



Technische
Universität
Braunschweig

Bachelorarbeit Nr. 811

Aktive Regelung einer Zylinderumströmung bei variierender Reynoldszahl durch bestärkendes Lernen

cand. mach. Fabian Gabriel

Ausgegeben: Prof. Dr.-Ing. R. Radespiel
Institut für Strömungsmechanik
Institutsleiter: Prof. Dr.-Ing. R. Radespiel
Technische Universität Braunschweig

Betreuer: Dr.-Ing. Andre Weiner (TU Braunschweig)

Veröffentlichung: 30.09.2021

Übersicht

Bei einer 2D-Zylinderumströmung entsteht im Nachlauf des Zylinders eine Kármánsche Wirbelstraße, die oszillierende Kräfte auf den Zylinder bewirkt. Durch einen Agenten, der mit einem tiefen, bestärkten Lernalgorithmus trainiert wurde, sollen diese Kräfte reduziert werden. Um dies zu erreichen rotiert der Agent basierend auf den Druckwerten auf der Zylinderoberfläche den Zylinder und regelt so aktiv die Strömung. Bei diesem Algorithmus kommt bei der Wahl der Drehgeschwindigkeit eine Wahrscheinlichkeitsverteilung zum Einsatz, die häufig eine Normalverteilung ist. Allerdings können dabei auch andere Wahrscheinlichkeitsverteilungen gewählt werden. In diesem Fall wurde der Unterschied in Trainingsgeschwindigkeit und -ergebnis zwischen der Normalverteilung, einer Normalverteilung, auf die eine Tangens hyperbolicus Funktion angewendet wurde, und der Beta-Verteilung untersucht. Die Beta-Verteilung erzielt dabei die besten Ergebnisse und kann die Belohnung des Agenten mit der Normalverteilung schon nach 13 Episoden, im Vergleich zu 86 Episoden, erreichen. Zudem kann die Beta-Verteilung die Reduktion des durchschnittlichen Widerstandsbeiwertes im Vergleich zur Normalverteilung um 0,43 Prozentpunkte auf 5,48 % steigern. Der Agent, der die Tangens hyperbolicus Funktion verwendete konnte ebenfalls bessere Ergebnisse als die unveränderte Normalverteilung erzielen, indem dieser die Reduktion des durchschnittlichen Widerstandsbeiwert um 0,1 Prozentpunkt auf 5,15 % steigerte.

Daraufhin wurde ein Agent mit einer Beta-Verteilung auf 3 verschiedene aber im Training konstante Reynoldszahlen trainiert und auf eine zeitvariante Strömung angewandt. Der maximale Widerstandsbeiwert konnte dabei um 14,43 % reduziert werden. Zeitweise konnte die Bildung der Kármánsche Wirbelstraße sogar vollständig verhindert werden. Währenddessen konnte der durchschnittliche Widerstandsbeiwert um 36,21 % gesenkt werden.

Abstract

In a 2D cylinder flow, a Kármán vortex street is created in the wake of the cylinder, which causes oscillating forces on the cylinder. An agent trained with a deep reinforcement learning algorithm is designed to reduce these forces. To achieve this, the agent rotates the cylinder based on the pressure values on the cylinder surface and thus actively regulates the flow. This algorithm uses a probability distribution, which is often a normal distribution, to select the rotation speed. However, other probability distributions can also be chosen. In this case, the difference in training speed and result between the normal distribution, a normal distribution to which a tangent hyperbolic function was applied, and the beta distribution was investigated. The beta distribution achieves the best results and can achieve the agent's reward with the normal distribution after only 13 episodes, compared to 86 episodes. In addition, the beta distribution can increase the reduction of the average drag by 0.43 percentage points to 5.48% compared to the normal distribution. The agent using the tangent hyperbolic function was also able to achieve better results than the unmodified normal distribution, reducing the average drag by 0.1 percentage point to 5.15%.

An agent with a beta distribution was then trained on 3 different but constant Reynolds numbers and applied to a time-varying flow. The maximum drag was reduced by 14.43 %. At times, the formation of the Kármán vortex street could even be completely prevented. Meanwhile, the average drag coefficient was reduced by 36.21 %.

Inhaltsverzeichnis

Nomenklatur	viii
1 Einleitung	1
1.1 Motivation	1
1.2 Stand der Technik	1
1.3 Vorgehen	6
2 Theorie	7
2.1 Zylinderumströmung	7
2.2 Wahrscheinlichkeitsverteilungen	8
2.2.1 Normalverteilung	8
2.2.2 Beta-Verteilung	11
3 Methodik	14
3.1 Formulierung des Strömungsproblems	14
3.2 Gitterabhängigkeitsstudie	16
3.2.1 Gittererstellung	16
3.2.2 Zeitschritt	17
3.2.3 Ergebnisse der Gitterabhängigkeitsstudie	17
3.3 Bestärkendes Lernen	18
3.3.1 Umgebung	18
3.3.2 Agent	19
3.3.3 Verbesserungen in der Umsetzung	21
3.4 Parallelisierung der Regelrandbedingung	21
4 Abhängigkeit der Strömungsbeiwerte von der Reynoldszahl	23
4.1 Strouhal-Zahl	23
4.2 Widerstandsbeiwerte	24
4.3 Auftriebsbeiwerte	25
5 Stabilisierungsmaßnahmen	26
5.1 Normalverteilung	26
5.2 Auswirkungen der Maßnahmen	32
5.3 Vergleich mit Literaturdaten	36
6 Zeitvariante Anströmung	37
6.1 Training	37
6.2 Veränderungen in der Simulationsumgebung	39
6.3 Ergebnis	39
7 Zusammenfassung	43
Literaturverzeichnis	45
Abbildungsverzeichnis	47
Tabellenverzeichnis	49

Nomenklatur

Lateinische Bezeichnungen

c_D	Widerstandsbeiwert
c_L	Auftriebsbeiwert
d	Durchmesser
f	Frequenz
G	Ertrag (Return)
H	Höhe
N	Anzahl
p	Druck
R	Belohnung (Reward)
Re	Reynoldszahl, Ul/ν
r	Radius
Sr	Strouhal-Zahl
t	Zeit
U, V	Geschwindigkeitskomponenten
x, y	kartesische Koordinaten

Griechische Bezeichnungen

α	1. Parameter der Beta-Verteilung
β	2. Parameter der Beta-Verteilung
γ	Modus, Reduktionsfaktor
σ	Standardabweichung
μ	Erwartungswert
ν	kinematische Viskosität
∇	Nabla Operator
ω	Winkelgeschwindigkeit
π	Regelwerk, policy
ρ	Dichte

Indizes

∞	ungestörte Aussenströmung; Anströmung
min	Minimum
max	Maximum
*	Entdimensionalisiert
D	Widerstand

L Auftrieb

Abkürzungen

CFD Computational Fluid Dynamics

CFL Courant Friedrichs Lewy

RL Reinforcement Learning

DRL Deep Reinforcement Learning

PPO Proximal Policy Optimization

Kapitel 1

Einleitung

1.1 Motivation

Durch den Klimawandel und den allgemeinen Kostendruck stehen Flugzeugentwickler aber auch Autobauer vor dem Problem, möglichst treibstoff- und damit energiesparende Fahrzeuge zu entwickeln. Naturgemäß ist bei einem Flugzeug der Luftwiderstand die wesentliche Widerstandskraft, die überwunden werden muss. Um diesen Luftwiderstand zu reduzieren und so Treibstoff zu sparen, gibt es verschiedene Möglichkeiten. Zum einen kann die Form des Flugzeugs und der Tragflächen optimiert werden. Zum anderen kann die Strömung selbst aktiv oder passiv kontrolliert werden. Eine passive Kontrolle kann dabei zum Beispiel über eine poröse Außenschicht [1] oder Haifischschuppen nachgeahmten Riplets [16] umgesetzt werden. Bei der aktiven Kontrolle wird externe Energie verwendet, um die Strömung zu beeinflussen. Ein Beispiel für eine solche Anwendung wäre das Ausblasen und Absaugen der Strömung an den Tragflügeln [18].

Die effiziente Kontrolle dieser Strömungen stellt allerdings ein großes Problem dar, da es keine einfache Möglichkeit gibt, anhand von einer Reihe von Druck- oder Geschwindigkeitsmessungen an der Oberfläche des Objekts eine Kontrollstrategie zu entwickeln. Die Hochdimensionalität und die daraus resultierenden hohen Berechnungskosten der Simulation machen eine analytische Lösung und eine Echtzeitvorhersage der Strömung aktuell unmöglich [7]. Allerdings gab es in den letzten Jahren erhebliche Fortschritte im Bereich des tiefen, bestärkenden Lernens im Bezug auf die Entwicklung von Kontrollstrategien, so dass dieses als eine vielversprechende Methode erscheint, eine aktive Strömungskontrolle umsetzen zu können. Dabei ist ein schnelles, effizientes und stabiles Training noch nicht selbstverständlich, so dass es hier noch Potential für Verbesserungen gibt.

1.2 Stand der Technik

Wenn sich die Grenzschicht an einem umströmten Zylinder ablöst, bildet sich ein Wirbel, wobei an der Oberfläche eine Strömung in entgegengesetzte Richtung entsteht. Der Wirbel zieht dabei auch die Scherschicht von der anderen Seite der Wirbelschleppe an und unterbricht dabei den weiteren Zufluss von Wirbelstärke zum Wirbel [17]. Der zurückbleibende Wirbel wird dabei Von-Kármán-Wirbel genannt. Die gegenläufigen und aufeinanderfolgenden Wirbel im Nachlauf werden eine Kármánsche Wirbelstraße genannt.

Die aktive Strömungskontrolle kann in zwei Bereiche aufgeteilt werden. Bei der open-loop-Kontrolle funktioniert die Einwirkung auf die Umgebung ohne Einbeziehung von Messdaten und Sensoren während der Kontrolle. Bei der closed-loop-Kontrolle hingegen entscheidet die Ausgabe der Sensoren über die Kontrollstrategie. Daher hat die closed-loop-Kontrolle zwar das größere Potential, die Möglichkeiten der aktiven Strömungskontrolle auszunutzen, ist aber auch in der Umsetzung deutlich aufwändiger [13].

Um eine closed-loop-Strategie zu entwickeln hat, sich das bestärkende Lernen (RL, von Englisch: reinforcement learning) als aussichtsreiche Methode etabliert. Bei diesem lernt der Software-Agent durch die Interaktion mit der Umgebung. Die Umgebung wird dabei dem Agenten häufig nicht direkt zur Verfügung gestellt, sondern beispielsweise durch Sensoren nur einzelne lokale Stellen dem Agenten übergeben. Diese Darstellung der Umgebung wird auch Zustand genannt. Dem Agenten wird ein Ziel vorgegeben, das er erreichen soll, und durch verschiedene ausnutzende oder erkundende Aktionen nimmt der Agent Einfluss auf die Umgebung. Die Aktionen werden dabei anhand des Zustandes nach einem Regelwerk (policy) ausgewählt. Die Umgebung wiederum gibt dem Agenten ein Belohnungssignal zurück, anhand dessen der Agent erkennen kann, ob er seinem Ziel näher kommt. Durch simples Ausprobieren (Erkunden) entdeckt der Agent dabei mögliche positive Aktionen und durch das Ausnutzen dieser Aktionen lernt der Agent, nach vielen Iterationen das Ziel zu erreichen. Dabei wird versucht, den Zustand auf den zu erwartenden Ertrag abzubilden (state value function). Diese state value function gibt also an, wie hoch der zu erwartende Ertrag bei dem aktuellen Zustand ist. Um auch den Einfluss der Aktion mitzubedenken, wird neben dem Zustand auch die Aktion auf den zu erwartenden Ertrag abgebildet (state-action value function oder auch Q-function). Beim policy-based RL wird direkt versucht, die optimale Strategie zu erlernen. Dabei wird der Zustand auch auf die optimale Aktion abgebildet [12]. Wenn dabei zur Approximation einer oder mehrerer dieser Abbildungen zudem künstliche neuronale Netze zum Einsatz kommen, spricht man von tiefem RL (DRL, von Englisch: deep reinforcement learning).

Neben der aktiven Strömungskontrolle, kann das DRL in der Strömungsmechanik auch für eine dynamische Formoptimierung eingesetzt werden. Dies sind hochdimensionale Optimierungsprobleme, für die das DRL besonders geeignet ist. Die Umgebung ist dabei meist eine Strömungssimulation als mathematisches Problem und dessen Lösung. Der Agent kann dann als weiteres Programm, typischerweise über eine Änderung einer Randbedingung, auf die Umgebung Einfluss nehmen. Das kann zum Beispiel durch die Rotation eines Zylinders [15] oder durch das Absaugen und Ausblasen an einer Düse [7] umgesetzt werden.

Als Algorithmus für das DRL in der Strömungsmechanik hat sich der proximal policy optimization (PPO) Algorithmus bewährt. Dabei handelt es sich um eine policy gradient method. Eine Policy ist die Abbildung von einem Zustand auf eine Aktion, so dass bei einem bestimmten Zustand eine bestimmte Aktion ausgewählt wird. Bei einer policy gradient method werden bei der Policy einzelne Aktions-Zustands-Paare so abgeändert, dass die zu erwartende Belohnung maximiert wird. Die Umgebung gibt dem Agenten einerseits dieses Belohnungssignal und andererseits den aktuellen Beobachtungsstatus zurück. In der Abbildung 1.1 ist das Schaubild eines solchen DRL-Systems zu sehen. Es gibt dabei zwei unterschiedliche neuronale Netze: Das eine ist der Akteur, der den Zusammenhang von Zustand und optimaler Aktion approximiert und versucht, anhand dessen die optimale Aktion zu wählen. Das Andere wiederum ist der Kritiker, der den Zusammenhang von Zustand und zu erwartende kumulative Belohnung approximiert. Dadurch versucht er zu berechnen, ob die gewählte Aktion des Akteurs einen positiven Effekt auf die zu erwartende kumulative Belohnung hat.

Durch diese zwei Netzwerke zählt PPO auch zu den actor-critic Algorithmen. Eine Besonderheit des PPO-Algorithmus im Gegensatz zu anderen DRL-Algorithmen ist, dass eine neu generierte Policy nur zu einem bestimmten, geringen Grad von der alten Policy abweichen darf. Dadurch stabilisiert sich das Training, da schon geringe Änderungen der Policy große Auswirkungen auf die Umgebung haben kann. Um auch eine Erkundung von möglicherweise besseren Aktionen zu ermöglichen, wird die Aktion im Training nicht deterministisch ausgewählt, sondern anhand einer Wahrscheinlichkeitsverteilung. Für kontinuierliche Aktionsräume wird dabei häufig eine Normalverteilung verwendet [10].

Bei einer solchen Normalverteilung können, wenn auch mit geringer Wahrscheinlichkeit, theoretisch unendlich große oder kleine Aktionen ausgewählt werden. In der Realität sind die Aktionen allerdings auf einen bestimmten Aktionsbereich limitiert, da zum Beispiel ein Motor nur eine begrenzte Kraft hat und sich daher nicht beliebig schnell drehen kann. 2018 haben Tuomas

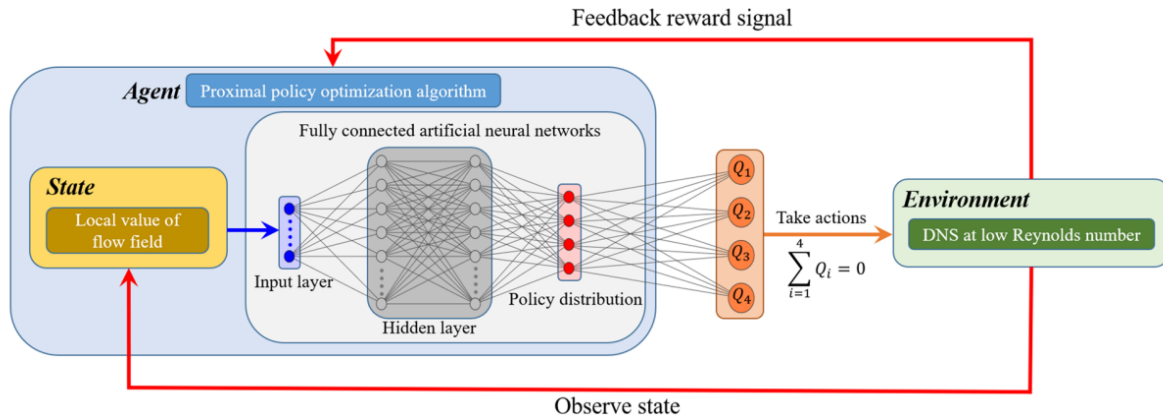


Abb. 1.1: Schaubild eines DRL-Systems. Diese Abbildung ist die Abbildung 4 des Artikels „Robust active flow control over a range of Reynolds numbers using an artificial neural network trained through deep reinforcement learning“ [6].

Haarnoja et al. [3] einen Artikel veröffentlicht, in dem sie eine Möglichkeit vorgestellt haben, eine Normalverteilung auf einen Aktionsbereich zu begrenzen. Dabei wird auf eine unbegrenzte Normalverteilung die Tangens hyperbolicus Funktion als Stauchungsfunktion angewendet. Dadurch entsteht eine Verteilung, die auf den Bereich von $[-1,1]$ begrenzt ist. Dieser Wert kann dann auf den gewünschten Aktionsbereich skaliert werden [3].

2020 haben Chloe Ching-Yun Hsu et al. [4] verschiedene Probleme des PPO-Algorithmus untersucht und einige mögliche Verbesserungen vorgeschlagen. Unter anderem haben sie festgestellt, dass die ursprünglich verwendete Normalverteilung Aktionen an den Rändern vor allem bei geringen Standardabweichungen übermäßig häufig auswählt. Dadurch entstehen übermäßig hohe Gradientenaktualisierungen, die unerwünscht sind. Ein zweites Problem ist ein mögliches suboptimales lokales Belohnungsmaximum, das erreicht werden kann, wenn die erste im Training gewählte Wahrscheinlichkeitsverteilung nicht gleichmäßig genug ist. Eine gleichmäßigere Verteilung ist bei Normalverteilungen generell schwierig, da durch die große Standardabweichung dann große Teile der Verteilung außerhalb der möglichen Aktionen platziert werden. Aus diesen Gründen schlagen die Autoren vor, statt einer Normalverteilung eine Beta-Verteilung zu verwenden. Einerseits lässt sich diese zu Beginn gleichmäßig verteilen, wenn die beiden Parameter $\alpha = \beta = 1$ sind, andererseits verhindert diese übermäßig große Gradientenveränderung, die durch Aktionen am Rande der Verteilung verursacht werden. Außerdem verhindert die Beta-Verteilung bei nicht unendlichen Aktionsräumen Verzerrungen am Rand des Aktionsraumes, die bei abgeschnittenen Normalverteilungen vorkommen. Insgesamt konnte auch ein zuverlässigeres Konvergenzverhalten festgestellt werden [4].

Die 2D-Simulation einer Zylinderumströmung ist in der Literatur eines der am häufigsten vorkommenden Benchmark-Probleme [7, 13, 6]. Durch die regelmäßige Wirbelablösung ist dieses Problem relativ einfach und eine instationäre Lösung lässt sich innerhalb weniger Minuten berechnen. Dadurch kann das Training des DRL-Agenten in wenigen Stunden oder Tagen absolviert werden. Trotzdem ist das zügige, stabile und vor allem auch erfolgreiche Training noch keine Selbstverständlichkeit.

2019 haben Jean Rabault et al. [7] zum ersten Mal ein künstliches neuronales Netz, das mit DRL trainiert wurde, auf eine 2D-Zylinderumströmung angewandt. Der Agent hatte dabei die Möglichkeit, über zwei gegenüberliegende Düsen auf dem Zylinder auf die Umgebung einzuwirken. Diese Düsen konnten dabei die Strömung sowohl absaugen als auch ausblasen. Es wurde außerdem darauf geachtet, dass der Durchfluss der beiden Düsen addiert 0 ergibt, um ein realistisches Ergebnis zu erzeugen und dem System keine Masse hinzuzufügen oder zu entfernen. Der Agent wurde mittels PPO optimiert, wobei das neuronale Netz aus zwei vollständig verbundenen Ebenen mit jeweils 512 Knoten bestand. Die möglichen Aktionen wurden dabei über eine Beta-

Verteilung ausgewählt, für die die Maximal- und Minimalwerte im Voraus bestimmt wurden. Die Sensoren wurden um den Zylinder und im Nachlauf positioniert und haben die Geschwindigkeit gemessen. Die Reynoldszahl war dabei mit 100 vergleichsweise niedrig eingestellt. Insgesamt konnte eine Reduktion im Widerstandsbeiwert von etwa 8 % erreicht werden. In der Abbildung 1.2 ist der Effekt dieser Kontrolle zu sehen. Durch die Einwirkung des DRL-Agenten stabilisiert sich die Strömung im direkten Nachlauf und sorgt so für einen geringeren Widerstandsbeiwert und deutlich reduzierte Schwankungen in Widerstand und Auftrieb [7].

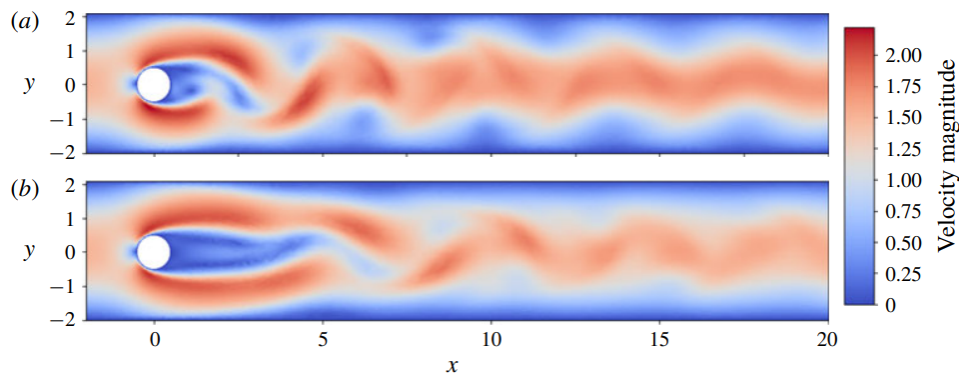


Abb. 1.2: Auswirkung der DRL-Kontrolle auf das Geschwindigkeitsfeld. (a) ist dabei der unkontrollierte Fall und (b) der aktiv kontrollierte. Diese Abbildung ist die Abbildung 4 des Artikels „Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control“ [7].

Aufbauend auf dieser Arbeit haben Jean Rabault und Alexander Kuhnle ebenfalls 2019 [8] eine weitere Arbeit veröffentlicht, in der sie die Simulation parallelisieren und so die benötigte Zeit für das Training deutlich reduzieren konnten. Dabei berechneten sie mehrere unabhängige Strömungssimulationen gleichzeitig und konnten die Ergebnisse aller Simulationen für den DRL-Agenten nutzen [8].

Ebenfalls aufbauend auf dieser Arbeit veröffentlichten Hongwei Tang et al. 2020 [13] eine Studie, in der einerseits die Anzahl an Düsen von zwei auf vier erhöht wurde und andererseits verschiedene Reynoldszahlen betrachtet wurden. Die Düsen waren dabei symmetrisch mit jeweils zwei Düsen oben und unten am Zylinder angebracht. Es wurden ebenfalls Geschwindigkeitssensoren verwendet, die um den Zylinder und im Nachlauf positioniert wurden. Der Agent wurde bei den Reynoldszahlen $Re = 100, 200, 300$ und 400 trainiert und erreichte dabei Widerstandsreduktionen in Höhe von 5,7 %, 21,6 %, 32,7 % und 38,7 %. Außerdem konnte der Agent auch bei Reynoldszahlen zwischen 60 und 400, auf die er vorher nicht trainiert wurde, die Strömung effizient kontrollieren und dabei vergleichbare Ergebnisse erzielen wie Agenten, die auf die jeweilige Reynoldszahl trainiert wurden [13].

Im Gegensatz zu den vorher genannten Arbeiten beeinflusst der DRL-Agent bei der 2020 veröffentlichten Studie von Mikhail Tokarev et al. [15] die Umgebung nicht durch das Einsaugen und Ausblasen der Strömung, sondern durch eine Rotation des Zylinders. Die Sensoren unterscheiden sich ebenfalls, denn es wurden Drucksensoren verwendet, die im Nachlauf nahe am Zylinder positioniert wurden. Ein weiterer wichtiger Unterschied ist, dass der Zylinder anders als in den zuvor genannten Studien nicht in einem Kanal, sondern freistehend positioniert wurde. Die Reynoldszahl liegt weiterhin bei $Re = 100$. Dabei konnte eine Reduktion von 14 % – 16 % im Widerstandsbeiwert erreicht werden. Ein weiteres Ergebnis dieser Studie ist, dass bei niedrigen Reynoldszahlen eine Verringerung der Amplitude und die Vermeidung von hochfrequenten Rotationsbewegungen bessere Ergebnisse liefert. Des Weiteren wurde hier ebenfalls die Simulation parallelisiert, in dem drei separate Strömungssimulationen gleichzeitig durchgeführt wurden [15]. Romain Paris et al. [6] haben 2021 eine Studie veröffentlicht, in der sie einerseits die Robustheit

des DRL-Agenten gegenüber unterschiedlichen Reynoldszahlen und andererseits die Auswirkungen in der Anzahl und Platzierung der Sensoren untersuchten. Die Sensoren wurden dabei als Drucksensoren ausgelegt und ebenfalls im Nachlauf des Zylinders platziert. Als Algorithmus wurde mit PPO-CMA eine Variante des PPO-Algorithmus verwendet, bei dem versucht wird, eine vorzeitige Verkleinerung der Erkundungsvarianz zu vermeiden. Der DRL-Agent wurde bei einer Reynoldszahl von $Re = 120$ trainiert und dann bei verschiedenen Reynoldszahlen im Bereich von $Re = 100 - 216$ evaluiert. Dabei wurde die Leistung mit DRL-Agenten verglichen, die auf die jeweilige Reynoldszahl trainiert waren. Es wurde festgestellt, dass der Agent sehr robust gegenüber der Veränderung der Reynoldszahl war und die Reduktion des Widerstandsbeiwertes nur wenig von den Werten der Vergleichsagenten abwich. Entscheidend dafür war laut den Autoren die Normierung der Referenzwerte, durch die die Konvektionszeit zwischen den Sensoren und dem Zylinder nahezu konstant blieb und auch die Varianzamplituden der Sensoren und Aktuatoren blieb zwischen den verschiedenen Reynoldszahlen vergleichbar. Nur die Größe und die Form der Wirbel unterschied sich, wodurch die Robustheit bei höheren Reynoldszahlen schlechter wird als bei niedrigeren. Ein Grund dafür könnte die Fixierung der Sensoren sein. Dadurch decken die Sensoren bei niedrigen Reynoldszahlen einen größeren Teil des Vortexes ab [6].

Die Anzahl und Platzierung der Sensoren wurde mit einer weiteren Abwandlung des PPO-CMA-Algorithmus untersucht. Beim S-PPO-CMA-Algorithmus werden nicht benötigte beziehungsweise redundante Sensorinformationen nicht berücksichtigt. Bis zu einer Reduzierung auf fünf Sensoren konnte keine signifikante Verschlechterung der Ergebnisse beobachtet werden. Dieser Effekt ist in Abbildung 1.3 zu sehen. Außerdem kann man dort erkennen, dass die Sensoren zentral hinter dem Zylinder wenig bis keinen Einfluss auf den Agenten zu haben scheinen, da der S-PPO-CMA-Algorithmus diese Sensoren nicht berücksichtigt hat. Eine mögliche Erklärung dafür ist, dass diese Sensoren keine Informationen über die Asymmetrie liefern können, da sie genau in der Symmetrielinie liegen [6].

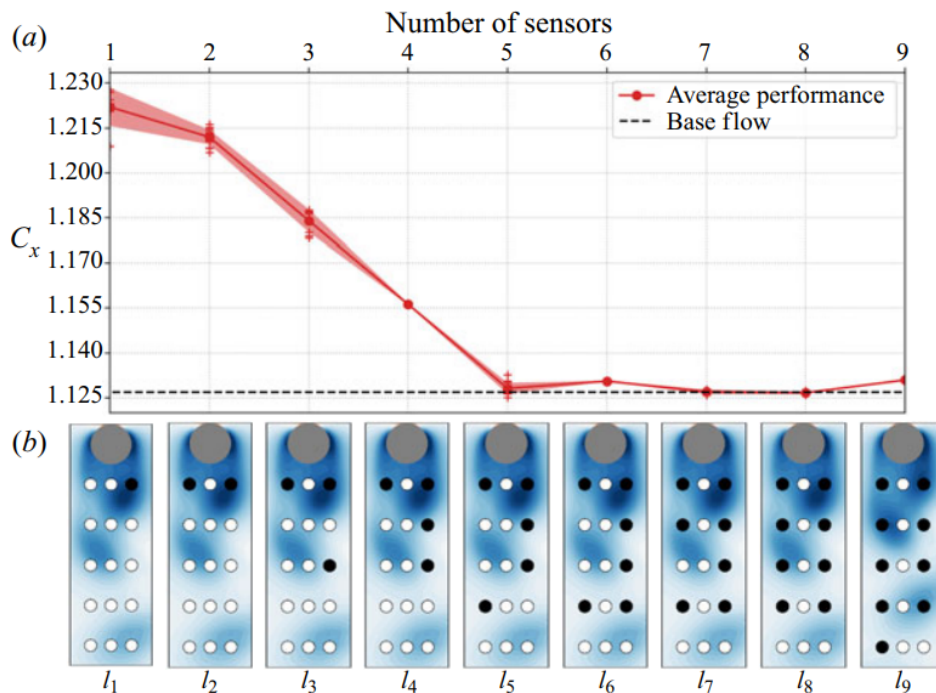


Abb. 1.3: Einfluss der Anzahl der Sensoren auf den DRL-Agenten bei optimaler Platzierung. Diese Abbildung ist die Abbildung 20 des Artikels „Robust flow control and optimal sensor placement using deep reinforcement learning“ [6].

1.3 Vorgehen

Diese Arbeit baut auf der Studienarbeit von Darshan Thummar „Active flow control in simulations of fluid flows based on deep reinforcement learning“ auf [14]. Daher wird der Simulationsaufbau größtenteils übernommen.

In dieser Arbeit sind die Sensoren ausschließlich auf der Oberfläche angebracht und als Drucksensoren ausgelegt. Diese Lage der Sensoren spiegelt einen etwas realitätsnäheren Simulationsaufbau wider, da in der Praxis bei zum Beispiel Flugzeugen oder Autos die Sensoren nur auf dem umströmten Objekt angebracht werden können. Als DRL-Algorithmus wird der PPO-Algorithmus mit zwei künstlichen Neuronalen Netzen, die zwei versteckten Schichten mit jeweils 64 Neuronen besitzen, verwendet. Der PPO-Agent kann dabei über die Rotation des Zylinders Einfluss auf die Strömung nehmen. Dies ist in der Simulation als Geschwindigkeitsrandbedingung, nachfolgend Regelrandbedingung genannt, umgesetzt. Als Programm für die Lösung des Strömungsproblems wird OpenFOAM verwendet [14]. Das Hauptziel dieser Arbeit ist es, den vorhandenen Trainingsablauf zu stabilisieren und zu optimieren. Dazu wird einerseits die Regelrandbedingung parallelisiert, damit die Zylinderoberfläche auf verschiedene Simulationsgebiete verteilt werden kann. Andererseits werden Maßnahmen zur Stabilisierung des Trainings getroffen. Dabei wird sichergestellt, dass die Regelgrenzen eingehalten werden und die Wahrscheinlichkeitsverteilung geändert, mit der die Regelung ausgewählt wird. Dieser überarbeitete Trainingsalgorithmus wird angewendet, um den DRL-Agenten bei verschiedenen Reynoldszahlen zu trainieren und anschließend bei einer Simulation mit zeitlich veränderlichen Reynoldszahlen anzuwenden.

Kapitel 2

Theorie

2.1 Zylinderumströmung

Die 2D-Simulation einer Zylinderumströmung wird in der Literatur als eines der häufigsten Benchmark-Probleme verwendet [7, 13, 6]. Wie in Abbildung 2.1 zu sehen ist, bildet sich bei vergleichsweise niedrigen Reynoldszahlen (Re) eine symmetrische Strömung aus, mit niedrigen Geschwindigkeiten im direkten Zylindernachlauf. Die Reynoldszahl ist die entdimensionalisierte Form der Geschwindigkeit, die sich aus der mittleren Einlassgeschwindigkeit u_∞ , der kinematischen Viskosität ν und dem Durchmesser des Zylinders d berechnet:

$$Re = \frac{u_\infty d}{\nu} \quad (2.1)$$

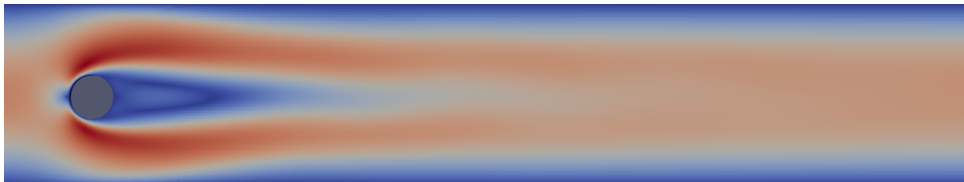


Abb. 2.1: Simulation einer 2D-Zylinderumströmung bei $Re = 50$. Dargestellt ist die Geschwindigkeit mit hoher Geschwindigkeit in rot und niedriger Geschwindigkeit in blau.

Bei größeren Reynoldszahlen löst sich die Grenzschicht am Zylinder ab und bildet einen Wirbel. Dabei entsteht an der Zylinderoberfläche eine Strömung in die entgegengesetzte Richtung. Dadurch wird auch die Scherschicht von der anderen Seite des Zylindernachlaufs angezogen, die den Zufluss von weiterer Wirbelstärke verhindert. Gleichzeitig entsteht dabei auch ein neuer Wirbel auf der anderen Seite des Zylindernachlaufs. Der zurückbleibende Wirbel wird auch von Kármán-Wirbel genannt. Der sich neubildende Wirbel zieht dann wie zuvor die Scherschicht von der anderen Seite des Zylindernachlaufs an, die wiederum ebenfalls den weiteren Zufluss von Wirbelstärke verhindert. Durch diese Effekte bilden sich sehr regelmäßige, gegenüberliegende Wirbel aus, die zusammen die Kármánsche Wirbelstraße bilden. Die Wirbel rotieren dabei abwechselnd in entgegengesetzte Richtungen [17].

Mit steigender Geschwindigkeit und damit auch Reynoldszahl nimmt die Frequenz, in der sich die Wirbel bilden, zu. Außerdem nimmt auch die Größe der Wirbel zu. In Abbildung 2.2 sind solche Kármánschen Wirbelstraßen dargestellt. Auf der linken Seite sind Wirbelstraßen mit moderaten Reynoldszahlen zu erkennen, während auf der rechten Seite Wirbelstraßen mit hohen Reynoldszahlen abgebildet sind. Bis zu einer Reynoldszahl von etwa 200 bleibt die Strömung dabei vollständig laminar. Bei höheren Reynoldszahlen geht die Strömung im weit entfernten

Zylindernachlauf in den turbulenten Zustand über. Je größer die Reynoldszahl wird, desto näher am Zylinder findet dieser Übergang statt. Bei sehr hohen Reynoldszahlen ist der Zylindernachlauf dann vollständig turbulent [17].

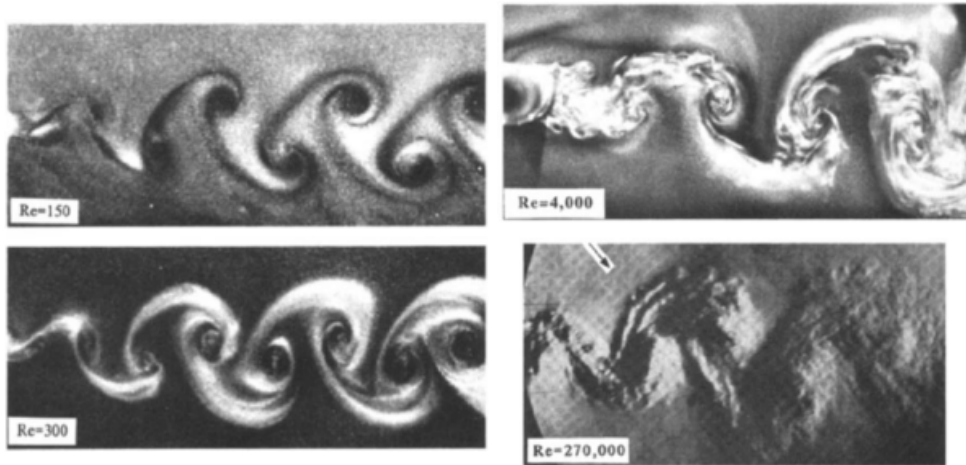


Abb. 2.2: Fotografien von Kármánschen Wirbelstraßen bei verschiedenen Reynoldszahlen. Diese Abbildung ist die Abbildung 1 des Artikels „Vortex dynamics in the cylinder wake“ [17].

2.2 Wahrscheinlichkeitsverteilungen

Um abschätzen zu können, ob eine andere Aktion möglicherweise eine erhöhte Belohnung bietet, muss der DRL-Agent im Training Aktionen zufällig auswählen. Es gibt dabei unterschiedliche Verteilungen nach denen eine Aktion ausgewählt werden kann. Die einfachste ist die Normalverteilung, die ursprünglich bei der Entwicklung des PPO-Algorithmus verwendet wurde [10] und auch erfolgreich auf die 2D-Zylinderumströmung angewendet wurde [6]. Diese weist allerdings einige Schwachstellen auf und daher wurde eine Beta-Verteilung als mögliche Verbesserung vorgeschlagen [4] und auch erfolgreich umgesetzt [7, 13].

2.2.1 Normalverteilung

Die Normalverteilung kann durch zwei wesentliche Parameter beschrieben werden: Einerseits durch den Erwartungswert μ , der sowohl Mittelwert und Modus als auch Median ist, und andererseits durch die Standardabweichung σ , die die Breite der Verteilung angibt. Die Dichtefunktion wird mit:

$$\varphi_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (2.2)$$

beschrieben [11]. Dabei liegt die Wahrscheinlichkeit bei 68,26 %, dass sich eine Zufallsgröße innerhalb eines Intervalls von $[\mu - \sigma, \mu + \sigma]$ befindet. Bei einem Intervall $[\mu - 2\sigma, \mu + 2\sigma]$ liegt diese Wahrscheinlichkeit bei 95,45 %.

Wie auch in Abbildung 2.3 zu sehen ist, gilt je größer die Standardabweichung ist, desto breiter ist auch die sogenannte Glockenkurve. Bei einer Erhöhung des Erwartungswertes verschiebt sich die Verteilung auf der x-Achse in die positive Richtung. Die Form bleibt dabei allerdings erhalten. Da jedes Ergebnis in dieser Verteilung möglich ist, ist die gesamte Fläche unter der Kurve gleich 1. Dadurch ergibt sich auch, dass die Glockenkurve bei gleichbleibendem Erwartungswert und steigender Standardabweichung nicht nur breiter, sondern auch niedriger wird. Da theoretisch, wenn auch mit sehr geringer Wahrscheinlichkeit, unendlich hohe oder niedrige Werte auftreten können, besitzt die Normalverteilung keinen kompakten Träger [11].

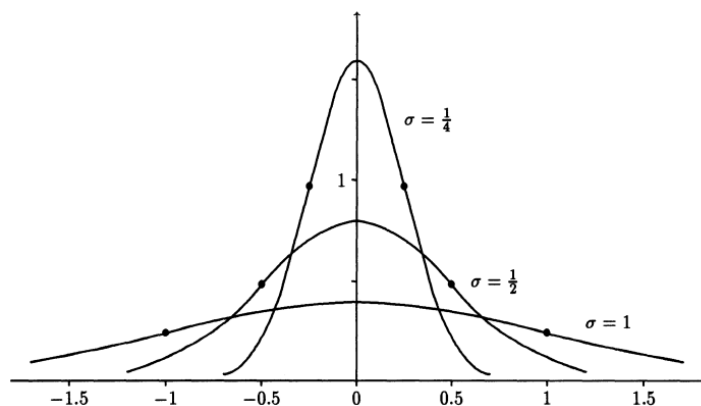


Abb. 2.3: Dichtefunktion der Normalverteilung bei verschiedenen Standardabweichungen. Die Abbildung stammt von der Seite 179 des Buches „Einführung in die mathematische Behandlung der Naturwissenschaften II“ [11].

Im Verbindung mit dem DRL kann das zu Problemen führen, da ein Agent, der mit einer Normalverteilung die Aktionen auswählt, unter Umständen einen Wert auswählt, der weit vom Erwartungswert entfernt liegt. Dadurch kann es zu übermäßig großen Gradientenaktualisierungen kommen [4].

Einhaltung der Bereichsgrenzen

In dieser Arbeit müssen die durch die Wahrscheinlichkeitsverteilungen gewählten Aktionen auf einen vorgegebenen Bereich begrenzt sein. Zum einen, weil eine zu hohe Rotationsgeschwindigkeit die Strömungssimulation zum Absturz bringen kann und zum anderen, weil auch in der Realität die Rotationsgeschwindigkeit durch die limitierte Kraft des Motors begrenzt wäre. Daher muss die Verteilung einen kompakten Träger besitzen, der bei der Normalverteilung durch eine Stauchungsfunktion erreicht werden kann [3]. Dabei wird die Zufallsvariable u mit der Standardabweichung s zunächst über eine unbegrenzte Normalverteilung mit der Dichtefunktion $\mu(\omega^{**}|s)$ ausgewählt und davon der Tangens hyperbolicus gebildet:

$$\omega^* = \tanh(\omega^{**}) \quad (2.3)$$

Die daraus entstehende Verteilung mit der Dichtefunktion $\pi(a|s)$ ist somit auf den Bereich $(-1,1)$ begrenzt. Die logarithmierte Dichtefunktion kann angegeben werden durch [3]:

$$\ln\pi(\omega^*|s) = \ln\mu(\omega^{**}|s) - \ln(1 - \tanh^2(\omega^{**})) \quad (2.4)$$

In Abbildung 2.4 wird dargestellt, dass die Stauchung der Verteilung dafür sorgt, dass sie nicht mehr symmetrisch zum Mittelwert ist, wenn dieser von 0 abweicht. Die Verteilung wird dann in Richtung 0 verzerrt. Die Stauchung gegen den Rand bewirkt außerdem, dass bei steigendem Erwartungswert, die Varianz der veränderten Normalverteilung sinkt, obwohl die Standardabweichung der unveränderten Normalverteilung gleich bleibt. In der Abbildung 2.4 ist das dadurch zu erkennen, dass die Dichtefunktion höher und gleichzeitig dünner wird. Des Weiteren ist in der Abbildung 2.4 zu sehen, dass die Stauchungsfunktion auch dann die Begrenzung sicherstellt, wenn der Erwartungswert der unveränderten Normalverteilung außerhalb des Bereiches liegt.

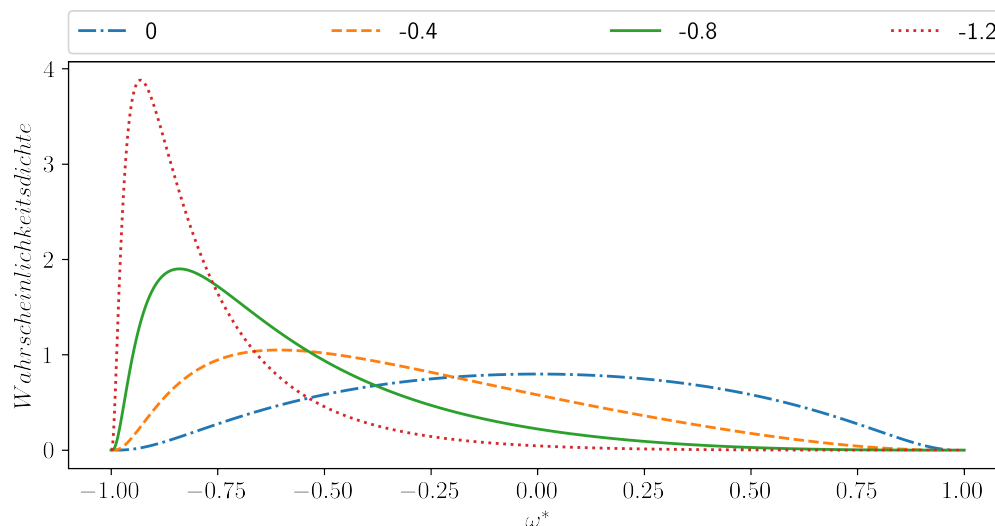


Abb. 2.4: Einfluss des Mittelwerts auf die veränderte Normalverteilung bei einer Standardabweichung von 0,5.

In Abbildung 2.5 ist der Einfluss der Standardabweichung auf die veränderte Normalverteilung zu sehen. Der Vergleich mit der Abbildung 2.3 zeigt, dass die Dichtefunktion der veränderten Normalverteilung schon bei geringeren Standardabweichungen breitere Verteilungen erzeugt. Zudem bildet sich bei einer Standardabweichung von 0,7 zwischen etwa -0,5 und 0,5 eine annähernde Gleichverteilung aus. Bei noch größeren Standardabweichungen verschiebt sich die Verteilung immer weiter zu den Rändern und sorgt damit dafür, dass die Werte an den Rändern der Verteilung wahrscheinlicher werden als der Erwartungswert. Dieses Verhalten ist beim DRL unerwünscht und daher wird die Standardabweichung der veränderten Normalverteilung im Folgenden auf maximal 0,7 begrenzt. Der Wert, der durch die Stauchung mit der Tangens hyperbolicus Funktion

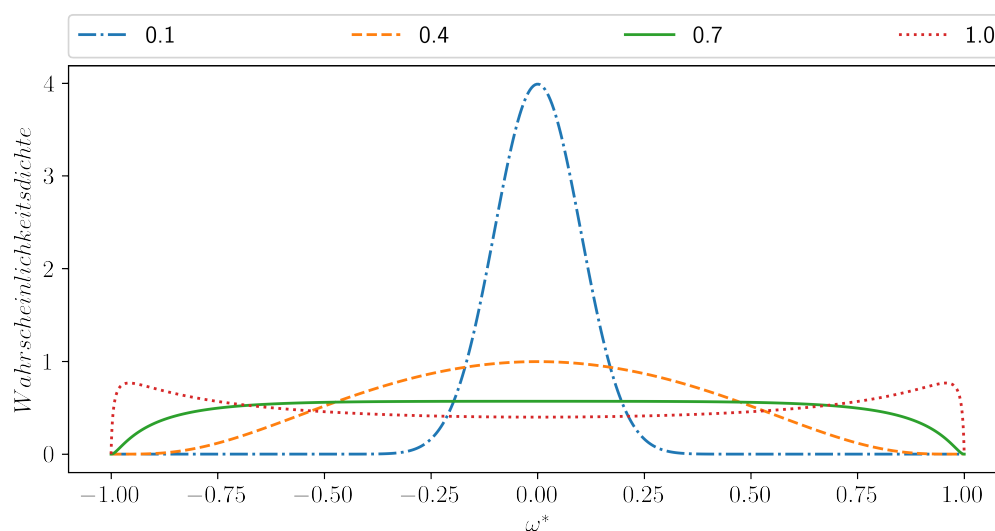


Abb. 2.5: Einfluss der Standardabweichung auf die veränderte Normalverteilung.

entsteht, muss dann auf den Aktionsbereich $[\omega_{min}, \omega_{max}]$ skaliert werden mit:

$$\omega = (\omega^* + 1) \cdot \frac{\omega_{max} - \omega_{min}}{2} + \omega_{min} \quad (2.5)$$

Wenn dabei der Mittelpunkt des Aktionsbereiches bei 0 liegt und somit gilt $-\omega_{min} = \omega_{max}$,

vereinfacht sich die Formel zu:

$$\omega = \omega^* \cdot \omega_{max} \quad (2.6)$$

2.2.2 Beta-Verteilung

Auch die Beta-Verteilung wird durch zwei Parameter (α und β) definiert. Im Gegensatz zur Normalverteilung besitzt die Beta-Verteilung allerdings einen kompakten Träger. Das bedeutet, außerhalb des möglichen Intervalls beträgt die Wahrscheinlichkeit gleich 0. Bei der Standard-Beta-Verteilung lautet die Dichtefunktion der Zufallsgröße $\bar{\omega}$ über das Intervall $(0,1)$ [5]:

$$Beta(\alpha, \beta) : \pi(\bar{\omega}|\alpha, \beta) = \frac{\bar{\omega}^{\alpha-1}(1-\bar{\omega})^{\beta-1}}{B(\alpha, \beta)} \quad (2.7)$$

Wobei $B(\alpha, \beta)$ die Beta-Funktion ist [5]:

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt \quad (2.8)$$

Diese dient der Normierung der Funktion. Die beiden Parameter müssen dabei größer 0 sein [5]. Wie in Abbildung 2.6 zu sehen ist, kann die Beta-Verteilung sehr unterschiedliche Formen annehmen, die nicht symmetrisch sein müssen. Wie auch bei der, durch die Tangens hyperbolicus Funktion gestauchten, Normalverteilung wird die Verteilung in den Mittelpunkt des Bereiches verzerrt. Dadurch gibt es große Ähnlichkeiten zwischen der Beta-Verteilung und der veränderten Normalverteilung. Bei einer Gleichheit der beiden Parameter ist die Dichtefunktion symmetrisch und hat dann große Ähnlichkeit zu einer unveränderten Normalverteilung. Wie ebenfalls in der Abbildung 2.6 zu sehen ist, verringert sich der Modus und der Erwartungswert, wenn β größer wird, und erhöht sich, wenn α größer wird.

Der Erwartungswert μ wird dabei berechnet mit [5]:

$$E(\bar{\omega}) = \mu = \frac{\alpha}{\alpha + \beta} \quad (2.9)$$

Der Modus, also die Spitze der Verteilung, liegt an einer der Kanten, wenn einer der Parameter kleiner als 1 ist. Wenn beide Parameter über 1 liegen wird der Modus berechnet durch [5]:

$$\gamma = \frac{\alpha - 1}{\alpha + \beta - 2} \quad (2.10)$$

und die Varianz durch [5]:

$$Var(\bar{\omega}) = \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (2.11)$$

Es ist also zu erkennen, dass die Varianz sinkt, je größer die Parameter α und β sind.

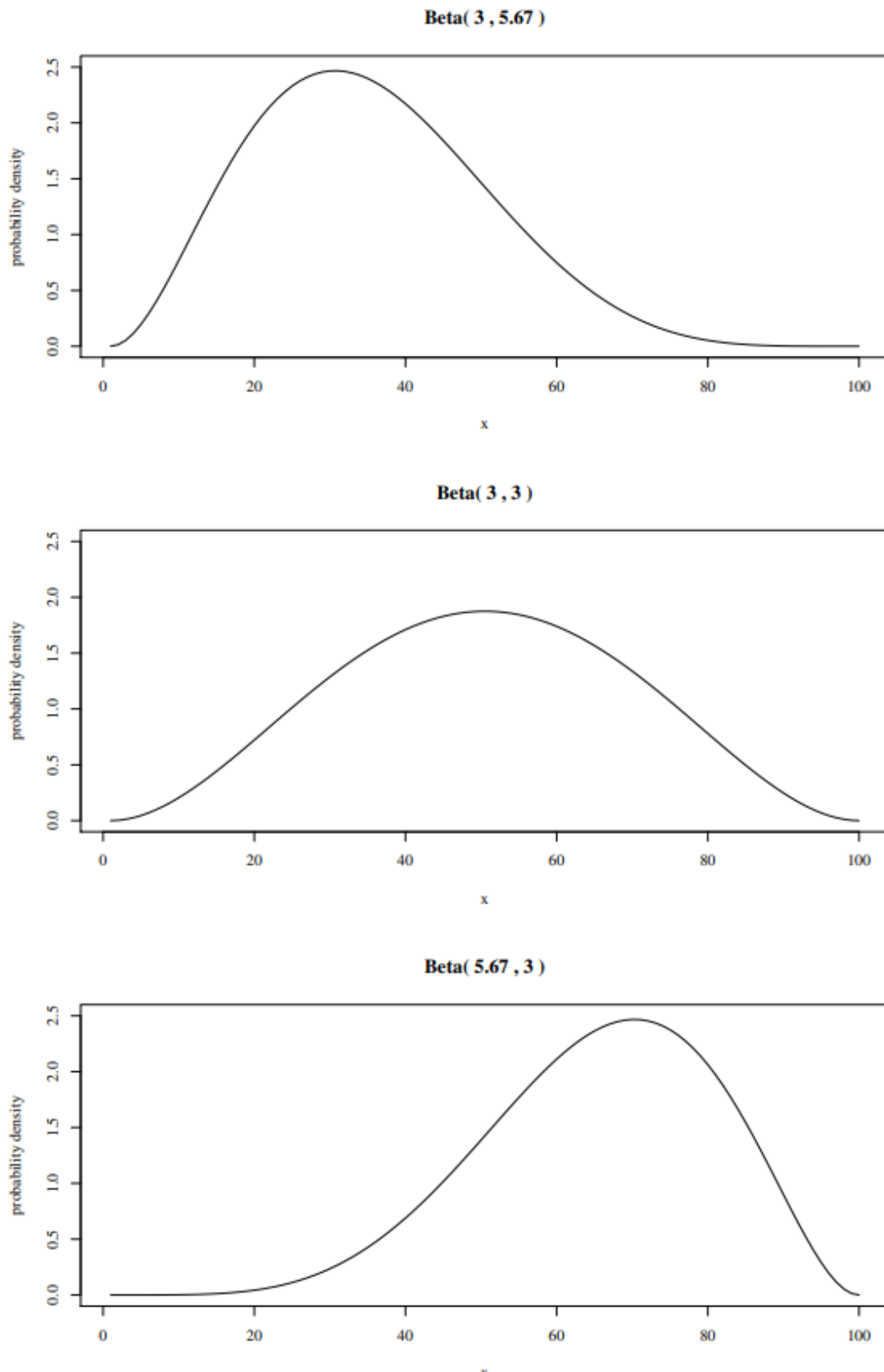


Abb. 2.6: Dichtefunktion der Beta-Verteilung mit unterschiedlichen Parametern. Diese Abbildung ist die Abbildung 4 des Artikels „Beta distribution“ [5].

Stetige Gleichverteilung

Abbildung 2.7 zeigt, dass bei der Beta-Verteilung auch eine stetige Gleichverteilung möglich ist, bei der alle Werte im Intervall gleich wahrscheinlich sind. Dies ist der Fall, wenn beide Parameter gleich 1 sind [5]. Bezogen auf das DRL kann eine solche stetige Gleichverteilung zu Beginn des Trainings große Vorteile bieten, da ein lokales, suboptimales Belohnungsmaximum erreicht werden kann, wenn dieses näher an dem Erwartungswert einer möglichen Anfangsnormalverteilung liegt [4].

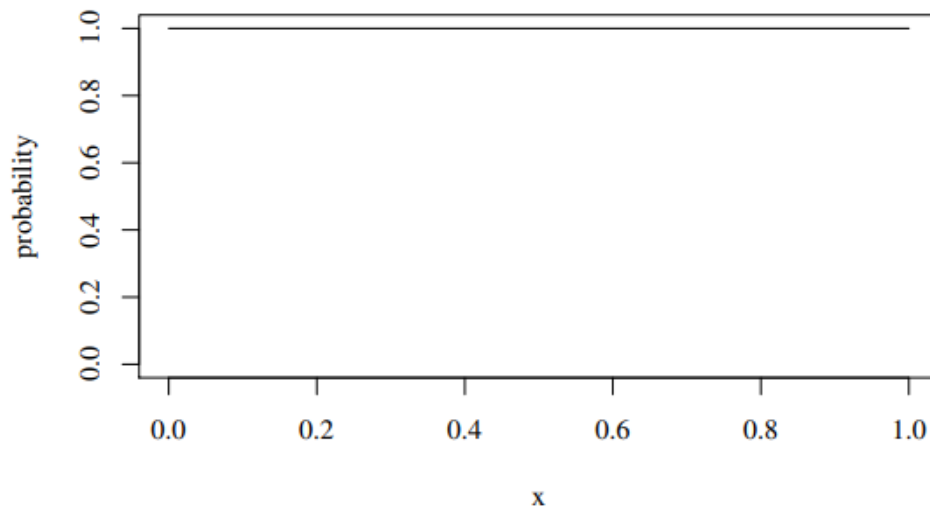


Abb. 2.7: Dichtefunktion der Beta-Verteilung als stetige Gleichverteilung. Diese Abbildung ist die Abbildung 3 des Artikels „Beta distribution“ [5].

Skalierung auf Aktionsbereich

Wie auch bei der Normalverteilung, auf die die Tangens hyperbolicus Funktion angewendet wurde, muss auch bei der Beta-Verteilung die ausgewählte Aktion (ω^*) auf den Aktionsbereich skaliert werden.

$$\omega = \omega^* \cdot (\omega_{max} - \omega_{min}) + \omega_{min} \quad (2.12)$$

Wenn der Mittelpunkt des Aktionsbereiches bei 0 liegt und somit gilt $-\omega_{min} = \omega_{max}$, vereinfacht sich die Formel zu:

$$\omega = (2\omega^* - 1) \cdot \omega_{max} \quad (2.13)$$

Kapitel 3

Methodik

3.1 Formulierung des Strömungsproblems

Um die Möglichkeiten einer aktiven Strömungskontrolle darstellen zu können und dabei gleichzeitig die Simulation und die Berechnungen vergleichsweise einfach zu halten, kann statt eines hochdimensionalen Strömungsproblems eine zweidimensionale, inkompressible Zylinderumströmung verwendet werden. Der Zylinder besitzt dabei einen kreisförmigen Querschnitt. Um eine gute Vergleichbarkeit zu anderen wissenschaftlichen Artikeln, wie zum Beispiel dem von Hongwei Tang [13], zu erreichen, wird im Folgenden der zweidimensionale Aufbau aus dem Artikel „Benchmark Computations of Laminar Flow Around a Cylinder“ von M.Schäfer und S. Turek [9] verwendet.

Geometrie

Der Durchmesser des Zylinders beträgt 0,1 m. Alle weiteren Längen werden über diesen Durchmesser des Zylinders entdimensionalisiert:

$$\tilde{x} = \frac{x}{d} \tag{3.1}$$

Die Länge des Simulationsbereiches beträgt dabei 22 und die Höhe beträgt 4,1. Wie in Abbildung 3.1 zu sehen ist, ist der Zylinder dabei nicht ganz mittig platziert, sondern mit einem Abstand 1,5 nach unten und 1,6 nach oben. Außerdem ist der Zylinder in einem Abstand von 1,5 von der linken Wand, dem Einlass, entfernt. Der Zylinder ist in der Mitte fixiert und drehbar gelagert. Dadurch verändern sich die Abstände im Laufe der Simulation nicht.

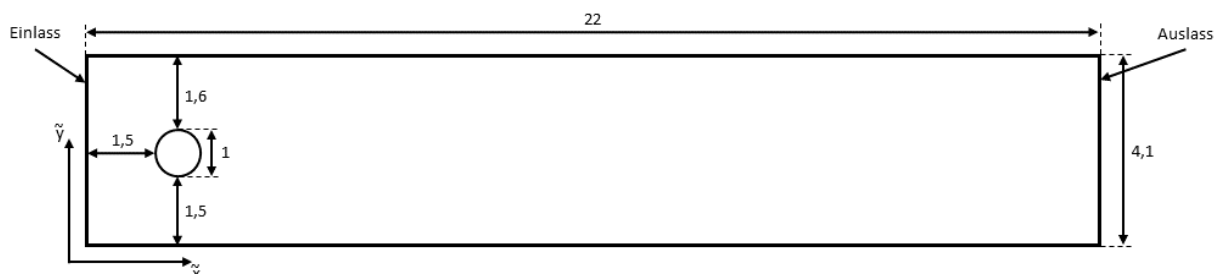


Abb. 3.1: Geometrie des Simulationsbereiches. Alle Längen sind über den Zylinderdurchmesser entdimensionalisiert.

Strömungseigenschaften

Bei den Simulationen wird ein inkompressibles, Newtonsches Fluid verwendet mit einer Dichte

von $\rho = 1,0 \text{ kg/m}^3$ und einer kinematischen Viskosität von $\nu = 10^{-3} \text{ m}^2/\text{s}$. Die mittlere Einlassgeschwindigkeit wird als u_∞ bezeichnet. Die Zeit t wird durch t^* :

$$t^* = \frac{t}{d} \cdot u_\infty \quad (3.2)$$

entdimensionalisiert. Die dimensionslosen Impuls- und Massenerhaltungsgleichungen werden desweiteren durch diese Parameter berechnet:

$$\mathbf{u}^* = \frac{\mathbf{u}}{u_\infty} \quad (3.3)$$

$$\nabla^* = d \nabla \quad (3.4)$$

$$p^* = \frac{p \cdot d}{\nu \cdot \rho \cdot u_\infty} \quad (3.5)$$

Wobei \mathbf{u}^* die entdimensionalisierte Form der Geschwindigkeit $\mathbf{u} = (u, v)^T$ ist, bei der u die Geschwindigkeit in x -Richtung und v die Geschwindigkeit in y -Richtung beschreibt. Außerdem sind p^* und ∇^* die dimensionslosen Parameter für p beziehungsweise ∇ . Die dimensionslosen Impuls- und Massenerhaltungsgleichungen lauten daher:

$$\nabla \cdot \mathbf{u}^* = 0 \quad (3.6)$$

$$\frac{\partial \mathbf{u}^*}{\partial t} + (\mathbf{u}^* \cdot \nabla^*) \mathbf{u}^* = -\nabla^* p^* + \frac{1}{Re} \nabla^{*2} \mathbf{u}^* \quad (3.7)$$

Die Ablösefrequenz wiederum wird als Strouhal-Zahl Sr entdimensionalisiert:

$$Sr = \frac{d}{u_\infty} f \quad (3.8)$$

Die Widerstands- und Auftriebskräfte werden als Integral über die Fläche des Zylinders S berechnet:

$$F_D = F_x = \int_S \left(\rho \nu \frac{\partial v_t}{\partial n} n_y - p n_x \right) dS \quad (3.9)$$

$$F_L = F_y = - \int_S \left(\rho \nu \frac{\partial v_t}{\partial n} n_x + p n_y \right) dS \quad (3.10)$$

n ist dabei eine Koordinate in einem Koordinatensystem, welches entlang der Oberfläche des Zylinders verläuft, und v_t ist die Tangentialgeschwindigkeit der Strömung. Entdimensionalisiert werden diese Kräfte zu den Widerstands- und Auftriebsbeiwerten c_D und c_L :

$$c_D = \frac{2F_D}{\rho u_\infty^2 d} \quad (3.11)$$

$$c_L = \frac{2F_L}{\rho u_\infty^2 d} \quad (3.12)$$

Randbedingungen

An der oberen und unteren Wand sowie am Zylinder wurde die Haftbedingung implementiert, sodass dort gilt:

$$u = v = 0 \quad (3.13)$$

$$\frac{dp^*}{d\mathbf{x}} = \mathbf{0}_{2,1} \quad (3.14)$$

Wobei \mathbf{x} der Positionsvektor $\mathbf{x} = (x, y)^T$ ist und $\mathbf{0}_{2,1}$ ein 2x1 Nullvektor. Am Einlass liegt eine parabolische Strömung vor, die über:

$$\begin{aligned} U(0, y) &= 4U_m y(H - y)/H^2 \\ V(0, y) &= 0 \\ \frac{dp^*}{d\mathbf{x}} &= \mathbf{0}_{2,1} \end{aligned} \quad (3.15)$$

definiert ist. U_m ist hierbei die Maximalgeschwindigkeit, die sich als $U_m = 1,5 \cdot u_\infty$ definiert, und H ist Höhe des Simulationsbereiches.

Der Auslass ist als Strömung mit Nullgradient implementiert:

$$\frac{d\mathbf{u}^*}{d\mathbf{x}} = \mathbf{0}_{2,1} \quad (3.16)$$

$$p^* = 0 \quad (3.17)$$

3.2 Gitterabhängigkeitsstudie

Um in der Strömungssimulation verlässliche Werte zu erhalten, sollte die Simulation möglichst unabhängig von der Beschaffenheit des Gitters sein. Das Gitter besteht aus einzelnen Zellen, bei denen der durchschnittliche Geschwindigkeits- oder Druckwert der Zelle als konstanter Wert für die gesamte Zelle gilt. Bei großen Zellen führt das unter Umständen zu größeren Abweichungen. Daher gilt: Je engmaschiger das Gitter, desto genauer ist die Simulation.

Allerdings steigt durch eine größere Zahl an Zellen auch die benötigte Rechenzeit an, weshalb kein beliebig engmaschiges Netz gewählt werden kann. Um nun ein optimales Rechenetz zu erstellen, wurde eine Gitterabhängigkeitsstudie durchgeführt.

3.2.1 Gittererstellung

Bevor die Gitterabhängigkeitsstudie durchgeführt werden kann, müssen die Gitter zuerst erstellt werden. Das geschieht über den Befehl `blockMesh` im Programm OpenFOAM. Dabei wird der Simulationsbereich in einzelne Unterbereiche aufgeteilt, die mit den gegebenen Parametern für die jeweiligen Unterbereiche mit Zellen gefüllt werden. Dabei entsteht ein Quadrat um den Zylinder, in dem die Zellen zum Zylinder hin immer kleiner werden (siehe Abbildung 3.2). Die Zellen direkt an der Zylinderwand haben dabei genau die halbe Dicke der Zellen an den Wänden. Diese Verteilung wurde gewählt, weil die Grenzschicht und die Kräfte, die untersucht werden sollen, direkt am Zylinder auftreten und daher die Auflösung dort am Besten sein sollte. Im Nachlauf des Zylinders nimmt die Zelldicke dann weiter zu bis sie das achtfache der Zelldicke am Zylinder erreicht.

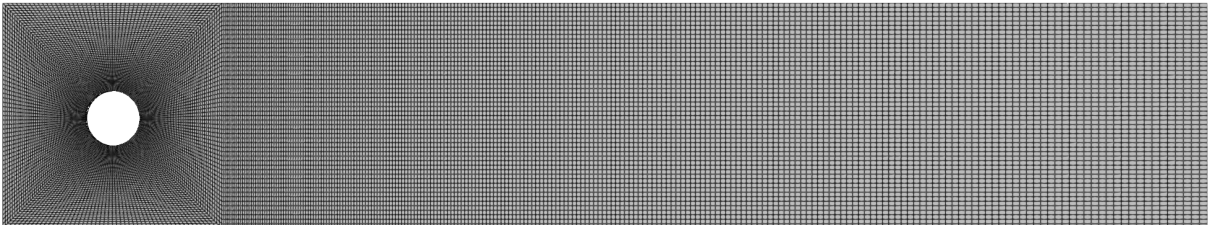


Abb. 3.2: Gitter mit 50 Zellen in y-Richtung.

In Abbildung 3.2 ist beispielhaft ein Gitter mit 50 Zellen in y-Richtung (N_y) und 275 Zellen in x-Richtung (N_x) zu sehen. N_x ist dabei abhängig von N_y über:

$$N_x = N_y + \frac{2,2 - 0,4}{0,4} \cdot N_y \quad (3.18)$$

Dadurch verdoppelt sich N_x bei einer Verdopplung von N_y und die Gitterverfeinerung hängt nur von einer variablen Größe ab. Diese Größe wird im Folgenden „Gittergröße“ N_y genannt.

3.2.2 Zeitschritt

Um die numerische Lösung des Strömungsproblems finden zu können, muss der Zeitschritt gesetzt werden. Der Zeitschritt gibt die Zeit an, die zwischen zwei diskreten Lösungen vergangen sein soll. Die Courant-Friedrichs-Lewy(CFL)-Bedingung gibt an, dass ein numerisches Verfahren nur dann stabil ist, wenn die CFL-Zahl geringer als 1 ist. Die CFL-Zahl berechnet sich durch:

$$CFL = \frac{u \cdot \Delta t}{\Delta x} \quad (3.19)$$

wobei Δx die Zellgröße in x-Richtung ist:

$$\Delta x = \frac{l_x}{N_x} \quad (3.20)$$

Für die Gittergröße 50 wurde bei einer Reynoldszahl von 100 initial ein Zeitschritt von $5 \cdot 10^{-4}$ gewählt. Dies entspricht einer maximalen CFL-Zahl von etwa 0,6. Bei einer Verdopplung der Gittergröße muss dementsprechend der Zeitschritt halbiert werden, um die CFL-Zahl annähernd konstant zu halten. Ebenso muss der Zeitschritt halbiert werden, wenn die Geschwindigkeit verdoppelt wird.

3.2.3 Ergebnisse der Gitterabhängigkeitsstudie

Für die Gitterabhängigkeitsstudie wurden drei Gitter untersucht, deren Gittergrößen jeweils um den Faktor 2 erhöht wurden:

Gittergröße (N_y)	Zellenanzahl	Anzahl an Punkten	Berechnungszeit (in s)
50	21250	43350	337,39
100	85000	171700	1.933,08
200	340000	683400	15.719,68

Tabelle 3.1: Zellen- und Punktanzahl sowie Berechnungszeit bei verschiedenen Gittergrößen.

Dabei wurde die Reynoldszahl $Re = 400$ gewählt, da die Genauigkeit der Simulation auch mit steigender Reynoldszahl abnimmt und daher möglichst die höchste Reynoldszahl, die im Folgenden untersucht werden soll, für die Gitterabhängigkeitsstudie verwendet werden sollte. Wie in Abbildung 3.3 zu sehen ist, ist die Änderung in den mittleren, maximalen und minimalen Widerstandsbeiwerten von der Gittergröße $N_y = 100$ zu der Gittergröße $N_y = 200$ relativ gering ($< 1\%$). Der Unterschied zwischen den Gittergrößen $N_y = 50$ und $N_y = 100$ ist hingegen sehr viel größer (ca. 3%). Da gleichzeitig die erforderliche Berechnungszeit für eine Sekunde in der Strömungssimulation zwischen den Gittergrößen $N_y = 100$ und $N_y = 200$ von 1.933,08 s auf 15.719,68 s ansteigt, wurde für alle folgenden Simulationen das Gitter der Größe $N_y = 100$ verwendet.

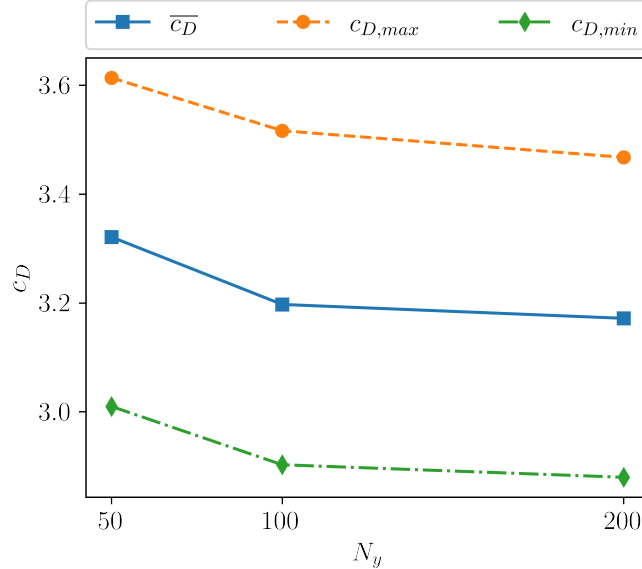


Abb. 3.3: Mittlere, maximale und minimale Widerstandsbeiwerte in Abhängigkeit der Gittergröße.

3.3 Bestärkendes Lernen

Um die Wirbelablösung zu kontrollieren und somit den Widerstandsbeiwert und die Schwankungen des Widerstands- und Auftriebsbeiwertes zu senken, wird ein Agent mit bestärkendem Lernen (RL, von Englisch: reinforcement learning) trainiert. Im folgenden Abschnitt wird zunächst die Interaktion des DRL-Agenten mit der Umgebung beschrieben und anschließend wird der Agent selbst und der Algorithmus, mit dem dieser trainiert wird, vorgestellt.

3.3.1 Umgebung

Die Strömungssimulation dient dem DRL-Agenten hier als Umgebung. Die Drehung des Zylinders als Aktion des Agenten wird durch das Verändern der Haftbedingung an der Oberfläche des Zylinders umgesetzt. Dabei wird die Oberflächengeschwindigkeit durch

$$u^* = \omega r \quad (3.21)$$

berechnet und auf der gesamten Oberfläche aufgetragen. Positive Drehgeschwindigkeiten ergeben dabei eine Rotation im Uhrzeigersinn und negative Drehgeschwindigkeiten eine Rotation gegen den Uhrzeigersinn.

Horizont

Eine einzelne Strömungssimulation wird als Trajektorie bezeichnet. Die Länge einer einzelnen Trajektorie wird als Horizont bezeichnet. Dabei kann zwischen einem begrenzten Horizont und einem unbegrenzten Horizont unterschieden werden. Bei einem unbegrenzten Horizont führt ein Agent so lange Aktionen aus bis ein bestimmtes Ziel erreicht wird. Wie viele Aktionen der Agent dafür braucht, unterscheidet sich von Trajektorie zu Trajektorie. Die Trajektorien sind daher bei einem unbegrenzten Horizont stets unterschiedlich lang und können beliebig lang werden. Im Gegensatz dazu ist die Trajektorie bei einem begrenzten Horizont auf eine bestimmte Länge festgelegt. Dabei versucht der Agent dem Ziel möglichst nahe zu kommen beziehungsweise die kumulative Belohnung zu maximieren. Da es bei der Senkung des Widerstandsbeiwertes und der Senkung der Schwankungen in den Beiwerten kein festes Ziel gibt, sondern nur die Belohnung maximiert werden kann, wurde in diesem Fall ein begrenzter Horizont gewählt.

Zustandserfassung

Dem DRL-Agenten wird die Umgebung in diesem Fall nicht direkt übergeben, sondern durch Drucksensoren auf der Oberfläche des Zylinders wird ein Teil der Umgebung erfasst. Dies nennt man den Zustand. Dabei wird jeweils ein Wert übergeben für jede Zelle des Gitters an der Oberfläche. Bei der gewählten Gittergröße 100 sind das insgesamt 400 Sensoren. Da diese Werte direkt aus dem Gitter entnommen werden, drehen die Sensoren sich nicht mit dem Zylinder mit, was es dem Agenten einfacher macht, den Zusammenhang zwischen Zustand und Aktion zu approximieren.

Belohnungsfunktion

Neben dem Zustand wird dem Agenten auch eine Belohnung zurückgegeben. Da das wesentliche Ziel des Agenten ist, den Widerstandsbeiwert zu reduzieren und die Schwankungen sowohl des Widerstandsbeiwertes als auch des Auftriebsbeiwertes zu reduzieren, sind diese beiden Werte die wesentlichen Bestandteile der Belohnungsfunktion:

$$R_t = 3 - (c_D + 0.1|c_L|) \quad (3.22)$$

Der zweite Wert ist dabei für die Gewichtung zwischen den beiden Beiwerten verantwortlich. Neben der Belohnung wird auch die kumulative Belohnung, der sogenannte Ertrag, übergeben. Bei diesem wird nicht nur die aktuelle Belohnung berücksichtigt, sondern auch alle nachfolgenden Belohnungen. Dadurch können auch die langfristigen Folgen einer Aktion berücksichtigt werden. Der Ertrag (G_τ) errechnet sich durch:

$$G_\tau = \sum_{t=0}^T \gamma^t R_t \quad (3.23)$$

Dabei ist γ der Reduktionsfaktor, der bestimmt, wie stark der Einfluss der nachfolgenden Belohnungen auf den Ertrag ist. In diesem Fall wird er auf 0,99 gesetzt.

3.3.2 Agent

Der Agent ist das Programm, das lernt, die Strömung zu kontrollieren. Über den Proximal policy optimization (PPO)-Algorithmus werden zwei separate Netzwerke trainiert, um einerseits den Zusammenhang zwischen Aktion und Ertrag und andererseits den Zusammenhang zwischen Zustand und Ertrag zu approximieren.

Algorithmus

In Abbildung 3.4 ist ein Flussdiagramm des hier verwendeten PPO-Algorithmus zu sehen. Dabei werden zuerst die für die Aktualisierung der Netzwerke nötigen Erfahrungen gesammelt, indem die policy π_k auf die Umgebung angewandt wird. Dies geschieht in zehn Umgebungen gleichzeitig, um das Sammeln der Trajektorien zu parallelisieren. Der Agent kann dabei die Erfahrungen von allen zehn Trajektorien nutzen. Während die Simulation läuft, wird jeweils die Wahrscheinlichkeit für die jeweilige Aktion berechnet und gespeichert. Nach Abschluss einer Trajektorie wird die jeweilige Belohnung für jede Aktion und der Ertrag berechnet.

Bei der Berechnung der Wahrscheinlichkeit einer Aktion durch das policy-Netzwerk kommen die Wahrscheinlichkeitsverteilungen aus Kapitel 2.2 zum Einsatz. Bei der, durch die Tangens hyperbolicus Funktion veränderten, Normalverteilung können die Wahrscheinlichkeiten über die Formel 2.4 berechnet werden. Bei der Beta-Verteilung wiederum können die Wahrscheinlichkeiten über die Formel 2.7 berechnet werden. Dazu muss die jeweilige Aktion allerdings erst wieder zurück auf den ursprünglichen Bereich skaliert werden. Dies geschieht über die Umkehrfunktion der Formel 2.5 beziehungsweise 2.12.

Weitere Ausführungen zum PPO-Algorithmus und dessen Umsetzung sind in der Arbeit von Darshan Thummar [14] zu finden, auf der diese Arbeit aufbaut.

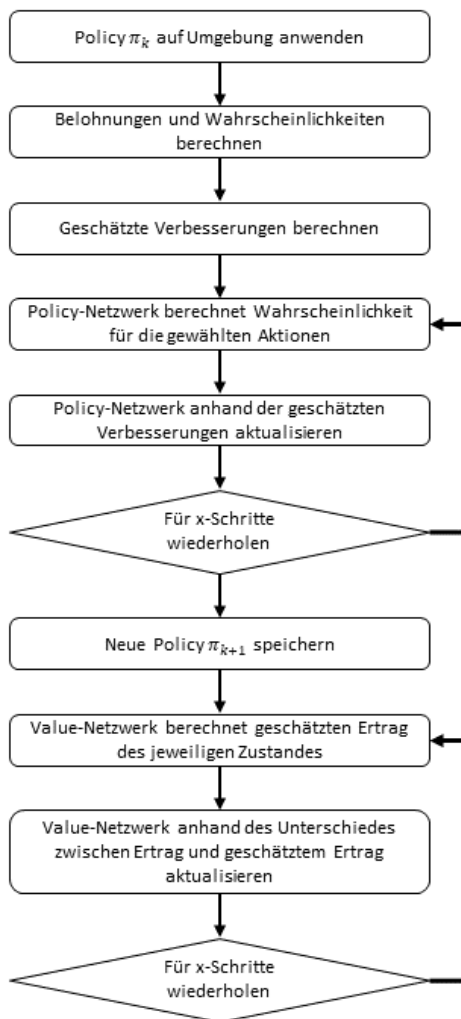


Abb. 3.4: Flussdiagramm des PPO-Algorithmus

Policy-Netzwerk

Das policy-Netzwerk besteht, wie in Kapitel 3.3.1 schon beschrieben, aus einer Eingangsschicht mit 400 Knoten. Zudem besteht es noch aus zwei verborgenen Schichten mit jeweils 64 Knoten und einer Ausgangsschicht mit zwei Knoten. Die Ausgangsschicht ist dabei unabhängig von der gewählten Wahrscheinlichkeitsverteilung. Allerdings unterscheidet sich die Interpretation der beiden Werte. Bei der Normalverteilung und der von der Tangens hyperbolicus Funktion veränderten Normalverteilung sind die beiden Werte der Erwartungswert und die Standardabweichung. Bei der Beta-Verteilung handelt es sich um die beiden Parameter α und β .

Value-Netzwerk

Das value-Netzwerk ist unabhängig von der gewählten Wahrscheinlichkeitsverteilung und unterscheidet sich im Aufbau von dem policy-Netzwerk nur in der Ausgangsschicht. Diese besteht beim value-Netzwerk nur aus einem Knoten, der dem berechneten Ertrag entspricht. Wie auch in Abbildung 3.4 zu sehen ist, werden die beiden Netzwerke allerdings unabhängig voneinander trainiert.

Zeitschritt

Der Zeitschritt des Agenten unterscheidet sich vom Zeitschritt der Strömungssimulation, da nicht bei jedem Zeitschritt eine neue Aktion ausgewählt werden soll. Dies würde das Training erheblich erschweren. Der Agent verändert die Drehgeschwindigkeit des Zylinders alle 20 Zeitschritte der

Strömungssimulation:

$$\Delta_{t,Agent} = 20 \cdot \Delta_{t,cfd} = 20 \cdot 2.5 \cdot 10^{-4} s = 5 * 10^{-3} s \quad (3.24)$$

Zu beachten ist hierbei, dass die Drehgeschwindigkeit kontinuierlich angepasst wird, um den Simulationsaufbau realistischer zu machen und sprunghafte Anstiege in den Beiwerten zu vermeiden.

3.3.3 Verbesserungen in der Umsetzung

Im Vergleich zu der Arbeit von Darshan Thummar [14] wurden einige kleinere Verbesserungen im Code durchgeführt, um einerseits das Training zu beschleunigen und andererseits die Ergebnisse miteinander vergleichbarer zu machen.

Verkürzung der Simulationszeit

Bei einer Reynoldszahl von 100 dauert es etwa 4,5 s bis sich die Kármánsche Wirbelstraße vollständig ausgebildet hat. Daher beginnt der Agent erst ab 4,5 s mit der Kontrolle der Strömung. Um alle möglichen Startzustände abdecken zu können, startet der Agent allerdings nicht bei genau 4,5 s, sondern zufällig in einem Bereich von 4,5 s bis 4,9 s. Die Strömungssimulation ist zuvor unabhängig von dem Agenten und muss daher nicht jedes Mal neu berechnet werden. Die Kontrollzeit wurde zudem von fünf auf zwei Sekunden reduziert. Insgesamt wurde so die Gesamt-simulationszeit von bis zu 8,5 s auf zwei Sekunden reduziert. Dies entspricht einer Reduktion von 76,5 %. Dadurch beträgt die Gesamtsimulationszeit nur noch etwa eine Stunde.

Erhöhung der Anzahl der parallelen Simulationen

Da die Simulation auf einem Hochleistungscluster berechnet wird, kann es passieren, dass bei einzelnen Strömungssimulationen die Ressourcen des belegten Rechners von einer anderen Aufgabe belegt werden und einzelne Simulationen so um ein vielfaches länger dauern als die restlichen. Um dies zu umgehen werden statt der eigentlich verwendeten zehn Strömungssimulationen zwölf gleichzeitig ausgeführt, von denen die ersten zehn für den Agenten verwendet werden. Da eine Strömungssimulation nach einer gewissen Zeit abgebrochen wird ohne die gesamte Zeit simuliert zu haben, wird so außerdem sichergestellt, dass der Agent für jede Iteration immer zehn verschiedene Strömungssimulationen verwenden kann. Dadurch werden die Ergebnisse verschiedener Trainings vergleichbarer, da sichergestellt wird, dass der Agent zu jedem Zeitpunkt gleich viele Erfahrungswerte sammeln kann.

3.4 Parallelisierung der Regelrandbedingung

Neben der Parallelisierung durch das gleichzeitige Ausführen mehrerer unabhängiger Strömungssimulationen kann auch jede einzelne dieser Strömungssimulationen parallelisiert werden, indem sie von mehreren Prozessorkernen gleichzeitig ausgeführt wird. Dafür wird das Rechengitter in einzelne Bereiche aufgeteilt. Jeder Prozessor berechnet einen dieser Bereiche und übergibt in den Rändern dieser Bereiche die Druck- und Geschwindigkeitswerte an die Prozessoren der anliegenden Bereiche weiter. Je mehr Prozessoren verwendet werden, desto mehr müssen die Prozessoren untereinander kommunizieren. Dies führt dazu, dass die insgesamt benötigte Simulationszeit nicht linear mit der Anzahl der Prozessoren abnimmt. Daher wird jede einzelne Simulation in den folgenden Abschnitten mit vier Prozessoren gleichzeitig durchgeführt.

Umsetzung in OpenFOAM

Wie im vorherigen Abschnitt schon erwähnt, kann jeder einzelne Prozessor nur auf die Druck- und Geschwindigkeitswerte zugreifen, die in dem Bereich liegen, der diesem Prozessor zugewiesen ist. Der DRL-Agent benötigt zur Berechnung der nächsten Aktion allerdings die Druckwerte auf der

gesamten Oberfläche des Zylinders. Wenn die Oberfläche des Zylinders auf mehrere Prozessoren verteilt ist, kann kein einzelner Prozessor auf alle diese Werte zugreifen. In OpenFOAM gibt es für das explizite Teilen von bestimmten Daten den `gather`-Befehl. Jeder Prozessor, der einen Teil der Oberfläche des Zylinders in seinem Bereich hat, sammelt die Druckwerte dieses Teils in einer Liste. Der Hauptprozessor sammelt die einzelnen Listen über den `gatherList`-Befehl ein und berechnet anhand dieser Werte die nächste Aktion. Anschließend muss diese neue Aktion wieder auf alle Prozessoren verteilt werden, da jeder Prozessor die Oberflächengeschwindigkeit auf dem Zylinder in seinem Bereich umsetzen muss. Dies geschieht in OpenFOAM über den `scatter`-Befehl.

Varianz des Agenten gegenüber der Prozessoranzahl

Da die Reihenfolge der Druckwerte in der Liste, die der Hauptprozessor von den anderen Prozessoren erhält, maßgeblich von der Anzahl der Prozessoren abhängt, ist der DRL-Agent nicht invariant gegenüber der Anzahl der Prozessoren. Wenn ein Agent mit Simulationen trainiert wurde, die mit vier Prozessoren durchgeführt wurden, kann er die Strömung auch nur in Simulationen kontrollieren, die mit vier Prozessoren durchgeführt werden. Wichtig ist dabei auch die Verteilung der Zellen auf die Prozessoren. Wenn die Test-Simulation mit mehr oder weniger Prozessoren berechnet wird oder die Zellen in einer anderen Weise auf die Prozessoren verteilt werden, ändert sich unter Umständen die Reihenfolge der Druckwerte und somit werden die falschen Druckwerte in die Eingangsknoten des policy-Netzwerks übergeben. Um diesen Effekt zu umgehen, müsste die Liste auf dem Hauptprozessor zuerst sortiert werden.

Kapitel 4

Abhängigkeit der Strömungsbeiwerte von der Reynoldszahl

Nachdem in Kapitel 3.2.3 das Gitter mit der Gittergröße 100 ausgewählt wurde, wird nun der Zusammenhang zwischen der Reynoldszahl und der Strouhal-Zahl untersucht. Des Weiteren werden die Widerstands- und Auftriebsbeiwerte in Abhängigkeit der Reynoldszahl betrachtet. Dabei wurden acht Simulationen beginnend bei $Re = 50$ in 50er-Schritten bis $Re = 400$ durchgeführt.

4.1 Strouhal-Zahl

Die hier untersuchte Strouhal-Zahl ist die entdimensionalisierte Frequenz (siehe Gleichung 3.8) der Kármánschen Wirbelstraße. Es wäre zu erwarten, dass die Strouhal-Zahl mit steigender Reynoldszahl zunimmt, da sich die Frequenz bei steigender Geschwindigkeit erhöht. Wie in Abbildung 4.1 zu sehen ist, trifft dies auch bei der Simulation zu. Bei einer Reynoldszahl von $Re = 50$ liegt die Strouhal-Zahl bei $Sr = 0,2671$. Zuerst steigt diese stärker an, bis sie ab $Re = 150$ etwas abflacht und ab dann flacher ansteigt.

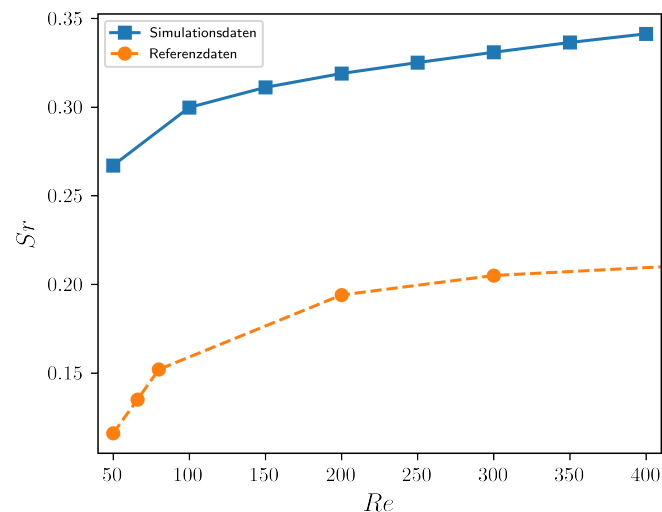


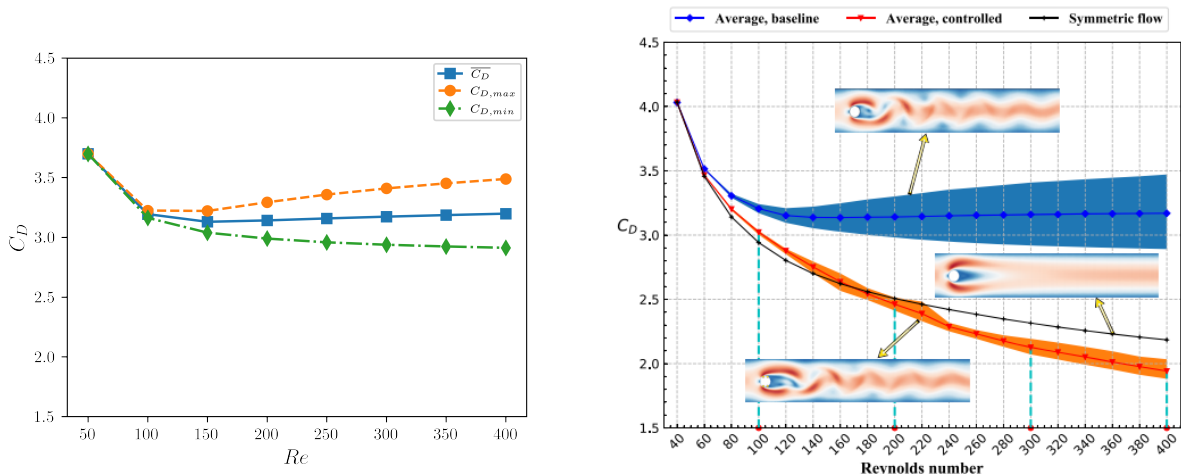
Abb. 4.1: Strouhal-Zahl in Abhängigkeit der Reynoldszahl. Die Referenzdaten stammen aus der Tabelle 1 aus dem Artikel „numerical calculation of laminar vortex shedding flow past cylinders“ [2].

Ein ähnlicher Trend lässt sich auch bei den Referenzdaten beobachten, die ebenfalls ab $Re = 50$ zunächst stärker ansteigen und dann abflachen. Da sich der Simulationsaufbau der Referenzdaten etwas von dem hier verwendeten Simulationsaufbau unterscheidet, differieren auch die Strouhal-Zahlen absolut deutlich voneinander. Der Verlauf ist allerdings sehr ähnlich, was die Simulationsergebnisse bestätigt.

4.2 Widerstandsbeiwerte

Der Widerstandsbeiwert ist die entdimensionalisierte Kraft, die in x-Richtung wirkt (Siehe Gleichung 3.11). Diese oszilliert um einen Mittelwert in Abhängigkeit von der Wirbelablösung. Je höher die Reynoldszahl ist, desto größer sind die Wirbel und damit auch die Amplitude der Schwingung. Daher ist zu erwarten, dass die maximalen und minimalen Widerstandsbeiwerte sich weiter vom Mittelwert entfernen, je größer die Reynoldszahl wird.

Wie in Abbildung 4.2 zu erkennen ist, ist das auch in der Simulation der Fall. Bei einer Reynoldszahl $Re = 50$ bildet sich die Kármánsche Wirbelstraße noch nicht aus. Daher weichen die maximalen und minimalen Widerstandsbeiwerte nicht signifikant vom Mittelwert ab. Der mittlere Widerstandsbeiwert liegt dabei bei 3,6985 und nimmt bis zur Reynoldszahl $Re = 150$ ab, wo er seinen Tiefpunkt bei 3,1298 erreicht. Ab dann steigt der Mittelwert zwar wieder leicht an, ist aber dennoch nahezu konstant. Die maximalen und minimalen Widerstandsbeiwerte entfernen sich dabei immer weiter vom Mittelwert und liegen bei $Re = 400$ etwa 0,29 über beziehungsweise unter dem Mittelwert. Das entspricht einem Unterschied von etwas mehr als 8 %. Der Vergleich mit den Literaturdaten in Abbildung 4.2b) zeigt einen sehr ähnlichen Verlauf. Da der Simulationsaufbau bei der Arbeit von Hongwei Tang et al. [13], aus der die Referenzdaten stammen, der gleiche ist wie in dieser Arbeit, stimmt der Verlauf sowohl quantitativ als auch qualitativ sehr gut mit den Referenzdaten überein.



a) Simulationsdaten

b) Literaturdaten aus der Abbildung 12 des Artikels „Robust active flow control over a range of Reynolds numbers using an artificial neural network trained through deep reinforcement learning“[13].

Abb. 4.2: Widerstandsbeiwert C_D in Abhängigkeit der Reynoldszahl.

4.3 Auftriebsbeiwerte

Die Kraft, die in y-Richtung wirkt, wird hier als Auftriebsbeiwert entdimensionalisiert. Wegen dem nahezu symmetrischen Aufbau der Simulation ist zu erwarten, dass der Mittelwert des Auftriebsbeiwertes in etwa 0 entspricht. Durch die Wirbelablösung oszilliert der Auftriebsbeiwert allerdings um diesen Mittelwert. Wie bei dem Widerstandsbeiwert, ist auch hier zu erwarten, dass sich die maximalen und minimalen Auftriebsbeiwerte weiter vom Mittelwert entfernen je größer die Reynoldszahl wird, da die Größe der Wirbel bei steigender Reynoldszahl ebenfalls zunimmt. Die Abbildung 4.3 zeigt, dass die Simulation diese Erwartungen erfüllt. Der mittlere Auftriebsbeiwert weicht nicht signifikant von 0 ab, während sich die maximalen und minimalen Auftriebsbeiwerte vom Mittelwert entfernen. Bei einer Reynoldszahl von 50 liegen die maximalen und minimalen Auftriebsbeiwerte, wie bei dem Widerstandsbeiwert, sehr nahe am mittleren Auftriebsbeiwert, da sich die Kármánsche Wirbelstraße noch nicht ausbildet. Bei steigender Reynoldszahl steigen die maximalen und minimalen Auftriebsbeiwerte zunächst stark an, so dass sie bei einer Reynoldszahl von 100 bei etwa 1 beziehungsweise -1 liegen. Dann flacht die Steigung etwas ab und scheint sich einem Grenzwert bei etwa 3 beziehungsweise -3 anzunähern.

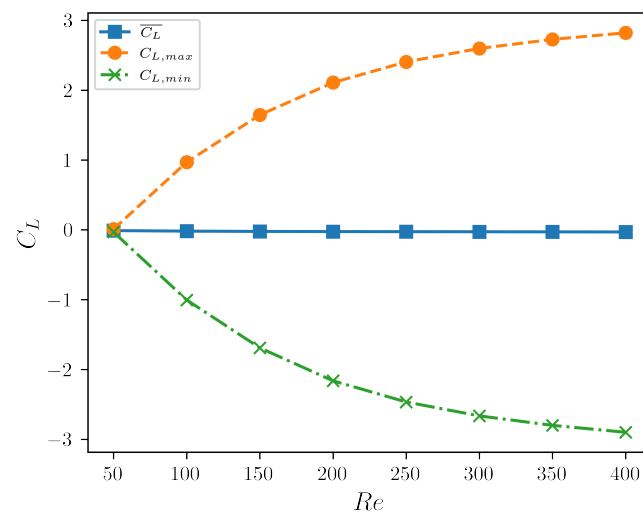


Abb. 4.3: Auftriebsbeiwert C_L in Abhängigkeit der Reynoldszahl.

Kapitel 5

Stabilisierungsmaßnahmen

Im folgenden Kapitel werden die Maßnahmen zur Stabilisierung des Trainings durch die Verwendung verschiedener Wahrscheinlichkeitsverteilungen dargestellt. Dazu wird zunächst der DRL-Agent mit einer unveränderten Normalverteilung betrachtet. Anschließend folgt der Vergleich mit einem Agenten, der eine Tangens hyperbolicus Funktion als Stauchungsfunktion verwendet, und einem Agenten, der eine Beta-Verteilung nutzt.

5.1 Normalverteilung

Das policy-Netzwerk des DRL-Agenten gibt hier einerseits den Erwartungswert und andererseits die Standardabweichung als Parameter aus. Während des Trainings wird zunächst die Aktion anhand einer Normalverteilung ausgewählt, nach dessen Ende bestimmt ausschließlich der Erwartungswert die Wahl der Aktion.

Belohnung

In Abbildung 5.1 ist die durchschnittliche Belohnung aller Trajektorien im Verlauf des Trainings zu erkennen. Da der DRL-Agent im Verlauf des Trainings lernt, die Wirbelbildung immer stärker zu unterdrücken und damit den Widerstandsbeiwert zu senken, steigt die durchschnittliche Belohnung deutlich an. Da die mögliche Reduktion begrenzt ist, flacht die Belohnungskurve zum Ende des Trainings ab, wenn der Agent sein Maximum erreicht hat. Ab Episode 80 ist kein signifikanter Anstieg der Belohnung mehr zu messen. Das Training ist daher nach 80 Episoden abgeschlossen.

Zu Beginn einer Trajektorie sind sowohl der Widerstandsbeiwert als auch der Auftriebsbeiwert auf dem Niveau des unkontrollierten Zustandes und die Belohnung ist somit zu Beginn der Trajektorie niedriger als am Ende. Dadurch ist die in Abbildung 5.1 dargestellte durchschnittliche Belohnung etwas geringer als die Belohnung gegen Ende einer Trajektorie.

Die Standardabweichung der Belohnungen steigt im Verlauf des Trainings ebenfalls an. Zum einen dadurch, dass sich zu Beginn einer Trajektorie der Widerstands- und Auftriebsbeiwert auf dem Niveau des unkontrollierten Zustands befinden und damit die Belohnung zu Beginn der Trajektorie sich von der Belohnung zum Ende der Trajektorie unterscheidet. Dieser Unterschied wird im Laufe des Trainings größer, da die Auftriebs- und Widerstandsbeiwerte durch die Aktionen des Agenten immer geringer werden. Zum anderen können gegen Ende des Trainings schon geringe Veränderungen der Aktionen zu Beginn einer Trajektorie große Auswirkungen auf die Belohnung haben, weshalb eine größere Streuung der einzelnen Trajektorien entsteht.

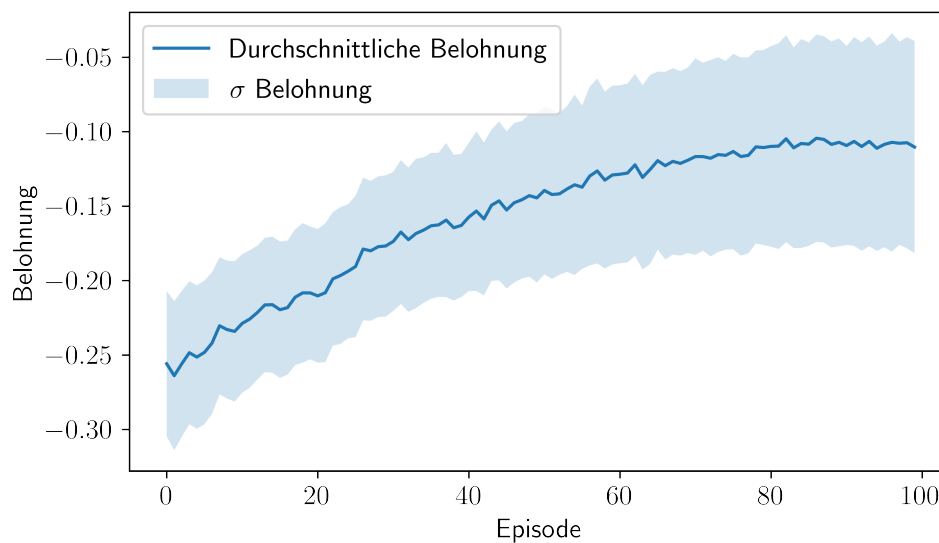


Abb. 5.1: Durchschnittliche Belohnung im Verlauf des Trainings.

Entwicklung des Agenten

Wie in Abbildung 5.2 zu sehen ist, hat der Agent zu Beginn noch einen sehr geringen Erwartungswert, der sich fast unabhängig von den Drucksensoren auf der Oberfläche verhält und daher nahezu konstant ist. Bis zur Episode 10 bildet sich ein sinus-förmiger Verlauf des Erwartungswertes aus, der etwa die gleiche Frequenz wie die Wirbelablösung hat. Im Verlauf der weiteren Episoden steigt die Amplitude dann immer weiter an.

Hierbei ist zu beachten, dass die in Abbildung 5.2 dargestellten Werte nur die Erwartungswerte darstellen und nicht die tatsächlich gewählte Aktion. Da nur die tatsächlich ausgewählten Aktionen die Simulationsumgebung beeinflussen, ergeben sich keine gleichförmigen Sinus-Verläufe. Wenn die Agenten der hier dargestellten Episoden keine Exploration hätten und zu jedem Zeitpunkt den Erwartungswert als Aktion auswählen würden, würde sich bei jedem der Agenten ein gleichmäßiger Sinus-Verlauf ausbilden.

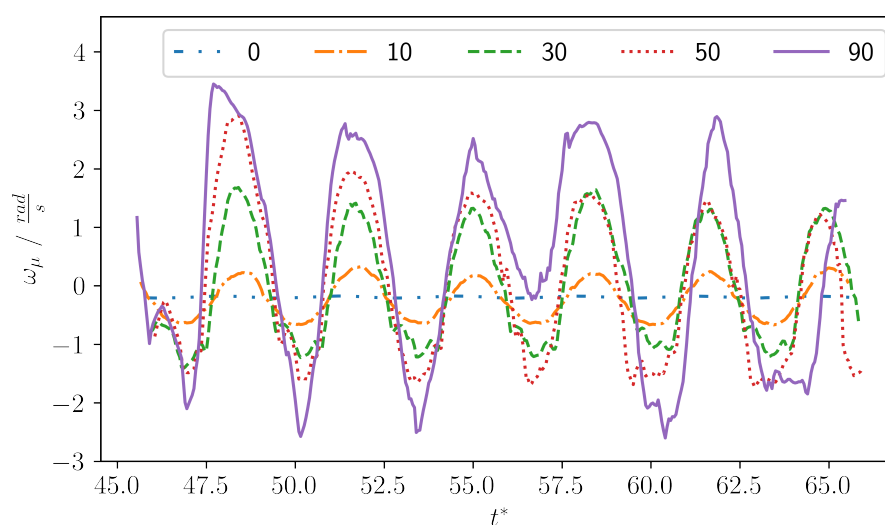


Abb. 5.2: Erwartungswert der Aktion im Verlauf einzelner, ausgewählter Episoden.

Die zweite Ausgabe des policy-Netzwerkes ist in Abbildung 5.3 zu sehen. Auffällig ist hierbei, dass die Standardabweichung bis zu Episode 30 im Verlauf der Episode nahezu konstant bleibt.

Bei Episode 50 und 90 hingegen ist deutlich zu erkennen, dass die Standardabweichung zu Beginn wesentlich geringer ist und im Verlauf der Trajektorie ansteigt. Es gilt also, dass zu Beginn, wo hauptsächlich der Widerstand reduziert wird, Abweichungen vom Erwartungswert große Auswirkungen auf die Belohnung haben. Im späteren Verlauf, wo die Wirbelablösung auf einem neuen, geringeren Niveau kontrolliert werden muss, ist die Standardabweichung hingegen vergleichsweise groß, was darauf schließen lässt, dass dort der exakte Wert der Aktion nicht mehr essentiell ist und daher mehr Wert auf die Erkundung neuer Aktionen gelegt werden kann.

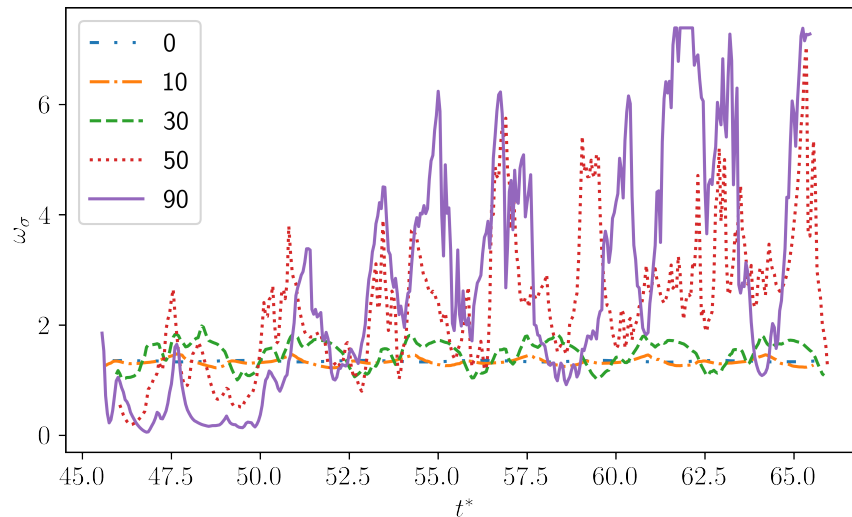


Abb. 5.3: Standardabweichung der Aktion im Verlauf einzelner, ausgewählter Episoden.

Ergebnis

Nach Beendigung des Trainings wird nun der Einfluss des DRL-Agenten auf die Strömungssimulation betrachtet. Der Agent hat dabei nur den Erwartungswert berücksichtigt und diesen als Aktion ausgewählt. In Abbildung 5.4 ist der Einfluss des Agenten auf den Widerstandsbeiwert dargestellt. Der Agent startet dabei bei der dimensionslosen Zeit 47. Zu Beginn kann der Agent den Widerstandsbeiwert deutlich reduzieren, bis dieser sich auf einem neuen, geringeren Niveau stabilisiert. Der durchschnittliche Widerstandsbeiwert liegt dann 5,05 % niedriger als im unkontrollierten Zustand. Die Amplitude der Schwingung des Widerstandsbeiwertes sinkt ebenfalls deutlich und wird insgesamt um 79,7 % reduziert.

Wie in Abbildung 5.5 ersichtlich wird, reduziert der Agent auch die Schwankungen im Auftriebsbeiwert. Nachdem sich der neue, stabile Zustand ausgebildet hat, wurde die Amplitude der Schwingung um 66,13% reduziert.

Der Vergleich mit der Arbeit von Darshan Thummar [14] zeigt sehr ähnliche Ergebnisse. Dort wurde mit einer sehr ähnlichen Simulationsumgebung und der gleichen Aktuation eine Reduktion von 5,10 % im Durchschnitt des Widerstandsbeiwertes erzielt. Außerdem wurde die Amplitude der Schwingung des Widerstandsbeiwertes um 72,30 % reduziert und die Amplitude der Schwingung des Auftriebsbeiwertes wurde um 67,18 % reduziert. In der Tabelle 5.1 sind die in diesem Abschnitt genannten Ergebnisse zusammengefasst.

Die Ergebnisse von Darshan Thummar konnten somit reproduziert werden und dienen im Folgenden als Referenz für die Überprüfung der Effektivität der Stabilisierungsmaßnahmen.

	\bar{c}_D	$c_{D_{max}}$	$c_{D_{min}}$	\hat{c}_D	$c_{L_{max}}$	$c_{L_{min}}$	\hat{c}_L
Unkontrollierter Zustand:	3,1941	3,2240	3,1639	0,0601	0.9682	-1.0045	1.9726
Kontrollierter Zustand:	3,0328	3,0381	3,0259	0,0122	0.1465	-0.5217	0.6682
Reduktion:	5,05 %			79,70 %			66,13 %
Reduktion in [14]:	5,10 %			72,30 %			67,18 %

Tabelle 5.1: Reduktion des Widerstandsbeiwertes, sowie Reduktion der Schwankungen im Auftriebs- und Widerstandsbeiwertes.

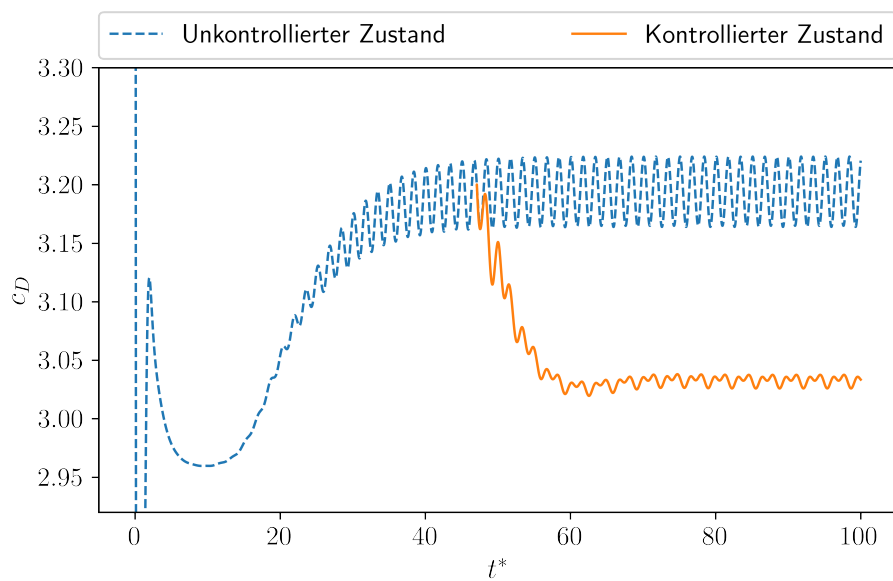


Abb. 5.4: Einfluss des Agenten auf den Widerstandsbeiwert.

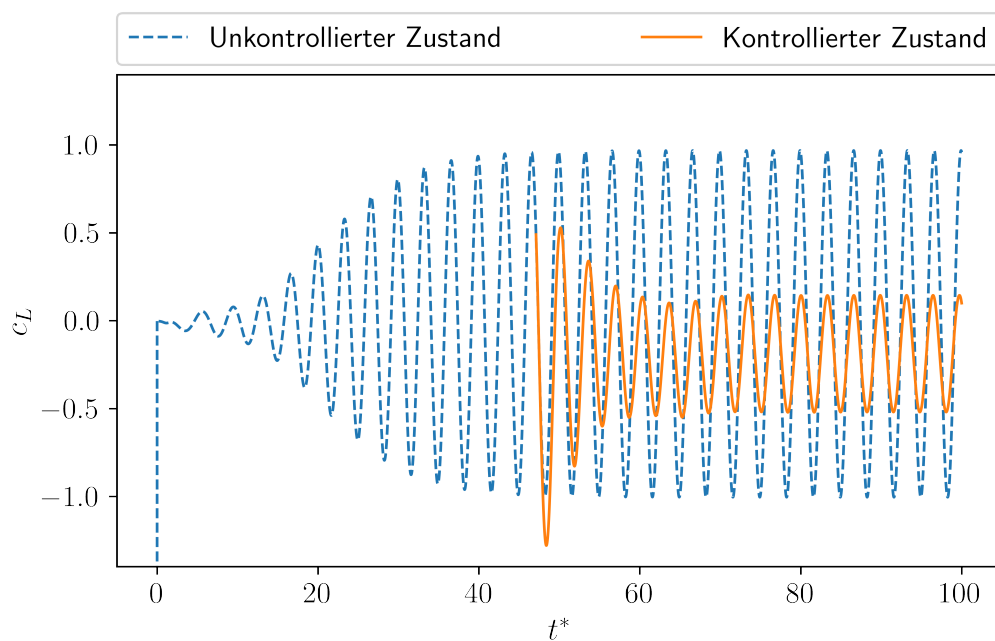


Abb. 5.5: Einfluss des Agenten auf den Auftriebsbeiwert.

Vergleich des Geschwindigkeitsfeldes

Wie in Abbildung 5.6 zu sehen ist, verändert der Agent durch die Aktuation das Geschwindigkeitsfeld so, dass die Blase direkt hinter dem Zylinder deutlich länger wird und dadurch die Wirbelablösung weiter von der Oberfläche des Zylinders entfernt liegt. Außerdem sind die Wirbel etwas weniger stark ausgeprägt. Die Strömung direkt über und unter dem Zylinder hat sich hingegen nicht signifikant verändert. Da die Frequenz der Wirbelablösung sich durch die Kontrolle des Agenten nicht ändert, ändert sich auch der Abstand zwischen den Wirbeln nicht signifikant.

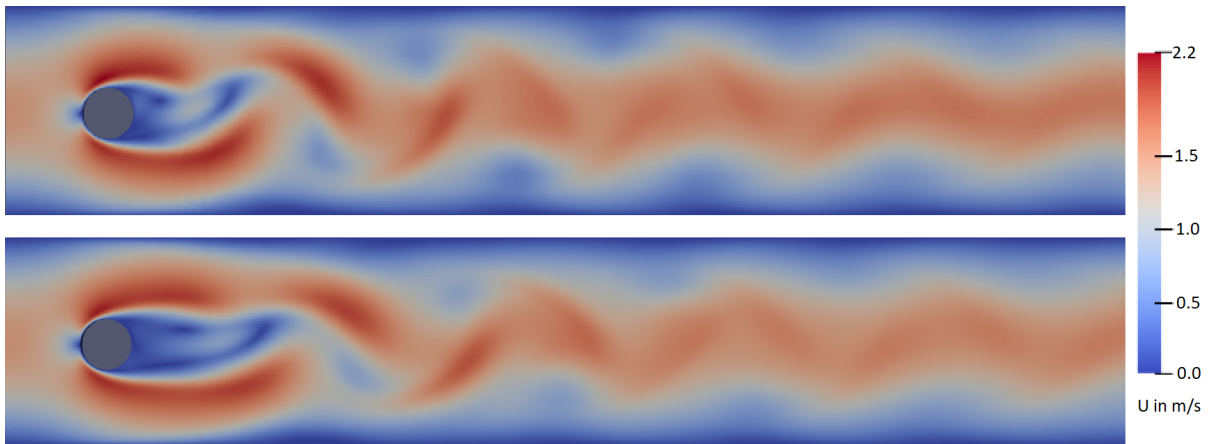


Abb. 5.6: Geschwindigkeitsfeld der Strömungssimulation nach acht Sekunden. Oben im unkontrollierten Zustand, unten im kontrollierten Zustand

Vergleich des Druckfeldes auf der Zylinderoberfläche

Die signifikante Senkung der Schwankungen der Widerstands- und Auftriebsbeiwerte ist vor allem auf die Reduzierung in den Schwankungen der Druckwerte auf der Oberfläche des Zylinders zurückzuführen. In Abbildung 5.7 sind einerseits die durchschnittlichen Druckwerte abgebildet und andererseits die Schwankungen von den Maximal- zu den Minimalwerten. Die Druckwerte sind dabei als Relativdruck zum Referenzdruck am Einlass und Auslass angegeben. Es ist deutlich zu erkennen, dass die Druckwerte im kontrollierten Zustand einerseits erheblich geringere Schwankungen besitzen und andererseits im Durchschnitt etwas höher liegen. Die Auftriebskraft entsteht vor allem durch den Druckunterschied zwischen der Oberseite und der Unterseite des Zylinders. Wenn auf der Oberseite ein Druckmaximum erreicht wird, entsteht auf der Unterseite ein Druckminimum. Die Kontrolle des Agenten bewirkt, dass die Druckunterschiede kleiner werden und somit die Auftriebskraft geringer wird. Im Gegensatz zum unkontrollierten Zustand liegt im kontrollierten Zustand das durchschnittliche Druckfeld allerdings nicht symmetrisch um die x-Achse des Zylinders. Der Druck auf der Oberseite des Zylinders ist im Durchschnitt etwas höher als auf der Unterseite des Zylinders. Die Auftriebskraft ist also im Mittel nicht mehr 0 sondern leicht negativ. Dies ist auch in der Abbildung 5.5 zu erkennen.

Die Widerstandskraft entsteht hauptsächlich durch den Unterschied des Druckes zwischen der linken Seite und der rechten Seite des Zylinders. Die reduzierten Schwankungen durch die Kontrolle des Agenten verringern somit auch die Schwankungen der Widerstandskraft. In Abbildung 5.7 ist außerdem zu sehen, dass die durchschnittlichen Druckwerte im kontrollierten Zustand auf der rechten Seite etwas größer sind als im unkontrollierten Zustand. Dies erklärt die Reduktion der durchschnittlichen Widerstandskraft.

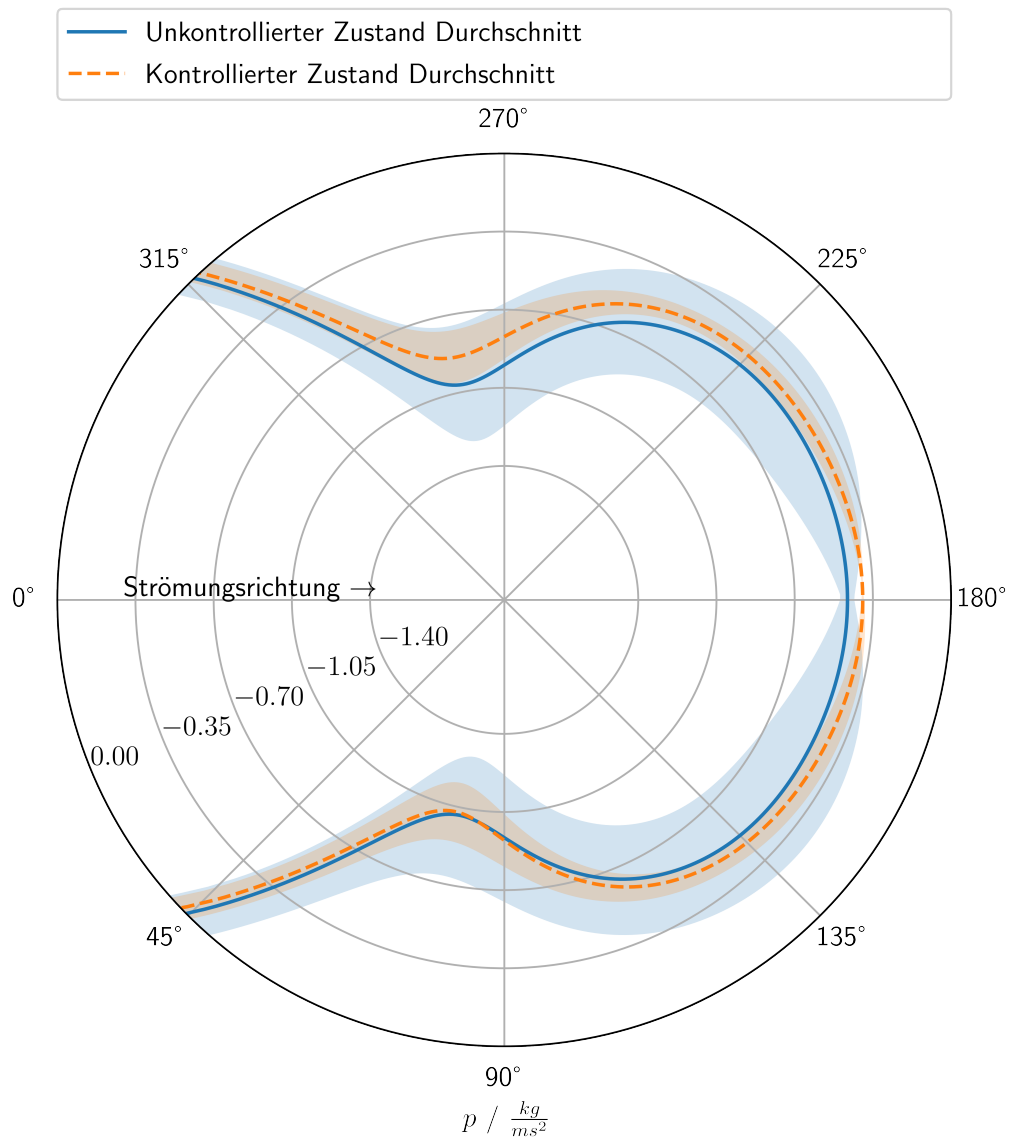


Abb. 5.7: Druckfeld der Strömungssimulation auf der Oberfläche des Zylinders nachdem der neue stabile Zustand eingetreten ist. Dargestellt ist der durchschnittliche Wert als Linie und alle Werte von Minimum bis Maximum als Fläche um die Linie. Der Druck ist als Relativdruck zum Referenzdruck an Einlass und Auslass angegeben.

5.2 Auswirkungen der Maßnahmen

Im folgenden Abschnitt werden die Auswirkungen der in Kapitel 2 vorgestellten Wahrscheinlichkeitsverteilungen dargestellt.

Belohnung

In Abbildung 5.8 ist die Belohnung der einzelnen Agenten zu sehen. Sowohl bei dem Agenten mit der Tangens hyperbolicus Funktion als Stauchungsfunktion, im Folgenden „Tanh“-Agent genannt, als auch bei dem Agenten mit der Beta-Verteilung, im Folgenden „Beta“-Agent genannt, steigt die Belohnung zu Beginn deutlich stärker an. Die maximale Belohnung der Normalverteilung von ca. $-0,104$ wird dabei schon nach 17 beziehungsweise 13 Episoden erreicht. Der Agent mit der Normalverteilung braucht dafür 86 Episoden. Der Tanh-Agent erreicht nach nur 25 Episoden sein Maximum bei etwa $-0,0873$. Danach nimmt die Belohnung wieder ab, so dass das Training nach 25 Episoden beendet ist. Dies entspricht einer Erhöhung um $16,06\%$ bei einer gleichzeitigen Zeitersparnis von $70,93\%$.

Im Vergleich dazu erreicht der Beta-Agent sein Maximum erst nach 81 Episoden, was einer Zeitersparnis von nur noch $5,81\%$ entspricht. Die Belohnung erhöht sich dabei auf $-0,0456$. Dies entspricht einem Anstieg von $56,15\%$. Demgegenüber wird die Belohnung des Tanh-Agent schon nach 22 Episoden übertroffen. Nach Episode 24 steigt die Belohnung allerdings nur noch sehr wenig an, sodass der Sprung von $-0,0706$ bei Episode 24 bis zu $-0,0456$ bei Episode 81 fast doppelt so lange dauert wie der Sprung von $-0,2598$ bei Episode 0 zu $-0,0706$ bei Episode 24.

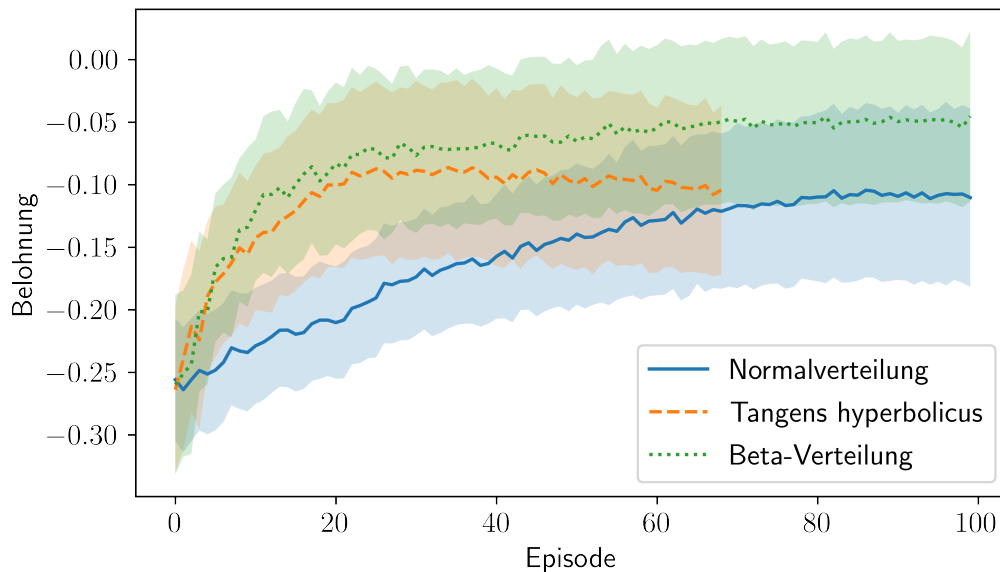


Abb. 5.8: Verlauf der durchschnittlichen Belohnung und der durchschnittlichen Belohnung $\pm\sigma$ im Verlauf des Trainings.

Ergebnisse

Wie in Abbildung 5.9 zu sehen ist, resultiert die höhere Belohnung auch in einer Reduktion des Widerstandsbeiwertes. Der Tanh-Agent reduziert dabei den Widerstandsbeiwert im Durchschnitt nur leicht um $5,15\%$, was im Vergleich zum Agenten mit der Normalverteilung einer weiteren Reduktion um $0,1$ Prozentpunkte entspricht. Die Amplitude der Schwingung des Widerstandsbeiwertes wird allerdings deutlich stärker reduziert. Diese sinkt im Vergleich zum unkontrollierten Zustand um $85,86\%$, was einer weiteren Reduktion um $6,16$ Prozentpunkte entspricht.

Der Beta-Agent erreicht insgesamt die größte Reduktion. Im Durchschnitt sinkt der Widerstandsbeiwert um 5,48 % und damit um weitere 0,43 Prozentpunkte im Vergleich zum Agenten mit der Normalverteilung und 0,33 Prozentpunkte im Vergleich mit dem Tanh-Agenten. Auch die Amplitude der Schwingung des Widerstandsbeiwertes wurde weiter reduziert und ist nun 87,85 % geringer als im unkontrollierten Zustand. Dies entspricht einer Reduktion von 8,15 Prozentpunkten im Vergleich zum Agenten mit der Normalverteilung und 1,99 Prozentpunkten im Vergleich zum Tanh-Agenten.

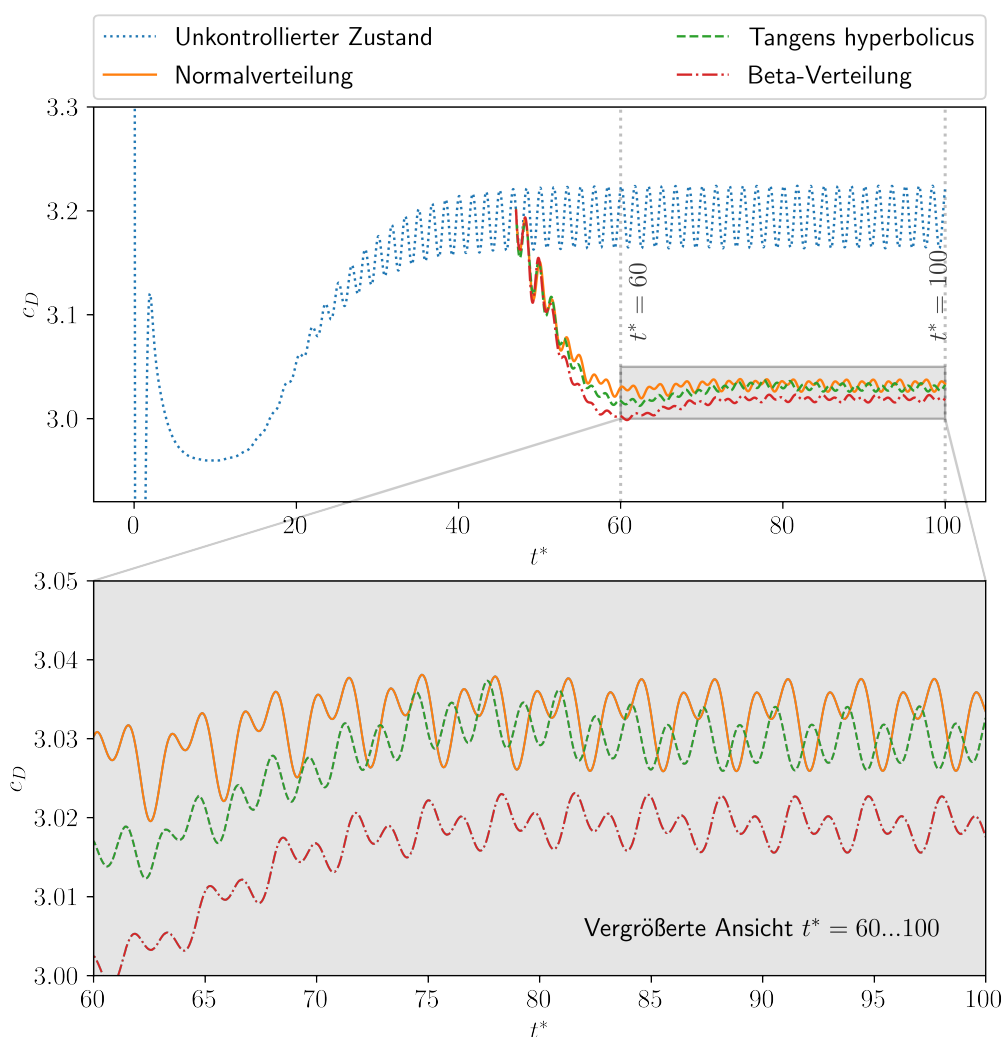


Abb. 5.9: Einfluss der einzelnen Agenten auf den Widerstandsbeiwert.

Neben der Amplitude der Schwingungen des Widerstandsbeiwertes sinkt auch die Amplitude der Schwingungen des Auftriebsbeiwertes durch die Kontrolle der Agenten weiter ab, wie in Abbildung 5.10 zu sehen ist. Der Tanh-Agent kann dabei mit 70,31 % die größte Reduktion verzeichnen. Dies entspricht einer Reduktion von weiteren 4,19 Prozentpunkten gegenüber dem Agenten mit der Normalverteilung. Der Beta-Agent kann hingegen nur eine Reduktion von 69,52 % verzeichnen. Damit liegt dieser 3,39 Prozentpunkte niedriger als der Agent mit der Normalverteilung, allerdings 0,8 Prozentpunkte höher als der Tanh-Agent. Insgesamt sind die Unterschiede zwischen den Agenten beim Auftriebsbeiwert also geringer als beim Widerstandsbeiwert. In Tabelle 5.2 sind alle absoluten Werte und die Reduktionen im Vergleich zum unkontrollierten Zustand noch einmal zusammengefasst.

	$\overline{c_D}$	$c_{D_{max}}$	$c_{D_{min}}$	\hat{c}_D	$c_{L_{max}}$	$c_{L_{min}}$	\hat{c}_L
Unkontrollierter Zustand:	3,1941	3,2240	3,1639	0,0601	0.9682	-1.0045	1.9726
Normalverteilung:	3,0328	3,0381	3,0259	0,0122	0.1465	-0.5217	0.6682
Reduktion:	5,05 %			79,70 %			66,13 %
Tangens hyperbolicus:	3,0297	3,0344	3,0259	0,0085	0.2378	-0.3479	0.5857
Reduktion:	5,15 %			85,86 %			70,31 %
Beta-Verteilung:	3,0191	3,0229	3,0156	0,0073	0.3468	-0.2545	0.6013
Reduktion:	5,48 %			87,85 %			69,52 %

Tabelle 5.2: Reduktion des Widerstandsbeiwertes, sowie Reduktion der Schwankungen im Auftriebs- und Widerstandsbeiwertes der verschiedenen Agenten.

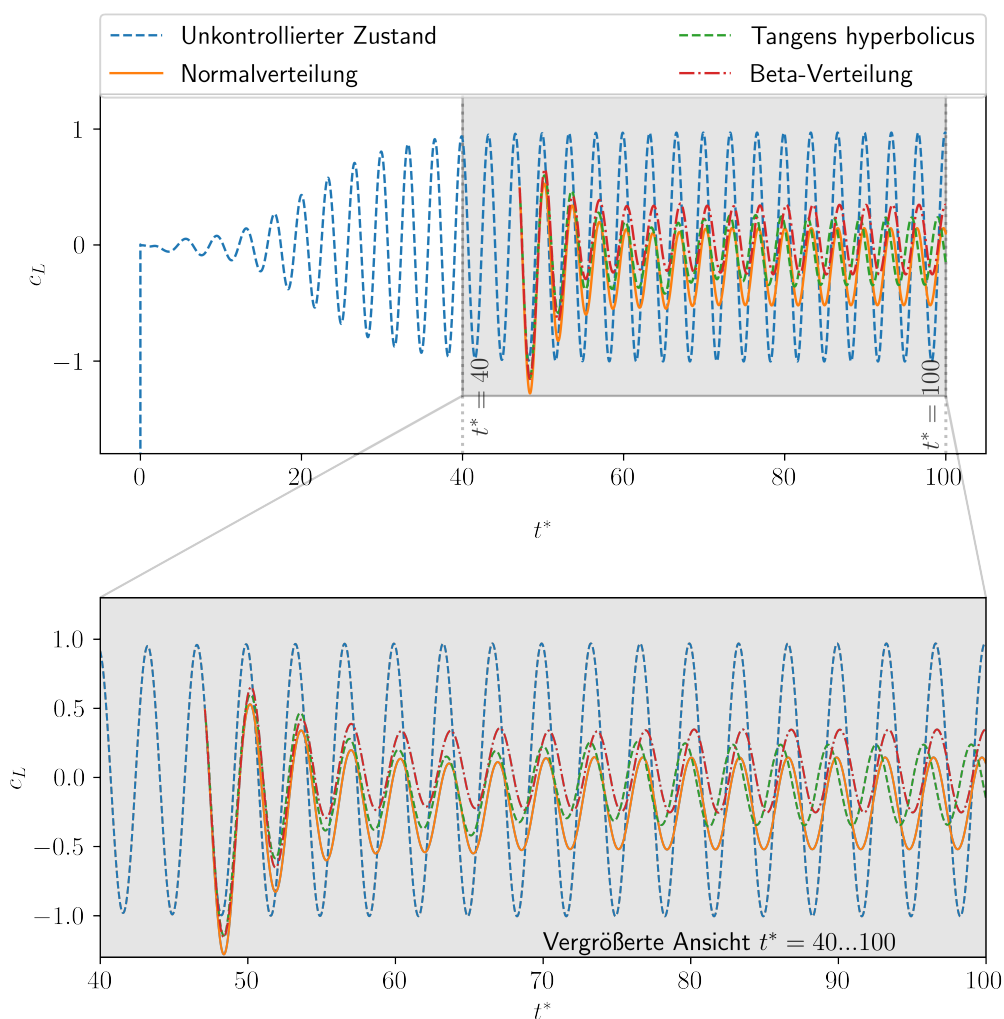


Abb. 5.10: Einfluss der einzelnen Agenten auf den Auftriebsbeiwert.

Verlauf der Erwartungswerte

Wie in Abbildung 5.11 zu erkennen ist, wählt sowohl der Beta-Agent als auch der Tanh-Agent zu Beginn der Kontrolle größere absolute Winkelgeschwindigkeiten aus. Nach der Einstellung des neuen stabilen Zustandes bildet sich jedoch ein deutlicher Unterschied in der Amplitude der Schwingung aus. Während der Beta-Agent und der Agent mit der Normalverteilung mit einem Unterschied von nur 1,05 % sehr vergleichbare Amplituden besitzen, liegt die Amplitude beim Tanh-Agent 38,33 % beziehungsweise 36,88 % höher. Da die Amplitude dieser Schwingung der Winkelgeschwindigkeit proportional zum Energieverbrauch eines Motors in einem Anwendungsfall ist, bedeutet das, dass der Energieverbrauch des Tanh-Agenten am größten ist. Alle hier betrachteten Agenten wurden allerdings nicht auf einen möglichst geringen Energieverbrauch, sondern nur auf eine möglichst starke Kontrolle trainiert, so dass selbst große Unterschiede hier nicht überraschend sind. Gleichzeitig wurde durch die Begrenzung des Aktionsraumes beim Tanh-Agenten und Beta-Agenten sichergestellt, dass der Motor die nötige Leistung erbringen könnte. Zu beachten ist hier auch, dass bei allen drei Agenten der Mittelwert der Winkelgeschwindigkeit von 0 abweicht. Das führt dazu, dass der Zylinder sich im Verlauf der Kontrolle insgesamt dreht.

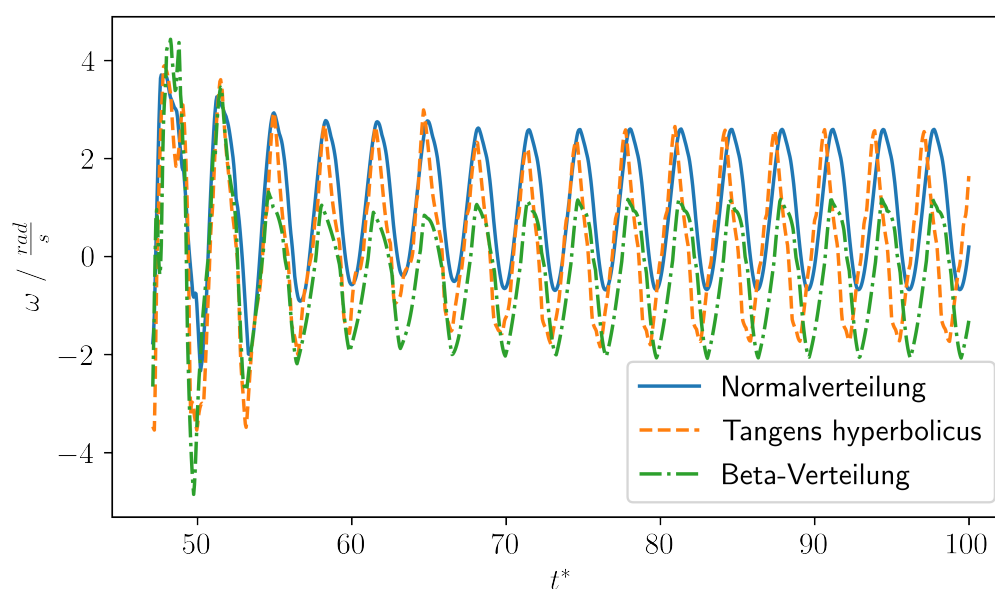


Abb. 5.11: Verlauf der Winkelgeschwindigkeit der einzelnen Agenten nach abgeschlossenem Training.

Verlauf der Standardabweichungen

In Abbildung 5.12 ist der Verlauf der Standardabweichungen der einzelnen Agenten nach Abschluss des Trainings dargestellt. Dabei wurde die Standardabweichung logarithmiert, um die einzelnen Verläufe besser darstellen zu können. Obwohl das Training der Agenten schon abgeschlossen ist und somit die Standardabweichung keine Auswirkung mehr auf die ausgewählten Aktionen hat, ist eindeutig zu erkennen, dass sowohl beim Beta-Agenten als auch beim Tanh-Agenten die Standardabweichung deutlich geringer ist. Das lässt darauf schließen, dass diese Agenten während des Trainings eine deutlich geringere Erkundung hatten und somit der Erwartungswert wesentlich wahrscheinlicher war. Auch die Schwankungen der Standardabweichung sind signifikant geringer, was zu einer gleichmäßigeren Erkundung im Verlauf einer Trajektorie führt.

Zu beachten ist auch, dass zu Beginn der Trajektorie die Standardabweichung niedrig ist und im Laufe der Trajektorie zuerst ansteigt und dann, wenn der neue kontrollierte Zustand erreicht

wird, auf ein mittleres Niveau sinkt. Diese Tendenz konnte schon in Kapitel 5.1 festgestellt werden und gilt in geringer Form auch für die beiden anderen Agenten.

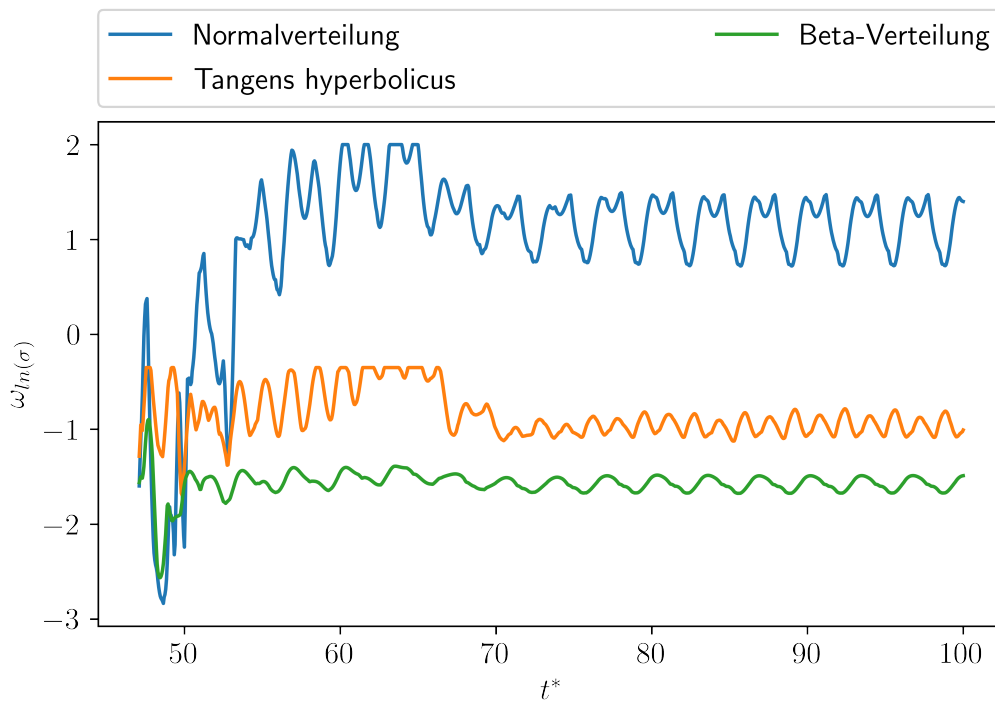


Abb. 5.12: Verlauf der Standardabweichungen der einzelnen Agenten nach abgeschlossenem Training in logarithmischer Darstellung.

5.3 Vergleich mit Literaturdaten

In Tabelle 5.3 sind die Reduktionen der einzelnen Agenten im Vergleich zu anderen Literaturdaten dargestellt. Der Simulationsaufbau von Hongwei Tang et al. [13] ist dem verwendeten Simulationsaufbau am ähnlichsten. Daher unterscheiden sich dort die Ergebnisse am wenigsten von den hier erzielten Ergebnissen. Durch unterschiedliche Simulationsaufbauten und unterschiedliche Aktuationen können große Unterschiede bei der Kontrolle der Wirbelablösungen erzielt werden. Dennoch liegen die hier erzielten Ergebnisse in einer ähnlichen Größenordnung wie die Literaturdaten.

Arbeit:	Reduktion $\overline{c_D}$	Reduktion \hat{c}_D	Reduktion \hat{c}_L
Rabault et al.[7]:	8 %	≈ 95 %	$\approx 82,5$ %
Tang et al. [13]:	5,7 %	-	-
Tokarev et al. [15]:	13,9 %	-	-
Normalverteilung:	5,05 %	79,70 %	66,13 %
Tangens hyperbolicus:	5,15 %	85,86 %	70,31 %
Beta-Verteilung:	5,48 %	87,85 %	69,52 %

Tabelle 5.3: Reduktion des Widerstandsbeiwertes, sowie Reduktion der Schwankungen im Auftriebs- und Widerstandsbeiwertes.

Kapitel 6

Zeitvariante Anströmung

In diesem Abschnitt wird nun der Agent mit verschiedenen Reynoldszahlen trainiert und dann auf eine zeitvariante Anströmung des Zylinders angewendet. Da im vorherigen Abschnitt der Beta-Agent die besten Ergebnisse erzielte, wird in diesem Abschnitt ausschließlich der Beta-Agent verwendet.

6.1 Training

Im Gegensatz zu Kapitel 5 wird das Training hier nicht nur bei der Reynoldszahl 100 durchgeführt, sondern auch gleichzeitig bei den Reynoldszahlen 200 und 400. Dazu werden bei jeder Episode jeweils 10 Simulationen zu jeder Reynoldszahl durchgeführt. Gemäß Formel 3.19 muss dazu auch der Zeitschritt angepasst werden und wird für die Reynoldszahl 200 halbiert und bei der Reynoldszahl 400 geviertelt. Da nach Formel 3.24 der Zeitschritt des Agenten vom Zeitschritt der Strömungssimulation abhängt, ändert sich damit auch der Horizont des Agenten bei veränderter Reynoldszahl. Um den Horizont konstant zu halten und die dimensionslose Simulationszeit nach Formel 3.2 dabei ebenfalls nicht zu verändern, wurde die reale Simulationszeit bei Reynoldszahl 200 auf eine Sekunde und bei Reynoldszahl 400 auf 0,5 Sekunden reduziert.

Belohnung

Der Verlauf der durchschnittlichen Belohnung während des Trainings (vgl. Abbildung 6.1) ist dem Verlauf des Beta-Agenten in Abbildung 5.8 sehr ähnlich. Zu Beginn ist die Belohnung etwas niedriger, was vor allem am größeren Aktionsbereich des neuen Agenten liegt. Die größeren Aktionen sorgen vor allem zu Beginn für deutlich erhöhte Schwankungen im Widerstands- und Auftriebsbeiwert, was wiederum die Belohnung senkt. Ab etwa Episode 30 steigt die Belohnung nur noch sehr leicht an und erreicht bei Episode 64 das Maximum von $+0,1586$. Sie liegt damit um einiges höher als noch in Abbildung 5.8, da die Reynoldszahlen 200 und 400 eine höhere durchschnittliche Belohnung besitzen. Die Standardabweichung ist über den gesamten Zeitraum nahezu konstant und liegt aufgrund der Unterschiede der Ergebnisse der einzelnen Reynoldszahlen etwas höher als zuvor.

In Abbildung 6.2 ist die durchschnittliche Belohnung nach Reynoldszahlen aufgeteilt und es ist klar zu erkennen, dass sich der Verlauf deutlich unterscheidet. Während der Verlauf bei der Reynoldszahl 100 dem Gesamtverlauf grob folgt, steigt die Belohnung bei der Reynoldszahl 200 zu Beginn deutlich stärker an. Bei dieser wird das Maximum schon nach etwa 20 Episoden erreicht und bleibt dann nahezu konstant. Auch bei der Reynoldszahl 100 bleibt die Belohnung nach Erreichen des Maximums bei etwa Episode 30 nahezu konstant. Bei der Reynoldszahl 400 steigt die Belohnung hingegen zu Beginn deutlich langsamer an. Bei etwa Episode 20 übersteigt die Belohnung dann die der Reynoldszahl 100 und steigt weiter an. Erst bei Episode 64 erreicht die durchschnittliche Belohnung bei der Reynoldszahl 400 ihr Maximum, was auch mit dem Ma-

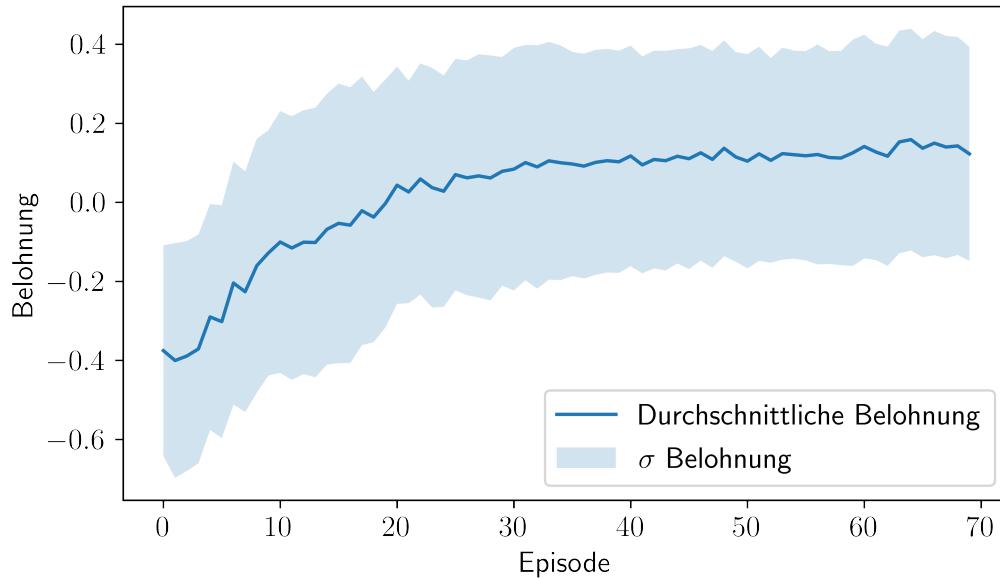


Abb. 6.1: Durchschnittliche Belohnung im Verlauf des Trainings bei variierender Reynoldszahl.

ximum der gesamten durchschnittlichen Belohnung übereinstimmt. Der langsame Anstieg der gesamten durchschnittlichen Belohnung lässt sich also darauf zurückführen, dass die Belohnung bei den Reynoldszahlen 100 und 200 nahezu konstant bleibt und lediglich bei der Reynoldszahl 400 ansteigt.

Zum Abschluss des Trainings ist die durchschnittlich Belohnung bei der Reynoldszahl 200 und 400 deutlich höher als bei der Reynoldszahl 100. Dies liegt vor allem an dem deutlich reduzierten durchschnittlichen Widerstandsbeiwert. Dass die Belohnung bei der Reynoldszahl 200 höher als bei der Reynoldszahl 400 ist, liegt vor allem daran, dass bei der Reynoldszahl 200 die Schwankungen im Widerstands- und Auftriebsbeiwert deutlich stärker reduziert werden.

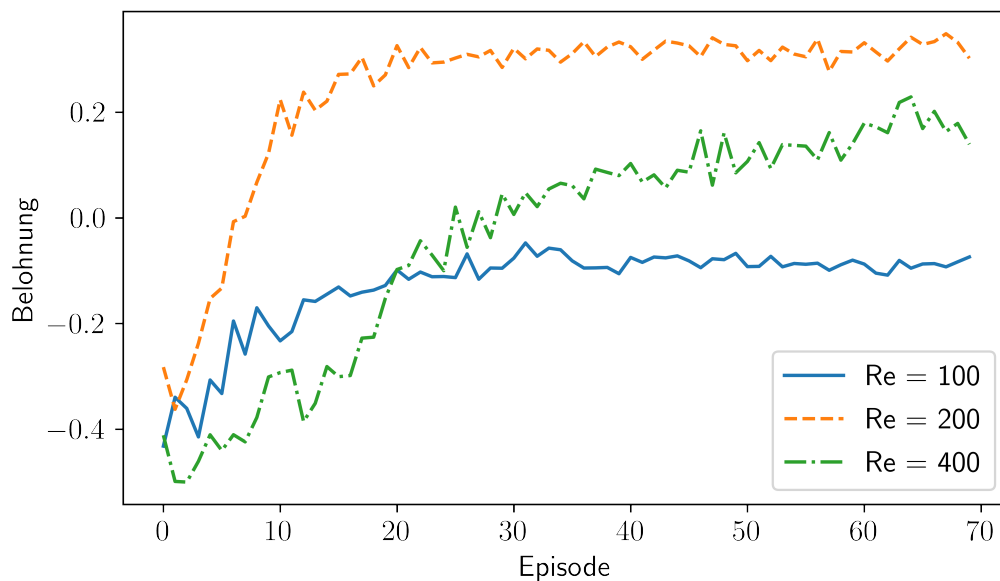


Abb. 6.2: Durchschnittliche Belohnung im Verlauf des Trainings, aufgeteilt nach den verschiedenen Reynoldszahlen.

6.2 Veränderungen in der Simulationsumgebung

Um eine zeitvariante Anströmung in der Simulationsumgebung zu erzeugen, muss die Einlassrandbedingung verändert werden. In Formel 3.15 wird die maximale Einlassgeschwindigkeit U_m mit $U_m = 1.5 \cdot u_\infty$ als zeitlich konstant festgelegt. Für die zeitvariante Anströmung wird die mittlere Einströmgeschwindigkeit u_∞ so verändert, dass gilt:

$$U_\infty = 2,5 \frac{m}{s} + 1,5 \frac{m}{s} \cdot \sin(\pi \cdot t) \quad (6.1)$$

Die mittlere Einströmgeschwindigkeit oszilliert mit einer Frequenz von 0,5 Hz und einer Amplitude von $1,5 \frac{m}{s}$ um einen Mittelwert von $2,5 \frac{m}{s}$. Für diese Simulationsumgebung bedeutet das eine Schwankung in der Reynoldszahl von 100 bis 400.

Bei der Entdimensionalisierung der Werte wo die mittlere Einlassgeschwindigkeit genutzt wird (siehe Kapitel 3.1), wird hier die durchschnittliche mittlere Einlassgeschwindigkeit von $2,5 \frac{m}{s}$ verwendet.

Da bei größeren Geschwindigkeiten auch größere Aktuatorenbereiche benötigt werden, wurde der Aktionsraum ebenfalls vergrößert. Der Agent kann in diesem Fall Winkelgeschwindigkeiten im Bereich von $[-40 \frac{rad}{s}, 40 \frac{rad}{s}]$ ausführen, was einer Oberflächengeschwindigkeit von $[-2 \frac{m}{s}, 2 \frac{m}{s}]$ entspricht.

6.3 Ergebnis

Durch den stark periodischen Charakter der mittleren Einlassgeschwindigkeit, ist auch der Widerstandsbeiwert stark periodisch (siehe Abbildung 6.3). Der gleitende Durchschnitt schwankt dabei zwischen etwa 0,59 im Minimum und 8,17 im Maximum. Im Minimum und in der ansteigenden Flanke bildet sich die Kármánsche Wirbelstraße nicht aus und der Widerstandsbeiwert schwankt dort aus diesem Grund nicht. Erst wenn der Widerstandsbeiwert einen Wert von etwa 5 übersteigt, bildet sich die Kármánsche Wirbelstraße. Da der Agent erst in der dritten Periode die Kontrolle vollständig ausbildet, wird im Folgenden nur diese Periode betrachtet. Der Agent schafft es zu Beginn der Wirbelbildung nicht, diese zu unterdrücken und die Schwankungen bleiben zunächst unverändert groß. Nach etwa vier Schwingungsperioden kontrolliert der Agent allerdings die sich bildenden Wirbel und schafft es dann vollständig, die Wirbelablösung zu verhindern. Statt der vergrößerten Blase im Nachlauf des Zylinders, wie in Kapitel 5 beschrieben, entstehen zwei stabile Wirbel im direkten Nachlauf mit einer vergleichsweise starken Rückflussströmung in der Mitte (vgl. Abbildung 6.4).

Dadurch, dass die Wirbelablösung verhindert wird, werden die Schwankungen im Widerstandsbeiwert fast vollständig eliminiert. In der Zeit zwischen etwa 6,45 Sekunden und 7,1 Sekunden, in der der Agent den Nachlauf stabil kontrolliert, sinkt der Widerstandsbeiwert im Durchschnitt um 36,21 % gegenüber dem gleitenden Durchschnitt. Durch die erheblichen Schwankungen zu Beginn der Wirbelbildung sinkt der maximale Widerstandsbeiwert allerdings nur um 14,43 % von 8,9086 auf 7,6235. Erst wenn der Widerstandsbeiwert einen Wert von 2 wieder unterschreitet, bildet sich wieder die Kármánsche Wirbelstraße aus, die sich allerdings fast sofort wieder auflöst, da dann das Minimum des Widerstandsbeiwertes erreicht wird.

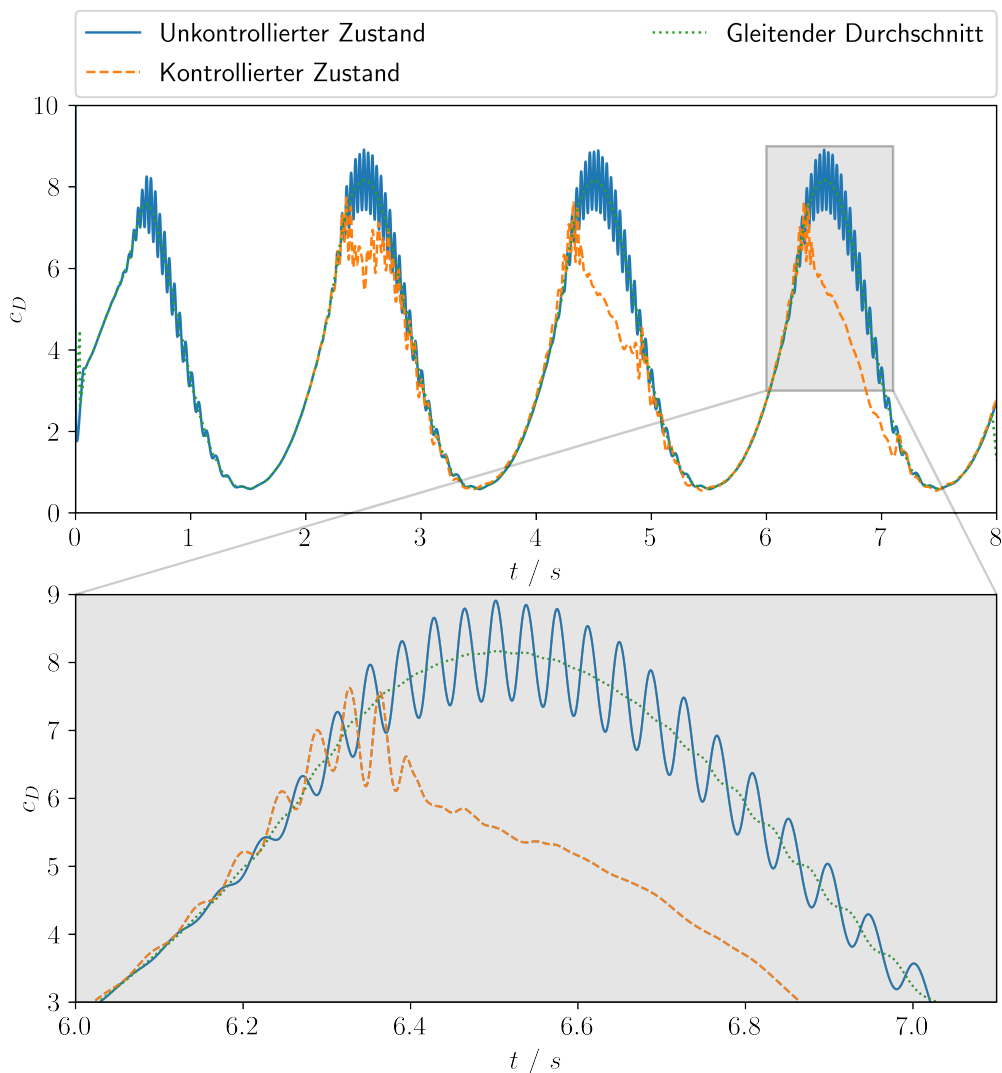


Abb. 6.3: Widerstandsbeiwert der Strömungssimulation mit zeitvarianter Anströmung im unkontrollierten und im kontrollierten Zustand.

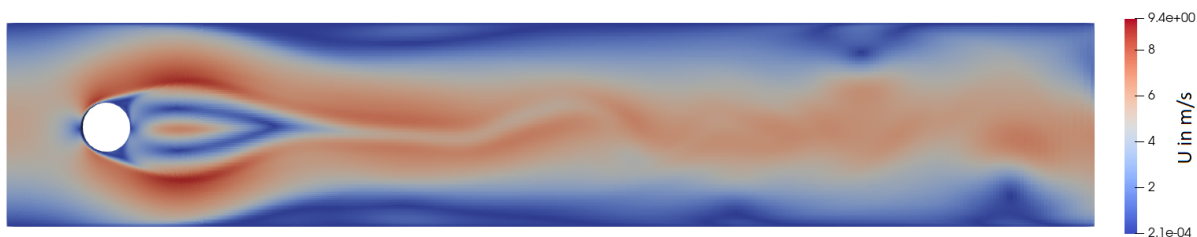


Abb. 6.4: Geschwindigkeitsfeld der kontrollierten Strömung nach 6,7 Sekunden.

Einfluss auf den Auftriebsbeiwert

Auch der Auftriebsbeiwert zeigt einen stark periodischen Charakter, allerdings bleibt hier der Mittelwert im unkontrollierten Zustand konstant bei 0. In Abbildung 6.5 ist zu sehen, dass der Agent wie beim Widerstandsbeiwert die Schwankungen zunächst nicht unterdrücken kann. Analog zum Widerstandsbeiwert werden die Schwankungen danach allerdings drastisch reduziert, bleiben aber noch leicht bestehen. Erst gegen Ende der Periode wird die Amplitude wieder größer, da der Nachlauf instabil wird und die Wirbelablösung wieder einsetzt.

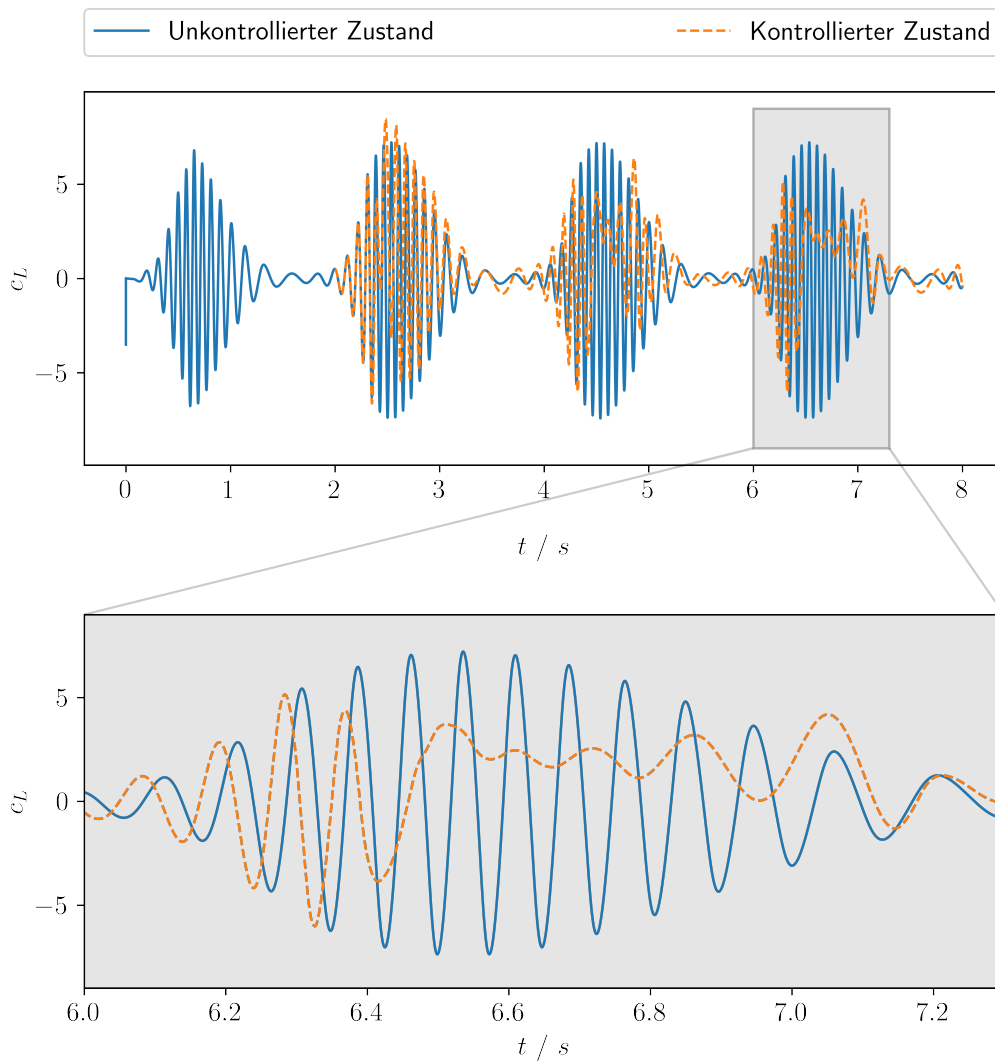


Abb. 6.5: Auftriebsbeiwert der Strömungssimulation mit zeitvarianter Anströmung im unkontrollierten und im kontrollierten Zustand.

Drehgeschwindigkeit des Agenten

Im Gegensatz zu den Agenten in Kapitel 5 bildet der Agent hier keinen gleichmäßigen Sinus-Verlauf der Drehgeschwindigkeit aus (siehe Abbildung 6.6). Zu Beginn jeder Periode gibt es große Ausschläge in positive und negative Richtung, die fast bis an die Grenzen des Aktionsraumes reichen. Danach liegt die Drehgeschwindigkeit in einem Bereich von etwa $[-30\frac{rad}{s}, 10\frac{rad}{s}]$ und ist dabei sehr unregelmäßig. Zum Ende jeder Periode steigt die Drehgeschwindigkeit noch einmal stark an, bis sie etwa $30\frac{rad}{s}$ erreicht. Im Minimum und in der ansteigenden Flanke, wo die Kármánsche Wirbelstraße sich nicht ausbildet, agiert der Agent auch nicht und verbraucht somit dort auch keine Energie. Dies war keine vorgegebene Maßnahme und entsteht, wie jede andere Aktion auch, alleine durch das policy-Netzwerk des Agenten.

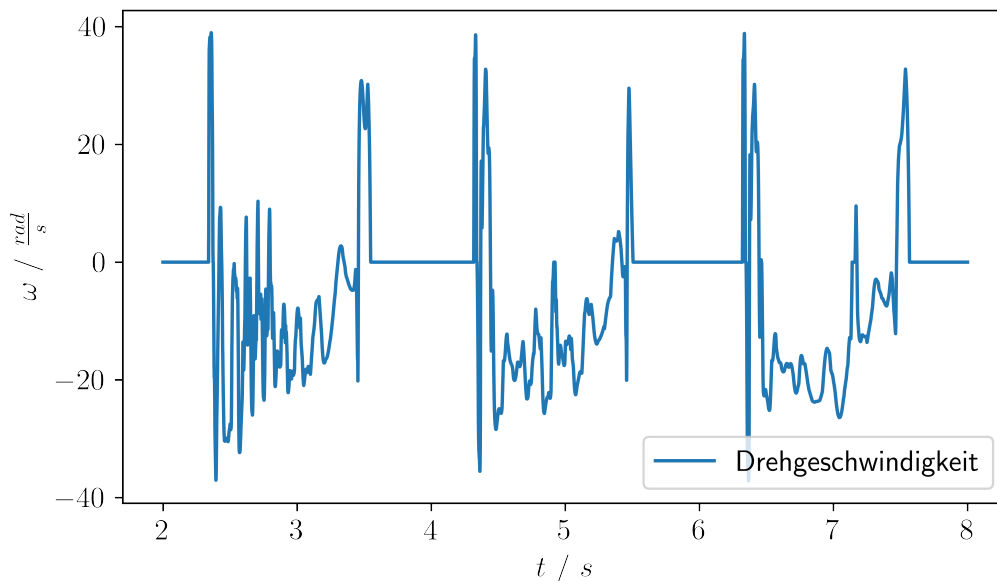


Abb. 6.6: Verlauf der Drehgeschwindigkeit bei der zeitvarianten Anströmung.

Verspätetes Einsetzen der Kontrolle

Eine mögliche Erklärung für das verspätete Eintreten der Kontrolle könnte sein, dass beim Agenten im Training der Startzeitpunkt der einzelnen Trajektorien so gewählt ist, dass sich zu Beginn der Kontrolle bereits eine Kármánsche Wirbelstraße voll ausgebildet hat. Eine Veränderung des Startzeitpunktes, so dass auch eine sich bildende Kármánsche Wirbelstraße kontrolliert werden kann, könnte dieses Verhalten verbessern.

Wie im vorherigen Abschnitt schon beschrieben, erreicht der Agent zu Beginn jeder Periode Drehgeschwindigkeiten die sehr Nahe an die Grenze des Aktionsraumes reichen. Eine weitere mögliche Erklärung für das verspätete Einsetzen könnte also sein, dass der Aktionsraum etwas zu klein gewählt wurde und eine Vergrößerung dieses Aktionsraumes eine verbesserte Kontrolle ermöglicht.

Fazit

Abschließend kann zusammengefasst werden, dass die Strategie den Agenten bei konstanten Reynoldszahlen zu trainieren, um eine zeitvariante Strömung zu kontrollieren, funktioniert hat. Der Agent hat es geschafft Auch die zeitvariante Strömung zu kontrollieren. Da diese Kontrolle allerdings erst etwas verspätet eintritt, gibt es hier noch Potential für weitere Verbesserungen.

Kapitel 7

Zusammenfassung

Um bei Flugzeugen und Fahrzeugen den Treibstoffverbrauch zu senken und insgesamt Kosten zu sparen, ist es von elementarer Bedeutung, die Widerstandskraft und speziell den Luftwiderstand zu reduzieren. Neben verschiedenen passiven Strömungskontrollmöglichkeiten kann eine Strömung um ein Objekt auch durch Zuführen externer Energie aktiv kontrolliert werden. Dazu wird allerdings eine effiziente Kontrollstrategie benötigt, die durch die Hochdimensionalität und den daraus resultierenden hohen Berechnungskosten der Simulation aktuell nicht analytisch bestimmt werden kann. Allerdings gab es im Bereich des tiefen, bestärkenden Lernens in den letzten Jahren große Fortschritte bei der Entwicklung dieser Kontrollstrategien. Bei der Stabilität und Geschwindigkeit des Training gibt es trotz vergleichsweise simplen Testfällen noch Potential für Verbesserungen.

Die 2D-Zylinderumströmung bietet aufgrund ihres relativ einfachen Aufbaus und damit verbundenen kurzen Simulationszeiten viele Vorteile für das Training eines DRL-Agenten. Es hat sich daher in der Literatur als eines der am häufigsten verwendeten Benchmark-Probleme entwickelt. Dabei lösen sich im Nachlauf des Zylinders sogenannte von Kármán-Wirbel ab, die einerseits den durchschnittlichen Widerstand erhöhen und andererseits große Schwankungen im Widerstand und Auftrieb erzeugen. Diese Wirbel zu kontrollieren ist das Hauptziel des DRL-Agenten. Um das Strömungsproblem numerisch lösen zu können, muss ein Rechengitter gefunden werden, das einerseits dicht genug ist, so dass die Strömungssimulation nicht signifikant von noch dichteren Gittern abweicht, und andererseits so wenig Zellen wie möglich hat, um eine möglichst schnelle Simulation zu erzeugen. Durch eine Gitterabhängigkeitsstudie wurde ein Gitter mit 85.000 Zellen gewählt, bei dem eine Sekunde einer Strömung mit einer Reynoldszahl von 400 in 1933 Sekunden oder 32 Minuten berechnet werden kann.

Der DRL-Agent wird über einen proximal policy optimization (PPO)-Algorithmus trainiert. Dabei werden zwei verschiedene neuronale Netze trainiert, um einerseits den Zusammenhang zwischen Aktion und Ertrag (policy-Netzwerk) und andererseits den Zusammenhang zwischen Zustand und Ertrag (value-Netzwerk) zu approximieren. Um dabei neue andere Aktionen zu erkunden und damit feststellen zu können, ob die Aktionen einen besseren Ertrag erzeugen würden, werden die Aktionen während des Trainings nicht deterministisch, sondern anhand einer Wahrscheinlichkeitsverteilung ausgewählt. Dabei wurde in der Vergangenheit häufig eine Normalverteilung verwendet. Da diese allerdings nicht auf einen Aktionsraum begrenzt ist, können theoretisch beliebig große Aktionen ausgewählt werden. Um dies zu vermeiden und ein potentiell stabileres Training zu erzeugen, wurde hier die Wahrscheinlichkeitsverteilung verändert und mit der Normalverteilung verglichen.

Der Agent, der eine Normalverteilung verwendet, erreicht nach 86 Episoden sein Maximum und kann den durchschnittlichen Widerstandsbeiwert um 5,05 % senken. Die Amplitude in den Schwankungen des Widerstands- und des Auftriebsbeiwertes können zudem um 79,7 % beziehungsweise 66,13 % reduziert werden.

Durch die Stauchung der Normalverteilung mit einer Tangens hyperbolicus Funktion wird diese auf den Bereich von $[-1,1]$ begrenzt. Die Aktion kann dann auf den gewünschten Aktionsbereich skaliert werden. Der Agent, der diese veränderte Normalverteilung verwendete, konnte den Widerstand um 5,15 % senken und somit 0,1 Prozentpunkte mehr als der Agent mit der Normalverteilung. Die Schwankungen des Widerstands- und des Auftriebsbeiwertes konnten ebenfalls weiter reduziert werden. Dort erreicht der Agent eine Reduktion von 85,86 % beziehungsweise 70,31 %. Dies entspricht einer weiteren Reduktion von 6,16 und 4,19 Prozentpunkten. Sein Maximum erreichte der Agent dabei schon nach 25 Episoden. Der Agent mit der Normalverteilung konnte wiederum schon nach 17 Episoden überboten werden.

Eine weitere Möglichkeit, den Aktionsbereich zu begrenzen bietet die Beta-Verteilung. Diese begrenzt den Bereich auf $[0,1]$ und kann dann ebenfalls auf den Aktionsbereich skaliert werden. Der Agent, der die Beta-Verteilung verwendete, konnte die Ergebnisse weiter verbessern. Der durchschnittliche Widerstandsbeiwert konnte um 5,48 % gesenkt werden und damit um 0,43 Prozentpunkte mehr gegenüber dem Agenten mit der Normalverteilung. Die Schwankungen des Widerstandsbeiwertes konnten ebenfalls weiter reduziert werden. Dort erreicht der Agent eine Reduktion von 87,85 %. Dies entspricht einer Reduktion von 8,15 Prozentpunkten gegenüber dem Agenten mit der Normalverteilung und 1,99 Prozentpunkten im Vergleich zum Tanh-Agenten. Lediglich die Schwankungen des Auftriebsbeiwertes konnten gegenüber dem Tanh-Agenten nicht weiter verbessert werden. Dort erreicht der Beta-Agent eine Reduktion von 69,52 %, was gegenüber dem Agenten mit der unveränderten Normalverteilung zwar einer weiteren Reduktion von 3,39 Prozentpunkten entspricht, aber gegenüber dem Tanh-Agenten einer Erhöhung von 0,8 Prozentpunkten.

Dieser Unterschied ist im Vergleich zu den sonstigen Verbesserungen des Beta-Agenten allerdings vernachlässigbar, sodass zusammenfassend gesagt werden kann, dass der Beta-Agent die besten Ergebnisse erzielt und bei der folgenden zeitvarianten Strömung als Agent verwendet wird.

Bei der zeitvarianten Strömung wurde der Agent mit den konstanten Reynoldszahlen 100, 200 und 400 trainiert und anschließend auf eine Strömungssimulation angewendet, bei der die mittlere Einlassgeschwindigkeit zwischen $1,0 \frac{m}{s}$ und $4,0 \frac{m}{s}$ mit einer Periode von 0,5 Hz oszilliert. Dabei konnte der Agent die Kármánsche Wirbelstraße erfolgreich kontrollieren und den maximalen Widerstandsbeiwert um 14,43 % senken. Zudem konnte der Agent zeitweise die Wirbelablösung vollständig verhindern und währenddessen den Widerstandsbeiwert um 36,21 % senken. Die Schwankungen des Widerstands- und Auftriebsbeiwertes konnten ebenfalls drastisch reduziert werden.

Zusammenfassend lässt sich also resümieren, dass durch die Wahl der Beta-Verteilung das Training deutlich verbessert werden kann. Zudem kann ein Agent, der während des Trainings bei konstanten Reynoldszahlen trainiert wurde, auch eine Strömung kontrollieren, deren Reynoldszahl sich im Laufe der Zeit ändert.

Um einer Anwendung in der Realität näher zu kommen, könnte getestet werden, ob ein DRL-Agent auch bei einem komplexeren Profil und einer dreidimensionalen Simulation die Strömung ebenfalls kontrollieren kann. Zudem könnte getestet werden, ob sich die Kontrollstrategie und die Ergebnisse auch auf die Realität übertragen lassen, indem zum Beispiel ein Zylinder in einem Windkanal einer solchen Strömung ausgesetzt wird und die Kräfte, die auf den Zylinder wirken gemessen werden.

Literaturverzeichnis

- [1] Charles-Henri Bruneau and Iraj Mortazavi. Numerical modelling and passive flow control using porous media. *Computers & Fluids*, 37(5):488–498, 2008. Special Issue Dedicated to Professor M.M. Hafez on the Occasion of his 60th Birthday.
- [2] R. Franke, Wolfgang Rodi, and Bernhard Schönung. Numerical calculation of laminar vortex-shedding flow past cylinders. *Journal of Wind Engineering and Industrial Aerodynamics*, 35:237–257, 1990.
- [3] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*, pages 1861–1870. PMLR, 2018.
- [4] Chloe Ching-Yun Hsu, Celestine Mender-Dünner, and Moritz Hardt. Revisiting design choices in proximal policy optimization. *CoRR*, abs/2009.10897, 2020.
- [5] Norman L Johnson, SI Kotz, and N Balakrishnan. Beta distributions. *Continuous univariate distributions. 2nd ed. New York, NY: John Wiley and Sons*, pages 221–235, 1994.
- [6] Romain Paris, Samir Beneddine, and Julien Dandois. Robust flow control and optimal sensor placement using deep reinforcement learning. *Journal of Fluid Mechanics*, 913:A25, 2021.
- [7] Jean Rabault, Miroslav Kuchta, Atle Jensen, Ulysse Réglade, and Nicolas Cerardi. Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *Journal of Fluid Mechanics*, 865:281–302, 2019.
- [8] Jean Rabault and Alexander Kuhnle. Accelerating deep reinforcement learning strategies of flow control through a multi-environment approach. *Physics of Fluids*, 31(9):094105, 2019.
- [9] M. Schäfer, S. Turek, F. Durst, E. Krause, and R. Rannacher. *Benchmark Computations of Laminar Flow Around a Cylinder*, pages 547–566. Vieweg+Teubner Verlag, Wiesbaden, 1996.
- [10] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms, 2017.
- [11] Hans Heiner Storrer. *Die Normalverteilung*, pages 179–194. Birkhäuser Basel, Basel, 1995.
- [12] Richard S. Sutton and Andrew G. Barto. Reinforcement learning: An introduction. 2012.
- [13] Hongwei Tang, Jean Rabault, Alexander Kuhnle, Yan Wang, and Tongguang Wang. Robust active flow control over a range of reynolds numbers using an artificial neural network trained through deep reinforcement learning. *Physics of Fluids*, 32(5):053605, 2020.
- [14] Darshan Thummar. Active flow control in simulations of fluid flows based on deep reinforcement learning, May 2021.
- [15] Mikhail Tokarev, Egor Palkin, and Rustam Mullyadzhhanov. Deep reinforcement learning

- control of cylinder flow using rotary oscillations at low reynolds number. *Energies*, 13(22), 2020.
- [16] P.R Viswanath. Aircraft viscous drag reduction using riblets. *Progress in Aerospace Sciences*, 38(6):571–600, 2002.
- [17] C H K Williamson. Vortex dynamics in the cylinder wake. *Annual Review of Fluid Mechanics*, 28(1):477–539, 1996.
- [18] Kianoosh Yousefi, Reza Saleh, and Peyman Zahedi. Numerical study of blowing and suction slot geometry optimization on naca 0012 airfoil. *Journal of Mechanical Science and Technology*, 28:1297–1310, 05 2014.

Abbildungsverzeichnis

1.1	Schaubild eines DRL-Systems. Diese Abbildung ist die Abbildung 4 des Artikels „Robust active flow control over a range of Reynolds numbers using an artificial neural network trained through deep reinforcement learning“ [6].	3
1.2	Auswirkung der DRL-Kontrolle auf das Geschwindigkeitsfeld. (a) ist dabei der unkontrollierte Fall und (b) der aktiv kontrollierte. Diese Abbildung ist die Abbildung 4 des Artikels „Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control“ [7].	4
1.3	Einfluss der Anzahl der Sensoren auf den DRL-Agenten bei optimaler Platzierung. Diese Abbildung ist die Abbildung 20 des Artikels „Robust flow control and optimal sensor placement using deep reinforcement learning“ [6].	5
2.1	Simulation einer 2D-Zylinderumströmung bei $Re = 50$. Dargestellt ist die Geschwindigkeit mit hoher Geschwindigkeit in rot und niedriger Geschwindigkeit in blau.	7
2.2	Fotografien von Kármánschen Wirbelstraßen bei verschiedenen Reynoldszahlen. Diese Abbildung ist die Abbildung 1 des Artikels „Vortex dynamics in the cylinder wake“ [17].	8
2.3	Dichtefunktion der Normalverteilung bei verschiedenen Standardabweichungen. Die Abbildung stammt von der Seite 179 des Buches „Einführung in die mathematische Behandlung der Naturwissenschaften II“ [11].	9
2.4	Einfluss des Mittelwerts auf die veränderte Normalverteilung bei einer Standardabweichung von 0,5.	10
2.5	Einfluss der Standardabweichung auf die veränderte Normalverteilung.	10
2.6	Dichtefunktion der Beta-Verteilung mit unterschiedlichen Parametern. Diese Abbildung ist die Abbildung 4 des Artikels „Beta distribution“ [5].	12
2.7	Dichtefunktion der Beta-Verteilung als stetige Gleichverteilung. Diese Abbildung ist die Abbildung 3 des Artikels „Beta distribution“ [5].	13
3.1	Geometrie des Simulationsbereiches. Alle Längen sind über den Zylinderdurchmesser entdimensionalisiert.	14
3.2	Gitter mit 50 Zellen in y-Richtung.	16
3.3	Mittlere, maximale und minimale Widerstandsbeiwerte in Abhängigkeit der Gittergröße.	18
3.4	Flussdiagramm des PPO-Algorithmus	20
4.1	Strouhal-Zahl in Abhängigkeit der Reynoldszahl. Die Referenzdaten stammen aus der Tabelle 1 aus dem Artikel „numerical calculation of laminar vortex shedding flow past cylinders“ [2].	23
4.2	Widerstandsbeiwert C_D in Abhängigkeit der Reynoldszahl.	24
4.3	Auftriebsbeiwert C_L in Abhängigkeit der Reynoldszahl.	25
5.1	Durchschnittliche Belohnung im Verlauf des Trainings.	27
5.2	Erwartungswert der Aktion im Verlauf einzelner, ausgewählter Episoden.	27
5.3	Standardabweichung der Aktion im Verlauf einzelner, ausgewählter Episoden.	28
5.4	Einfluss des Agenten auf den Widerstandsbeiwert.	29
5.5	Einfluss des Agenten auf den Auftriebsbeiwert.	29
5.6	Geschwindigkeitsfeld der Strömungssimulation nach acht Sekunden. Oben im unkontrollierten Zustand, unten im kontrollierten Zustand	30

5.7	Druckfeld der Strömungssimulation auf der Oberfläche des Zylinders nachdem der neue stabile Zustand eingetreten ist. Dargestellt ist der durchschnittliche Wert als Linie und alle Werte von Minimum bis Maximum als Fläche um die Linie. Der Druck ist als Relativdruck zum Referenzdruck an Einlass und Auslass angegeben.	31
5.8	Verlauf der durchschnittlichen Belohnung und der durchschnittlichen Belohnung $\pm\sigma$ im Verlauf des Trainings.	32
5.9	Einfluss der einzelnen Agenten auf den Widerstandsbeiwert.	33
5.10	Einfluss der einzelnen Agenten auf den Auftriebsbeiwert.	34
5.11	Verlauf der Winkelgeschwindigkeit der einzelnen Agenten nach abgeschlossenem Training.	35
5.12	Verlauf der Standardabweichungen der einzelnen Agenten nach abgeschlossenem Training in logarithmischer Darstellung.	36
6.1	Durchschnittliche Belohnung im Verlauf des Trainings bei variierender Reynoldszahl.	38
6.2	Durchschnittliche Belohnung im Verlauf des Trainings, aufgeteilt nach den verschiedenen Reynoldszahlen.	38
6.3	Widerstandsbeiwert der Strömungssimulation mit zeitvarianter Anströmung im unkontrollierten und im kontrollierten Zustand.	40
6.4	Geschwindigkeitsfeld der kontrollierten Strömung nach 6,7 Sekunden.	40
6.5	Auftriebsbeiwert der Strömungssimulation mit zeitvarianter Anströmung im unkontrollierten und im kontrollierten Zustand.	41
6.6	Verlauf der Drehgeschwindigkeit bei der zeitvarianten Anströmung.	42

Tabellenverzeichnis

3.1	Zellen- und Punktzahl sowie Berechnungszeit bei verschiedenen Gittergrößen. .	17
5.1	Reduktion des Widerstandsbeiwertes, sowie Reduktion der Schwankungen im Auftriebs- und Widerstandsbeiwertes.	29
5.2	Reduktion des Widerstandsbeiwertes, sowie Reduktion der Schwankungen im Auftriebs- und Widerstandsbeiwertes der verschieden Agenten.	34
5.3	Reduktion des Widerstandsbeiwertes, sowie Reduktion der Schwankungen im Auftriebs- und Widerstandsbeiwertes.	36