbereitung von gescannten Röntgenbildern zur filmlosen Diagnostik

95-09 *Hans-Jörg Burtschick, Wolfgang Lindner*
On Sets Turing Reducible to P-Selective Sets

95-10 *Boris Hartmann*
Berücksichtigung lokaler Randbedingung bei globaler Zieloptimierung mit neuronalen Netzen am Beispiel Truck Backer-Upper

Ulmer Informatik-Berichte

ISSN 0939-5091

Herausgeber: Fakultät für Informatik

Universität Ulm, Oberer Eselsberg, D-89069 Ulm