CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL
UNIDAD ZACATENCO

DEPARTAMENTO DE CONTROL AUTOMÁTICO

Control adaptable para estados restringidos de un robot bípedo usando aprendizaje por refuerzo

T E S I S

Que presenta

KARLA RINCON MARTINEZ

Para obtener el grado de

DOCTORA EN CIENCIAS

En la especialidad de

CONTROL AUTOMÁTICO

Directores de tesis:

DR. WEN YU LIU

DR. JORGE ISAAC CHAIREZ ORIA

Ciudad de México Abril, 2023

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL
UNIDAD ZACATENCO

DEPARTAMENTO DE CONTROL AUTOMÁTICO

Adaptive state-restricted control of a Biped Robot using Reinforcement Learning

D I S S E R T A T I O N

Presented by

KARLA RINCON MARTINEZ

In candidacy for the Degree of

DOCTOR OF SCIENCE

In the specialization of

AUTOMATIC CONTROL

Thesis Advisors:

DR. WEN YU LIU

DR. JORGE ISAAC CHAIREZ ORIA

Mexico City April 2023

The scientist finds his reward in what Henri Poincare calls the joy of comprehension, and not in the possibility of application to which any discovery may lead. — Albert Einstein [a]

[a] *The Ultimate Quotable Einstein*, Princeton University Press

Acknowledgement

"We must find time to stop and thank the people who make a difference in our lives." — John F. Kennedy

I want to thank all the people who have been at my side giving me the courage to achieve the end of this phase in my academic journey.

Particularly to my family: Tere y Ale, Soco, Mario y María. Your support and acceptance mean everything to me.

To my beloved Diana, who has changed so many things, including the way that I see both myself, and the future I can shape, all the words in all the languages all around the world cannot be enough to say how grateful I am to be part of each other's life. I can't wait to see what the future holds for us.

To my friends who always find time to hear me or advise me on my challenges, setbacks, and venting. Rebeca, for your friendship, advice, and coffee dates, particularly in this 4-year marathon. Pris, Pau & Clar, because time and distance won't corrupt our bond, every time we see us is like time doesn't go. The Other Side girls, always listening to my eternal whining, I couldn't have survived the pandemic doctorate era without you.

To my compis and acquaintances at the Signals Processing and Medical Robotics Lab, you have made so much difference in the last stretch of this journey.

A profound acknowledgment to Dr. Isaac and Dr. Wen, for your time and advice, your patience with me, and all the knowledge shared. It has been such an honor and I will always be grateful and proud to have been your student.

To CINVESTAV, especially to the Automatic Control Department, for all the resources, knowledge, and facilities along my Ph. D., for being The Best research institution.

Last but not least, I am deeply grateful to CONACYT for the scholarship granted (CVU 745450) to support my graduate studies.

Y soporté... 

Resumen

En la actualidad, el desarrollo de robots bípedos es cada vez más relevante debido a que son los más adecuados para desplazarse en zonas transitadas por humanos. Esto ha generado un creciente interés en diseñar y controlar robots bípedos capaces de realizar tareas en entornos con restricciones de espacio y movilidad, lo que requiere precisión en la planificación y control de los movimientos del robot.

En esta tesis se propone el diseño de un robot bípedo con medidas antropométricas para la implementación de un controlador que cubra los requerimientos mencionados anteriormente. El robot cuenta con 8 grados de libertad simétricamente distribuidos en ambas piernas, lo que permite su movimiento en dos planos anatómicos (frontal y sagital). Los movimientos del robot se basan en el uso de actuadores lineales cuyo desplazamiento se transforma en movimientos angulares gracias a su disposición mecánica, lo que es una aportación novedosa de esta tesis y tiene mayor semejanza con la estructura anatómica humana.

Se propone un nuevo algoritmo de control basado en Funciones de Lyapunov tipo Barrera cuya estructura obedece a una función tangencial asimétrica y que utiliza el algoritmo de super-twisting como diferenciador robusto y exacto, lo que es una propuesta novedosa para este tipo de estructuras robóticas. Se eligió este tipo de controlador debido a las restricciones angulares inherentes al desplazamiento bípedo. Además, se presenta el análisis de estabilidad demostrando que se garantiza el seguimiento de trayectorias de referencia. Se evaluó el desempeño del controlador de forma numérica utilizando el software Matlab®en el ambiente Simulink®y posteriormente se evaluó en el robot suspendido.

La sintonización de las variables del controlador se realizó utilizando el esquema de Apren-

dizaje Reforzado Multiagente para garantizar el seguimiento de las trayectorias de referencia considerando restricciones en las articulaciones. El esquema multiagente se compone de dos partes: la primera implementa una estrategia de programación dinámica adaptable basada en la aproximación de la ecuación de Hamilton-Jacobi-Bellman y la segunda implementa una metodología repetitiva basada en el algoritmo de Gradiente de Política Determinista Profunda, similar a las cadenas de Markov, permitiendo ajustar tanto el seguimiento de la trayectoria como la trayectoria en sí misma, respectivamente. El uso de este esquema disminuyó el consumo energético en comparación con un controlador clásico por retroalimentación.

A partir de la propuesta de diseño presentada para el robot, se construyó la estructura mecánica en la que se realizaron las pruebas del controlador propuesto. La estructura, ensamblada en aluminio, incluyó instrumentación electrónica digital y de potencia. La instrumentación permitió el monitoreo y la conversión analógica-digital de las señales obtenidas de resistores variables insertos en los actuadores lineales, así como la generación de las señales de control basadas en PWM con voltaje de 0 a 3.3V. La segunda utilizó los pulsos como entradas para un puente H que garantizó el seguimiento de las trayectorias, asegurando al mismo tiempo, la protección de la sección lógica mediante optoacopladores.

Abstract

Currently, the development of bipedal robots becomes increasingly relevant due to their suitability for navigating environments originally intended for humans. This has led to a growing interest in the design and control of bipedal robots that can perform tasks in constrained spaces, requiring precise planning and control of robot movements.

This thesis proposes the design of an anthropomorphic bipedal robot as a platform for implementing a controller. The robot features 8 degrees of freedom symmetrically distributed in both legs, allowing movement in two anatomical planes (frontal and sagittal). The robot's movements rely on the use of linear actuators assembled on aluminum channels, with their linear displacement transformed into angular movements thanks to the mechanical arrangement in which they were placed. This mechanical arrangement is a contribution of the present thesis and has a greater resemblance to the human anatomical structure.

An algorithm based on Barrier Lyapunov Functions, whose structure follows an asymmetric tangent function, is presented for control. This controller was chosen due to the angular restrictions inherent in bipedal locomotion. The performance of the controller was numerically evaluated using Matlab software in the Simulink environment, and subsequently evaluated on the suspended robot.

The controller's variable tuning was performed using a Multi-Agent Reinforcement Learning scheme, which sought to ensure the tracking of reference trajectories while considering joint constraints. The multi-agent scheme considers two agents: the first implements an adaptable dynamic programming strategy based on approximating the Hamilton-Jacobi-Bellman equation, and the second implements a repetitive methodology based on the Deep Deterministic Policy Gradient algorithm, similar to Markov chains, allowing adjustment of both trajectory tracking and trajectory itself,

respectively.

From the proposed robot design, the mechanical structure was constructed, on which the proposed controller was tested. The aluminum structure included digital and power electronic instrumentation, where the former allowed for monitoring and subsequent analog-digital conversion of signals obtained from variable resistors inserted in the linear actuators, such as generating control signals based on PWM in voltage pulses from 0 to 3.3V, and the latter used the pulses as inputs for an H-bridge that ensured trajectory tracking, ensuring the protection of the logic section through optocouplers.

Contents

List of Figures

List of Tables

Acronyms

Abd-Add Abduction-Adduction.

AI Artificial Intelligence.

AOFC Adaptive output-feedback controller.

ANN Artificial Neural Network.

BF Barrier Function.

bLf Barrier Lyapunov Function.

BR Biped Robot.

BRD Biped Robotic Device.

BW Biped Walking.

CoM Center of Mass.

CoP Center of Pressure.

DDPG Deep Deterministic Policy Gradient.

DOF Degrees of Freedom.

DPG Deterministic Policy Gradient.

EMA Electromechanical Actuator.

FE Flexion-Extension.

GC Gait Cycle.

HJB Hamilton-Jacobi-Bellman.

HG Human Gait.

IC Integrated Circuit.

LA Linear Actuator.

LAd Limb Advancement.

MDP Markov Decision Process.

ML Machine Learning.

MARL Multi-agent Reinforcement Learning.

NN Neural Network.

PID Proportional–Integral-Derivative.

RED Robust Exact Differentiator.

RL Reinforcement Learning.

RL-PID Reinforcement Learning-based PID.

SLS Single Limb Support.

SL Supervised Learning.

STA Super-Twisting Algorithm.

UL Unsupervised Learning.

WA Weight Acceptance.

ZMP Zero Moment Point.

Chapter 1

Introduction

This chapter presents the general introduction, the motivation of the research associated with the proposed theoretical result on the development of the automatic controller for the biped robot, the main objectives, and the formal definitions of the main concepts upon which this work is based, that is, Biped Walking (BW), RL, and the Barrier Lyapunov Functions (bLfs). In addition, state of art and the description of the thesis structure are presented at the end of this chapter.

1.1 Fundamentals of automatic safe walking biped robots

The development of BRs implies gathering several approaches (such as mechanical design, electrical instrumentation, and automatizing, among others) in technological development due to the necessity of overcoming the many challenges its complex and hybrid nature implies. The humanoid robots are different due to the tasks they were designed to make, thus leading to different capabilities, including rescue, assisting, education, healthcare, etc. Many of them are capable of walking stably on even terrains and grass, yet in rough environments; not only walking but also climbing stairs [1, 2] or obstacles [3], jumping [3, 4] and walking on slopes [5, 6]. Aiming to encourage developers and researchers, there are many international competitions such DARPA [7–10], FIRA [10, 11], HuroCup [12] and the RoboCup Federation [13], showing up the current relevance these robotic devices have today.

Side research on the development of robotic walking devices refers to the evaluation of locomotion that in humans would be unfeasible to test [14]. Overcoming all these tasks leads to the development of an integrated robotic device that can achieve regular, stable, and robust displacement in almost all types of environments. There exist various stability criteria for bipedal walking that relays on the fact that the robot is moving slowly: the first one is the position of the Center of

Mass (CoM) and the Center of Pressure (CoP) in the human-like robot, considered stable if any of both is within an area called Support Polygon [15]. The following criterion is based on the Zero Moment Point (ZMP) concept, which refers to a point on the ground where all the forces that interact in the robotic motion are zero. Hence, if the vertical projection of this point on the supporting surface stays within the support polygon area results in a stable gait. A third criterion is based on the concept of passive dynamic walking, called Dynamic Walk. All these concepts involve a completely actuated-legged robot, and the control problem is reduced to a joint tracking control that fulfills all the previously mentioned criteria [16].

Another main property of automatic robotic walking is the dependency on appropriate coordination of all the joints at every moment since one unsuitable position of any articulation could lead to the falling of the structure and, consequently, structural damage. To avoid these undesirable situations, it is necessary to consider the term *safety* in robotic walking. This term refers to avoiding incorrect robot configuration during the gait cycle [17] and, in this case, shunning joints to exceed certain predefined working regions. Handling the safety requirement implies introducing stability conditions in the control action, which can be solved using the Barrier Function (BF) concept, which can contribute to the design of the controller gains. A BF is a continuous function that meets this requirement by approaching its value to infinity when the state vector approaches the boundary of a predefined (usually considered safe) set in the task or the joint spaces. The inclusion of BF prevents the escape of the trajectories from the safeness condition enforced by the safe set. This method offers a strategy to ensure the safe motion of bipedal devices, including the design of adaptive controllers based on the application of BLfs [18–21]. This approach guarantees that the states remain inside a convex hull previously defined in the controller implementation [22], which can be related to the safe set to ensure a stable walking process.

The application of BLfs must include a consequent control design that could guarantee the joints' operation within the predefined safe set. If this control action considers a robust strategy, the proper motion for all the joints could be highly conservative, leading to unfeasible biped robot operation. On the other hand, if the controller has an adaptive structure with state-dependent gains, then the controller could offer a less conservative outcome but require more complex online implementation due to the algorithms needed to adjust the control gains. This problem can be solved considering the available technology on digital boards that can instrument complex control configurations.

1.2 Biped walking

Biped locomotion is considered one of the most significant evolution adaptations in nature [23]. It is a highly specialized and non-common way of locomotion [24] that in humans is one of their main characteristics. Two phases characterize the modern human walking cycle: the stance phase, when one leg works like a pivot on the ground and supports the person's weight, and the swing phase, when the leg is off the ground. Anatomically, most of the human gait is made along the sagittal plane. However, it is a three-dimensional movement. Since walking is a series of cyclical events, its fundamental unit is a Gait Cycle (GC). The cycle starts at the point of initial contact of one limb (used as reference) [25–27] ending with the following contact to the ground of the same foot. A healthy GC starts with the heel contact or heel strike[1]. Traditionally a GC or stride is subdivided from 0% to 100%, being the initial contact 0% and the next contact of the same extremity (100%). For description purposes, the whole GC is divided, as mentioned, into two main phases, which in turn are divided into events that enable the limb to achieve three basic tasks: Weight Acceptance (WA), Single Limb Support (SLS) and Limb Advancement (LAd). Within a stride, the body goes through two periods of double-limb support and two periods of single-limb support. There exist eight significant events that are used to distinguish seven periods for the GC (Figure 1.1)

WA starts the stance phase and includes the first initial periods (initial contact and loading response). The stance continues during SLS with the following two periods of the GC: mid stance and terminal stance. Finally, LAd begins with the pre-swing period through the end of the swing phase. A thorough overview of each task is given next.

1.2.1 Task: Weight acceptance

According to [27], this is the most challenging task over the three that constitute a regular GC. This is because, in this period, a sudden transference of the weight from one to the other limb that just finished the swing forward occurs, thus having an unstable alignment. This task demands three main functional patterns: shock absorption, initial limb stability, and preservation of progression.

Initial contact

This stage includes the moment when the foot touches the surface. Depending on the joints' position, the limb's loading response changes its pattern.

[1] As not all people have this capability, it is preferred the generic term *initial contact*

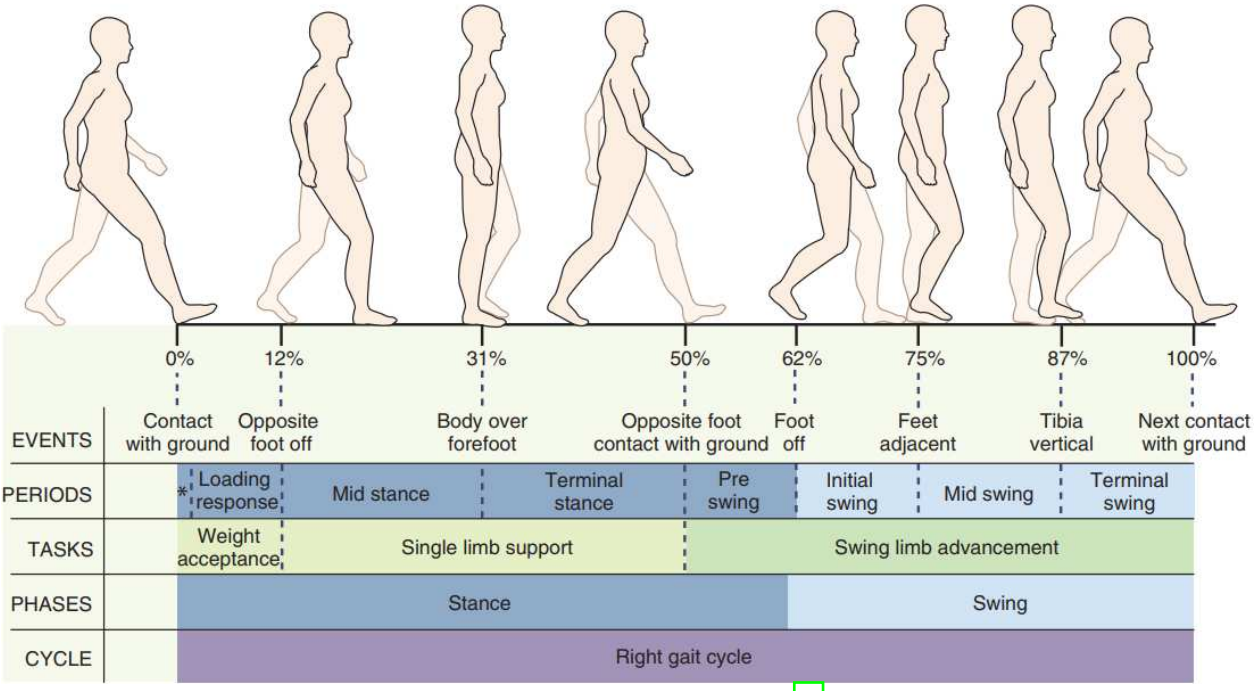


Figure 1.1: Gait cycle divisions [28]

.

The articulation disposition of the reference limb includes flexion of the hip, an extender knee, and the ankle is dorsiflexed to neutral. Meanwhile, the other limb is at the end of the terminal stance (Figure 1.2).

Interval: 0 to 2% of GC

Objective: Placing the limb to start the stance with a heel rocker.

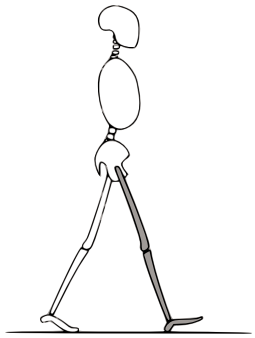


Figure 1.2: Initial contact. Adapted from [27]

Loading response

It is the initial double stance period. It begins with the first ground contact continuing to the lifting of the other foot until the swing phase starts.

The body weight is transferred onto the forward limb. The heel functions as a pivot; the knee is flexed to absorb the impact shock, and the ankle plantarflexion limits the heel rocker by forefoot contact with the floor. The opposite limb is in the pre-swing period. This is depicted in Figure 1.3.

Interval: 0 to 10% of GC

Objectives:

- Shock
- Weight-bearing stability
- Progression maintenance

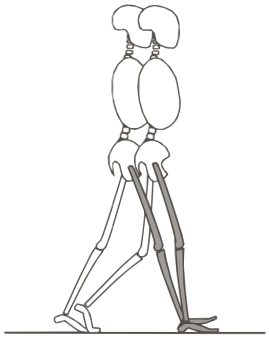


Figure 1.3: Loading response. Adapted from [27]

1.2.2 Task: Single limb support

The second task interval of the GC for the reference limb begins when the non-reference foot is lifted for swing, finishing when the opposite foot contacts the ground again. While the interval last, the whole support of the body weight, in both sagittal and coronal planes, relies exclusively on one limb ensuring that the progression continues. Two periods are involved: mid stance and terminal stance, differentiated by the progression mechanisms in the bipedal walking.

Mid stance

This stage begins as the non-reference foot is elevated and keep going up to the body weight has been aligned over the forefoot by the movement.

Referring to the disposition of the reference limb during this period, the foot is fixed to the ground while the torso moves forward to it by ankle dorsiflexion, and the knee and hip are extended. The opposite limb is advancing in its mid-swing period (Figure 1.4).

Interval: 10 to 30% of GC

Objectives:

- Passing reference foot
- Stability of limb and trunk

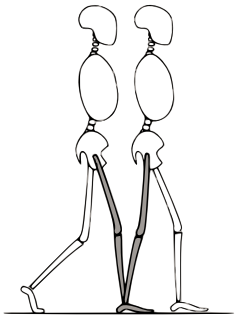


Figure 1.4: Mid stance. Adapted from [27]

Terminal stance

The terminal period finalizes the single limb support (Figure 1.5). The heel rising marks the period's beginning, lasting until the other foot touches the ground. During this period, the weight moves ahead of the forefoot.

This is the second half of the single limb support. At this time, the heel rises, and the reference limb advances over the forefoot pivot. Biomechanically speaking, the knee spreads and then starts to bend barely. The increased hip extension allows putting the limb in a more trailing position. The non-reference limb is in the terminal swing.

Interval: 30 to 50% of GC

Objective: Progression of the body beyond the supporting foot

1.2.3 Task: Limb advancement

To fulfill the task of advancing the limb, prelim posturing starts at stance. The limb swings through three postures while it lifts itself, advances, and prepares for the next stance phase. It is composed of four periods: pres-swing[2], initial swing, mid swing, and terminal swing.

[2] *Weight release* and *weight transfer* are other names given to this period

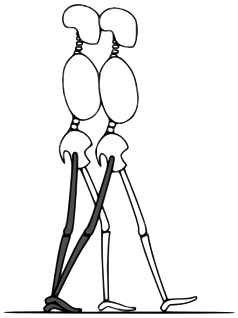


Figure 1.5: Terminal stance. Adapted from [27]

Pre-swing

This is the last double stance interval in the GC; it begins with the initial contact and ends with the toe-off of the non-reference limb. Though the sudden weight relocation immediately discharges the limb, it does not actively contribute to the event. Instead, the unloaded limb gets prepared for the swing.

Referring to the posture, the floor contact by the non-reference limb has started terminal double support being in loading response as the reference limb increased ankle plantar flexion, widened knee flexion, and loss of hip extension (Figure 1.6).

Interval: 50 to 60% of GC

Objective: Put the limb for swing

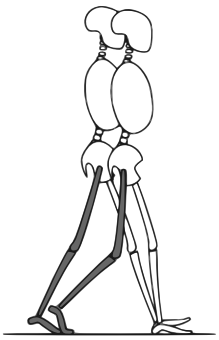


Figure 1.6: Pre-swing. Adapted from [27]

Initial swing

This period is approximately a third part of the swing phase, beginning with lifting the foot from the ground and ending when the swinging limb is contrary to the stance foot.

Regarding posture, the foot is elevated and the limb fared by hip flexion; the non-reference limb is starting mid-stance (Figure 1.7).

Interval: 60 to 73% of GC

Objectives:

- Foot clearance of the floor.
- Advancement of the limb.

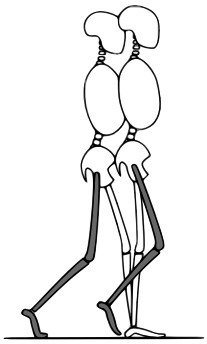


Figure 1.7: Initial swing. Adapted from [27]

Mid swing

This second part of the swing phase begins while the swinging limb is opposite the stance limb, as shown in Figure 1.8, and finishes when the tibia is vertical and the swinging limb is forward, meaning that the hip and knee are in equal flexion postures. The advancement of the reference limb, anterior to the body weight line, is because of wider hip flexion. The knee is passively extended, responding to gravity as the ankle keeps dorsiflexing to neutral. The non-reference limb is the last part of the mid stance (Figure 1.8).

Interval: 73 to 87% of GC

Objectives:

- Limb progression.
- Foot lifting from the floor.

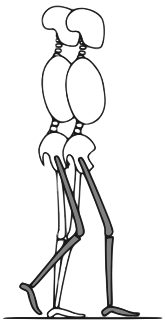


Figure 1.8: Mid swing. Adapted from [27]

Terminal swing

This stage starts with a vertical tibia, ending when the foot touches the floor (see Figure 1.9). It is completed once the shank moves ahead of the thigh. Describing the posture, the limb advancement is completed by the knee extension; the hip keeps its flexion, and the neutral dorsiflexion of the ankle remains. The non-reference limb is in a terminal stance.

Interval: 87 to 100% of GC

Objectives:

- Complete limb progression.
- Prepare the limb for stance.

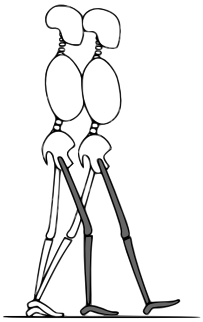


Figure 1.9: Terminal swing. Adapted from [27]

1.2.4 Joint trajectories

Based on the description of the stages and tasks carried out during walking, each joint can be kinematically described on different anatomical planes, allowing the characterization of movement amplitudes at a given moment in the cycle. In this study, movements on the sagittal plane of the hip, knee, and ankle, as well as on the frontal plane of the hip, were considered.

Sagittal plane

The displacements of the articulations in the sagittal plane are illustrated in the Figure 1.10. Each movement is described bellow.

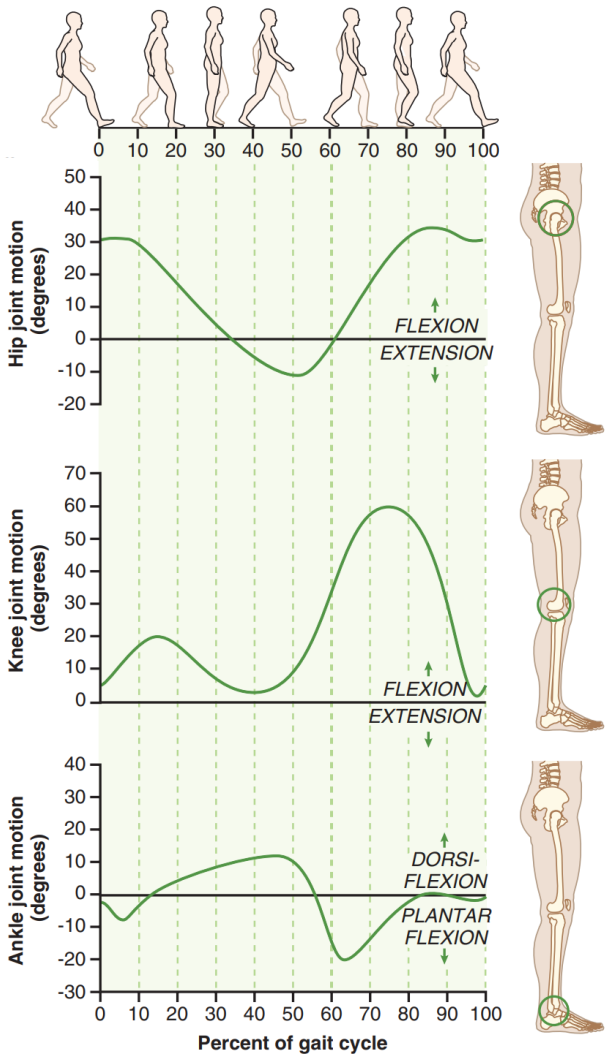


Figure 1.10: Sagittal plane angular rotation. Adapted from [28]

Hip

At a typical walking speed, the hip joint is flexed around 30 degrees at the moment of heel contact, as illustrated in Figure 1.10. As the body moves forward over the fixed foot, the hip extends, achieving maximum extension of approximately 10 degrees just before toe off. The flexion of the hip joint is initiated during push off, and the hip joint reaches about 0 degrees of flexion at toe off, which corresponds to 60% of the gait cycle.

In the swing phase, the hip joint further flexes to move the lower extremity forward for the next foot placement. Maximum flexion, which is slightly over 30 degrees, is achieved just before heel contact. It should be noted that at heel contact, the hip joint has already begun to extend in preparation for weight acceptance. Generally, normal walking requires approximately 30 degrees of flexion and 10

degrees of extension at the hip joint, measured from the anatomic neutral position. As with all lower extremity joints, the extent of hip joint movement is proportional to walking speed.

Knee

The knee undergoes a flexion of around 5 degrees upon initial heel contact, followed by an additional 10 to 15 degrees during the first 15% of the gait cycle. This slight knee flexion serves to absorb shock and distribute weight as the body load is transferred to the lower extremity. Subsequently, the knee nearly fully extends until the heel-off phase (occurring between 30% and 40% of the gait cycle), when the knee begins to flex again, reaching approximately 35 degrees of flexion at toe-off (60% of the gait cycle). At the beginning of mid-swing (75% of the gait cycle), the knee experiences a maximum flexion of approximately 60 degrees. During early swing, knee flexion shortens the length of the lower extremity, facilitating toe clearance. In mid and late swing, the knee extends almost to full extension before slightly flexing again in preparation for heel contact. Thus, the normal function of the knee during level walking necessitates a range of motion between almost full extension and approximately 60 degrees of flexion.

Ankle

The ankle undergoes a slightly plantar flexed position (between 0 and 5 degrees) at heel contact. Following heel contact (within the first 8% of the gait cycle), the ankle dorsiflexors eccentrically control a movement of plantar flexion to position the foot flat on the ground. During stance, the ankle dorsiflexes up to 10 degrees as the tibia moves forward over the foot, which remains in contact with the ground (from 8% to 45% of the gait cycle). After heel-off (between 30% and 40% of the gait cycle), the ankle plantar flexes, reaching a maximum of 15 to 20 degrees of plantar flexion just after toe-off. During the swing phase, the ankle dorsiflexes to a neutral position to allow for toe clearance. Ambulation at an average speed requires approximately 10 degrees of dorsiflexion and 20 degrees of plantar flexion.

Frontal plane

The displacements of the hip in the frontal plane is illustrated in the Figure 1.11. The movement is described bellow.

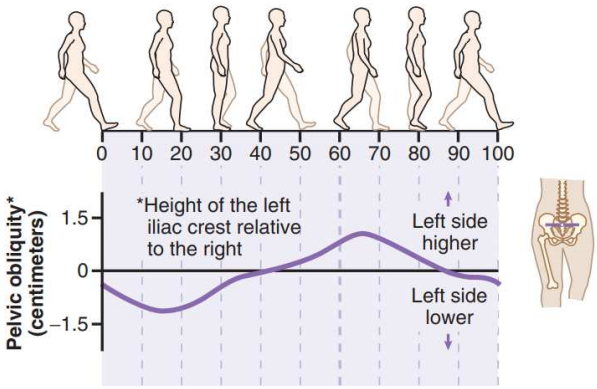


Figure 1.11: Frontal plane angular rotation. Adapted from [28]

Hip

In the stance phase of walking, the pelvis-on-femoral kinematics are the primary contributor to frontal plane motion, while a smaller and less consistent amount of frontal plane motion is attributed to femoral-on-pelvic kinematics. This motion can be observed through medial-lateral movement of the knee. During the swing phase, the pelvis and the advancing femur work together to return the hip to its neutral frontal plane position.

1.3 Automatic learning in biped walking

Human walking abilities inspire biped robots. Hence, it is common that biped robots are expected to develop gait cycles imitating the regular human gait cycle. According to this operative scenario, the biped robot motion can be handled using the effective integration of mechanical designs, electrical instrumentation, and reliable automatic controllers. This strategy allows many biped robots to work accordingly to predefined sequences of expected motions. Nonetheless, this strategy limits the capacity of the biped robot to perform its tasks in non-previously revisited scenarios. This is contrary to what human beings do to *learn* how to move their bodies with the objective of complete walking processes under diverse scenarios, including irregular terrains, discontinuous surfaces, and many others.

The progressive knowledge humans acquire to complete complex walking processes is a consequence of their learning capabilities. Human learning is based on neural motor control, proprioception, and biomechanical regulation of muscles and bones. Moreover, the walking process's progression is mostly based on an ongoing trial and error recursive procedure performed by humans at early ages.

Given the *method* used by humans during the walking learning process, it seems reasonable to use it as an inspiration to provide more independence to biped robots with the aim of improving how they perform the gait cycle under different conditions. Such a complex task can be solved using the artificial learning methodologies that are now popular in the machine learning discipline.

Due to new computing technologies, machine learning today is not like it was in the past. It was born out of pattern recognition and the theory that computers can learn without being programmed to do specific tasks. Researchers interested in artificial intelligence wanted to know if computers could learn from data just as in the human case when they try to learn walking. The iterative aspect of machine learning is important because as models are exposed to new data, it can adapt independently, then they learn from previous calculations to produce reliable and repeatable decisions and results.

Machine learning is based on the concept of automatic learning that considers the use of available information to correct a system performance considering input-output relationships (supervised), characteristics of input data (unsupervised) and the environment feedback consequence of system action (reinforcement), among others. The application of machine learning for enforcing the bioinspired action by biped robots during the walking process is highly related to the characteristics of reinforcement learning.

The design of better controllers for regulating biped robot motion that relies on Artificial Intelligence (AI) techniques has been increased in the past years, developing a whole agenda for control theory [29], giving more autonomy to many types of dynamic systems, and putting aside the time-consuming tuning for algorithms that implicitly look for the optimal values. In the vast field of study of AI, Machine Learning (ML) is a sub-field of the computer sciences that deals with "*the question of how to develop software agents that improve automatically with experience*"[30].

As we have claimed there exists three main paradigms of ML [31]:

- Supervised Learning (SL)
- Unsupervised Learning (UL)
- Reinforcement Learning

SL learns from a labeled training set of examples that are given by an external supervisor. Meanwhile, UL looks for a concealed structure in data without labels [32]. Finally, RL maps situations to actions by maximizing a reward signal.

RL possesses unique features that distinguish it from others. First of all, the trade-off between exploration and exploitation, [31–33], means that a learning agent must exploit its experiences in order to get a reward but, on the same time, exploring to *decide* better actions forward. However, neither exploration nor exploitation could be left behind, and the algorithm must seek a wide variety of actions to choose the best. Secondly, RL considers the problem as a whole, taking into account

the fact that the agent is interacting with its environment. Finally, it uses the information acquired from the training to evaluate the actions made, rather than learn from given certain actions.

1.4 Reinforcement learning

All so-called living superior organisms in nature learn by exerting repetitive actions of their physiological activities (including functional movements) and evaluating the environmental response by analyzing outcomes and internalizing the obtained events (reward stimulus or looseness inhibition). Such recurrent behavior enforces the continuous adjustment of the organism's actions and decisions to get the best attainable reward or reduce the looseness within a given period.

In the evolutionary biological theory and the learning sciences, getting knowledge based on developing complex associations by chaining actions to positive or negative rewards is called trial-and-error learning [32]. Most studies in the psychology of animal learning have proven that the so-called superior living organisms mainly use such learning methods. Essentially, this method considers that repeated actions, followed by the association of positive or negative rewards, reflect the construction of practical knowledge [34].

The foundations of trial-and-error learning were originally proposed by Edward Thorndike [35], who called these outcomes *Law of Effect* due to the relationship to reinforcing acts and how it changes the affinity to make certain actions. The so-called *Law of Effect* is based on the consolidation of two critical aspects: selection and associative. On one hand, the selection phase refers to the fact that the election of diverse alternatives is conducted by the preliminary (based on the own experience) association between them and their consequences. On the other hand, the associative phase relates to the establishment of a suitable association between elected alternatives with particular situations. The combination of both elements is the essence of the Law of Effect and trial-and-error learning. The idea of learning based on the continued interaction between the environment and the organism (an agent related to its environment modifying its actions based on stimuli received in response to them) settles down the base concept for developing *Reinforcement Learning* algorithms.

1.4.1 Basic concepts of reinforcement learning theory

Since the main idea of RL is to maximize (or minimize) a reinforcement signal by means of remembering positive control decisions, i.e., a cumulative reward in the long run. RL is connected from a theoretical point of view with both adaptive control and optimal control methods.

Elements of RL

Reinforcement learning is a way to map situations to actions [32, 36] via a *policy* $\pi : s \rightarrow a$ such that a numerical value called *reward* is maximized (or minimized). This result is obtained assuming the agent is functioning despite the environmental recognizance affronted. In addition to the agent and the environment, for all reinforcement learning systems, it is possible to identify four arch components: a policy, a reward signal, a value function, and an environmental model.

First of all, consider the establishment of the interaction that exists among all the mentioned components: the agent is interacting with its environment through observation and action which can be seen as a *decision-maker*. For every step of the so-called interaction, the agent obtains as input from the current state, $s_t \in \mathcal{S}$, some information from the environment, consequently, the agent chooses an action at time t, $a_t \in \mathcal{A}(s_t)$, producing the associated output. Then, this action changes the state of the environment (s_{t+1}), and the worth of this change is sent to the agent using a reinforcement signal, or reward, $r_{t+1} \in \mathbb{R}$ [33, 37] (Figure 1.12).

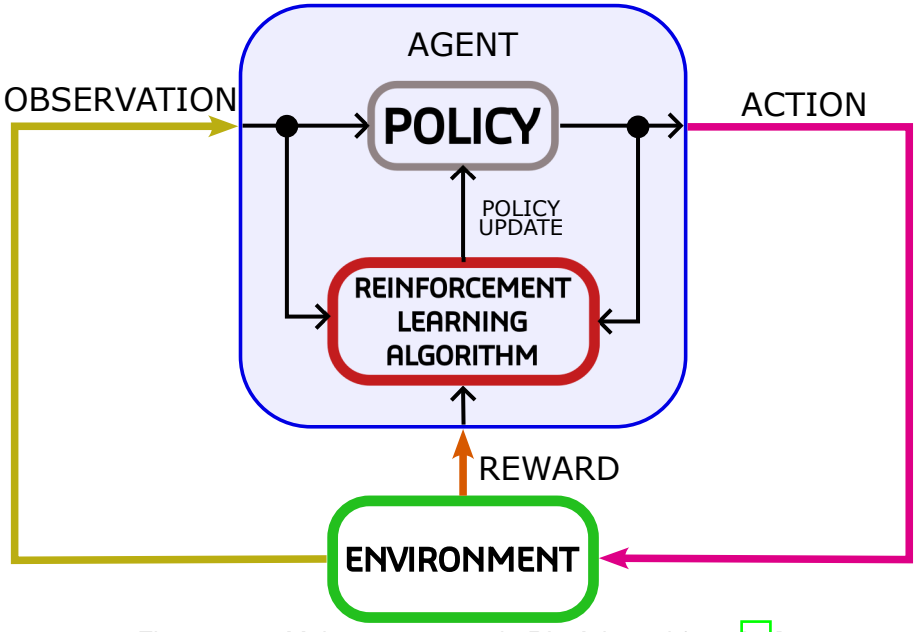


Figure 1.12: Main components in RL. Adapted from [38].

The strategy followed by the agent is the *policy* π_t. This policy maps the states to the option of choosing each possible action. Hence, finding the best policy can be considered as an optimization problem.

One of the key methodologies to derive RL approaches is the well-known Markov Decision Process. This technique has been exploited for many years with the potential inconvenience of pre-defining state transition laws.

Markov Decision Process

Formally, one may describe the RL as a Markov Decision Process (MDP), which provides a mathematical framework for studying optimization problems using dynamic programming. A MDP can represented as a 6-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \rho_0, \gamma)$, with $\mathcal{S}$ being the set of all the states between the environment and the agent. Also, the term $\mathcal{A}$ defines the set of all the actions produced by the agent. The symbol $\mathcal{R}$ describes the accumulation of immediate reward function $r(s_t, a_t)$ in the current state s_t and action a_t in time-step t. The scalar $mathcalP$ denotes the transition probability $p(s_{t+1}|s_t, a_t)$ which maps the pair state-action at time-step t to the next one $t+1$. The distribution of the initial states is depicted as ρ_0, and γ is a discount factor that satisfies $\gamma \in [0, 1]$. If gamma is closer to 0, the agent is biased towards immediate reward maximization, leading to short-term oriented decision-making. In other words, the agent focuses on achieving quick gains rather than considering the long-term consequences of its actions. This approach may be useful in scenarios where short-term benefits are more critical than long-term gains. Meanwhile, agents with a discount factor closer to 1 seeks an optimal policy of far-sighted rewards [39], meaning they prioritize the accumulation of rewards over an extended period of time. This strategy can be advantageous in tasks that require strategic planning and decision-making, as it allows the agent to consider the potential long-term consequences of its actions. Then, finding the best policy can be managed as an optimal mathematical problem.

We can say that the action-value function $Q(s, a)$, and the value function $V(s)$ are defined as the expected discounted rewards given a current policy π in the states s_t and action a_t, respectively.

$$Q_\pi(s, a) = \mathbb{E}_{\pi,p}[G_t | s_t = s, a_t = a] \tag{1.1}$$

$$V_\pi(s) = \mathbb{E}_{\pi,p}[G_t | s_t = s] \tag{1.2}$$

Where $G = \sum_{t=0}^{\infty} \gamma^t r_t$ is a total discounted reward from time step t, which means that the expectation over all actions according to the policy π is considered. $Q_\pi(s_t, a_t)$ is the *state action* value function, also known as the quality function. It is the expected return starting from state s_t, taking action a_t, then following policy π. It is focused on the particular action at the particular state. The expected value is denoted by $\mathbb{E}_{\pi,p}$, where π is the current policy and p is the transition probability distribution that describes the likelihood of the agent transitioning from state s_t to state s_{t+1} after taking action a_t.

Thus, the optimal policy can be obtained by maximizing (1.2) and (1.1) via the optimal Bellman equations:

$$V_{\pi^*}(s_t) = \max_{a_t} \mathbb{E}_{s_{t+1},p}[r(s_t, a_t) + \gamma V_{\pi^*}(s_t)] \tag{1.3}$$

$$Q_{\pi^*}(s_t, a_t) = \mathbb{E}_{s_{t+1},p}\left[r(s_t, a_t) + \gamma \max_{a_{t+1}} Q_{\pi^*}(s_t, a_t)\right] \tag{1.4}$$

The goal of RL is to *learn* the optimal policy that maximize [3] the expected return from the start distribution. This goal is formalized in (1.5), where J represents the expected return:

$$J = \mathbb{E}ri, s_i \backsim E, a_i \backsim \pi[R_1] \tag{1.5}$$

The expected return $\mathbb{E}$ is calculated as the sum of rewards obtained by executing the policy π from the initial state s_i to the final state s_f, denoted as R_1. The expected value is computed by averaging the return over all possible state and action sequences (r_i, s_i, a_i) drawn from the probability distribution defined by the environment E and the policy π. Here, $\backsim$ denotes random sampling. The aim is to identify the policy that maximizes J, which corresponds to maximizing the expected cumulative reward over all feasible trajectories.

With the aim of providing a link between RL and Control Theory, it is necessary to present a mathematically idealized form of RL called MDP which allows that many dynamical decision processes can be modeled as MDP. In this framework, it is considered a delayed reward and the necessity to balance immediate and delayed reward [32]. The part that learns and makes decisions is the *agent* which interacts with the *environment* [4].

The interaction between the agent and the environment consists of the election of actions by the agent and the response of the environment to these actions. This result is performed by changing and giving feedback to the agent through rewards, which in turn, tries to maximize them over time by selecting "the best" attainable actions [5]. The interaction of both elements is depicted in Figure 1.13.

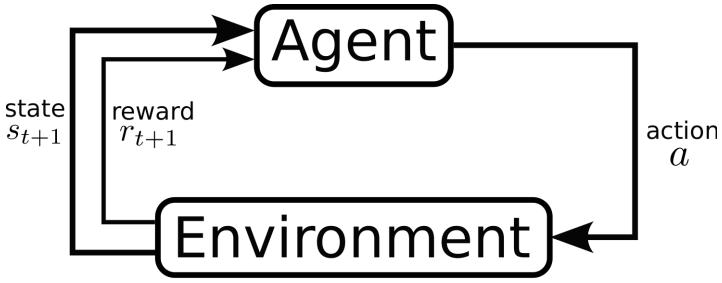


Figure 1.13: Interaction in a MDP

[3] or minimize, depending on the problem to be solved

[4] The environment can be seen as everything that is not the agent.

[5] Analogously, the *agent* can be seen as a *controller*, the *environment* as the *plant*, and the *action* is equivalent to the *control signal*

Formally, we can consider the MDP as a 4-tupla $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$, where $\mathcal{S}$ is a set of states, and $\mathcal{A}$ a set of admissible actions/controls, [40, 41]. There exists a probability of transitions $\mathcal{P} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow [0, 1]$ which describes from each state $s \in \mathcal{S}$ and each action $a \in \mathcal{A}$, the conditional probability

$$\mathcal{P}^{a}_{s,s_{t+1}} = Pr\{s_{t+1}|s, a\}$$

of transitioning from the current state s to the next one $s_{t+1} \in \mathcal{S}$ given an action a. The costs "paid" after this states transition is described through a cost function $\mathcal{R} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow R$ which is the expected immediate cost $\mathcal{R}^{a}_{s,s_{t+1}}$ after the change of states.

The main property in the MDP is the fact that these probabilities of transition relay only in the current state s without taking into account any of the previous actions. The principal issue for the agent in a MDP is to find a map $\pi : s \times a \rightarrow [0, 1]$ such that for each state s and action a gives the conditional probability $\pi(s, a) = Pr\{a|s\}$ of taking the action a since the MDP is currently in the state s. This mapping is called *policy*.

Every MDP has at least one policy $\pi^* : \mathrm{S} \rightarrow \mathrm{A}$ that does not change as a time function, which action is chosen always in $\mathcal{S}$ and undominated, in other words, a stationary, deterministic optimal policy [39]. This technique has many applications in diverse areas of science and technology, but fails when it is proposed as a potential solution to the self-learned gait cycle of humanoid robots due to the complex sequence of stages during the gait cycle. Hence, a more deterministic formulation is required. This is one of the main justification of the control algorithm developed in this study, that takes advantage of the deep-learning methodologies.

1.4.2 Deep reinforcement learning

Deep Reinforcement Learning (DRL) is the study of RL using neural networks as function approximation, existing a wide variety of choices of what approximate such as policies, environment models, action-value functions, value functions, etc. [42]. There are also many algorithms to reach this goal, each of them with diverse properties addressing different problems. A non-exhaustive classification map is shown in Figure 1.14.

The two main branches in DRL are the distinction if the agent needs to learn or has access to an environmental model, i.e., a function that can predict rewards as state transitions. While this brings many advantages to simple systems, for complex environments this yields in a real challenge, due to the bias in the model can be exploited by the agent. Such fact leads to good performance with respect to the learned model, but possibly terrible in a real environment.

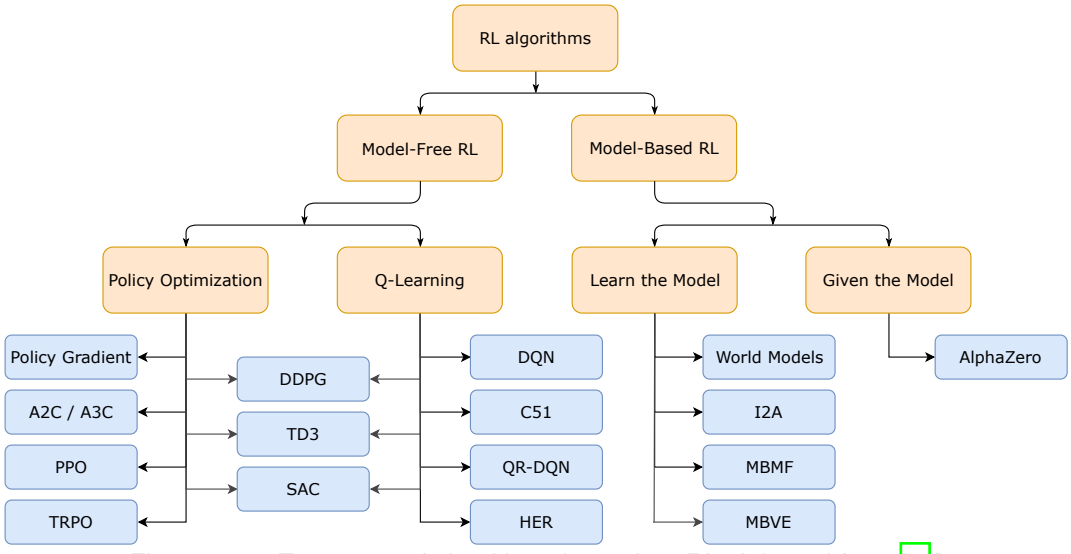


Figure 1.14: Taxonomy of algorithms in modern RL. Adapted from [43].

When a model-free RL algorithm is selected, we can approach to represent and train agents from two main perspectives:

- Policy Optimization. In this branch, a policy is represented explicitly, optimizing its parameters in an on-policy way, i.e., each update uses the collected data acting in accordance with the most recent version of the policy. Two examples of policy optimization methods are:
  - A2C / A3C (Asynchronous Advantage Actor-Critic [44])
  - PPO (Proximal Policy Optimization [45])
- Q-Learning. These methods learn an approximator for the action-value function, using typically an objective function based on Bellman equation, optimization realized off-policy, this means that each update can use the data obtained at any time during training, regardless of how the agent chose to explore the environment when obtaining the data. Some examples are:
  - DQN (Deep Q-Networks [46])
  - C51 (Categorical 51-Atom DQN [47])

The two described branches should not be seen as extreme opposites. There exist algorithms that reach both and can trade off the advantages and disadvantages of both. One of them is the DDPG algorithm.

Deep Deterministic Policy Gradient (DDPG) is an algorithm that learns a Q-function and a policy at the same time [48], using off-policy data and the Bellman equation to learn the Q-function. At the same time, uses the Q-function to learn the policy. This algorithm is closely connected to Q-learning, meaning that while you know the optimal action-value function $Q_*(s_t, a_t)$. Then for any state, we

can find the optimal action $a^*(s)$ solving

$$a^*(s) = \arg\max_a Q_*(s_t, a_t)$$

One advantage of this algorithm is that is adapted for environments with continuous action spaces, modeling the policy as a deterministic policy instead of a stochastic policy [49].

Given the benefits that deep-learning may have, then it is feasible to work with it to improve control actions using these tools. Nevertheless, there is still the necessity of including the natural motion restrictions to the RL approaches. Such solution can be gotten using the barrier functions that are now significant tools in the design of state restricted control designs for systems as the biped system considered in this thesis.

1.5 Barrier functions

As it has been described in subsection 1.2, during the gait process, it is necessary to develop every articular displacement to a specific range at each articulation in order to assure a *safe* motion for the device. Thus, the design of controllers must consider both stability while observing the constraints. A widely used technique to achieve the former objective is by implementing Lyapunov's method which provides a way to determine stability without explicit cognizance of the system [50]. This method has been used as one of the foremost strategies among researchers. Such fact relies on their intuition and expertise to construct stable Lyapunov functions to reach stable control designs. However, this can not guarantee that the designed Lyapunov function suits the needs of the system, especially when there are output constraints [51].

The constraints have been considered in many approaches, including model predictive control [52, 53], predictive safety filters [54], adaptive fuzzy tracking control [55], particle swarm optimization [56], among others.

In [57] it is proposed the use of bLfs for nonlinear systems with an output constraint by designing the control so that the time derivative of the bLf is negative semidefinite, keeping the Lyapunov function of the barrier bounded in the closed loop, ensuring that the constraints are not transgressed. This is a significant contribution to the class of solutions generated by the application of RL using deep-learning tools with the consideration of state restrictions.

1.6 General problem statement

The proper motion of robots which are designed to move around in biped-oriented environments is still open, due to the complexity of such task in terms of mechanical design, electrical and electronic instrumentation and automatic control designs. The coordination of movements along all the robotic system is not just desirable, but necessary.

This property implies as a fundamental tool the application of appropriate control schemes. However, such strategy may need the use of techniques such as RL to find the optimal set of gains to reach proper movements for the articulations of biped robots. Hence, this thesis considers as the key problem to solve the design of an automatic controller with state dependent gains that can be adjusted online considering the presence of state restrictions, and the recursive modifications, based on RL techniques with the aim of optimizing a given operative functional.

The proposed strategy considers a hybrid strategy that brings into the play the adjustment of the gains in a double rate performance, with the aim of optimizing a convex functional with respect to the tracking error for all the joints of the legs in the biped robot.

1.7 Objectives

1.7.1 Main

Mobilize a regulated biped robot following a walking pattern, based on the reference trajectories following a human gait, using an adaptive controller with state dependent gains and a recursive strategy based on RL that considers, from its structure, the angular restrictions of the mechanical configuration.

1.7.2 Specifics

- Design a controller structure which considers constrained states.
- Verify that the constraints in the mechanical structure are met by building an electrical system that will digitally control the system.
- Develop a two-legged robotic structure using modular aluminum channels and linear actuators as testing platform for the controller.
- Mathematically describe the mechanical structure of the robot using linear actuators.
- To propose a Multi-agent Reinforcement Learning (MARL) scheme to adapt the reference trajectories to the biped system while tuning the gains of a controller.

- Simulate the biped robotic device using the MARL scheme.

1.8 State of art

The development of humanoid robotic devices has motivated significant research efforts, at least since the 70s of the past century [58]. The inspiration for the design comes from the anthropomorphic distribution itself, promoting the idea that the robot concept is equivalent to the humanoid representation. In general, humanoid robotic devices are classified according to two common categories: appearance and behavior [59]. Since biped robots could be classified in both terms, it is evident that they are one of the most exciting examples of modern technology that aggregate disciplines, such as mechanical design, advanced electrical and electronic instrumentation methods, and automatic control designs.

The current versions of bipeds robots are characterized by efficient, light, and human-like mechanical structures. Their instrumentation approaches have aggregated diverse variants of actuators and sensors, that have allowed the realization of complex controlled movements, including regular gait cycle, sitting down and standing up, climbing stairs, or walking in unstructured environments.

The technical literature published over the last thirty years is full of exceptional biped robots. Most of these devices are fully controlled to exert a well-defined sequence of movements. When the biped robot is aimed to make autonomous activities independently of predefined tasks, there is a necessity of introducing a degree of short-time adaptation (based on the automatic control design) depending on the current robot-environment interaction conditions and long-term learning depending on the *experience* acquired by the robotic device using machine learning tools in an iterative way. Most of the existing machine learning options have taken their inspiration from living organisms.

There exist different systems which nature makes it very difficult to figure beforehand an accurate mathematical model. Since many control techniques are oriented to solve the reaching of a given level or following a previously defined trajectory for a structure, these problems assume previous knowledge of both the goal and the system. Nevertheless, the construction of the reference trajectories is also one of the main problems for navigation tasks in a variety of mobile robots. As an example of a workable agent, that could be described as an entity that perceives and act in consequence [60], biped robots are structures that can be designed to walk independently in certain terrains. Designing a controller for this kind of robot, which must be able to do bipedal displacement while maintaining equilibrium, is a major task due to the complexity of the system.

The use of RL in complex systems, particularly in biped robots, has been exploited in past years because the algorithm does not receive direct instruction about which joint needs to be modified [61].

1.9 Outline

The general outline of the chapters that make up this thesis, which are organized according to the main tasks carried out for the development of the project is presented here. Each chapter will present the problem statement to be addressed, as well as the detailed process undertaken to solve it. Additionally, the conclusions obtained for each chapter will be included. With this approach, we aim to offer a clear and coherent structure that facilitates the understanding and follow-up of the results and conclusions obtained throughout the research.

- Chapter 2 presents the mathematical model obtained from the structure that incorporates linear actuator. This controller is designed to take into account the joint constraints and actuator dynamics in order to achieve a more accurate approximation to the physical reality of the robot developed in this work.
- Chapter 3 presents a detailed proposal for a Lyapunov barrier function controller that considers the constraints inherent in this type of robot. The chapter also presents the results of implementing the proposed controller in a simplified simulated version of the robot, and compares them against those obtained from a classical controller.
- Chapter 4 describes the use of a multi-agent reinforcement learning paradigm to tune the controller gains and adjust the reference trajectories to match the robot's structure. The chapter also presents the results obtained from the reinforcement learning implementation and compares them against those obtained from a classical controller.
- Chapter 5 provides a detailed description of the physical structure proposed in this work. The chapter presents the requirements considered for the dimensions of the structure, as well as the type and characteristics of the components used in the robot. The physical structure is described in three main elements: mechanical structure, power electronics, and digital electronics.
- The general conclusions and final remarks of this thesis are presented in Chapter 6. Additionally, future research directions are discussed.

Chapter 2

Mathematical Modeling and Control of a Biped Robot driven by Linear Actuators

The aim of this chapter is to realize the mathematical model of the biped robot structure taking into account that each articulation is mobilized by a Linear Actuator (LA). Each extremity of the Biped Robotic Device (BRD) has three articulations with a LA enforcing the controlled motion for each articulation. A control problem that considers the task of tracking reference trajectories that define a regular gait cycle is presented. Taking into account the forward complete setting of the proposed BRD, an output feedback formulation of the an adaptive controller is also developed. A virtual dynamic representation of the BRD is used to test the proposed controller over the LAs mathematical model using a distributed implementation of the adaptive controller. This chapter is derived from [62].

2.1 Introduction

Over the last three decades, significant progress has been made in the development of bipedal robots (BRDs). Advances in mechanical design, electrical instrumentation, and automatic control methods have enabled BRDs to perform increasingly complex activities, such as autonomous walking, stair climbing, and jumping. There is currently a scientific and technical trend towards developing more efficient BRD configurations, which involves the use of lighter and stronger materials, faster digital processors, and advanced electronic sensors and actuators. Actuators are particularly important for BRDs, as they define their mobility capabilities. Several actuator options are available for BRDs, including permanent magnet and brushless motors, steppers, pneumatic muscles, brushless motors [63], multi-motor drive systems [64], and, more recently, LAs.

LAs are becoming increasingly popular in robotics due to their unique characteristics, including device resizeability, setup versatility, stack setting, higher torque, electrical equivalence to DC motors, and remarkable motion robustness with respect to external perturbations and long periods of activity. These devices are comprised of an electrical motor (usually DC with permanent magnet) attached to a ball-screw driving system (as illustrated in Figure 2.1). However, despite their advantages, there are several drawbacks that need to be addressed before considering LAs a viable option for driving the movement of BRDs. Firstly, the design and manufacturing complexity often results in neglecting the internal dynamics of the LAs during modeling. Moreover, LAs have a limited range of motion, which can result in restricted articulation angular displacement. Nevertheless, LAs present multiple benefits for the mobilization of BRDs compared to other actuation methods. For instance, the spatial separation between the drive and joint enables lighter movable parts. Additionally, when the leg link and LA form a closed kinematic chain, the stiffness of the joint and surrounding structure increases. Furthermore, backlash can be minimized by utilizing a synchronous drive belt to transmit actuator power to the joint. Also, a crank mechanism on the thigh can reduce leg inertia and allow for the installation of heavy objects, such as actuators, near the hip joint [65]. Lastly, given its type of movement, they can mimic the human muscular system [66].

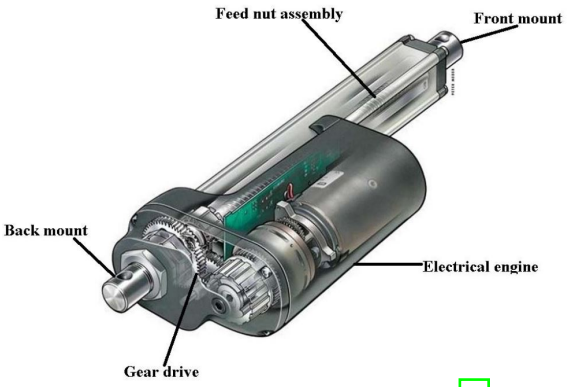


Figure 2.1: General LA structure [67]

Several studies have emphasized the benefits of using LAs as driving elements for BRDs. For instance, the Technical University of Munich has developed the humanoid robot LOLA [68], which uses LAs based on a ball screw drive attached to the main structure by Cardan joints. In this design, the LA is also used to maintain the moment of inertia of the thigh. Additionally, the WL family of robots, designed by Waseda University, employs LAs in its Stewart platform-like mechanism that is used as a novel biped walking type wheelchair [69, 70]. In another study, [71] proposed a double linear ball-screw drive structure for the ankle joint. In comparison, [72] presented a structure in which both the knee and ankle possess a single and double LAs structure, respectively. Overall,

these studies demonstrate the versatility and adaptability of LAs in designing BRDs.

Previous studies on biped robots utilizing LAs as active driving elements have overlooked the importance of analyzing their electromechanical dynamics, including the temporal evolution of the motor and its connection to the ball-screw configuration. Additionally, these studies often assume that LAs have an unbounded range of linear displacement. In order to fill this gap in the existing literature and offer a more comprehensive mathematical model, the current chapter aims to analyze the electromechanical characteristics and employ them as constraints in the design of automatic controls, with the objective of ensuring a successful execution of the gait cycle, even under conditions where the BRD is suspended.

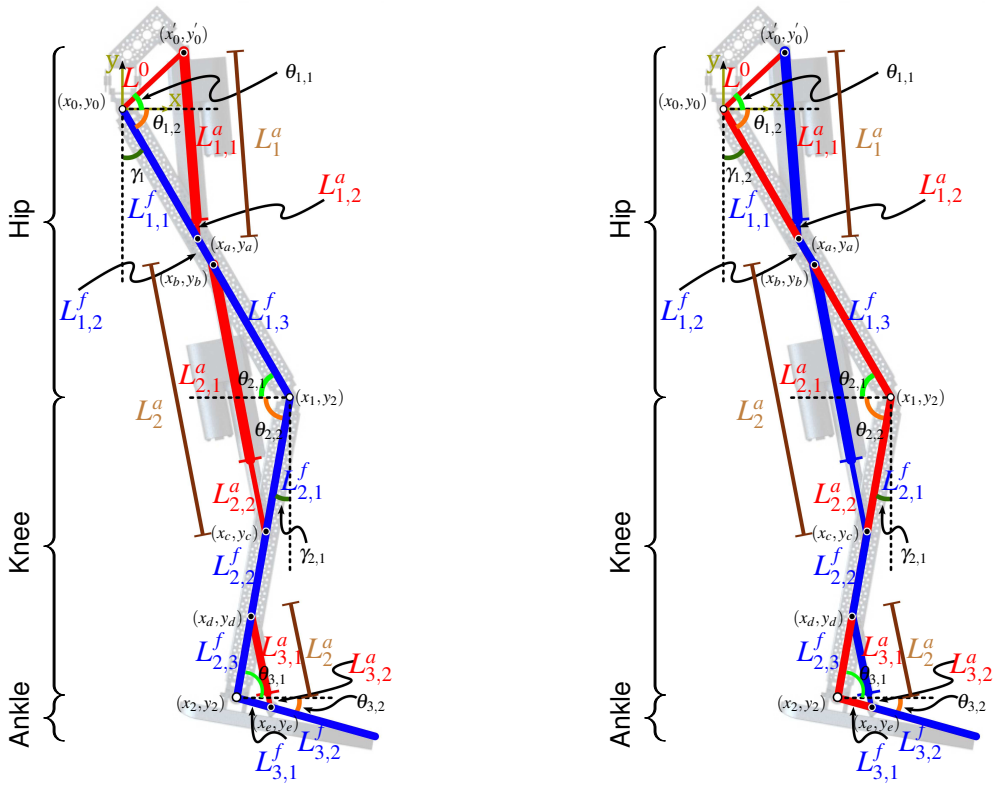
Incorporating linear actuator systems (LA) into biped robot designs presents challenges for both structural mathematical modeling and automatic control design [73]. Traditionally, BRDs are dynamically modeled as robotic manipulators with an open kinematic chain structure [74]. However, the addition of LAs alters the architecture of the BRD from a mere manipulator, resulting in changes to mass distribution, damping, and limitations on angular displacement at each joint of the BRD (see Figure 2.2) [75]. These mechanical changes, combined with the dynamic nature of LA, create challenges for designing control techniques, necessitating new design paradigms that consider the mechanical role of the actuator in conjunction with its driving function [76, 77].

The limited velocity of the LA and restricted screw movement range increase the number of states in the actuated BRD, and consequently, bound the control actions. The control design for tracking reference trajectories that define required articulation movements must consider these features, in addition to modeling imprecision and external uncertainties. Various techniques have been proposed to address these issues in the control design, such as penalty state-dependent functions, convex set projections, composite state convex sets, and adaptive gain bLfs. This thesis is focused on developing the latter technique.

This chapter is organized as follows. The modeling strategy presented in Section 2.2 details the method to include the LA presence for each articulation in the BRD. Section 2.3 describes the mathematical model of the BRD considering that the main structure are restricted by the movement of the actuator. Section 2.4 considers the dynamics of the BRD when the LA form part of the main section in the device. In Section 2.5the LA mathematical model is detailed. Section 2.6 presents an adaptive controller based on the application of logarithmic bLfs. Section 2.7 details the implementation strategy and implementation of the output feedback controller . Section **??** details the virtual BRD designed to test the proposed controller over the LA formulation which was implemented in Sim-Mechanics/Matlab. The numerical evaluations of the controller considering the mathematical formulation of the actuator dynamics are presented in Section **??**, where the proposed controller is compared with a classical PID controller.

2.2 Mathematical modeling strategy

The integration of LAs into a biped robot structure offers two possible modeling strategies. The distinction between these strategies is illustrated in Figure 2.2, including the angular and linear variables used for each model.



(a) Unactuated mechanical structure for first mathematical modeling

(b) Actuated mechanical structure for second mathematical modeling.

Figure 2.2: The two modeling strategies considered incorporating linear actuators into the structure of the BRD.

The first strategy involves modeling the unactuated mechanical structure as the core element, with the LAs considered only as a section that provides mobilization to the mechanical linkages. In Figure 2.2a, the blue lines represent the modeled section of the robot without considering the displacement of the LAs, but rather the angular movement in the links. The second strategy involves integrating the LAs as part of the mechanical structure itself. In Figure 2.2b, the blue lines represent the corresponding modeled section of the robot, now considering the displacement of the LAs as

part of the kinematics. The class of control strategies is also taken into account in the comparison of the two models.

In this section, two modeling strategies are presented, each sharing the following common characteristics:

- The biped robot is capable of moving its legs on the sagittal plane.
- There is no dynamic contact between the actuator tip and the structural section of the biped, implying no stiffness-damping effect.
- Each articulation's LA is constructed with a lead screw and nut mechanism, enabling linear motion.
- A DC motor is employed to activate the actuator and is incorporated into the control design.
- All articulations in the biped robot are stationary based on a rotational joint, which is justified by the motion in the sagittal plane.
- The biped robot is suspended and does not make contact with the support surface.

These shared features are crucial for both modeling strategies and provide a common foundation for the development of robust controllers.

Considering that $q \in \mathcal{Q} \subset \mathbb{R}^n$ is the vector of generalized coordinates and $\dot{q} \in \mathcal{T}q\mathcal{Q} \subset \mathbb{R}^n$, then the kinetic energy $T : \mathcal{TQ} \to \mathbb{R}^+$, $T = T(q, \dot{q})$ and the potential energy $U : \mathcal{Q} \to \mathbb{R}^+$, $U = U(q)$ for the entire system.

The energy contributions of the individual subsystems can be calculated independently. Hence, the equations of motion are obtained with the calculation of the Euler-Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_i}\mathcal{L}\left(q, \frac{d}{dt}q(t)\right)\right) - \frac{\partial}{\partial q_i}\mathcal{L}\left(q, \frac{d}{dt}q(t)\right) = \tau_i - \Xi_i \tag{2.1}$$

with the generalized forcing and dissipation terms $\tau_i \in \mathbb{R}$, $\Xi_i \in \mathbb{R}$ as well as the Lagrangian: $\mathcal{L}(q, \dot{q}) = T(q, \dot{q}) - U(q)$. The derivation and principles behind this formalism can be found in any textbook on classical mechanics. Except for the differentiations that can become cumbersome for higher-order systems, the Euler-Lagrange equations formalism is straightforward to apply because it is based on universal scalar energy functions.

Based on the calculus for the Lagrangian and the application of the Euler-Lagrange equations, the complete model for a single leg of the biped robot satisfies the following differential equation:

$$M(q(t))\frac{d^2}{dt^2}q(t) + C\left(q(t), \frac{d}{dt}q(t)\right)\frac{d}{dt}q(t) + G(q(t)) = \tau(t) - \Xi(t) \tag{2.2}$$

Where $q \in \mathbb{Q} \subset \mathbb{R}^3$ is the vector of positions including all joints of a single leg in the exoskeleton.

The inertia matrix is defined as $M : \mathbb{R}^3 \to \mathbb{R}^{3\times3}$, $C : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^{3\times3}$ is the matrix of Coriolis and centrifugal forces, $G : \mathbb{R}^3 \to \mathbb{R}^3$ has the terms associated to the gravitational forces and the term $\Xi \in \mathbb{R}^3$ is the vector that represents the effect of viscous and dried frictions, external perturbations and modeling imprecision.

2.3 Development of Mathematical Model 1

The mechanical section of the biped robot depicted in Figure 2.2a is made up of fixed links which are interconnected by the LA structures. This figure shows the configuration of the articulations in the BRD which are restricted by the linear movement of the actuator.

The expression of the kinetic energy of this first model corresponds to

$$T = \sum_{k=1}^{3} T_{m_k^f}\left(\theta_{1,2}, \theta_{2,2}, \theta_{3,2}, \dot{\theta}_{1,2}, \dot{\theta}_{2,2}, \dot{\theta}_{3,2}\right) + \sum_{k=1}^{3} T_{m_{k,1}^a}\left(\gamma_1, \gamma_2, \gamma_3, \dot{\gamma}_1, \dot{\gamma}_2, \dot{\gamma}_3\right) + \sum_{k=1}^{3} T_{m_{k,2}^a}\left(\gamma_1, \gamma_2, \gamma_3, \dot{\gamma}_1, \dot{\gamma}_2, \dot{\gamma}_3\right) \tag{2.3}$$

Where $\theta_{j,k}$ corresponds to the angular movement at each joint with respect to the horizontal plane, with k denoting positive or negative angle, and γ_j is used to denote vertical movement. The nomenclature corresponds to the Figure 2.2. Here the individual kinetic energies for the masses in the biped robot are presented below. The first section of individual kinetic energies correspond to the masses of the fixed mechanical sections of the BRD (solid links), that $T_{m_1^f}$, $T_{m_2^f}$ and $T_{m_3^f}$ respectively for the three sections of the device:

$$T_{m_1^f} = \frac{1}{2} m_1^f \left(0.5 L_1^f\right)^2 \left(\dot{\theta}_{1,2}\right)^2 \tag{2.4}$$

$$T_{m_2^f} = \frac{1}{2} m_2^f \left(L_1^f\right)^2 \left(\dot{\theta}_{1,2}\right)^2 + \frac{1}{2} m_2^f \left(\frac{L_2^f}{2}\right)^2 \left(\dot{\theta}_{2,2}\right)^2 - m_2^f L_1^f L_2^f cos\left(\theta_{1,2} - \theta_{2,2}\right) \dot{\theta}_{1,2} \dot{\theta}_{2,2} \tag{2.5}$$

$$
\begin{aligned}
T_{m_3^f} = \frac{1}{2} m_3^f \left(L_1^f\right)^2 \left(\dot{\theta}_{1,2}\right)^2 + \frac{1}{2} m_3^f \left(L_2^f\right)^2 \left(\dot{\theta}_{2,2}\right)^2 \\
+ \frac{1}{2} m_3^f \left(\frac{L_3^f}{2}\right)^2 \left(\dot{\theta}_{3,2}\right)^2 - \\
2 m_3^f L_1^f L_2^f cos\left(\theta_{1,2} - \theta_{2,2}\right) \dot{\theta}_{1,2} \dot{\theta}_{2,2} + \\
m_3^f L_1^f L_3^f cos\left(\theta_{1,2} - \theta_{3,2}\right) \dot{\theta}_{1,2} \dot{\theta}_{3,2} - \\
m_3^f L_2^f L_3^f cos\left(\theta_{2,2} - \theta_{3,2}\right) \dot{\theta}_{2,2} \dot{\theta}_{3,2}
\end{aligned}
\tag{2.6}
$$

The second section of individual kinetic energies correspond to the fixed mechanical sections of each LA in the BRD, that $T_{m_{j,1}^a}, j = 1,2,3$, respectively for the three actuators of the device:

$$
\begin{aligned}
T_{m_{j,1}^a} = T_{m_{j,1}^a}^r + \frac{1}{4} m_{j,1}^a \left(L_{j,1}^a\right)^2 \left(\left(\dot{\gamma}_j\right)^2\right) + \\
m_{j,1}^a L_{j,1}^a \left(\dot{x}_{j,1}^{a,r} cos\left(\gamma_j\right) - \dot{y}_{j,1}^{a,r} sin\left(\gamma_j\right)\right)
\end{aligned}
\tag{2.7}
$$

where $T_{m_{j,1}^a}^r$ is

$$
T_{m_{j,1}^a}^r = \frac{1}{2} m_{j,1}^a \left(\left(\dot{x}_{j,1}^{a,r}\right)^2 + \left(\dot{y}_{j,1}^{a,r}\right)^2\right) \tag{2.8}
$$

The third section of individual kinetic energies correspond to the mobile sections of each LA in the BRD, that $T_{m_{j,2}^a}, j = 1,2,3$, respectively for the three actuators:

$$
\begin{aligned}
T_{m_{j,2}^a} = T_{m_{j,2}^a}^r + \frac{1}{4} m_{j,2}^a \left(L_{j,1}^a\right)^2 \left(\left(\dot{\gamma}_j\right)^2\right) + \frac{1}{8} m_{j,2}^a \left(\dot{L}_{j,2}^a\right)^2 + \frac{1}{8} m_{j,2}^a \left(L_{j,2}^a\right)^2 \left(\dot{\gamma}_j\right)^2 + \\
m_{j,2}^a L_{j,1}^a \left(\dot{x}_{j,1}^{a,r} cos\left(\gamma_j\right) - \dot{y}_{j,1}^{a,r} sin\left(\gamma_j\right)\right) + \frac{1}{2} m_{j,2}^a \dot{L}_{j,2}^a \left(\dot{x}_{j,1}^{a,r} sin\left(\gamma_j\right) + \dot{y}_{j,1}^{a,r} cos\left(\gamma_j\right)\right) + \\
\frac{1}{2} m_{j,2}^a L_{j,2}^a \dot{\gamma}_j \left(\dot{x}_{j,1}^{a,r} cos\left(\gamma_j\right) - \dot{y}_{j,1}^{a,r} sin\left(\gamma_j\right)\right) + \frac{1}{2} m_{j,2}^a L_{j,1}^a \dot{L}_{j,2}^a \left(sin(\gamma_j - \gamma_j)\right) + \\
\frac{1}{2} m_{j,2}^a L_{j,1}^a L_{j,2}^a \left(cos(2\gamma_j)\right) + \frac{1}{4} m_{j,2}^a L_{j,2}^a \dot{L}_{j,2}^a \left(sin(\gamma_j - \gamma_j)\right) \dot{\gamma}_j
\end{aligned}
\tag{2.9}
$$

where $T_{m_{j,2}^a}^r$ corresponds to

$$
T_{m_{j,2}^a}^r = \frac{1}{2} m_{j,2}^a \left(\left(\dot{x}_{j,2}^{a,r}\right)^2 + \left(\dot{y}_{j,2}^{a,r}\right)^2\right) \tag{2.10}
$$

Correspondingly, the formal expression of the potential energy is defined as follows

$$V = \sum_{k=1}^{3} V_{m_k^f}(\theta_{1,2}, \theta_{2,2}, \theta_{3,2}) + \sum_{k=1}^{3} V_{m_{k,1}^a}(\gamma_1, \gamma_2, \gamma_3) + \sum_{k=1}^{3} V_{m_{k,2}^a}(\gamma_1, \gamma_2, \gamma_3) \tag{2.11}$$

The first section of individual potential energies correspond to the masses of the fixed mechanical sections of the BRD (solid links), that $V_{m_k^f}$ respectively for the three sections of the device:

$$V_{m_k^f} = m_k^f g y_k^f \tag{2.12}$$

The second section of individual potential energies correspond to the fixed mechanical sections of each LA in the BRD, that is $V_{m_{j,1}^a}, j = 1,2,3$, respectively for the three actuators of the device:

$$V_{m_{j,1}^a} = m_{j,1}^a g y_{ja}^a \tag{2.13}$$

The third section of individual potential energies correspond to the mobile mechanical sections of each LA in the BRD, that is $V_{m_{j,2}^a}, j = 1,2,3$, respectively for the three actuators of the device:

$$V_{m_{j,2}^a} = m_{j,2}^a g y_{jb}^a \tag{2.14}$$

The calculus of both the kinetic and the potential energies in the BRD uses the estimation of the centers of masses of the fixed sections in the robot x_j^f and y_j^f that correspond to

$$\begin{aligned} x_1^f &= \frac{1}{2} L_1^f cos(\theta_{1,2}) \\ y_1^f &= \frac{1}{2} L_1^f sin(\theta_{1,2}) \end{aligned} \tag{2.15}$$

$$\begin{aligned} x_2^f &= L_1^f cos(\theta_{1,2}) - \frac{1}{2} L_2^f cos(\theta_{2,2}) \\ y_2^f &= L_1^f sin(\theta_{1,2}) - \frac{1}{2} L_2^f sin(\theta_{2,2}) \end{aligned} \tag{2.16}$$

$$\begin{aligned} x_3^f &= L_1^f cos(\theta_{1,2}) - L_2^f cos(\theta_{2,2}) + \frac{1}{2} L_3^f cos(\theta_{3,2}) \\ y_3^f &= L_1^f sin(\theta_{1,2}) - L_2^f sin(\theta_{2,2}) + \frac{1}{2} L_3^f sin(\theta_{3,2}) \end{aligned} \tag{2.17}$$

To connect the active length of the actuators with the angular motion obtained at each articulation in the legs of the BRD, the LAs are configured as shown in Figure 2.2. The following relationships represent these connections:

$$\begin{aligned}\left(L^f_{1,1}\right)^2 &= \left(L^0\right)^2 + \left(L^a_1\right)^2 - 2L^0L^a_1 cos\left(\gamma^c_1\right)\\ \left(L^f_{2,1}\right)^2 &= \left(L^f_{1,3}\right)^2 + \left(L^a_2\right)^2 - 2L^f_{1,3}L^a_2 cos\left(\gamma^c_2\right)\\ \left(L^f_{3,1}\right)^2 &= \left(L^f_{2,3}\right)^2 + \left(L^a_3\right)^2 - 2L^f_{2,3}L^a_3 cos\left(\gamma^c_3\right)\end{aligned} \tag{2.18}$$

Complementary, the angles $\gamma_j = \pi/2 - \gamma^c_j$ with $j = 1,2,3$. The corresponding coordinates for the fixed sections of the LAs are $\left(x^a_{ja}, y^a_{ja}\right)$, while the coordinates for the mobile section defined by $\left(x^a_{jb}, y^a_{jb}\right)$ which can be expressed as follows

$$\begin{aligned}x^a_{ja} &= x^{a,r}_{j,1} + \frac{1}{2}L^a_{j,1} sin\left(\gamma_j\right)\\ x^a_{jb} &= x^{a,r}_{j,1} + L^a_{j,1} sin\left(\gamma_j\right) + \frac{1}{2}L^a_{j,2} sin\left(\gamma_j\right)\\ y^a_{ja} &= y^{a,r}_{j,1} + \frac{1}{2}L^a_{j,1} cos\left(\gamma_j\right)\\ y^a_{jb} &= y^{a,r}_{j,1} + L^a_{j,1} cos\left(\gamma_j\right) + \frac{1}{2}L^a_{j,2} cos\left(\gamma_j\right)\end{aligned} \tag{2.19}$$

where $x^{a,r}_{1,1} = x'_0$, $y^{a,r}_{1,1} = y'_0$; $x^{a,r}_{2,1} = x_b$, $y^{a,r}_{2,1} = y_b$ and $x^{a,r}_{3,1} = x_d$, $y^{a,r}_{3,1} = y_d$. The relationships for the coordinates of the fixed and mobile sections requires the analysis of the connections between the lengths of the actuators and the angular motions, which are

$$\begin{aligned}\left(L^0\right)^2 &= \left(x'_0\right)^2 + \left(y'_0\right)^2\\ \left(L^f_{1,1} + L^f_{1,2}\right)^2 &= \left(x_b\right)^2 + \left(y_b\right)^2\\ \left(L^f_1\right)^2 + L^f_1\left(L^f_{2,1} + L^f_{2,2}\right) cos\left(\theta_{1,2} - \theta_{2,2}\right) +&\\ \frac{1}{4}\left(L^f_{2,1} + L^f_{2,2}\right)^2 &= \left(x_d\right)^2 + \left(y_d\right)^2\end{aligned} \tag{2.20}$$

The corresponding velocities for the fixed sections of the links in the BRD are the following:

$$\begin{aligned} \dot{x}_1^f &= -\frac{1}{2}L_1^f sin\left(\theta_{1,2}\right)\dot{\theta}_{1,2} \\ \dot{y}_1^f &= \frac{1}{2}L_1^f cos\left(\theta_{1,2}\right)\dot{\theta}_{1,2} \end{aligned} \tag{2.21}$$

$$\begin{aligned} \dot{x}_2^f &= -L_1^f sin\left(\theta_{1,2}\right)\dot{\theta}_{1,2} + \frac{1}{2}L_2^f sin\left(\theta_{2,2}\right)\dot{\theta}_{2,2} \\ \dot{y}_2^f &= L_1^f cos\left(\theta_{1,2}\right)\dot{\theta}_{1,2} - \frac{1}{2}L_2^f cos\left(\theta_{2,2}\right)\dot{\theta}_{2,2} \end{aligned} \tag{2.22}$$

$$\begin{aligned} \dot{x}_3^f &= -L_1^f sin\left(\theta_{1,2}\right)\dot{\theta}_{1,2} + L_2^f sin\left(\theta_{2,2}\right)\dot{\theta}_{2,2} - \\ &\quad \frac{1}{2}L_3^f sin\left(\theta_{3,2}\right)\dot{\theta}_{3,2} \\ \dot{y}_3^f &= L_1^f cos\left(\theta_{1,2}\right)\dot{\theta}_{1,2} - L_2^f cos\left(\theta_{2,2}\right)\dot{\theta}_{1,2} + \\ &\quad \frac{1}{2}L_3^f cos\left(\theta_{3,2}\right)\dot{\theta}_{3,2} \end{aligned} \tag{2.23}$$

The corresponding velocities of fixed and mobile masses for the LAs are given by

$$\dot{x}_{ja}^a = \dot{x}_{j,1}^{a,r} + \frac{1}{2}L_{j,1}^a cos\left(\gamma_j\right)\dot{\gamma}_j$$

$$\begin{aligned} \dot{x}_{jb}^a &= \dot{x}_{j,1}^{a,r} + L_{j,1}^a cos\left(\gamma_j\right)\dot{\gamma}_j + \frac{1}{2}\dot{L}_{j,2}^a sin\left(\gamma_j\right) + \\ &\quad \frac{1}{2}L_{j,2}^a cos\left(\gamma_j\right)\dot{\gamma}_j \end{aligned} \tag{2.24}$$

$$\begin{aligned} \dot{y}_{ja}^a &= \dot{y}_{j,1}^{a,r} - \frac{1}{2}L_{j,1}^a sin\left(\gamma_j\right)\dot{\gamma}_j \\ \dot{y}_{jb}^a &= \dot{y}_{j,1}^{a,r} - L_{j,1}^a sin\left(\gamma_j\right)\dot{\gamma}_j + \frac{1}{2}\dot{L}_{j,2}^a cos\left(\gamma_j\right) - \\ &\quad \frac{1}{2}L_{j,2}^a sin\left(\gamma_j\right)\dot{\gamma}_j \end{aligned}$$

Considering the basis of the LAs, the following restrictions hold for the relationships between the

time-varying lengths of the LAs:

$$\left(\left(\dot{x}_{1,1}^{a,r}\right)^2+\left(\dot{y}_{1,1}^{a,r}\right)^2\right)=0$$

$$\left(\left(\dot{x}_{2,1}^{a,r}\right)^2+\left(\dot{y}_{2,1}^{a,r}\right)^2\right)=\left(L_{1,1}^f+L_{1,2}^f\right)^2\left(\dot{\theta}_{1,2}\right)^2$$

$$\left(\left(\dot{x}_{3,1}^{a,r}\right)^2+\left(\dot{y}_{3,1}^{a,r}\right)^2\right)= \tag{2.25}$$

$$\left(L_1^f\right)^2\left(\dot{\theta}_{1,2}\right)^2+\left(L_{2,1}^f+L_{2,2}^f\right)^2\left(\dot{\theta}_{2,2}\right)^2-$$

$$2L_1^f\left(L_{2,1}^f+L_{2,2}^f\right)cos\left(\theta_{1,2}-\theta_{2,2}\right)\dot{\theta}_{1,2}$$

The expressions of the positions and velocities are gotten using the configurations in the BRD and considering the relative motions of the LAs with respect to the fixed sections of the BRD.

2.4 Development of Mathematical Model 2

To integrate the dynamic movement of the actuators with the mechanical section of the biped robot, the model shown in Figure 2.2 is considered. In this section, the geometric analysis that combines the mechanical section with the actuator dynamic movement is presented. Specifically, for the hip section of the model in Figure 2.2, the following relationships are valid:

$$\theta_1=\theta_{1,1}+\theta_{1,2} \tag{2.26}$$

The fixed relation between (x_0,y_0) and (x_0',y_0') yields

$$\theta_{1,1}=atan\left(\frac{y_0'-y_0}{x_0'-x_0}\right) \tag{2.27}$$

Now considering that the length of the selected LA for the hip section L_1^a is given by $L_1^a=L_{1,1}^a+L_{1,2}^a$. In view of the mechanical relation between (x_0,y_0), (x_0',y_0') and (x_a,y_a) yields the following equation:

$$(L_1^a)^2=\left(L^0\right)^2+\left(L_{1,1}^f\right)^2-2L^0L_{1,1}^f cos\left(\theta_1\right) \tag{2.28}$$

In view (2.28), $\gamma_{1,2} = 0.5\pi - \theta_{1,2}$, the following relations are valid:

$$\begin{aligned} cos\,(\gamma_{1,2}) &= \frac{y_0 - y_a}{L^f_{1,1}} = \frac{y_0 - y_b}{L^f_{1,1} + L^f_{1,2}} = \frac{y_0 - y_1}{L^f_1} \\ sin\,(\gamma_{1,2}) &= \frac{x_0 - x_a}{L^f_{1,1}} = \frac{x_0 - x_b}{L^f_{1,1} + L^f_{1,2}} = \frac{x_0 - x_1}{L^f_1} \end{aligned} \tag{2.29}$$

Using a similar geometrical study for the knee study, the following angular equation is valid:

$$\theta_2 = \theta_{2,1} + \theta_{2,2} \tag{2.30}$$

The static relation given for (x_b, y_b) and (x_1, y_1) yields

$$\theta_{2,1} = atan\left(\frac{y_b - y_1}{x_b - x_1}\right) \tag{2.31}$$

Taking into account the length of the selected LA for the knee section L^a_2 is given by $L^a_2 = L^a_{2,1} + L^a_{2,2}$. In view of the mechanical relation between (x_b, y_b), (x_1, y_1) and (x_c, y_c) yields the following equation:

$$(L^a_2)^2 = \left(L^f_{1,3}\right)^2 + \left(L^f_{2,1}\right)^2 - 2L^f_{1,3}L^f_{2,1}cos\,(\theta_2) \tag{2.32}$$

In view (2.32), $\gamma_{2,2} = 0.5\pi - \theta_{2,2}$, the following relations are valid:

$$\begin{aligned} cos\,(\gamma_{2,2}) &= \frac{y_1 - y_c}{L^f_{2,1}} = \frac{y_1 - y_d}{L^f_{2,1} + L^f_{2,2}} = \frac{y_1 - y_2}{L^f_2} \\ sin\,(\gamma_{2,2}) &= \frac{x_1 - x_c}{L^f_{2,1}} = \frac{x_1 - x_d}{L^f_{2,1} + L^f_{2,2}} = \frac{x_1 - x_2}{L^f_2} \end{aligned} \tag{2.33}$$

For the ankle analysis, the corresponding angular equation is valid:

$$\theta_3 = \theta_{3,1} + \theta_{3,2} \tag{2.34}$$

The static relation given for (x_d, y_d) and (x_2, y_2) yields

$$\theta_{3,1} = atan\left(\frac{y_d - y_2}{x_d - x_2}\right) \tag{2.35}$$

In view of the length of the selected LA for the ankle section L^a_3 corresponds to $L^a_3 = L^a_{3,1} + L^a_{3,2}$.

Considering the mechanical relation between (x_d, y_d), (x_2, y_2) and (x_e, y_e) yields:

$$(L_3^a)^2 = \left(L_{2,3}^f\right)^2 + \left(L_{3,1}^f\right)^2 - 2L_{2,3}^f L_{3,1}^f cos(\theta_3) \tag{2.36}$$

Accordingly (2.36), $\gamma_{3,2} = 0.5\pi - \theta_{3,2}$, the following relations are valid:

$$\begin{aligned} cos(\gamma_{3,2}) &= \frac{x_2 - x_e}{L_{3,1}^f} = \frac{x_2 - x_3}{L_3^f} \\ sin(\gamma_{3,2}) &= \frac{y_2 - y_e}{L_{3,1}^f} = \frac{y_2 - y_3}{L_3^f} \end{aligned} \tag{2.37}$$

According to the mechanical relations described in this section, the kinetic and potential energies can be obtained in the form presented in the Mathematical Model 1. The calculus is omitted just to avoid unnecessary repetition of similar information.

2.5 Tendon-Driven Actuation Principle

The operation of the linear screw-lead and nut device as LA on the biped robot is described as follows: First, a DC electrical motor rotates a ball-screw, and the nut is locked in rotation but free to move along the screw axis with the help of an anti-rotation mechanism. This configuration has been utilized in various LAs and is similar to the one used in this chapter. The induced linear displacement of the nut is transmitted over a specific link, which eventually drives the rotational joint with the associated torque. This configuration enables a single-acting actuation architecture that reproduces tendon routing, thus reducing friction. It is worth noting that friction due to force-transmitting tendons is a significant source of friction in complex mechanical biped robot configurations. In this chapter's biped robot, the LA operates at the level of rotational torque that mobilizes each articulation. As each articulation is driven by two primary forces, the LA and gravitational, the following relation is considered:

$$\tau_i = \left(\vec{r}_i \times m_i^a \frac{d^2}{dt^2}\vec{x}_s\right) - \sum_{j=1}^{N_i^a} \left(\vec{r}_{CM,j} \times m_j \vec{g}\right) \tag{2.38}$$

where $\vec{r}_i$ is the vector from the articulated joint and the point of actuator insertion on the robot link, m_i^a is the mass of the mobile section in the LA, $\vec{x}_s$ is the displacement of the same mobile section, N_i^a is the number of the links placed below the analyzed joint which induce torque action, $\vec{r}_{CM,j}$ is the vector connecting the coordinates of the joint and those of the center of mass of each link j, m_j

is the mass of the link and $\vec{g}$ is the gravity vector.

2.5.1 Actuator Dynamics

An Electromechanical Actuator (EMA) consists of an electric direct current motor (electrical section) and an associated mechanical gear-box (mechanical section). EMAs are categorized into rotary and linear classes. LA is further classified into geared and direct-driven linear EMA. Generally, a geared linear electromechanical actuator comprises a DC motor with a lead-screw (or ball-screw or roller-screw) and a nut assembly. The motor shaft connects to the lead-screw with the aid of a shaft-coupling or a gear-box. The DC motor converts electrical energy into the rotational motion of the shaft, and the lead-screw converts this rotary motion into linear motion. The motor shaft rotates in either the clockwise or anti-clockwise direction, depending on the input polarity. Once the rotations of the shaft are transferred to the lead-screw, the nut makes a translational motion either in the forward or backward direction, depending on the rotational direction of the shaft.

The direct-driven linear electromechanical actuator (EMA) is characterized by the nut serving as the rotational element and the lead-screw acting as the translational element, with the lead-screw arrangement being accommodated inside the motor itself. Specifically, the internally threaded nut is fixed inside the hollow portion of the rotor. When the motor is supplied with electrical energy, both the rotor and the nut rotate, and the rotational motion of the nut facilitates the linear travel of the lead-screw. Additionally, the shaft attached to the end of the lead-screw has slotted grooves that constrain the rotational motion of the lead-screw along with the nut. In terms of function, these two actuators are not significantly different, but the direct-driven EMA boasts a more compact size than the geared EMA. This feature renders the direct-driven EMA particularly well-suited for aerospace industries, robotics, and other industrial machinery applications.

Modeling a single EMAis essential for constructing a multi-element actuator like High-redundancy Actuators (HRA). In this section, the mathematical model of a direct-driven linear EMAis described. The electric circuit implies that the applied electrical power with a voltage (V_a) is opposed by the conducting path resistance (R_a), inductance (L_a) and a back e.m.f (V_b) which is equal to $K_e \frac{d\theta_m}{dt}$. The voltage can then be expressed as:

$$K_e \frac{d}{dt}\theta_m + L_a \frac{d}{dt} i_a + R_a i_a = V_a \tag{2.39}$$

where K_e is the back e.m.f constant and θ_m is the angular displacement of the motor. In the mechanical system, damping, inertia and frictional properties are concentrated quantities. When the electromagnetic torque generated by the electric circuit is applied to the mechanical system's (lead-

screw and nut) rotary element, it rotates at a speed of $\frac{d}{dt}\theta_m$ which has a moment of inertia J_m. The rotational speed of both the rotor and the nut will be the same because they both are attached to each other. The variables x_s and x_l are the linear displacements of the lead-screw and the load respectively. K_s represents the equivalent stiffness that exists between lead-screw and nut and the damping that exists between bearing and rotor is represented by C_s. The generated torque τ_m is opposed by damping torque τ_d, load torque τ_l and an inertial torque τ_i. Hence, the generated electromagnetic torque is balanced by the other three torques [78], that is:

$$\begin{gathered} \tau_m = \tau_d + \tau_l + \tau_i \\ \tau_m = K_t i_a \quad \tau_d = D\frac{d}{dt}\theta_m \quad \tau_i = J\frac{d^2}{dt^2}\theta_m \\ \tau_l = \frac{C_s \cdot l}{2\pi}(\dot{x}_s - \dot{x}_l) + \frac{K_s \cdot l}{2\pi}(x_s - x_l) \end{gathered} \tag{2.40}$$

where D is the equivalent viscous damping coefficient at the armature. The torque τ_m generated by the current i_a which is passing through the armature is given by the expression (2.40). K_t represents the motor torque constant. Similarly, the damping, inertial and load torques are given in the subsequent expressions in (2.40) respectively. In that equation, the term $\frac{l}{2\pi}$ is used for converting the angular displacement of the nut to linear displacement of the lead-screw, where l is the screw lead. The substitution of τ_m, τ_d, τ_l and τ_i in (2.40) and bringing the electrical relationship in (2.39) yield:

$$\begin{gathered} K_t i_a = D\frac{d}{dt}\theta_m + \frac{C_s \cdot l}{2\pi}(\dot{x}_s - \dot{x}_l) + \frac{K_s \cdot l}{2\pi}(x_s - x_l) + J\frac{d^2}{dt^2}\theta_m \\ K_e\frac{d}{dt}\theta_m + R_a i_a + L_a\frac{d}{dt}i_a = V_a \end{gathered} \tag{2.41}$$

with $x_s = (l \cdot \theta)/(2 \cdot \pi)$ the relationship between the angular displacement and the displacement of the LA (considering the lead screw configuration). In terms of the movement efficiency the following relationship is also valid $\frac{d}{dt}x_s = \Gamma_r\frac{d}{dt}\theta\frac{1}{\eta \cdot \tau_i}$ where Γ_r is the resistive torque and η is the ball-screw efficiency. The application of the assumption describing the class of interaction between the tip of the actuator and the link in the biped robot, then $x_s = x_l$.

Using the state variable theory for each actuator τ_i with the definitions $\chi^a_{1,i} = \theta_{m,i}$, $\chi^a_{2,i} = \dot{\theta}_{m,i}$ and $\chi^a_{3,i} = i_{a,i}$ and the corresponding parameters (labeled with a superscript i) yields to the following model for the actuator:

$$\begin{aligned}\frac{d}{dt}\chi^a_{1,i} &= \chi^a_{2,i}\\ \frac{d}{dt}\chi^a_{2,i} &= \frac{D_i}{J_i}\chi^a_{2,i} + \frac{K_{t,i}}{J_i}\chi^a_{3,i}\\ \frac{d}{dt}\chi^a_{3,i} &= -\frac{K_{e,i}}{L_{a,i}}\chi^a_{2,i} - \frac{R_{a,i}}{L_{a,i}}\chi^a_{3,i} + \frac{V_{a,i}}{L_{a,i}}\end{aligned} \tag{2.42}$$

The pushing force of the LA can be estimated as $F^a_i = \left(m^a_i \frac{d}{dt}\chi^a_{2,i}\right)$, therefore the torque $\tau_i = L^f_{i,1} \cdot F^a_i \cdot sin(\delta_i)$ produced by this force over the joint i corresponds to $\delta_i = \pi - \theta_{m,i} - \gamma^c_i$.

Actuator effect over the second Mathematical model

Notice that integrating the actuator as an active element of the mechanical section simplifies the analysis of the actuator effect. Indeed, one may notice that mathematical modeling of the actuator dynamics for this case corresponds to τ_m in (2.40), with $\tau_d = \tau_l = 0$. Therefore, the model for the actuation system inside the actuator (which is operating here as a prismatic joint) is satisfying the dynamics presented in (2.42) with $D_i = 0$.

In spite that this second model appears as a more simple version of the mathematical description of the BRD, it could be more complicated to be used in the control design formulation.

2.6 Feedback control of biped robot with linear actuators

The use of bLf has become a popular choice in addressing articulation restricted movement. The benefits of bLf in improving transient and steady operation of restricted systems have contributed to the growing number of its applications in recent years. The theory of bLf dates back to the mid-20th century [79], with its main variants being logarithmic [80–82] and tangent [83, 84].

Biped robotic devices controlled by adaptive gains based on the application of bLf have demonstrated significant advantages over systems using traditional control approaches such as state feedback, extended state approaches, and sliding modes, among others. Although adaptive state-dependent gains have been applied to proposed controllers in a growing number of applications, these gains are calculated using controlled bLf. However, no study has been reported proposing the application of adaptive controllers based on bLf for LAs in BRD configurations. Thus, this section presents a novel design for a state-dependent adaptive controller for a BRD with LAs in its structure. The control design considers state restrictions in both the LA and all BRD articulations while tracking reference trajectories to develop a regulated gait cycle. The dynamics of the DC motor that drives

the LA motion are also included in the study. This section provides an implementation strategy for LAs as part of the biped robot structure.

2.6.1 Control implementation for BR using linear ball screw drivers as actuation devices

The dynamic model of a single leg presented in (2.2) can be represented (with $\zeta_\alpha := q$ and $\zeta_\beta := \frac{d}{dt}q$) as follows [85, 86]:

$$\begin{gathered}\frac{d}{dt}\zeta_\alpha(t)=\zeta_\beta(t)\\ \frac{d}{dt}\zeta_\beta(t)=f(\zeta(t),t)+g(\zeta_\alpha(t))\,\tau(t)+\Upsilon(\zeta(t),t)\\ y(t)=\zeta_\alpha(t)\end{gathered} \tag{2.43}$$

with $\zeta^\top = [\zeta_\alpha^\top, \zeta_\beta^\top]$, where $\zeta \in X \subset \mathbb{R}^{2n}$, X an open subspace of feasible positions, $\zeta_\alpha \in \mathbb{R}^n$ is the vector of angles that describes the movements at each joint of the biped robot. The vector $\zeta_\beta \in \mathbb{R}^n$ represents the time derivative of the angular displacements for each joint. The variable y represents the available output information that can be measured from the biped robot. The function $f : X \times \mathbb{R}^+ \to \mathbb{R}^n$ is a vector field that represents the internal dynamics of the biped robot, its structure can be obtained by using Euler-Lagrange method. In this chapter it is not thoroughly studied the interactions between the contact surface and the biped structure. The properties of the f function guarantee validity of the locally Lipschitz condition presented here [87]:

$$\begin{gathered}\|f(\zeta^1)-f(\zeta^2)\| \le L_f\|\zeta^1-\zeta^2\|,\\ \zeta^1 \in X, \zeta^2 \in X, L_f \in \mathbb{R}^+\end{gathered} \tag{2.44}$$

The function $g : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ describes how the input vector affects the dynamics of the biped robot. This structure owns a well defined inverse since the inertia matrix of the biped robot is positive definite and satisfies the inequality (2.45) uniformly on $t \ge 0$, where it is used the Frobenius matrix norm:

$$\begin{gathered}0 < g^- \le \|g(\zeta_\alpha)\|_F \le g^+ < +\infty\\ g^- \in \mathbb{R}^+, g^+ \in \mathbb{R}^+\end{gathered} \tag{2.45}$$

The term $\Upsilon : X \times \mathbb{R}^+ \to \mathbb{R}^n$ represents the uncertainties/perturbations, and it belongs to the following set:

$$\Xi = \left\{ \Upsilon \Big| \sup_{t\ge 0,\ \zeta\in\mathbb{R}^n} \|\Upsilon\|^2 \le \Upsilon_0 + \Upsilon_1\|\zeta\|^2 \right\} \tag{2.46}$$

Because of the nature of the mechanical structure, it must satisfy the angular displacement restrictions for all the components of the state vector:

$$-\infty < \zeta_{\alpha,i}^{-} \leq \zeta_{\alpha,i} \leq \zeta_{\alpha,i}^{+} < +\infty, \;\; \zeta_{\alpha,i}^{-}, \zeta_{\alpha,i}^{+} \in \mathbb{R} \tag{2.47}$$

There exists an upper limit ζ^{+}, which is defined as:

$$\frac{(\max\limits_{i} \zeta_i^{+})}{\lambda_{min}\{P\}} = \zeta^{+} \tag{2.48}$$

where ζ_i defines the restriction for the i^{th} angular displacement and $P \in \mathbb{R}^{n \times n}$ is a positive definite weighting matrix. Such matrix can be used to restrict some specific articulations. Then, ζ^{+} defines the bounds for the states accordingly to the following inequality:

$$\| \zeta \|_P^2 \leq \zeta^{+} \tag{2.49}$$

The term Υ corresponds to $\Upsilon = M^{-1}\Xi$. A biped robot actuated with linear ball screw drivers gain postural force. Usually, an individual driver is placed at each articulation forming the configuration observed in Figure 2.2.

2.7 Output feedback control

This section details all the elements in the automatic control design including the implementation of the output feedback form, the implementation strategy as well as a brief sketch of the stability proof for the close loop dynamics.

2.7.1 Problem Statement for the Control Design

The aim of this chapter is to develop an output feedback control $u = u(t)$ for a biped robot such that

$$\| \zeta_a(t) - \zeta_a^{*}(t) \| \leq \beta_0 \quad , \quad \forall t \geqslant T > 0 \tag{2.50}$$

Here T is a given time, the positive scalar β_0 defines the quality of the trajectory tracking for the

reference trajectory ζ_a^*, which is satisfying the following ordinary differential equation:

$$\begin{aligned} \frac{d}{dt}\zeta_\alpha^*(t) &= \zeta_\beta^*(t) \\ \frac{d}{dt}\zeta_\beta^*(t) &= h\left(\zeta_\alpha^*(t), \zeta_\beta^*(t), t\right) \end{aligned} \tag{2.51}$$

Here the vector field $h : X \times X \times \mathbb{R}^+ \to \mathbb{R}^n$ is a locally Lipschitz function, which characterizes the reference trajectories. The references trajectories are proposed (refer to 1.2.4) to satisfy the same bounds that are admitted by the angular displacements of each joint, that is

$$-\infty < \zeta_{\alpha,i}^- \leq \zeta_{\alpha,i}^* \leq \zeta_{\alpha,i}^+ < +\infty, \;\; \zeta_{\alpha,i}^-, \zeta_{\alpha,i}^+ \in \mathbb{R} \tag{2.52}$$

Defining the trajectory tracking error $\delta(t) \in \mathbb{R}^{2n}$ which is given by $\delta = \zeta - \zeta^*$, where $\zeta^* = [(\zeta_\alpha^*)^\top, (\zeta_\beta^*)^\top]^\top$. Hence, the dynamics of $\delta(t)$ is given by:

$$\begin{aligned} \frac{d}{dt}\delta(t) &= A\delta(t) + B\left(f\left(\zeta(t),t\right) + g\left(\zeta_\alpha(t)\right)\tau(t)\right) + B\left(\Upsilon\left(\zeta(t),t\right) - h\left(\zeta_\alpha^*(t), \zeta_\beta^*(t), t\right)\right) \\ A &= \begin{bmatrix} 0_n & I_n \\ 0_n & 0_n \end{bmatrix} \quad B = \begin{bmatrix} 0_n \\ I_n \end{bmatrix} \end{aligned} \tag{2.53}$$

The matrices A and B satisfies a companion controllable form with the correct dimensions.

The controller design considered in this chapter assumes that the control form satisfies the following linear form with a state dependent gain:

$$\tau(t) = g^{-1}(\zeta_a(t))\left(K(t)\hat{\delta}(t) - h\left(\zeta_\alpha^*(t), \zeta_\beta^*(t), t\right) - f\left(\zeta(t),t\right)\right) \tag{2.54}$$

Notice that this formulation appears to be a class of adaptive state dependent feedback for the class of biped robot. In the controller form, $\hat{\delta}$ is the tracking error using the estimates of ζ_β. The gain of the suggested controller is defined as $K(t) = [K_P(t) \;\; K_D(t)]$ with $K_P \in \mathbb{R}^{n\times n}$ and $K_D \in \mathbb{R}^{n\times n}$. The well-known efficient tracking characteristics of the linear control form, its implementation needs of measuring $\frac{d}{dt}y(t)$. The possibility of getting the estimation of such derivative on-line usually implies important technical resource such additional sensors. A feasible alternative is instrumenting a class of Robust Exact Differentiator (RED) which can yield an accurate estimation of the time derivative of $\frac{d}{dt}y(t)$ [88]. This study uses the significant benefits of the Super-Twisting Algorithm (STA) method that has proven to be one of the more efficient alternatives to get an accurate estimation of the time derivative of the output information, while it can be used as part of output feedback control forms [89]. For a scalar variable, the STA operating as a RED can be detailed as follows. Consider that

$w_1(t) = r(t)$ with $r(t) \in \mathbb{R}$ being the signal which must be differentiated, $w_2(t) = \dfrac{d}{dt}r(t)$ describes its time-derivative. Assuming that $\left|\dfrac{d^2}{dt^2}r(t)\right| \leq r^+$, it is feasible to propose the following auxiliary dynamic representation

$$\frac{d}{dt}w_1(t) = w_2(t), \qquad \frac{d}{dt}w_2(t) = \frac{d^2}{dt^2}r(t) \tag{2.55}$$

The set of differential equations (2.55) consists of the state-space description for $r(t)$. The STA used to obtain the derivative of $r(t)$ satisfies

$$\begin{aligned} \frac{d}{dt}\bar{w}_1(t) &= \bar{w}_2(t) - \lambda_1 |\tilde{w}_1(t)|^{1/2} \operatorname{sign}(\tilde{w}_1(t)) \\ \frac{d}{dt}\bar{w}_2(t) &= -\lambda_2 \operatorname{sign}(\tilde{w}_1(t)) \\ \tilde{w}_1 = \bar{w}_1 - w_1;\ \ d(t) &= \frac{d}{dt}\bar{w}_1(t) \end{aligned} \tag{2.56}$$

where $\bar{w}_1$ and $\bar{w}_2$ are the STA variables and $\lambda_1, \lambda_2 > 0$ are its gains [88]. The time-dependent signal, $d(t)$ represents the output of the RED in (2.56). Here,

$$\operatorname{sign}(v) = \begin{cases} 1 & if \ \ v > 0 \\ [-1,1] & if \ \ v = 0 \\ -1 & if \ \ v < 0 \end{cases} \tag{2.57}$$

2.7.2 Adaptive output-feedback controller (AOFC) including the joint restrictions and actuator dynamics

The dynamics of each articulation in system (2.43) can be presented as the composition of n-subsystems as follows:

$$\begin{aligned}
\frac{d}{dt}\zeta_{\alpha,i}(t) &= \zeta_{\beta,i}(t) \\
\frac{d}{dt}\zeta_{\beta,i}(t) &= f_i(\zeta(t),t) + g_{i,i}(\zeta_\alpha(t))\,\tau_i(t) + \sum_{j=1,j\neq i}^{n} g_{i,j}(\zeta_\alpha(t))\,\tau_j(t) + \Upsilon_i(\zeta(t),t) \\
\tau_i(t) &= L_{i,1}^{f}\cdot\left(m_i^a\frac{d}{dt}\chi_{2,i}^a\right)\cdot sin\left(\pi-\chi_{1,i}^a-\gamma_i^c\right)
\end{aligned} \tag{2.58}$$

$$\begin{aligned}
\frac{d}{dt}\chi_{1,i}^a &= \chi_{2,i}^a \\
\frac{d}{dt}\chi_{2,i}^a &= \frac{D_i}{J_i}\chi_{2,i}^a + \frac{K_{t,i}}{J_i}\chi_{3,i}^a \\
\frac{d}{dt}\chi_{3,i}^a &= -\frac{K_{e,i}}{L_{a,i}}\chi_{2,i}^a - \frac{R_{a,i}}{L_{a,i}}\chi_{3,i}^a + \frac{V_{a,i}}{L_{a,i}}
\end{aligned}$$

where $\zeta_{a,i}$ and $\zeta_{b,i}$ are the $i-th$ and $(n+i)-th$ states of (2.43). The nonlinear forms $f_i(\cdot)$ and $g_{i,i}(\cdot)$ are the corresponding functions corresponding to the states $\zeta_{a,i}$ as well as $\zeta_{b,i}$. Additionally, $\gamma_i(\cdot,\cdot)$ describes the corresponding uncertain section of the studied subsystem. The positioning of the LA and the movement dynamics of each articulation in the biped robot (2.43) presented in Figure 2.2, establishes a relationship between torques at the analyzed i$^t h$ joint. This fact is represented in (2.58). Based on the definition for the tracking error introduced in (2.53), the individual tracking error for each angular joint in the biped robot satisfies:

$$\begin{aligned}
\frac{d}{dt}\delta_i(t) &= \delta_{i+n}(t) \\
\frac{d}{dt}\delta_{i+n}(t) &= f_i(\zeta(t),t) + g_{i,i}(\zeta_\alpha(t))\,\tau_i(t) - \\
h_i\left(\zeta_a^*(t),\zeta_b^*(t)\right) &+ \sum_{j=1,j\neq i}^{n} g_{i,j}(\zeta_\alpha(t))\,\tau_j(t) + \Upsilon_i(\zeta(t),t)
\end{aligned} \tag{2.59}$$

Here, the Lipschitz function $h_i\left(\zeta_a^*(t),\zeta_b^*(t)\right)$ is the $i-th$ component of $h\left(\zeta_a^*(t),\zeta_b^*(t)\right)$. Notice that each tracking error δ_i admits the following bounds

$$\delta_i^- := \zeta_{\alpha,i}^- - \zeta_{\alpha,i}^+ < \delta_i < \zeta_{\alpha,i}^+ - \zeta_{\alpha,i}^- =: \delta_i^+ \tag{2.60}$$

Based on the restrictions estimated for each δ_i, the application of asymmetric output based bLf yields the construction of the proposed controller that drives the corresponding tracking error towards the invariant zone near the origin.

2.7.3 Control design for the biped robot device

The convergence for the trajectory tracking error δ in (2.59) can be justified by the arguments described in the next theorem:

Theorem 1. *Consider the BRD system (2.43), with the control law proposed in (2.54) and the gains adjusted as in*

$$\begin{aligned}\frac{d}{dt}K_{P,i}(t) &= -\lambda_{P,i}P_{1,i}\delta_i|\delta_i| - \alpha_i\left(K_{P,i}(t) - K^*_{P,i}\right)\\ \frac{d}{dt}K_{D,i}(t) &= -\lambda_{P,i}P_{1,i}\frac{d}{dt}\hat{\delta}_i|\delta_i| - \alpha_i\left(K_{D,i}(t) - K^*_{D,i}\right)\end{aligned} \tag{2.61}$$

where $\hat{\delta}_i$ is the estimated derivative of δ_i. Hence, the controller uses the time-derivative of the tracking error obtained by the implementation of the STA as RED given in (2.56). If there are positive constants α_i, $\lambda_{1,i}$, $\lambda_{2,i}$, then the Lyapunov inequalities have positive-definite solutions $P_{1,i}$ matrices:

$$\begin{gathered}A_{1,i}^\top P_{1,i} + P_{1,i}A_{1,i} \leq -Q_{1,i}\\ A_{1,i} = \begin{bmatrix} -\lambda_{1,i} & 1 \\ -2\lambda_{2,i} & 0 \end{bmatrix}, \quad Q_{1,i} = Q_{1,i}^\top > 0, \quad Q_{1,i} \in \mathbb{R}^{2\times 2}\end{gathered} \tag{2.62}$$

and if there are positive gains $K^*_{P,i}$, $K^*_{D,i}$ that can be selected in such a way that the Riccati equations described by

$$P_{2,i}(A_{2,i} + \alpha_i I) + (A_{2,i} + \alpha_i I)^\top P_{2,i} + P_{2,i}R_{2,i}P_{2,i} + Q_{2,i} \leq 0 \tag{2.63}$$

have positive definite $P_{2,i}$ matrices as solutions with

$$\begin{gathered}A_{2,i} = \begin{bmatrix} 0 & 1 \\ -K^*_{P,i} & -K^*_{D,i} \end{bmatrix}, \quad R_{2,i} = \Lambda_{a,i} + \Lambda_{b,i}\\ Q_{2,i} = 4\lambda_{\max}\left\{\Lambda_{b,i}^{-1}\right\} I_{2\times 2} + \bar{\Lambda}_{a,i},\\ \bar{\Lambda}_{a,i} = L_i\Lambda_{a,i}, \quad L_i \in \mathbb{R}^+\\ \Lambda_{a,i} = \Lambda_{a,i}^\top > 0, \quad \Lambda_{b,i} = \Lambda_{b,i}^\top > 0\\ \Lambda_{a,i}, \Lambda_{b,i} \in \mathbb{R}^{2\times 2} \quad \alpha_i \in \mathbb{R}^+\end{gathered} \tag{2.64}$$

hence $\delta^\top = [\delta_1, ..., \delta_n, \delta_{n+1}, ..., \delta_{2n}]$ *is globally ultimately bounded with the ultimate bound given by*

$$\overline{\lim_{t\to\infty}} E^\top(t) P_2 E(t) \leq \sum_{i=1}^{n} \frac{\gamma_i}{\alpha_i} \tag{2.65}$$

where

$$P_2 = \begin{bmatrix} P_{2,1} & 0_{2\times 2} & \cdots & 0_{2\times 2} \\ 0_{2\times 2} & P_{2,2} & \cdots & 0_{2\times 2} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{2\times 2} & 0_{2\times 2} & \cdots & P_{2,n} \end{bmatrix} \tag{2.66}$$

and $\gamma_i = 2\lambda_{\max}\left\{\Lambda_{b,i}^{-1}\right\}(\Upsilon_0)$

Proof. Consider the Lyapunov candidate function given by

$$\begin{aligned} V\left(\xi, E, \bar{k}_1, \bar{k}_2\right) &= \sum_{i=1}^{n} V_i\left(\xi_i, E_i, \bar{k}_{1,i}, \bar{k}_{2,i}\right) \\ V_i\left(\xi_i, E_i, \bar{k}_{1,i}, \bar{k}_{2,i}\right) &:= V_{1,i}(\xi_i) + V_{2,i}(E_i) + V_{3,i}\left(\bar{k}_{1,i}, \bar{k}_{2,i}\right) \end{aligned} \tag{2.67}$$

with $V_{1,i}(\xi_i) = \xi_i^\top P_{1,i}\xi_i$, $V_{2,i}(E_i) = E_i^\top P_{2,i} E_i$ and $V_{3,i}\left(\bar{k}_{1,i}, \bar{k}_{2,i}\right) = \pi_{1,i}\bar{k}_{1,i}^2 + \pi_{2,i}\bar{k}_{2,i}^2$ where $\bar{k}_{1,i} = K_{P,i}(t) - K_{P,i}^*$ and $\bar{k}_{2,i} = K_{D,i}(t) - K_{D,i}^*$. The term referred to as ξ_i is described by $\xi_i = \left[|\delta_{1,i}|^{1/2}\text{sign}(\delta_{1,i}) \quad \delta_{2,i}\right]^\top$. The time derivative of each function $V_i\left(\xi_i, e_i, \bar{k}_{1,i}, \bar{k}_{2,i}\right)$ is:

$$\frac{d}{dt}V_i(t) = 2\xi_i^\top(t) P_{1,i}\frac{d}{dt}\xi_i(t) + 2E_i^\top P_{2,i}\frac{d}{dt}E_i + 2\pi_{1,i}\bar{k}_{1,i}\frac{d}{dt}\bar{k}_{1,i} + 2\pi_{2,i}\bar{k}_{2,i}\frac{d}{dt}\bar{k}_{2,i} \tag{2.68}$$

Notice that $V_{1,i}(\xi_i)$ is continuous, however it is not differentiable if $\delta_{1,i} = 0$. The ideas proposed by [89] are reconsidered here for handling the class of no regular Lyapunov functions. Let continue the analysis of the second stage $V_{2,i}(E_i)$. The calculus of the full-time derivative for the second function, one gets $\frac{d}{dt}V_{2,i}(t) = 2E_i^\top(t) P_{2,i}\frac{d}{dt}E_i(t)$ using the Dini's derivative. The direct substitution

of $\frac{d}{dt}\xi_i$ and $\frac{d}{dt}E_i$ in the expression corresponding to the full-time derivative of $V_i(t)$ yields:

$$\frac{d}{dt}V_i(t) \leq \xi_i^\top(t)\left[P_{1,i}A_{i,1}+A_{i,1}^\top P_{1,i}\right]\xi_i + E_i^\top(t)\left[P_{2,i}(A_{2,i}+\alpha_i I)+(A_{2,i}+\alpha_i I)^\top P_{2,i}\right]E_i(t) +$$

$$E_i^\top(t)\left[P_{2,i}R_{2,i}P_{2,i}+Q_{2,i}\right]E_i(t) - \alpha_i V_i(t) + \gamma_i + 2\bar{k}_{1,i}(t)\left\{\pi_{1,i}\frac{d}{dt}\bar{k}_{1,i}(t) + e_i(t)M_a^\top P_{2,i}E_i(t)\right\} +$$

$$2\bar{k}_{2,i}(t)\left\{\pi_{2,i}\frac{d}{dt}\bar{k}_{2,i}(t) + e_{i+n}(t)M_b^\top P_{2,i}E_i(t)\right\} \tag{2.69}$$

The differential inclusion appears after the application of the Young's inequality $X^\top Y + Y^\top X \leq X^\top\Lambda^{-1}X + Y^\top\Lambda Y$ $X,Y \in \mathbb{R}^{s\times q}$ with $0<\Lambda=\Lambda^\top \in \mathbb{R}^{s\times s}$ [90].

Considering the satisfaction of the condition for the existence of positive definite solutions for the given Lyapunov equations as well as the Riccati algebraic equations given in the theorem statement and in consideration of the adjusting laws for the gains $\bar{k}_{1,i}(t)$ and $\bar{k}_{2,i}(t)$ in (2.61), leads to the transformation of inclusion into: $\frac{d}{dt}V_{2,i}(t) \leq -\alpha_i V_{2,i}(t) + \gamma_i$.

Considering that the solution of the inclusion for $\frac{d}{dt}V_{2,i}(t)$, and the application of the comparison Lemma, it lead straightforwardly to $V_{2,i}(t) \leq V_{2,i}(0)\,e^{-2\alpha_i t} + \frac{\gamma_i}{\alpha_i}\left(1-e^{-2\alpha_i t}\right)$. Considering the case when the inclusion transforms to the equivalent ordinary differential equation. Redefining $V_i(t)$ as $V_i^{eq}(t)$, it can be proven that

$$V(t) \leq \sum_{i=1}^{n}\sqrt{V_{1,i}(0)} - 2t\sum_{i=1}^{n}\kappa_{1,i} + \sum_{i=1}^{n}\left(V_{2,i}(0)\,e^{-\alpha_i t} + \frac{\gamma_i}{\alpha_i}\left(1-e^{-\alpha_i t}\right)\right) \tag{2.70}$$

Clearly, when $t \geq T$ with $T = \dfrac{\sum_{i=1}^{n}\sqrt{V_{1,i}(0)}}{2\sum_{i=1}^{n}\kappa_{1,i}}$ the following inequality can be justified

$$V(t) \leq \sum_{i=1}^{n}\left(V_{2,i}(T)\,e^{-\alpha_i(t-T)}\right) + \sum_{i=1}^{n}\frac{\gamma_i}{\alpha_i}\left(1-e^{-\alpha_i(t-T)}\right) \tag{2.71}$$

If the upper limit is calculated, with $t\to\infty$ yields to $\overline{\lim_{t\to\infty}}V(t) \leq \sum_{i=1}^{n}\frac{\gamma_i}{\alpha_i}$. This result finishes the proof.

□

2.8 Actuator dynamics and voltage determination

In view of the relationships presented for the LA and the control actions obtained in Theorem 1 yields to:

$$\begin{aligned}\tau_i = L_{i,1}^f \cdot F_i^a \cdot sin(\delta_i) = L_{i,1}^f \cdot m_i^a \frac{d}{dt}\chi_{2,i}^a \cdot sin(\delta_i) = \\ L_{i,1}^f \cdot m_i^a \left(\frac{D_i}{J_i}\chi_{2,i}^a + \frac{K_{t,i}}{J_i}\chi_{3,i}^a\right) \cdot sin(\delta_i)\end{aligned} \tag{2.72}$$

Using the STA, one may substitute $\chi_{2,i}^a$ by its estimate $\hat{\chi}_{2,i}^a$, then the proposed controller may be represented as follows:

$$\hat{\tau}_i = L_{i,1}^f \cdot m_i^a \left(\frac{D_i}{J_i}\hat{\chi}_{2,i}^a + \frac{K_{t,i}}{J_i}\chi_{3,i}^a\right) \cdot sin(\delta_i) \tag{2.73}$$

To obtain the requested voltage in the LA, let introduce the difference between the needed torque $\Delta_{\tau_i} = \tau_i^* - \hat{\tau}_i$ which satisfies the following dynamics

$$\frac{d}{dt}\Delta_{\tau_i} = \frac{d}{dt}\tau_i^* - \frac{d}{dt}\left(L_{i,1}^f \cdot m_i^a \left(\frac{D_i}{J_i}\hat{\chi}_{2,i}^a + \frac{K_{t,i}}{J_i}\chi_{3,i}^a\right) \cdot sin(\delta_i)\right) \tag{2.74}$$

The estimation of the derivative applied in the equation $\frac{d}{dt}\Delta_{\tau_i}$ yields to the following representation:

$$\begin{aligned}\frac{d}{dt}\Delta_{\tau_i} = \frac{d}{dt}\tau_i^* - \left(L_{i,1}^f \cdot m_i^a \left(\frac{D_i}{J_i}\frac{d}{dt}\hat{\chi}_{2,i}^a + \frac{K_{t,i}}{J_i}\frac{d}{dt}\chi_{3,i}^a\right) \cdot sin(\delta_i)\right) + \\ \left(L_{i,1}^f \cdot m_i^a \left(\frac{D_i}{J_i}\hat{\chi}_{2,i}^a + \frac{K_{t,i}}{J_i}\chi_{3,i}^a\right) \cdot cos(\delta_i)\frac{d}{dt}\delta_i\right)\end{aligned} \tag{2.75}$$

The expression given in (2.75) can be alternatively expressed as follows

$$\frac{d}{dt}\Delta_{\tau_i} = \Phi_i^a(t) + \Gamma_i^a(t)V_{a,i}(t)$$

where the expressions for $\Phi_i^a(t)$ and $\Gamma_i^a(t)$ are given here

$$\begin{aligned}\Phi_i^a(t) = \frac{d}{dt}\tau_i + \left(\frac{L_{i,1}^f m_i^a \left(J_i K_{t,i} K_{e,i} - D_i\right)}{J_i L_{a,i}} \sin(\delta) + \frac{L_{i,1}^f m_i^a D_i}{J_i} \cos(\delta) \frac{d}{dt}\delta\right)\hat{\chi}_{2,i}^a \\ + \left(\frac{L_{i,1}^f m_i^a \left(K_{t,i} R_{a,i} - J_i D_i K_{t,i}\right)}{J_i L_{a,i}} \sin(\delta) + \frac{L_{i,1}^f m_i^a K_{t,i}}{J_i} \cos(\delta) \frac{d}{dt}\delta\right)\chi_{3,i}^a\end{aligned} \tag{2.76}$$

and

$$\Gamma_i^a(t) = \frac{L_{i,1}^f m_i^a K_{t,i}}{J_i L_{a,i}} \sin(\delta) \tag{2.77}$$

Thus, the dynamic evolution of the error can be written as:

$$\frac{d}{dt}\Delta_{\tau_i} = F_i^a\left(\delta, \hat{\chi}_{2,i}^a, \chi_{3,i}^a, t\right) + \Gamma_i^a(\delta, t) V_{a,i} \tag{2.78}$$

Then $V_{a,i}(t)$ is:

$$V_{a,i}(t) = -\left[\Gamma_i^a(t)\right]^{-1}\left[k_i^a sign(\Delta_{\tau_i}) + F_i^a\left(\delta, \hat{\chi}_{2,i}^a, \chi_{3,i}^a, t\right)\right] \tag{2.79}$$

Notice that it is straightforward to prove that the proposed voltage is enforcing the finite-time reproduction of the needed torque for controlling the proposed BRD.

Chapter 3

Adaptive controller design with finite and fixed time convergence

This chapter introduces a novel trajectory tracking robust controller for a biped robotic system with restricted articulations. The robotic device comprises a suspended biped configuration. The proposed robust control approach employs state-dependent gains, which ensure finite-time convergence of the tracking deviation. The implementation of controlled tangent barrier Lyapunov function is used to estimate the controller gains that fulfill the state restrictions. Stability analysis for the tracking error leads to the explicit design of the state-dependent gains. The forward complete characteristic of the rehabilitation device allows for the use of a finite-time convergent super-twisting-based differentiator, which enables output feedback realization of the proposed controller. A computerized model of the proposed rehabilitation robot serves as a reliable testing platform for the suggested robust controller. Numerical evaluations validate the effectiveness of the proposed design by demonstrating tracking error convergence, satisfaction of the articulation restrictions, and the efficacy of the gain optimization design. For comparison purposes, the regular state feedback control design is considered as the benchmark. The faster convergence of the mean square estimation of the tracking error justifies the design of the proposed control, while the state feedback structure proves the origin is a fixed-time stable equilibrium point for the space of tracking error. This chapter is derived from [91].

3.1 Introduction

Controlling biped robotic devices has been a subject of interest in various fields, including robotics, automatic control, and medical rehabilitation. As the applications of biped robots increase, new automatic controllers that consider the unique characteristics of these devices are being developed.

Most control designs use the regular human gait as a reference (see Section 1.2.4) to prioritize the safe headway of the biped robotic device. The ZMP [92, 93] is a crucial criterion for postural stability in biped robots. This criterion considers the aggregation of gravitational and inertial forces and aims to force a null moment of inertia. Ensuring the ZMP condition requires a robust controller that can efficiently track reference trajectories [94, 95]. The references define a sequence of articulation movements that must be restricted to the allowable articular displacements of the biped robot. These restrictions can form part of the design of the stable GC for the considered autonomous bipedal robotic device.

Despite the numerous remarkable designs of robust controllers for biped robots inspired by the ZMP criterion, many apply asymptotic controllers [96–100] and do not take into account the angular restrictions for angular movements in biped robot designs. Consequently, the gain tuning process may take a long time [101, 102], during which the biped robot may suffer undesired effects, including mechanical damages due to the transient period of the controller action. Therefore, it is desirable to implement finite-time controllers that consider the state restrictions of the biped robot [103–105].

Finite and fixed time convergent controllers are desirable for state-restricted robotic systems, such as biped robotic devices [106], due to the strict time constraints that may be present in real-world applications with predefined finite-duration working time windows [107]. Recently, set-valued functions such as the sign function have been proposed as the key component of the control design for finite and fixed time convergent controllers [108]. These controllers often employ discontinuous control functions and are typically based on sliding-mode theory [109–111]. The sliding-mode controllers drive the trajectories of the robotic device to a predefined subspace, even for high-order variants that consider augmented variables [112, 113]. Nonlinear state combinations can be used to create sliding surfaces that enforce predefined convergence velocities of states to the origin once they are on the sliding manifold [114, 115]. In addition to discontinuous controllers, recent studies have proposed diverging gains for certain classes of linear controllers and observers to ensure finite time convergence, which may also be suitable for biped robots. These systems can be regulated by diverse sliding-mode or gain-divergent control designs while still ensuring either finite or fixed time convergent approaches, provided they are state or output feedback linearizable systems with well-defined finite relative degree with respect to the available online measurable outputs [116, 117].

Theoretical results on developing feasible controllers, including gain tuning and estimation of convergence time, for the analyzed class of systems are mature. However, some studies describe the methodology for implementing these controllers with state constraints within fixed bounds known in advance [103, 118]. Currently, significant advances have been made in developing controllers that can consider state bounds [119]. Among the existing techniques, the application of bLfs is an option to consider the restricted movements for all articulations included in the biped robotic device

under study. The use of BLFs has grown in popularity in recent years due to their relevance in implementing complex nonlinear controllers on real plants. Different variations of BLFs have been proposed, including logarithmic form [57], integral structures [120], and tangent [121] functions. However, when bLfs are used as the key stability tool in the second stability method of Lyapunov, there is a need to ensure finite-time convergence of the system trajectories. One alternative is to apply polynomial-form control designs that force a non-Lipschitz form around the origin [122] or to use adaptive variants such as neural networks and fuzzy systems [123, 124]. Although these control designs may not exhibit robustness with respect to non-vanishing external perturbations, they still provide reliable transient evolution in a closed set centered on the equilibrium point of the tracking error space [125]. This study proposes a methodology for implementing a class of tangent barrier Lyapunov function with diverging gain, which yields finite-time convergence of the tracking error while satisfying the states restrictions all the time. This is a significant contribution to the control of biped robotic devices, considering their characteristics.

This chapter presents three principal contributions. Firstly, a robust controller based on a novel stability analysis is developed using a controlled tangent bLf. Secondly, the proposed tangent BLF includes fractional and high order polynomial composition of the tracking error, which justifies the existence of a finite-time equilibrium point of the tracking error dynamics between the articulation angles and their associated references. Finally, a set of numerical studies is conducted to test the suggested controller on a virtual representation of the bipedal robotic device, which confirms that the proposed controller enforces the stability conditions produced by the application of the suggested tangent BLF.

The chapter is structured as follows. In Section 3.2, it is described the simplified bipedal robotic structure. Section 3.3 provides details on the adaptive control design based on a novel tangent-type bLf and presents the corresponding proof. The proposed controller's numerical evaluations are shown in Section 3.4, where we compare it with a classical controller. Finally, in Section 3.6, we present the experimental evaluation of the controller over the mechanical structure.

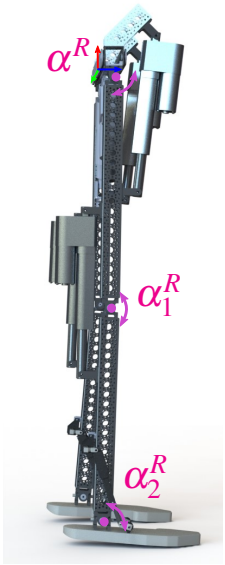
3.2 Mathematical model

Since in Chapter 2 it was demonstrated that the integration of the bLf controller into the electromechanical dynamics of the LA is possible, and the simulations will be carried out on a robot simulation platform (Simulink®), a simplified version of the model is presented in which the robot is considered as an open kinematic chain. The biped robot, which is designed for this study, is depicted in Figure 3.1. The proposed design has three links for each leg, which allows for mobilization on the sagittal

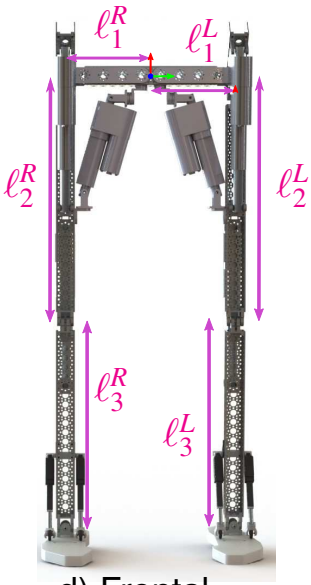
plane. The simulated robot assumes dimensions appropriate for an adult male with a height of ‘ “175 cm and weights appropriate for an aluminum structure with linear actuators attached to it (see Sections 5.4.6 and A for more details).

a) Isometric

b) Superior



c) Right



d) Frontal

of

Figure 3.1: Mechanical diagram of the proposed biped robot detailing the links and rotating joints for each leg.

According to the Euler-Lagrange methodology, the Lagrangiang function was estimated consid-

ering the calculus of kinetic $K(\dot{q},q)$ and potential energies $V(q)$ with $q \in \mathbb{R}_6$, the vector of generalized coordinates with $q = [\left(q^L\right)^\top, \left(q^R\right)^\top]^\top$, $q^L \in \mathbb{R}_3$, $q^R \in \mathbb{R}_3$. In view of the symmetry and relative independence of both legs, each of the legs can be modeled independently.

In principle, the Lagrangian of the biped robot $\mathcal{L}(\dot{q},q)$ can be calculated as the sum of the Lagrangian of the right $\mathcal{L}^R\left(\dot{q}^R,q^R\right)$ and left $\mathcal{L}^L\left(\dot{q}^L,q^L\right)$ legs, that is $\mathcal{L}(\dot{q},q) = \mathcal{L}^R(\dot{q}^R,q^R) + \mathcal{L}^L(\dot{q}^L,q^L)$. In consequence, let consider the modeling process of a single leg which can be easily adjusted for each leg. Without lost of generality, let consider the Lagrangian for the left leg:

$$\mathcal{L}^L(\dot{q}^L,q^L) = K\left(\dot{q}^L,q^L\right) - V\left(q^L\right) \tag{3.1}$$

With the application of the Euler-Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}^L}\mathcal{L}^L(\dot{q}^L,q^L)\right) - \frac{\partial}{\partial q^L}\mathcal{L}^L(\dot{q}^L,q^L) = \tau^L - Q^L_{n-i} \tag{3.2}$$

Here $\tau^L \in \mathbb{R}^3$ is the set of torques realized by the linear actuators, while $Q^L_{n-i} \in \mathbb{R}^3$ is the set of non-conservative forces (except the torques) that includes friction, damping and some others. The matrix form corresponding to the calculus of the Euler-Lagrange equations of the left leg in the biped robot corresponds to the form

$$D^L\left(q^L\right)\ddot{q}^L + C^L\left(\dot{q}^L,q^L\right)\dot{q}^L + G^L\left(q^L\right) = \tau^L + Q^L_{n-i}\left(\dot{q}^L,q^L,t\right) \tag{3.3}$$

where the inertia matrix $D^L\left(q^L\right)$ is given by

$$D^L = \begin{bmatrix} D^L_{11} & D^L_{12} & D^L_{13} \\ D^L_{12} & D^L_{22} & D^L_{23} \\ D^L_{13} & D^L_{23} & D^L_{33} \end{bmatrix}$$

The components of inertia matrix D^L are described by the following expressions:

$$\begin{aligned} D^L_{11} &= \frac{1}{4}m_1 l_1^2 + m_2 l_1^2 + \frac{1}{4}m_2 l_2^2 + m_3 l_1^2 + m_3 l_2^2 + \frac{1}{4}m_3 l_2^2 \\ &\quad + m_3 l_1 l_3 c^L_{23} + m_3 l_2 l_3 c^L_3 + m_2 l_1 l_2 c^L_2 + 2m_3 l_1 l_2 c^L_2 \\ D^L_{12} &= -\frac{1}{4}m_2 l_2^2 - m_3 l_2^2 - \frac{1}{4}m_3 l_3^2 - m_3 l_2 l_3 c^L_3 - \frac{1}{2}m_2 l_1 l_2 c^L_2 \\ &\quad - m_3 l_1 l_2 c^L_2 - \frac{1}{2}m_3 l_1 l_3 c^L_{23} \end{aligned}$$

$$
\begin{aligned}
D_{13}^{L} &= \frac{1}{4}m_3 l_3^2 + \frac{1}{2}m_3 l_1 l_3 c_{23}^{L} + \frac{1}{2}m_3 l_2 l_3 c_3^{L} \\
D_{22}^{L} &= m_3 l_2^2 + \frac{1}{4}m_3 l_3^2 + \frac{1}{4}m_2 l_2^2 + m_3 l_2 l_3 c_3^{L} \\
D_{23}^{L} &= -\frac{1}{4}m_3 l_3^2 - \frac{1}{2}m_3 l_2 l_3 c_3^{L} \\
D_{33}^{L} &= \frac{1}{4}m_3 l_3^2
\end{aligned}
$$

For the Coriolis matrix $C^{L}\left(\dot{q}^{L}, q^{L}\right)$, its formal expression is given as follows:

$$
C^{L}\left(q^{L}, q^{L}\right) = \begin{bmatrix} C_{11}^{L} & C_{12}^{L} & C_{13}^{L} \\ C_{21}^{L} & C_{22}^{L} & C_{23}^{L} \\ C_{31}^{L} & C_{32}^{L} & C_{33}^{L} \end{bmatrix}
$$

The components of $C^{L}\left(\dot{q}^{L}, q^{L}\right)$ are given by:

$$
\begin{aligned}
C_{11}^{L} &= m_1 g\frac{l_1}{2}s_1^{L} + m_2 g l_1 s_1^{L} + m_3 g l_1 s_1^{L} + m_2 g\frac{l_1}{2}s_{12}^{L} \\
&\quad + m_3 g l_2 s_{12}^{L} + m_3 g\frac{l_3}{2}s_{123}^{L} - m_3 l_1 l_3 s_{23}^{L}\dot{q}_2^{L} \\
&\quad - m_3 l_1 l_3 s_{23}^{L}\dot{q}_3^{L} \\
C_{12}^{L} &= m_3 l_2 l_3 \dot{q}_3^{L} s_3^{L} + \frac{1}{2}m_2 l_1 l_2 \dot{q}_2^{L} s_2^{L} + m_3 l_1 l_2 \dot{q}_2^{L} s_2^{L} \\
&\quad + \frac{1}{2}m_3 l_1 l_3 s_{23}^{L}\dot{q}_2^{L} - \frac{1}{2}m_3 l_1 l_3 s_{23}^{L}\dot{q}_3^{L} - m_2 l_1 l_2 \dot{q}_1^{L} s_2^{L} \\
&\quad - 2m_3 l_1 l_2 \dot{q}_1^{L} s_2^{L} \\
C_{13}^{L} &= -\frac{1}{2}m_3 l_1 l_3 s_{23}^{L}\dot{q}_2^{L} + \frac{1}{2}m_3 l_1 l_3 s_{23}^{L}\dot{q}_3^{L} - m_3 l_2 l_3 \dot{q}_1^{L} s_3^{L} \\
&\quad - \frac{1}{2}m_3 l_2 l_3 \dot{q}_3^{L} s_3^{L} \\
C_{21}^{L} &= \frac{1}{2}m_3 l_1 l_3 \dot{q}_1^{L} s_{23}^{L} + \frac{1}{2}m_2 l_1 l_2 \dot{q}_1^{L} s_2^{L} + m_3 l_1 l_2 \dot{q}_1^{L} s_2^{L} \\
&\quad + \frac{1}{2}m_3 l_1 l_3 s_{23}^{L}\dot{q}_2^{L} - \frac{1}{2}m_3 l_1 l_3 s_{23}^{L}\dot{q}_3^{L} \\
C_{22}^{L} &= \frac{1}{2}m_2 l_1 l_2 \dot{q}_1^{L} s_2^{L} - m_3 l_1 l_2 \dot{q}_1^{L} s_2^{L} - \frac{1}{2}m_3 l_1 l_3 \dot{q}_1^{L} s_{23}^{L} \\
&\quad + \frac{1}{2}m_2 l_1 l_2 \dot{q}_1^{L} s_2^{L} + m_3 l_1 l_2 \dot{q}_1^{L} s_2^{L} \\
C_{23}^{L} &= \frac{1}{2}m_3 l_1 l_3 \dot{q}_1^{L} s_{23}^{L} + m_3 l_2 l_3 \dot{q}_1^{L} s_3^{L} - m_3 l_2 l_3 \dot{q}_2^{L} s_3^{L}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2}m_3 l_2 l_3 \dot{q}_3^L s_3^L \\
C_{31}^L &= \frac{1}{2}m_3 l_2 l_3 s_3^L \dot{q}_1^L - \frac{1}{2}m_3 l_1 l_3 q_1^L s_{23}^L - \frac{1}{2}m_3 l_1 l_3 s_{23}^L \dot{q}_2^L \\
&+\frac{1}{2}m_3 l_1 l_3 s_{23}^L \dot{q}_3^L \\
C_{32}^L &= \frac{1}{2}m_3 l_2 l_3 s_3^L \dot{q}_2^L - m_3 l_2 l_3 \dot{q}_1^L s_3^L + \frac{1}{2}m_3 l_1 l_3 \dot{q}_1^L s_{23}^L \\
C_{33} &= -\frac{1}{2}m_3 l_1 l_3 \dot{q}_1^L s_{23}^L
\end{aligned}
$$

In the case of the gravitational term G, its expression corresponds to

$$
G^L\left(q^L\right) = \begin{bmatrix} 0 \\ -m_2 g\frac{l_2}{2} - m_3 g l_2 s_{12}^L - m_3 g\frac{l_3}{2}s_{123}^L \\ m_3 g\frac{l_3}{2}s_{123}^L \end{bmatrix}
$$

In the last terms, the following nomenclature is used: $s_i^L = \sin(q_i^L)$, $c_i^L = \cos(q_i^L)$ with $i = \{1,2,3\}$ $c_{23}^L = cos(q_2^L - q_3^L)$, $s_{23}^L = \sin(q_2^L - q_3^L)$, $s_{12}^L = sin(q_1^L - q_2^L)$ and $s_{123}^L = \sin(q_1^L - q_2^L + q_3^L)$.

3.3 Output feedback control

The bipedal robotic device considered in this study, represented by system (3.3), can be mathematically described using the state variable theory [85, 86]. Let $\zeta^E \alpha = q^E$ and $\zeta^E \beta = \dot{q}^E$, where $E = \{L, R\}$. Then, the mathematical description can be written as::

$$
\begin{aligned}
\frac{d}{dt}\zeta_\alpha^E(t) &= \zeta_\beta^E(t) \\
\frac{d}{dt}\zeta_\beta^E(t) &= f^E\left(\zeta^E(t),t\right) + g^E\left(\zeta_\alpha^E(t)\right)u^E(t) + \Upsilon^E\left(\zeta^E(t),t\right)
\end{aligned}
\tag{3.4}
$$

where $\left(\zeta^E\right)^\top = [\left(\zeta_\alpha^E\right)^\top, \left(\zeta_\beta^E\right)^\top]$. Here $\zeta^E \in X^E \subset \mathbb{R}^6$,where X^E is an open subspace of the state variable space. $\zeta_\alpha^E \in \mathbb{R}^3$ represents the vector of angular displacements at each joint for each leg of the biped robot, while $\zeta_\beta^E \in \mathbb{R}^3$ describes the full-time differentiation of the angular variations measured at each joint in the bipedal robotic device. The function $f^E : X^E \times \mathbb{R}^+ \to \mathbb{R}^3$ is the vector field used to represent the drift term for the bipedal robotic device, which can be obtained by applying the Euler-Lagrange representation over (3.3) and corresponds to $f^E = -\left(D^E\left(q^E\right)\right)^{-1}\left(C^E\left(\dot{q}^E, q^E\right)\dot{q}^E + G^E\left(q^E\right)\right)$.

This study does not analyze the contact interactions between the supporting surface and the foot of the biped robot. The given properties for the f vector function guarantee the validity of the locally Lipschitz condition, which is given by:

$$\begin{gathered}\|f^E\left(\zeta_1^E\right)-f^E\left(\zeta_2^E\right)\|\leq L_f\|\zeta_1^E-\zeta_2^E\|,\\ \zeta_1^E\in X^E,\zeta_2^E\in X^E,L_f\in\mathbb{R}^+\end{gathered}\tag{3.5}$$

The function $g^E:\mathbb{R}^3\to\mathbb{R}^{3\times3}$ defines the effect of the input vector over the bipedal robotic device and could be estimated as $g^E=\left(D^E\left(q^E\right)\right)^{-1}$. This structure considers a well-defined inverse in view of the inertia (positive definite) matrix associated to the bipedal robotic device. The term $\Upsilon^E:X^E\times\mathbb{R}^+\to\mathbb{R}^n$ represents the uncertainties/perturbations. This vector function can be obtained using the mechanical form (3.3) and leads to $\Upsilon^E=\left(D^E\left(q^E\right)\right)^{-1}Q_{n-i}^E\left(\dot{q}^E,q^E,t\right)$. The term Υ^E belongs to the following set:

$$\Xi=\left\{\Upsilon^E|\sup_{t\geq0,\ \zeta^E\in\mathbb{R}^3}\|\Upsilon^E\|^2\leq\Upsilon_0+\Upsilon_1\|\zeta^E\|^2\right\}\tag{3.6}$$

In view of the nature of the biped robot (mechanical structure), each angular displacement is restricted according to the following structure:

$$-\infty<\zeta_{\alpha,i}^{E,-}\leq\zeta_{\alpha,i}^E\leq\zeta_{\alpha,i}^{E,+}<+\infty,\ \ \zeta_{\alpha,i}^{E,-},\zeta_{\alpha,i}^{E,+}\in\mathbb{R}\tag{3.7}$$

According to the restrictions for the states, there is an upper limit $\zeta^{E,+}$ defined as:

$$\frac{(\max_i\zeta_i^{E,+})3}{\lambda_{min}P^E}=\zeta^{E,+}\tag{3.8}$$

Here $\zeta_i^{E,+}$ is giving the restriction for the i^{th} angular displacement and $P^E\in\mathbb{R}^{6\times6}$ is a positive definite weighting matrix. That matrix is going to be used to restrict some specific articular sections. Then, $\zeta^{E,+}$ characterizes the bounds for the generalized coordinates according to the following inequality:

$$\|\zeta^E\|_{P^E}^2\leq\zeta^{E,+}\tag{3.9}$$

3.3.1 Problem Statement for the controller design

The objective of the controller developed in this chapter is designing a controller $u^E = u^E(t)$ for the bipedal robotic device in such a way that

$$\|\zeta_a^E(t) - \zeta_a^{E,*}(t)\| \leq \varsigma^E \quad , \quad \forall t \geqslant T : 0 \tag{3.10}$$

Here, T denotes the fixed time for convergence, ς^E is the tracking error associated with the reference trajectories, ζ_a^E corresponds to the given position of all joints as explained previously, and $\zeta_a^{E,*}$ represents the target reference trajectories. The reference trajectories are determined based on the expected motion of the biped robot, often following a bio-inspired gait cycle observed in human beings. It is important that such reference trajectories are differentiable.

The generalized coordinate ζ_a^E is defined as follows

$$\zeta_a^E = |\zeta_{a,i}^E|_{i=1:3}$$

3.3.2 Control design

The control design strategy considers that the proposed reference trajectories satisfies the following dynamic form

$$\begin{aligned} \frac{d}{dt}\zeta_\alpha^{E,*}(t) &= \zeta_\beta^{E,*}(t) \\ \frac{d}{dt}\zeta_\beta^{E,*}(t) &= h^E\left(\zeta_\alpha^{E,*}(t), \zeta_\beta^{E,*}(t), t\right) \end{aligned} \tag{3.11}$$

The vector function $h^E : X^E \times X^E \times \mathbb{R}^+ \to \mathbb{R}^n$ is locally Lipschitz. Such function characterizes the reference states.

Taking into consideration that tracking error vector $\delta^E(t) \in \mathbb{R}^6$ is defined as $\delta^E = \zeta^E - \zeta^{E,*}$, where $\zeta^{E,*} = [(\zeta_\alpha^{E,*})^\top, (\zeta_\beta^{E,*})^\top]^\top$. Then, $\delta^E(t)$ dynamical behavior is given by:

$$\begin{aligned} \frac{d}{dt}\delta^E(t) = A\delta^E(t) + B\left(f^E\left(\zeta^E(t), t\right) + g^E\left(\zeta_\alpha^E(t)\right)u^E(t)\right) + \\ B\left(\Upsilon^E\left(\zeta^E(t), t\right) - h^E\left(\zeta_\alpha^{E,*}(t), \zeta_\beta^{E,*}(t), t\right)\right) \end{aligned} \tag{3.12}$$

Both system matrices A and B satisfy the companion controllable form of appropriate dimensions. Consider the adaptive controller (with diverging gain) that satisfies the following nonlinear structure form

$$u^E(t) = \left(g^E(\zeta_a^E(t))\right)^{-1} K^E(t)\delta^E(t) \tag{3.13}$$

The selected time-varying gain satisfies the following conditions:

$$\begin{gathered}\frac{d}{dt}K^E(t) = -2\lambda^{-1}c_1^E\tilde{K}^E(t)\left(\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right)^{-\frac{1}{p}} - 2\lambda^{-1}c_2^E\tilde{K}^E(t)\left(\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right)^{\frac{1}{p}} - \\ 2\lambda^{-1}\frac{p\pi}{2\zeta^{E,+}}\cdot\Gamma(\delta^E(t))B^\top P^E\delta^E(t)\delta^E(t)^\top \\ \Gamma(\zeta^E(t)) = tan^{p-1}(z(\delta^E(t))) + tan^{p+1}\left(z(\delta^E(t))\right) \\ z(\delta^E(t)) = \pi\frac{\|\delta^E(t)\|^2_{P^E}}{4\zeta^{E,+}} \qquad \tilde{K}^E(t) = K^E(t) - K^{E,*}\end{gathered} \tag{3.14}$$

with $\lambda \in \mathbb{R}^+$, the constants $c_1^E = \left(\frac{\lambda\rho p\pi}{\zeta^{E,+}}\right)^{\frac{p-1}{p}}$, $c_2^E = \left(\frac{\lambda\rho p\pi}{\zeta^{E,+}}\right)^{\frac{p+1}{p}}$, $\rho \in \mathbb{R}^+$ and $p \in \mathbb{R}$, $p \in (0,1]$, $K^E(0) = |k_s^E|_{s=1,\ldots,6}, -\infty < k_s^E < 0$. The matrix $K^{E,*} \in \mathbb{R}^{3\times 6}$ should be chosen considering that $A + BK^{E,*}$ must be a Hurwitz matrix, and taking into account that matrix A is gotten from the sublinear section of the bipedal robotic mathematical representation while the matrix B characterizes the input injection section in the robotic dynamics. Using the traditional extended Ackerman method for multivariable systems [126], the matrix $K^{E,*}$ can be estimated.

The next theorem (Theorem 2) introduces the principal result of this study associated to the gains selection for the controller based on the suggested technique which may guarantee the satisfaction of the state (angular displacement) restrictions for the bipedal robotic device operation while the finite time convergence for the tracking error to the origin is ensured.

Theorem 2. *Taking into account the type of nonlinear mechatronic systems presented in (3.12) driven by the controller form $u^E(t)$ proposed in eq. (3.13) armed with the state dependent gain $K^E(t)$ satisfying the adaptation structure in (3.14). Assume that the uncertain part of the mechatronic model Υ^E belongs to the class of nonlinear forms proposed in (3.6). If there exists a positive scalar constant γ^E and a positive definite and symmetric matrix P^E for the following matrix inequality*

$$\begin{gathered}\Sigma^E(P^E) < 0 \ \ subjected\ \ to\ \ (4) \\ \Sigma^E(P^E) = \begin{bmatrix}\Sigma_{11}^E(P^E) & P^E D \\ D^\top P^E & -Q_\Upsilon^E\end{bmatrix}\end{gathered} \tag{3.15}$$

where $\Sigma_{11}^E(P^E) = P^E(A_{K^E} + \gamma I_n) + (A_{K^E}^\top + \gamma I_n)P^E + \Upsilon_1 Q_\zeta^E$, $D = [0_{3\times 3}, I_{3\times 3}]$, $K^{E,*} \in \mathbb{R}^{m\times n}$, then the origin in the tracking error is a practically fixed-time stable equilibrium point for $\delta^E(t)$ with the

invariant and attractive zone corresponding to $\|\delta^E(t)\|^2_{P^E} \geq \varsigma^E = 2\dfrac{\Upsilon^E_0}{\alpha^E}$ *given that* $\alpha^E > 0$.

Proof. Consider the suggested Lyapunov candidate function (satisfying the tangent form of barrier Lyapunov structure) as follows:

$$\begin{gathered} V := V^L(\delta^L(t)) + V^R(\delta^R(t)) \\ V^E = V^R \ \ or \ \ V^E = V^L \\ V^E(\delta^E(t)) = tan^p\left(z(\delta^E(t))\right) + \lambda tr\left\{\left(\tilde{K}^E(t)\right)^\top \tilde{K}^E(t)\right\} \end{gathered} \tag{3.16}$$

The time derivative (Dini's type) of the Tangent Barrier Lyapunov-like function corresponds to

$$\frac{d}{dt}V^E(\delta^E(t)) = p\cdot tan^{p-1}\left(z(\delta^E(t))\right)\frac{d}{dt}tan\left(z(\delta^E(t))\right) + 2\lambda tr\left\{\left(\tilde{K}^E(t)\right)^\top \frac{d}{dt}\tilde{K}^E(t)\right\} \tag{3.17}$$

where

$$\frac{d}{dt}tan\left(z(\delta^E(t))\right) = \frac{\pi sec^2\left(z(\delta^E(t))\right)}{2\left(\zeta^{E,+}\right)}\left(\delta^E(t)\right)^\top P^E\frac{d}{dt}\delta^E(t) \tag{3.18}$$

Consequently, the following differential equation is well posed for the time derivative of the barrier Lyapunov-like function V^E associated to:

$$\frac{d}{dt}V^E(\delta^E(t)) = \frac{p\pi}{2\zeta^{E,+}}\cdot\Gamma(\zeta^E(t))\left(\delta^E(t)\right)^\top P^E\frac{d}{dt}\delta^E(t) + 2\lambda tr\left\{\left(\tilde{K}^E(t)\right)^\top \frac{d}{dt}K^E(t)\right\} \tag{3.19}$$

The term $\left(\delta^E(t)\right)^\top P\dfrac{d}{dt}\delta^E(t)$ needs the substitution of $\delta^E(t)$ dynamics proposed in (3.12), yielding

$$\left(\delta^E(t)\right)^\top P\frac{d}{dt}\delta^E(t) = \left(\delta^E(t)\right)^\top P^E\left(A\delta^E(t) + B\left(w^E(t) + g^E\left(\zeta^E_\alpha(t)\right)u^E(t)\right)\right)\delta^E(t) \tag{3.20}$$

where $w^E(t) = f^E\left(\zeta^E(t),t\right) + \Upsilon^E\left(\zeta^E(t),t\right) - h^E\left(\zeta^{E,*}_\alpha(t),\zeta^{E,*}_\beta(t),t\right)$. The equation (3.20) admits the following representation:

$$\begin{gathered} \left(\zeta^E\right)^\top P^E\frac{d}{dt}\zeta^E = \frac{1}{2}\left(\eta^E(t)\right)^\top\begin{bmatrix} P^EA^E_K + \left(A^E_K\right)^\top P^E & P^ED \\ D^\top P^E & 0_{n\times n}\end{bmatrix}\eta^E(t) + \\ \left(\delta^E(t)\right)^\top P^EB\tilde{K}^E(t)\delta^E(t) \end{gathered} \tag{3.21}$$

where $\eta^E = \begin{bmatrix}\left(\zeta^E\right)^\top & \left(\Upsilon^E\right)^\top\end{bmatrix}^\top$ and $A^E_K = A + BK^{E,*}$. Including the upper-bound of the Υ^E introduced in (3.6) and considering that $\sigma_1^\top = \sigma_2 = tr\left\{\sigma_2\sigma_1^\top\right\}$ with $\sigma_1 \in \mathbb{R}^n$ and $\sigma_2 \in \mathbb{R}^n$ transforms

the expression (3.21) into

$$\left(\zeta^E\right)^\top P^E \frac{d}{dt}\zeta^E \leq \frac{1}{2}\left(\eta^E\right)^\top \Sigma^E\left(P^E\right)\eta^E + \Upsilon_0^E + tr\left\{\tilde{K}^E(t)B^\top P^E \delta^E(t)\left(\delta^E(t)\right)^\top\right\} \tag{3.22}$$

Taking into account the obtained result in (3.22) and substituting it in the differential equation (3.19) yields

$$\begin{aligned}\frac{d}{dt}V^E(\delta^E(t)) \leq \frac{p\pi}{2\zeta^{E,+}}\cdot\Gamma(\delta^E(t))\left[\frac{1}{2}\left(\eta^E\right)^\top\Sigma^E\left(P^E\right)\eta^E + \Upsilon_0^E - \alpha^E\left(\delta^E(t)\right)^\top P^E\delta^E(t)\right] + \\ tr\left\{\tilde{K}^E(t)\left(\frac{p\pi}{2\zeta^{E,+}}\cdot\Gamma(\delta^E(t))B^\top P^E\delta^E(t)\left(\delta^E(t)\right)^\top + 2\lambda\frac{d}{dt}K^E(t)\right)\right\}\end{aligned} \tag{3.23}$$

Reorganizing the expression (3.23) allows to introduce the adaptive law for K^E as follows

$$\begin{aligned}\frac{d}{dt}V^E(\zeta^E(t)) \leq \frac{p\pi}{2\zeta^{E,+}}\cdot\Gamma(\delta^E(t))\cdot \\ \left[\frac{1}{2}\left(\eta^E\right)^\top\Sigma^E\left(P^E\right)\eta^E + \Upsilon_0^E - \alpha^E\left(\delta^E(t)\right)^\top P^E\delta^E(t)\right] + \\ tr\left\{\left(\tilde{K}^E(t)\right)^\top\left(\frac{p\pi}{2\zeta^{E,+}}\cdot\Gamma(\delta^E(t))B^\top P^E\delta^E(t)\left(\delta^E(t)\right)^\top + 2\lambda\frac{d}{dt}K^E(t)\right)\right\} + \\ tr\left\{c_1^E\left(\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right)^{\frac{p-1}{p}} + c_2^E\left(\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right)^{\frac{p+1}{p}}\right\} \\ -c_1^E tr\left\{\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right\}^{\frac{p-1}{p}} + c_2^E tr\left\{\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right\}^{\frac{p+1}{p}}\end{aligned} \tag{3.24}$$

Considering the result proposed in (3.15) and using the state restriction given as follows $\|\delta^E\|^2_{P^E} \geq 2\frac{\Upsilon_0^E}{\alpha^E}$, yielding (with $\rho > 0$)

$$\begin{aligned}\frac{d}{dt}V^E(\delta^E(t)) \leq -\frac{\rho p\pi}{2\zeta^{E,+}}\cdot\Gamma(\delta^E(t)) - c_1^E tr\left\{\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right\}^{\frac{p-1}{p}} - \\ c_2^E tr\left\{\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right\}^{\frac{p+1}{p}} + tr\left\{c_1^E\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\left(\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right)^{\frac{-1}{p}}\right\} + \\ tr\left\{c_2^E\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\left(\left(\tilde{K}^E(t)\right)^\top\tilde{K}^E(t)\right)^{\frac{1}{p}}\right\} + \\ tr\left\{\left(\tilde{K}^E(t)\right)^\top\left(\frac{p\pi}{2\zeta^{E,+}}\cdot\Gamma(\zeta^E(t))B^\top P^E\delta^E(t)\left(\delta^E(t)\right)^\top + 2\lambda\frac{d}{dt}K^E(t)\right)\right\}\end{aligned} \tag{3.25}$$

Taking into account the barrier Lyapunov type function in (3.16) leads to:

$$\frac{d}{dt}V^E(\delta(t)) = \frac{p\pi}{2\zeta^{E,+}} \cdot \left(\delta^E(t)\right)^\top P^E \delta^E(t) \left(\left(V^E\right)^{\frac{p-1}{p}} \left(\zeta^E(t)\right) + \left(V^E\right)^{\frac{p+1}{p}} \left(\delta^E(t)\right) \right) + 2\lambda tr\left\{ \left(\tilde{K}^E(t)\right)^\top \frac{d}{dt} K^E(t) \right\} \tag{3.26}$$

Considering the triangle inequality for the p-norm, and using structure of the energetic function proposed (3.16) yields

$$\frac{d}{dt}V^E(\delta^E(t)) \leq \frac{-\rho p\pi}{2\zeta^{E,+}} \left(\left(V^E\right)^{\frac{p-1}{p}} \left(\delta^E(t)\right) + \left(V^E\right)^{\frac{p+1}{p}} \left(\delta^E(t)\right) \right) +$$

$$tr\left\{ c_1^E \left(\tilde{K}^E(t)\right)^\top \tilde{K}^E(t) \left(\left(\tilde{K}^E(t)\right)^\top \tilde{K}^E(t) \right)^{\frac{-1}{p}} + c_2^E \left(\tilde{K}^E(t)\right)^\top \tilde{K}^E(t) \left(\left(\tilde{K}^E(t)\right)^\top \tilde{K}^E(t) \right)^{\frac{1}{p}} \right\} +$$

$$tr\left\{ \tilde{K}^\top(t) \left(\frac{p\pi}{\zeta^{E,+}} \cdot \Gamma(\zeta^E(t)) B^\top P^E \delta^E(t) \left(\delta^E(t)\right)^\top + 2\lambda \frac{d}{dt} K^E(t) \right) \right\} \tag{3.27}$$

If the adaptation law for $\tilde{K}^E(t)$ (which is described in 3.14) is substituted in (3.27), yileds

$$\frac{d}{dt}V^E(\delta^E(t)) \leq \frac{-\rho p\pi}{2\zeta^{E,+}} \left(\left(V^E\right)^{\frac{p-1}{p}} \left(\delta^E(t)\right) + \left(V^E\right)^{\frac{p+1}{p}} \left(\delta^E(t)\right) \right) \tag{3.28}$$

Noticing that, $0 < \frac{p-1}{p} < 1$ and $\frac{p+1}{p} > 1$, hence the origin of the tracking error space is a finite-time practically stable equilibrium point considering the preliminaries presented in [127]. □

The fixed-time stability analysis outlined in Theorem 2 is a result of the fact that the temporal evolution of $K^E(t)$ increases when $\delta^E(t)$ approaches the boundary of the invariant set. It may seem counterintuitive to use unbounded gains when aiming for fixed-time regulation, especially when dealing with unbounded inputs. However, upon examining the control design formulation, we can see that the product of the gain and the tracking error $K \cdot \delta$ can be viewed as a scaled state that remains bounded due to the fixed-time convergence, as well as the stability analysis based on the barrier-like energetic function.

Remark 1. *: The proposed controller design utilizes a time-varying gain, denoted as K^E, which is computed based on the developed adaptive law. The adaptive law involves three free parameters, namely λ, c_1^E, and c_2^E, which act as adjusting gains for the time evolution of K^E. These parameters*

are selected in a manner such that the reference trajectory can be accurately tracked within the finite time period, provided that the matrix inequality specified in Theorem 2 is satisfied. The scalar p also plays a crucial role in the tracking performance, as it controls the rate of decay of the difference between time-varying K^E and $K^{E,*}$.

Remark 2. *As Υ_0^E varies with the set Ξ and is equal to 0, it follows that the origin of the tracking error is a finite-time stable equilibrium point. This result is directly related to the proposition presented in Theorem 2 in this chapter.*

Remark 3. *In this study, we have taken into consideration the effect of perturbations satisfying $\|\delta^E\|_{P^E}^2 \geq 2\dfrac{\Upsilon_0^E}{\alpha^E}$, which could appear as a restrictive condition. Notice that if bounded perturbations may be considered for the class of biped robots, then the controller could be modified to include a pure sliding mode controller that can reject such class of perturbations. Then, the classical of the sliding mode theory can be useful to reject a more general class of perturbations including both the variant presented in this study as well as the bounded case.*

3.3.3 Output feedback form for the control design proposed for the bipedal robotic device

The sliding mode based super-twisting algorithm (STA) has been applied as a reliable method to obtain the on-line estimation for the time-derivative for continuous variables. This algorithm is defined by a couple of differential equations [89, 128] governed by:

$$\begin{aligned} \frac{d}{dt}\hat{z}_a^i(t) &= \hat{z}_b^i(t) + k_1^i |e^i(t)|^{\frac{1}{2}} sign(e^i(t)) \\ \frac{d}{dt}\hat{z}_b^i(t) &= k_2^i sign(e^i(t)) \end{aligned} \tag{3.29}$$

Here $\hat{z}_a^i$ and $\hat{z}_b^i$ are the state variables of the differentiator. The index i characterizes the variable that must be differentiated ($i = 1, \ldots, 3$) means to derive each of the angular displacements in order to get the corresponding angular velocities. The differentiator error is represented by the variable e^i that corresponds to $e^i = z_{\alpha,i}^E - \hat{z}_{\alpha,i}^E$. In view of the angular position variables could be measured continuously, these differentiators working in parallel would yield the estimation of the angular velocities that can be inserted as part of the proposed controllers. For the purposes of this study, the sign

function satisfies $|sign(z)| \leq 1, \ \forall z \in \mathbb{R}$. According to the result presented by [89], it can be proven that

$$\left|e^i\right| = 0, \ \left|\frac{d}{dt}e^i\right| = 0 \ \forall t \geqslant T^{*,i} = \sqrt{\frac{0.5\mathcal{S}^i(0)}{2}} \tag{3.30}$$

The function $\mathcal{S}^i$ satisfies the expression $\mathcal{S}^i(\xi^i) = \left[\xi^i\right]^\top \mathcal{R}^i \xi^i$ with $\left[\xi^i\right]^\top = \left[|e_1^i|^{\frac{1}{2}} sign(e_1^i) \ \ e_2^i\right]$. Each matrix $\mathcal{R}^i \in \mathbb{R}^{2\times 2}$ is a positive definite and symmetric matrix. The application of STA as part of the closed loop output feedback control realization can be consulted in several studies including [129, 130]. Please see these studies and the references therein for more details.

3.4 Virtual platform

The first experiments that were suggested to evaluate the proposed controller used a virtualized version of the biped robot (Figure 3.2). The virtualized mechanical model was proposed in SolidWorks® the computer assisted design software. The assembled model corresponding to the bipedal robotic device was transferred to Matlab® applying the Sim-Mechanics Toolbox®. Such numeric simulations were significant to evaluate the control design developed in this study. This virtual model representation allowed to evaluate the effectiveness of the suggested control structure.

The proposed virtualized structure considered all the mechanical sections in the biped robot, as well as the main movement elements. Such realistic configuration permitted to evaluate the suggested controller and comparing it with a classical state feedback approach, which is usually known as proportional-differential control. The virtual robot considers the actual masses measured in the experimental version of the robotic device (see next subsection). Also, the effect of the actuator was considered implementing the electromechanical characteristics of such device.

The controller operation over the biped robotic device can be implemented considering the structural form depicted in Figure 3.3.

3.5 Numerical evaluations

A simplified (considering the proposed articulations) representation of the biped robotic device, that is designed using the Computer Aided Design software SolidWorks® taking into account the parameters of the physical mechanical structure. The designed model is instrumented as testing numerical model for the proposed controller. The control design eq. (3.13) was tested using the simplified version in MATLAB® in the numerical simulated environmental Simulink®.

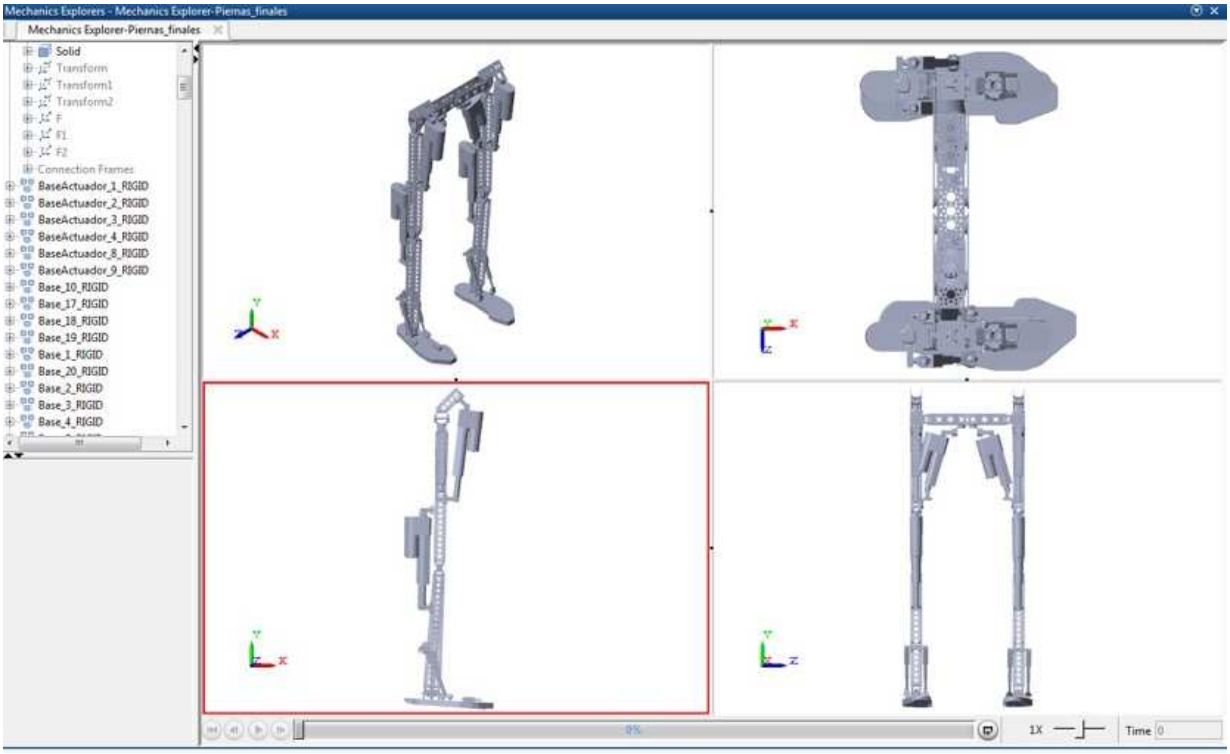


Figure 3.2: Simulated version of the biped robot exported to Matlab from the SolidWorks based on the application of the Sim-Mechanics Toolbox.

The evaluated settings of the numerical simulation were: fixed integration step at $10\mu\ s$, considering integration solver method was selected as Runge-Kutta. With the aim of comparing the performance of the suggested controller, a classical state feedback control form was also evaluated with a decentralized structure. The comparison of the trajectory tracking error is carried out in a single lower limb for evaluation purposes. The reference states were proposed with the aim of guaranteeing the fulfillment of ZMP criterion.

Figure 3.4 depicts all results of the numerical simulation that evaluates the tracking of reference states for the hip joint in the bipedal robotic device. Noticing that in comparison to the classical state feedback control form, the tracking of the reference trajectories is achieved within a time window of two seconds earlier. Nevertheless, state feedback control is not able for providing (theoretically formalized), the finite-time robust convergence for the reference tracking deviation to the estimated invariant and attractive set detailed in the main outcome proposed in this study. A general comparison for a period of 10 seconds is shown in Figure 3.4a. With the aim of demonstrating the

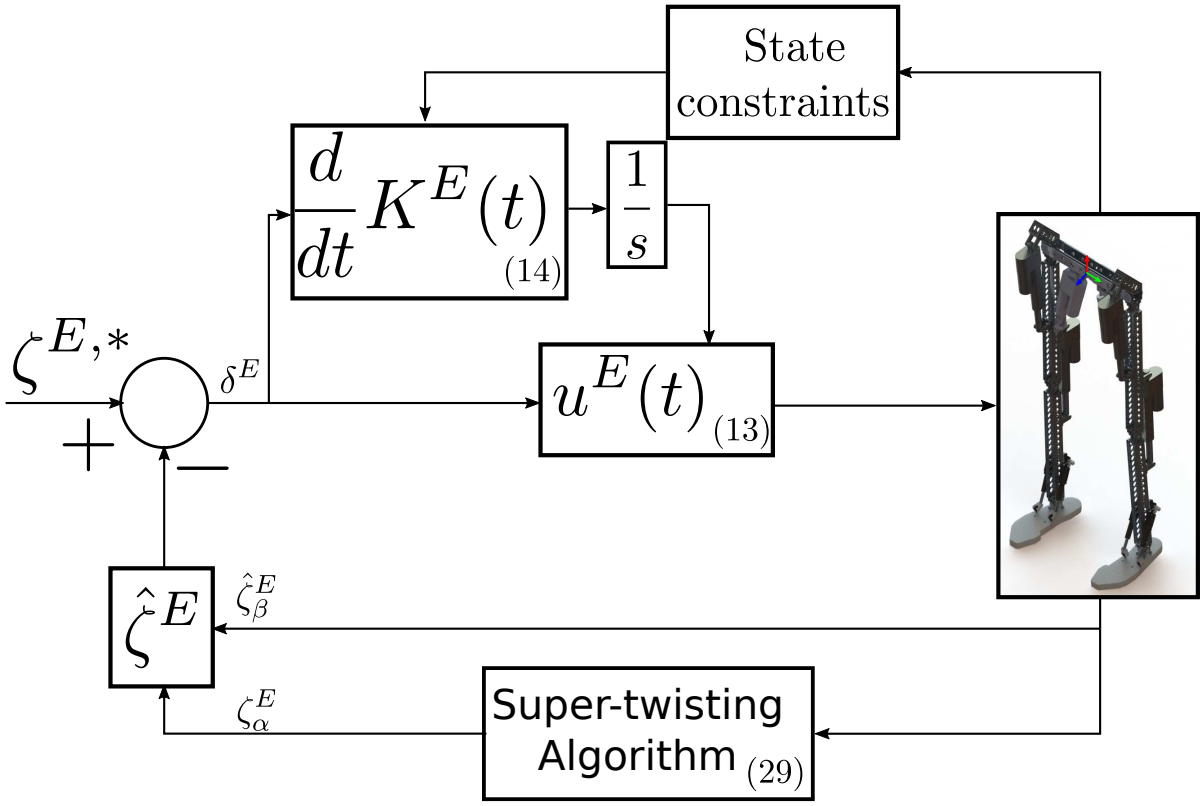
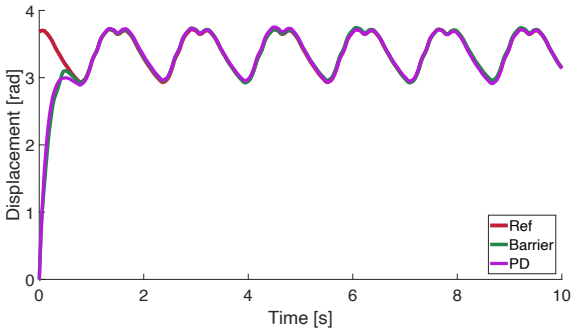


Figure 3.3: Operational diagram of the proposed finite-time controller based on the application of tangent barrier Lyapunov function.

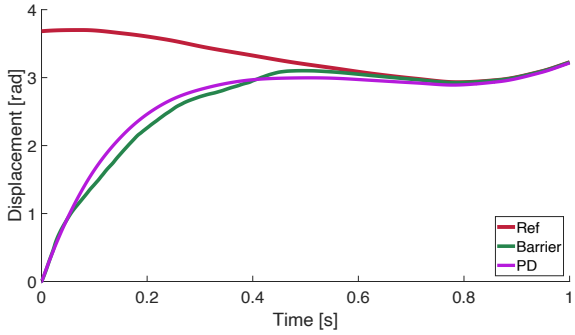
benefits of using the proposed controller, Figure 3.4b demonstrates the comparison of the tracking trajectories during the first second of the numerical simulation. A comparison of the reference states as well as the controlled trajectories during the period between 4.3 to 5.0 seconds demonstrates the steady state tracking evaluation of the diverging gain control action with respect to the state feedback formulation (Figure 3.4c).

Figure 3.5 shows the outcomes of the numerical simulation that analyzes the tracking performance for the reference states for the knee joint of the right leg of the bipedal robotic device. Noticing that the comparison for the regular state feedback control action, the tracking of reference state is achieved a couple of seconds later. A general comparison for a period of 10 seconds is shown in Figure 3.5a. With the aim of demonstrating the benefits of using the proposed controller, Figure 3.5b demonstrates the contrast between the tracking trajectories for the knee during the first two seconds of the numerical simulation. A comparison of reference as well controlled states for the knee during the period between 3.9 to 4.8 seconds demonstrates the steady state tracking evaluation of the diverging gain control action with respect to the state feedback formulation (Figure 3.5c).

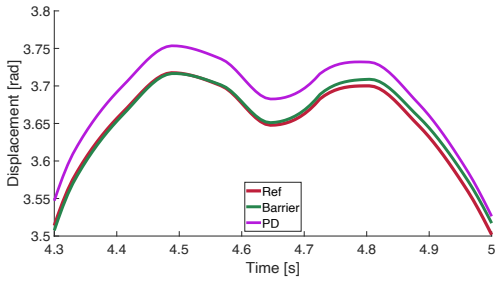
Figure 3.6 shows the comparison of the controller actions obtained with the proposed form and



(a) Evaluation of trajectory tracking between the suggested finite-time and state feedback control at the hip joint in the bipedal robotic device



(b) Evaluation of trajectory tracking between the suggested finite-time and state feedback control at the hip joint in the bipedal robotic device, closed view: 0.0-1.0 s

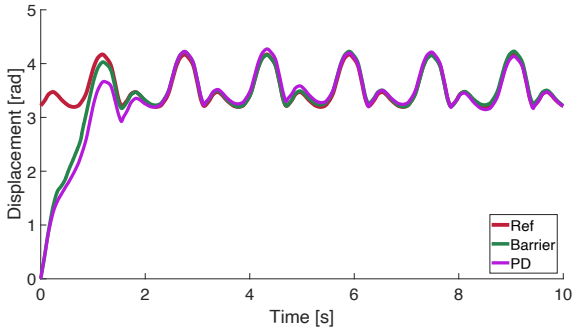


(c) Evaluation of trajectory tracking between the suggested finite-time and state feedback control at the hip joint in the bipedal robotic device, closed view: 4.3-5.0 s

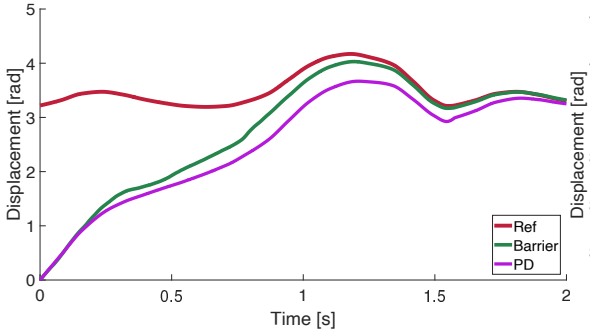
Figure 3.4: Trajectory tracking evaluation for the hip joint with the proposed controller and the state feedback formulation.

the state feedback structure. Figures 3.6a as well as 3.6b show the comparison of temporal evolution for logarithmic norm (control power) needed to develop tracking of the reference states for hip as well as knee articulations. The comparison confirms an equivalent control power consumption for these articulations in the lower limb. Nonetheless, taking into account that the proposed control based on the suggested finite-time convergent barrier form consumes less energy if for the moments when the reference states change the movement direction that is a significant characteristic with respect to the actual operation for the bipedal robotic device.

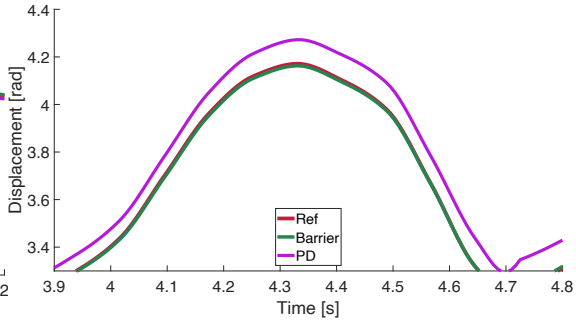
Figure 3.7 depicts the norm of the tracking error for the proposed controller and its comparison with the state feedback structure. Figures 3.7a and 3.7b depict the temporal evolution of the norms for the tracking errors between reference and biped robot trajectories for hip and knee joints, respectively. Such simultaneous evaluations confirm that the suggested control design endorses the finite-time convergence of the actual robot states to the reference trajectories as well as the state



(a) Evaluation of trajectory tracking between the suggested finite-time and state feedback control at the knee joint in the bipedal robotic device



(b) Evaluation of trajectory tracking between the suggested finite-time and state feedback control at the knee joint in the bipedal robotic device, closed view: 0.0-1.0 s



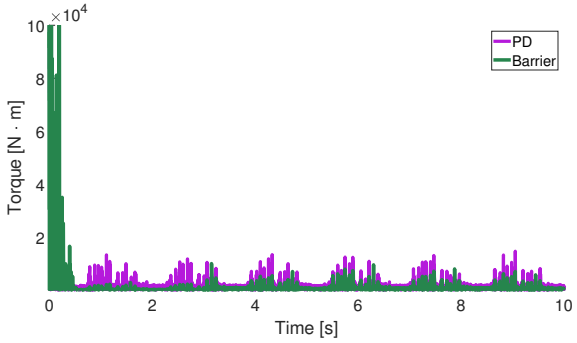
(c) Evaluation of trajectory tracking between the suggested finite-time and state feedback control at the knee joint in the bipedal robotic device, closed view: 4.3-5.0 s

Figure 3.5: Trajectory tracking evaluation for the knee joint with the proposed controller and the state feedback formulation.

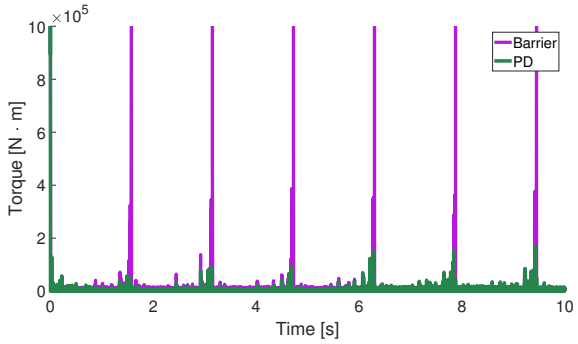
feedback does. However, the proposed approach ensures the finite-time convergence while the state restrictions are justified. Notice that this condition restricts the consumed control energy which is needed to track the reference trajectories corresponding to a regular gait cycle for the bipedal robot. In spite this negative condition, the suggested control design succeeds to solve the tracking of the references in finite-time.

With the aim of evaluating the performance of both controllers, this study evaluates the L_2 norm corresponding to the control function as well as the tracking error for each articulation in the bipedal robotic device (Table 3.1).

The evaluation of the performance index J_e in Table II was calculated as $J_e = \int_0^{t_f} |e^2(t)|dt$. Here e corresponds to the vector of tracking errors measured at both articulations, hip and knee. Likewise, J_c was calculated as $J_c = \int_0^{t_f} |u^2(t)|dt$ with u is the control vector formed with the control signals

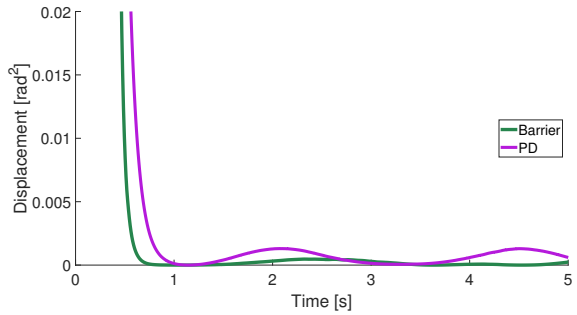


(a) Comparison of the logarithmic of the norm for the control signal: the suggested finite control and the PD controllers for the hip articulation.

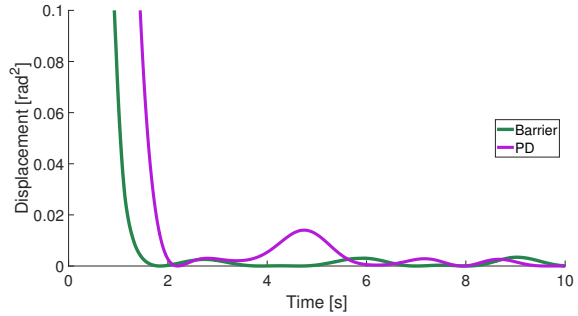


(b) Comparison of the logarithmic of the norm for the control signal: the suggested finite control and the PD controllers for the knee articulation.

Figure 3.6: Logarithmic norm of the estimated controllers for the hip and the knee joints: the proposed controller and the state feedback formulations.



(a) Comparison of the error norm signal between the proposed and the PD controllers for the hip articulation.



(b) Comparison of the error norm signal between the proposed and the PD controllers for the knee articulation.

Figure 3.7: Comparison of the Euclidean norms of the tracking errors for the knee joint with the proposed controller and the state feedback form.

Table 3.1: Value of the L_2 norm corresponding to the tracking error and control actions of both robotic joints.

Characteristics	Controller	
	Barrier	PD
$J_{e,h}$	2.5892×10^4	9.0636×10^3
$J_{c,h}$	7.4752×10^6	9.6471×10^6
$J_{e,k}$	1.8016×10^4	1.1642×10^3
$J_{c,k}$	8.1580×10^5	1.0912×10^6

at the hip and the knee implementing the proposed and the state feedback forms. The norm $J_{c,h}$ associated to the control signal using the state feedback realization corresponds to 1.2905 greater than the result obtained with the proposed finite-time controller. Alike, the L_2 norm signal of the state feedback controller for the knee articulation is 1.3376 times larger than the evaluated norm

calculated by the controller forced by the barrier Lyapunov function. The tracking errors for both the hip and the knee articulations, $J_{e,h}$ and $J_{e,k}$ respectively, produced by the controller based on the barrier Lyapunov function are 2.8567 and 15.4752 times larger than the one produced by the state feedback control action at the given articulation.

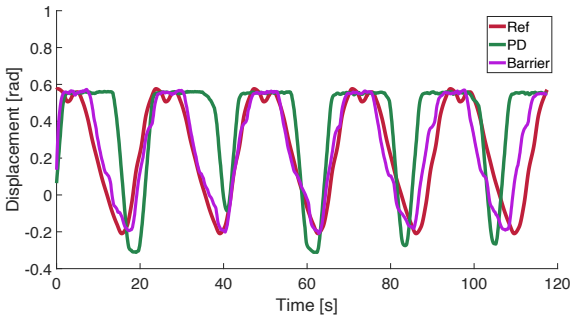
3.6 Experimental evaluations over the mechanical structure

An experimental evaluation was conducted on a self-developed biped robot to test the proposed controller, which is compatible with the SolidWorks® CAD software. The physical parameters of the bipedal robotic device were taken into account, including information on the aluminum links and actuators. The suggested control law in equation eq. (3.13) was tested using a discretized version implemented in the TIVA board, with a fixed-step of 10μs seconds and an Euler discretization methodology. To compare the performance of the suggested controller, the proposed state feedback form was also implemented in the experimental structure of the biped robot, and a direct comparison of the trajectory tracking was carried out for both lower limbs. Reference trajectories were proposed to ensure continuous satisfaction of the ZMP criterion, which was also met by the states of the bipedal robotic device implementing the suggested controller.

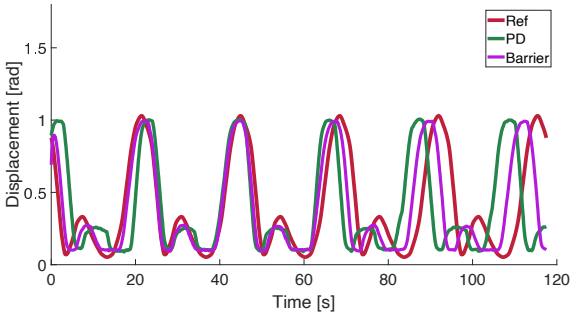
The experimental results presented in Figure 3.8 validate the effectiveness of the proposed controller in tracking the reference states for the hip (shown in Figure 3.8a) and knee (shown in Figure 3.8b) joints of the right leg of the biped robot. It can be observed that the actual robotic state does not converge to the reference when using the traditional state feedback controller, while the proposed controller ensures accurate tracking of the reference trajectory. Additionally, the state feedback implementation experiences delays and is unable to continuously track the reference. This comparison highlights the significant improvement in tracking performance provided by the proposed controller, at least for the right leg of the robot.

Figure 3.9 displays the experimental results of tracking the reference trajectories for the left leg of the biped robot. Specifically, Figure 3.9a shows the tracking of the reference trajectory for the hip articulation, while Figure 3.9b shows the tracking of the reference trajectory for the knee articulation. Comparing the results of the proposed controller with the traditional state feedback controller, the actual robotic state using the latter does not converge to the reference, while the proposed controller accurately enforces the tracking to the reference. These results demonstrate that, at least for the left leg, the proposed controller significantly improves the tracking performance compared to the traditional state feedback realization.

Figure 3.10 shows a comparison between the norms of the controller actions obtained with the

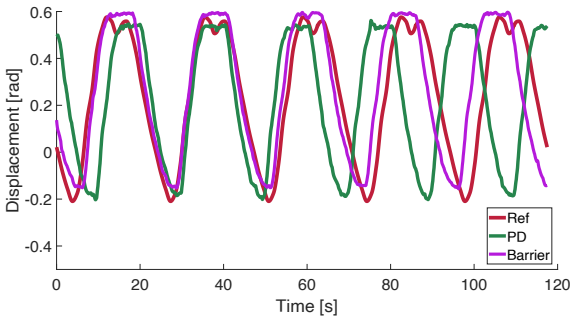


(a) Comparison of trajectory tracking performance between the proposed and PD controllers at the right hip joint in the biped robot

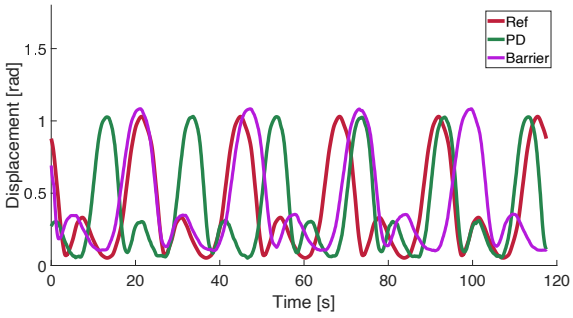


(b) Comparison of trajectory tracking performance between the proposed and PD controllers at the right knee joint in the biped robot

Figure 3.8: Trajectory tracking for the hip and the knee joints of the right leg with the proposed and the state feedback formulations. Comparison of the trajectories forced by the suggested controllers including the evaluation of both the transient and the steady state periods.



(a) Comparison of trajectory tracking performance between the proposed and PD controllers at the left hip joint in the biped robot

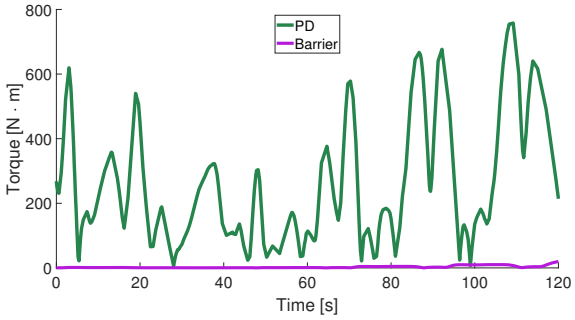


(b) Comparison of trajectory tracking performance between the proposed and PD controllers at the left knee joint in the biped robot

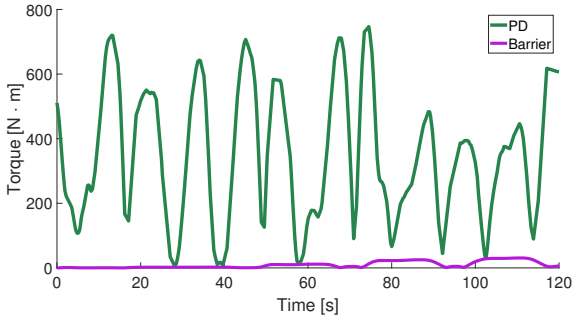
Figure 3.9: Trajectory tracking for the hip and the knee joints of the left leg with the proposed and the state feedback formulations.

proposed controller and the state feedback controller. Figures 3.10a and 3.10b depict the temporal evolution of the control power required to achieve the reference tracking for the hip and the knee joints. This comparison shows that the control norm for both joints is significantly different between the two controllers. The suggested barrier Lyapunov controller consumes smaller amounts of energy compared to the state feedback controller, which is a significant advantage for the operation of biped robots in real-world applications.

Figure 3.11 presents the norm of the tracking errors of the proposed controller and compares it with the state feedback structure. Figures 3.11a and 3.11b show the temporal evolution of the norms of the tracking errors considering the reference and experimental trajectories of the biped robot. These simultaneous evaluations confirm that the proposed control action supports the reference



(a) Comparison of the control norm signal between the proposed and PD controllers at the right leg in the biped robot

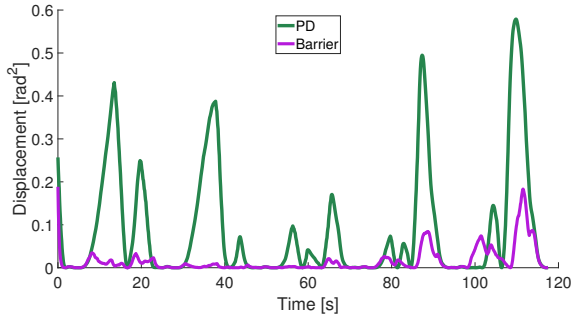


(b) Comparison of the control norm signal between the proposed and PD controllers at the left leg in the biped robot

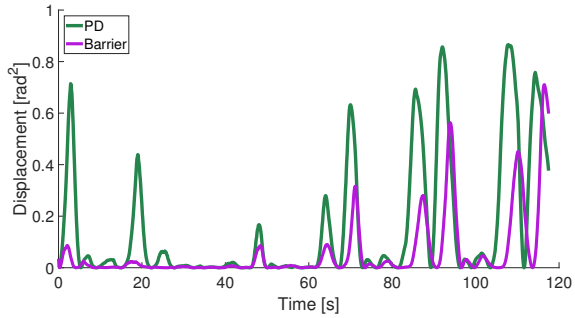
Figure 3.10: Norm of the estimated controllers for the right and left limbs: the proposed controller and the state feedback formulations.

tracking as well as the traditional state feedback does, but with the additional benefits of satisfying the finite time convergent and the state restrictions.

It should be noted that the proposed state feedback controller does not ensure the attraction of the origin for the tracking error dynamics. Additionally, both controllers exhibit a similar rate of convergence to zero at the beginning of the numerical evaluations, which is an interesting characteristic of the proposed controller even with the incorporation of an adaptive gain.



(a) Contrasting between tracking errors measured by the norm between the proposed and PD controllers at the right leg in the bipedal robotic device.



(b) Contrasting between tracking errors measured by the norm between the proposed and PD controllers at the left leg in the bipedal robotic device.

Figure 3.11: Contrasting between tracking errors measured by the norm between the proposed and PD controllers at both legs in the bipedal robotic device.

Chapter 4

Adaptive Barrier Control with Reinforcement Learning Gains Adjustment

This chapter describes a RL strategy to improve trajectory tracking of bipedal robotic systems with articulation restrictions, using robust controllers. The RL approach builds on state-dependent gains developed earlier in the chapter. The strategy involves two levels of adjustment: first, an adaptive dynamic programming approach that approximates Bellman's function using neural networks, and second, an iterative methodology based on deep deterministic policy gradient with a similar strategy as Markov chains. The proposed controller optimizes a cost function based on tracking trajectories for a system with an uncertain mathematical model that satisfies nonlinear perturbed systems that match the biped robot's dynamics. The dynamic programming formulation enables the robust control of bounded modeling uncertainties and disturbances. The HJB equation's value function is approximated by a class of neural network, allowing for the estimation of the closed-loop controller formulation. Numerical evaluations validate the convergence of the tracking error. The regular state feedback control design is used as a benchmark for comparison purposes.

4.1 Introduction

Designing RL methods involves characterizing and controlling data-driven approximate models of input-output or input-state relationships that can optimize a predefined reward functional. In its most fundamental way, RL schemes are sampled-discrete or sampled-continuous closed-loop, assuming that sampling occurs at the moment when a reward-penalty evaluation updates the control realization for the discrete or continuous system under analysis. Indeed, the learning algorithm does not have any predefined action to take aiming for optimization of the performance functional. In the last

years, two main trends have appeared to update the controller action at the given sampling time: a) based on machine learning tool to estimate the update based on parametric optimization methods or, b) discovering what actions lead to the best attainable reward by trying and testing them out under safe operative scenarios. In many interesting and challenging RL cases, the updated actions could affect not only the immediate reward (evaluated through the proposed functional), but also the next situation and, through that, all subsequent rewards. These characteristics can be framed as part of a closed-loop sub-optimal solution essentially.

Essentially, a walking robot created by humans is a robotic manipulator that has a movable and detachable base. The design of bipedal robots has been heavily influenced by the human body, which is the most sophisticated and versatile biped known to mankind. As a result, most of the models or machines developed exhibit a strong resemblance to the human body, with two lower limbs connected through a central member being a common characteristic of almost any model or machine.

RL is significantly different from supervised learning, chiefly an extended form of learning in most current investigations in the machine learning framework. Supervised learning acquires knowledge from a given training set that includes labeled instances of the input-output relationships. These labels are supplied by an external supervisor, which may not have the option to change such labels online. Each instance gathers a complete description of an operating situation associated with a labeled *correct* action that should be taken into account by the system or more properly, by the related control action.

The object of the learning is to create a kind of interpolation between the instances in the set and eventually to provide the extrapolation beyond the set boundaries (usually known as a generalization), which forces the learner to act correctly in situations that have not been considered in the provided training set. This is the most developed kind of learning, but it is not suitable for exerting learning activities from environment-robot interactions. In a such class of scenarios, it is impractical to capture examples of the given desired behavior that are both correct and representative of all the conditions in which the biped robot device has to exert its activities.

RL is distinct from unsupervised learning, which seeks to uncover structure in unlabeled data. While RL does not rely on examples of correct behavior, it is fundamentally different in that it seeks to maximize a reward signal rather than discover hidden patterns. While discovering structure in an agent's experience can aid in RL, it does not directly address the problem of maximizing a reward signal. Therefore, we consider RL to be a distinct machine learning paradigm, in addition to supervised and unsupervised learning, and potentially others.

Bipedal locomotion systems can range from simple structures with three-point masses connected by massless links [131] to highly complex structures that mimic the human body [132]. De-

spite their differences in structure, both types of systems are capable of walking several steps. The robotics community has been actively engaged in modeling and controlling bipedal locomotion systems for many years.

Bipedal robots present several challenges and constraints to the design of control systems due to their exceptional characteristics. Firstly, they are highly nonlinear and naturally unstable systems, which renders the well-developed classical control theories for linear systems inapplicable. Secondly, each walking cycle consists of two different situations in a sequence: the statically stable double-support phase (both feet in contact with the ground) and the statically unstable single-support phase (only one foot contact with the ground). Suitable control strategies are required for step-to-step transitions, adding to the challenge. Thirdly, walking robots are unconnected to the ground, which means that even if all joints of a bipedal robot are perfectly controlled, it is still not enough to completely control all the degrees of freedom (DOFs) of the robot. Fourthly, walking systems usually have many DOFs, especially in 3-D spaces, making the coordination of multijoint movements and the interactions between DOFs a difficult control problem. Finally, bipedal robots have to be adaptive to uncertainties and respond to environmental changes correctly, such as uneven, elastic, sticky, soft, or stiff ground, or obstacles on the ground. Bipedal robots must adjust their control strategies fast enough to respond to such environmental changes.

4.2 Problem statement for the reinforcement learning scheme

Notice that the hybrid formulation considered for the class of problem studied in this thesis requires a first result that modifies the control action within each continuous cycle of adjustment for the control action. This result is based on the application of the following continuous adaptive dynamic programming.

To solve this problem, lets consider the functional $J(u(\cdot))$ related with the tracking error presented here as the variable Δ. Hence the problem studied here corresponds to optimize (minimize indeed) the next functional.

$$J(u(\cdot)) = h_0(T) + \int_{t=0}^{T} h(\Delta(t), u(t))dt$$

where the terminal and the ongoing sections of the functional are given by

$$\begin{aligned} h_0(T) &= \Delta^\top(T)P(T)\Delta(T) \\ h(\Delta(t), u(t)) &= \Delta^\top(t)R(t)\Delta(t) + u^\top(t)Q(t)u(t) + \end{aligned}$$

Here $P(T) \in \mathbb{R}^{n\times n}$ is a positive definite matrix ($P > 0$) related to the final term, $R(t) \in \mathbb{R}^{n\times n}$ is a semi-positive definite matrix that characterizes the effect of states on the Lagrange section of the functional, while $Q(t) \in \mathbb{R}^{n\times n}$ is a positive definite matrix that defines the effect of input function on the integral section of the functional. Introduce $\Delta^\top = \begin{bmatrix} \Delta_a^\top & \Delta_b^\top \end{bmatrix}$ to represent the tracking trajectory error where

$$\begin{aligned} \Delta_a &: = x_a^* - x_a \\ \Delta_b &: = x_b^* - x_b \end{aligned}$$

The functional $J(u(\cdot))$ is subjected to the dynamics of the biped robot, which satsifes the following dynamic expression given by

$$M(q)\ddot{q} + C(\dot{q},q)\dot{q} + G(q) + \lambda^\top H(\dot{q},q) + \Psi(\dot{q},q) = \tau$$

Here $q \in \Omega \subset \mathbb{R}^n$, $M : \mathbb{R}^n \to \mathbb{R}^{n\times n}$, $C : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$, $G : \mathbb{R}^n \to \mathbb{R}^n$, $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^p$, $\lambda \in \mathbb{R}^{p\times n}$, $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$. All these components are corresponding to the model developed in the second chapter of this thesis. Notice that here the effect of motion restrictions are considered in the model of the biped robot.

Based on the state variable theory, lets introduce the following auxiliary vector corresponding to the angular displacements and velocities of the joints in the biped robot

$$\begin{aligned} x_a &= q \\ x_b &= \dot{q} \end{aligned}$$

Based on the proposed auxiliary variables, the dynamics of the biped robot can be presented as follows

$$\begin{aligned} \frac{d}{dt}x_a &= x_b \\ \frac{d}{dt}x_b = \ddot{q} &= M^{-1}(q)\left[\tau - \left(C(\dot{q},q)\dot{q} + G(q) + \lambda^\top H(\dot{q},q) + \Psi(\dot{q},q)\right)\right] \end{aligned}$$

Considering uncertainties in the mathematical representation of the biped robot model, one gets

$$\begin{aligned} \frac{d}{dt}x_a &= x_b \\ \frac{d}{dt}x_b &= (M_N(q) + M_U(q))^{-1}\left[\tau - ((C_N(\dot{q},q) + C_U(\dot{q},q))\dot{q})\right] - \\ &(M_N(q) + M_U(q))^{-1}\left[G_N(q) + G_U(q) + \lambda^\top H(\dot{q},q) + \Psi(\dot{q},q)\right] \end{aligned}$$

Considering the presence of the uncertain terms that agreggates the modeling imprecision and

the external perturbations, one gets for the dynamics of the angular velocity for all the joints

$$\frac{d}{dt}x_b = F(x_a,x_b) + G(x_a)u(t) + \Phi(t,x_a,x_b) =$$
$$F_N(x_a,x_b) + F_U(x_a,x_b) + (G_N(x_a) + G_U(x_a))\,u(t) + \Phi(t,x_a,x_b)$$
$$u(t) = \tau$$

Based on the uncertain representation of the inertia matrix, the following description takes place

$$(M_N(q) + M_U(q))^{-1} = M_N^{-1}(q) - M_N^{-1}(q)\left(M_U^{-1}(q) + M_N^{-1}(q)\right)^{-1} M_N^{-1}(q)$$

The more realistic description of the biper robotic system under analysis obeys the following ordinary differential equation

$$\frac{d}{dt}x_a = x_b$$
$$\frac{d}{dt}x_b = F_N(x_a,x_b) + F_U(x_a,x_b) + (G_N(x_a) + G_U(x_a))\,u(t) + \Phi(t,x_a,x_b)$$

Here $\frac{d}{dt}x_b = F_N(x_a,x_b) + G_N(x_a)u(t) + \eta(t,x_a,x_b,u)$ is valid for the simplified representation of the biped robot dynamics.

In view of the composition prepared before using the known and the unknown sections of the mechanical system, after implementing the matrix inversion lemma and some algebraic manipulations, the term η is

$$\eta(t,x_a,x_b,u) = [F_U(x_a,x_b) + \Phi(t,x_a,x_b) + G_U(x_a)u(t)]$$

The term η belongs to the following set $\eta(t,x_a,x_b,u) \in \Omega_\eta$

$$\Omega_\Psi = \left\{ \begin{array}{c} \eta \,|\, \eta : \{\mathbb{R}^+ \cup \{0\}\} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, \\ \|\eta(t,x_a,x_b,u)\|^2 \le \eta_0 + \eta_1 \|q\|^2 + \eta_2 \|\dot{q}\|^2 + \eta_3 \|u\|^2, \\ \forall (\dot{q},q) \in \Omega_{dq} \oplus \Omega_q, u \in U_{adm} \end{array} \right\}$$

where

$$U_{adm} = \left\{ u \,|\, u, \|u\|^2 \le u_0 + u_1 \|q\|^2 + u_2 \|\dot{q}\|^2, \;\; \forall (\dot{q},q) \in \Omega_{dq} \oplus \Omega_q \right\}$$

Here U_{adm} is the set of admissible controllers. $\oplus$ represents Minkowski addition.

4.3 Properties of the proposed biped robot

The following properties are considered valid for the class of systems that represent the biped robot.

a) The inertia matrix $M : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is not singular for all possible configurations in q, that is

$$\det(M(q)) \neq 0, \ \ \forall q \in \Omega_q$$

b) The inertia matrix can be decomposed as the combination of a known section named $M_N : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and the section with uncertain mathematical model referred to as $M_U : \mathbb{R}^n \to \mathbb{R}^{n\times n}$. The known section represents the nominal values of the inertia matrix, while the uncertain section accounts for the deviations from the nominal values due to modeling uncertainties.

$$M(q) = M_N(q) + M_U(q)$$

.

c) The Coriolis matrix defined by $C : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$

$$C(\dot{q}, q) = C_N(\dot{q}, q) + C_U(\dot{q}, q)$$

where $C_N : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$, $C_U : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ are the section of the Coriolis matrix with known and uncertain mathematical model.

d) The term $G : \mathbb{R}^n \to \mathbb{R}^n$ defines the gravitational effects over the dynamics of the biped robot, and it is defined as follows

$$G(q) = G_N(q) + G_U(q)$$

where $G_N : \mathbb{R}^n \to \mathbb{R}^n$, $G_U : \mathbb{R}^n \to \mathbb{R}^n$ are the section of the gravitational term with known and uncertain mathematical model.

e) The term $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ characterizes the effect of modeling imprecion and external perturbation on the dynamics of the biped robot.

$$\begin{aligned} \Psi(\dot{q}, q) &\in \Omega_\Psi \\ \Omega_\Psi &= \left\{ \Psi | \, \Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n, \|\Psi(\dot{q}, q)\|^2 \leq \Psi_0 + \Psi_1 \|q\|^2 + \Psi_2 \|\dot{q}\|^2, \forall (\dot{q}, q) \in \Omega_{dq} \oplus \Omega_q \right\} \end{aligned}$$

4.4 State variable representation of the tracking trajectory problem

The dynamics of the reference trajectories can be described by

$$\begin{aligned}\frac{d}{dt}x_a^* &= x_b^* \\ \frac{d}{dt}x_b^* &= h(t)\end{aligned}$$

where x_a^* is the vector of reference trajectories. Here the function h is a time dependent vector function such that its second derivative coincides with it.

Given the reference trajectories, the dynamics of Δ corresponds to

$$\begin{aligned}\frac{d}{dt}\Delta_a &= \Delta_b \\ \frac{d}{dt}\Delta_b &= h(t) - (F_N(x_a,x_b) + G_N(x_a)u(t) + \eta(t,x_a,x_b,u))\end{aligned}$$

The controller u is proposed to satisfy the following composite structure

$$u(t) = u_c(t) + u_o(t)$$

Here the following definitions are considered to compose the suggested control action $u_c \in \mathbb{R}^m$ and $u_o \in \mathbb{R}^m$, the compensating and the sub-optimal sections, respectively, are given by:

$$\begin{aligned}u_c(t) &= G_N^{-1}(x_a)(-F_N(x_a,x_b) + h(t) - K_a\Delta_a - K_b\Delta_b) \\ u_o(t) &= G_N^{-1}(x_a)v\end{aligned}$$

Notice that the first part is a class of state feedback control formulation. Using the proposed controller on the trajectory tracking error dynamics leads to

$$\begin{aligned}\frac{d}{dt}\Delta_a &= \Delta_b \\ \frac{d}{dt}\Delta_b &= -K_a\Delta_a - K_b\Delta_b + v + \eta(t,x_a,x_b,u)\end{aligned}$$

Using the extended state representation based on Δ, leads to the following ordinary differential

equation for the tracking error dynamics

$$\frac{d}{dt}\Delta = A\left(K_a, K_b\right)\Delta + Bv + B\eta\left(t, x_a, x_b, u\right)$$

$$A\left(K_a, K_b\right) = \begin{bmatrix} 0_n & I_n \\ K_a & K_b \end{bmatrix}, \; B = \begin{bmatrix} 0_n \\ I_n \end{bmatrix}$$

The online optimization problem is to design the control action v such that the following problem can be solved

$$\begin{aligned} v^*(\cdot) &\rightarrow \arg\min_{v \in V_{adm}} \left\{ \max_{\eta \in \Omega} J((k-1)T, kT, v) \right\} \\ J((k-1)T, kT, v) &= h_0(\Delta(kT)) + \int_{t=(k-1)T}^{kT} h(\Delta(t), v(t))dt \\ h(\Delta(t), v(t)) &= \|\Delta\|_Q^2 + \|v\|_R^2 \qquad Q = Q^\top, Q \geq 0; \quad R = R^\top, R > 0 \end{aligned}$$

Notice that the evolution of the functional is evaluated over a given time period defined by the recurrent cycles of training. This hybrid strategy is defining the application of the RL strategy developed in this thesis. Observe that this optimization problem is proposed in terms of a finite horizon related to the concept of epoch used in artificial neural networks theory.

The solution of the optimization problem can be performed using the adaptive dynamic programming theory with the guaranteed performance (considering the effect of the uncertainties or perturbations). Hence, it is feasible to consider the application of the Hamilton-Jacobi-Bellman (HJB) equation with the robustness condition using the min max formulation. Such version of the HJB corresponds to

$$-\frac{\partial V(\Delta, t)}{\partial t} = \max_{v \in V_{adm}} \left\{ \min_{\eta \in \Omega} H\left(-\frac{\partial V(\Delta, t)}{\partial x}, \Delta, v, \eta \right) \right\}$$

In view of the form of $J((k-1)T, kT, v)$, the expression corresponding to the Hamiltonian is represented by $H\left(-\frac{\partial V(\Delta,t)}{\partial x}, \Delta, v, \eta\right)$ and is defined as follows:

$$H\left(-\frac{\partial V(\Delta, t)}{\partial x}, \Delta, v, \eta \right) = \left[-\frac{\partial V(\Delta, t)}{\partial \Delta} \right]^\top \frac{d}{dt}\Delta - h(\Delta(t), v(t))$$

The first part of the solution for the HJB with the admissible perturbations and control forms

requires solving the next optimization problem

$$\begin{aligned}\min_{\eta\in\Omega} H\left(-\frac{\partial V(\Delta,t)}{\partial x},\Delta,v,\eta\right) &= \min_{\eta\in\Omega}\left(\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\frac{d}{dt}\Delta - h(\Delta(t),v(t))\right)\\ &= \min_{\eta\in\Omega}\left(\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\frac{d}{dt}\Delta\right)\end{aligned}$$

The previous equation is valid in view of the function $h(\Delta(t),v(t))$ is independent of the perturbations/uncertainties term.

$$\begin{aligned}\min_{\eta\in\Omega} H\left(-\frac{\partial V(\Delta,t)}{\partial x},\Delta,v,\eta\right) &= \min_{\eta\in\Omega}\left(\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\left(A\left(K_a,K_b\right)\Delta + Bv + B\eta\left(t,x_a,x_b,u\right)\right)\right)\\ &= \min_{\eta\in\Omega}\left(\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\left(B\eta\left(t,x_a,x_b,u\right)\right)\right)\end{aligned}$$

The optimization problem with respect to the action of uncertainties and perturbations corresponds to

$$\min_{\eta\in\Omega}\left(\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\left(B\eta\left(t,x_a,x_b,u\right)\right)\right) = \max_{\eta\in\Omega}\left(\left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\left(B\eta\left(t,x_a,x_b,u\right)\right)\right)$$

Using the Cauchy-Schwarz inequality, the following upper bound takes place for the right-hand side of the previous inequality

$$\left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\left(B\eta\left(t,x_a,x_b,u\right)\right) \leq \left\|B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]\right\|\cdot\left\|\eta\left(t,x_a,x_b,u\right)\right\|$$

In view of the admissible class of perturbations, the following upper estimation for the term that should be maximized is obtained

$$\left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top}\left(B\eta\left(t,x_a,x_b,u\right)\right) \leq \sqrt{\eta_0}\left\|B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]\right\|$$

The maximum that should be calculated for the left-hand side in the previous inequality can be estimated when the equation is evaluated, that is

$$\left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^{\top} B\eta^{*}\left(t,x_a,x_b,u\right) = \sqrt{\eta_0}\left\|B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]\right\|$$

This relation is true if the vectors considered in the vector product are parallel each other. Based on this argument, the following identity can be used to estimate the maximum admissible value for the perturbations/uncertainties term η^*, that is

$$B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right] \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^\top B\eta^*(t,x_a,x_b,u) = \sqrt{\eta_0}\left\|B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]\right\| B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]$$

The corresponding *worst* possible term η^* is given by

$$\eta^*(t,x_a,x_b,u) = \sqrt{\eta_0}\left\|B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]\right\| \left[B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right] \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^\top B\right]^{-1} B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]$$

An algebraic manipulation on the previous expression leads to

$$\eta^*(t,x_a,x_b,u) = \frac{B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]}{\left\|B^\top \left[\frac{\partial V(\Delta,t)}{\partial \Delta}\right]\right\|}$$

Using this result of the optimization problem for the uncertainties/perturbations, in the calculus related to the Hamiltonian yields

$$\begin{aligned}
&\max_{v\in V_{adm}} H\left(-\frac{\partial V(\Delta,t)}{\partial x},\Delta,v,\eta^*\right) \\
&= \max_{v\in V_{adm}} \left\{\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^\top \left(A(K_a,K_b)\Delta + Bv + B\eta^*(t,x_a,x_b,u)\right) - h(\Delta(t),v(t))\right\} \\
&= \max_{v\in V_{adm}} \left\{\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^\top \left(A(K_a,K_b)\Delta + Bv + B\eta^*(t,x_a,x_b,u)\right) - \left(\|\Delta\|_Q^2 + \|v\|_R^2\right)\right\}
\end{aligned} \tag{4.1}$$

Performing the estimation of maximum of the Hamiltonian with respect to the input, that is $\max_{v\in V_{adm}} H\left(-\frac{\partial V(\Delta,t)}{\partial x},\Delta,v,\eta^*\right)$, that can be calculated using the gradient rule, yielding

$$\begin{aligned}
&\max_{v\in V_{adm}} H\left(-\frac{\partial V(\Delta,t)}{\partial x},\Delta,v,\eta^*\right) <=> \\
&\nabla_v \left\{\left[-\frac{\partial V(\Delta,t)}{\partial \Delta}\right]^\top \left(A(K_a,K_b)\Delta + Bv + B\eta^*(t,x_a,x_b,u)\right) - \left(\|\Delta\|_Q^2 + \|v\|_R^2\right)\right\} = 0
\end{aligned} \tag{4.2}$$

Considering that the admissible class of the controls v is the entire $\mathbb{R}^n$, hence $V_{adm} = \mathbb{R}^m$, hence

the following gradient operation lead to estimate the formal expression of the control action

$$\nabla_v \left\{ \left[-\frac{\partial V(\Delta,t)}{\partial \Delta} \right]^\top \left(A\left(K_a,K_b\right)\Delta + Bv + B\eta^*\left(t,x_a,x_b,u\right)\right) - \left(\|\Delta\|_Q^2 + \|v\|_R^2 \right) \right\} =$$

$$\nabla_v \left\{ \left[-\frac{\partial V(\Delta,t)}{\partial \Delta} \right]^\top Bv - \|v\|_R^2 \right\} =$$

$$\left(\left[-\frac{\partial V(\Delta,t)}{\partial \Delta} \right]^\top B \right)^\top - 2Rv^* = 0$$

Based on the results presented above, it is feasible to calculate the optimal controller v^* accordingly to the relationship given by

$$2Rv^* = \left(\left[-\frac{\partial V(\Delta,t)}{\partial \Delta} \right]^\top B \right)^\top$$

The simple solution to the algebraic relationship presented above leads to

$$v^* = -\frac{1}{2} R^{-1} B^\top \frac{\partial V(\Delta,t)}{\partial \Delta}$$

In view of the difficulties to get the solution of the previous partial differential equation, this study proposes to get an approximate solution in the following form

$$V(\Delta,t) = V_N(\Delta,t) + \tilde{V}(\Delta,t)$$

In the approximate solution to the Bellman´s function, we consider the following representation of that bounded function considering the universal approximation result developed by Cybenko, that is

$$\begin{aligned} V_N(\Delta,t) &= \Delta^\top P(t)\Delta + \sigma^\top(\Delta,\Theta) W^\top(t) \Lambda W(t) \sigma(\Delta,\Theta) \\ &= \|\Delta\|_{P(t)}^2 + \|W(t)\sigma(\Delta,\Theta)\|_\Lambda^2 \end{aligned}$$

Also, the term $\tilde{V}(\Delta,t)$ is the approximation error which appears considering a bounded number of sigmoidal activation functions in the neural network representation. This term is bounded with respect to its two arguments considering the Cybenko´s theorem.

Using the HJB (with the corresponding optimal values estimated above), the following differential

relationships can be gotten

$$\frac{\partial V_N(\Delta,t)}{\partial t} = \Delta^\top \left(\frac{d}{dt}P(t)\right)\Delta + 2\sigma^\top(\Delta,\Theta)W^\top(t)\Lambda\left(\frac{d}{dt}W(t)\right)\sigma(\Delta,\Theta)$$

$$\frac{\partial V_N(\Delta,t)}{\partial \Delta} = 2P(t)\Delta + 2\left(\nabla_\Delta \sigma(\Delta,\Theta)\right)^\top W^\top(t)\Lambda W(t)\sigma(\Delta,\Theta)$$

The substitution of these results on the HJB equation implies the following equation

$$-\Delta^\top \frac{d}{dt}P(t)\Delta - 2\sigma^\top(\Delta,\Theta)W^\top(t)\Lambda\left(\frac{d}{dt}W(t)\right)\sigma(\Delta,\Theta) - \frac{\partial \tilde{V}(\Delta,t)}{\partial t} =$$

$$\left[2P(t)\Delta + 2\left(\nabla_\Delta \sigma(\Delta,\Theta)\right)^\top W^\top(t)\Lambda W(t)\sigma(\Delta,\Theta)\right]^\top \cdot$$

$$\left(A(K_a,K_b)\Delta + B\left(-R^{-1}B^\top P(t)\Delta\right) + B\eta^*(t,x_a,x_b,u)\right) +$$

$$\left[-\frac{\partial \tilde{V}(\Delta,t)}{\partial \Delta}\right]^\top \left(A(K_a,K_b)\Delta + B\left(-R^{-1}B^\top P(t)\Delta\right) + B\eta^*(t,x_a,x_b,u)\right) -$$

$$\left(\|\Delta\|_Q^2 + \left\|\frac{1}{2}R^{-1}B^\top\left(2P(t)\Delta + 2\left(\nabla_\Delta \sigma(\Delta,\Theta)\right)^\top W^\top(t)\Lambda W(t)\sigma(\Delta,\Theta)\right)\right\|_R^2\right)$$

Using the result estimated for the worst potential effect of the uncertainties and perturbations yields to

$$-\Delta^{\top}\frac{d}{dt}P(t)\Delta-2\sigma^{\top}(\Delta,\Theta)W^{\top}(t)\Lambda\left(\frac{d}{dt}W(t)\right)\sigma(\Delta,\Theta)-\frac{\partial\tilde{V}(\Delta,t)}{\partial t}=$$

$$\left[2P(t)\Delta+2\left(\nabla_{\Delta}\sigma(\Delta,\Theta)\right)^{\top}W^{\top}(t)\Lambda W(t)\sigma(\Delta,\Theta)\right]^{\top}\cdot$$

$$\left(A(K_a,K_b)\Delta+B\left(-R^{-1}B^{\top}P(t)\Delta\right)+B\frac{B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial\Delta}\right]}{\left\|B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial\Delta}\right]\right\|}\right)+$$

$$\left[-\frac{\partial\tilde{V}(\Delta,t)}{\partial\Delta}\right]^{\top}\left(A(K_a,K_b)\Delta+B\left(-R^{-1}B^{\top}P(t)\Delta\right)+B\eta^{*}(t,x_a,x_b,u)\right)-$$

$$\left(\|\Delta\|_{Q}^{2}+\left\|\frac{1}{2}R^{-1}B^{\top}\left(2P(t)\Delta+2\left(\nabla_{\Delta}\sigma(\Delta,\Theta)\right)^{\top}W^{\top}(t)\Lambda W(t)\sigma(\Delta,\Theta)\right)\right\|_{R}^{2}\right)$$

The expression described above can be reordered in the following manner, using the relationships based on the tracking errors and the components having the weights

$$\Delta^{\top}\left(-\frac{d}{dt}P(t)+Ric(R,Q)\right)+$$

$$tr\left\{W^{\top}(t)\left(-2\frac{d}{dt}W(t)\Lambda\sigma(\Delta,\Theta)\,\sigma^{\top}(\Delta,\Theta)+LL(\Delta,P(t))\right)\right\}=$$

$$\frac{\partial\tilde{V}(\Delta,t)}{\partial t}+\left[-\frac{\partial\tilde{V}(\Delta,t)}{\partial\Delta}\right]^{\top}\left(A(K_a,K_b)\Delta+B\left(-R^{-1}B^{\top}P(t)\Delta\right)+B\eta^{*}(t,x_a,x_b,u)\right)$$

Here the following expressions can be calculated for the differential form of the Riccati equation and the corresponding learning law for the weights W:

$$Ric(R,Q)=-PA(K_a,K_b)-A^{\top}(K_a,K_b)P+PB\left(R^{-1}+\frac{2}{\left\|B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial\Delta}\right]\right\|}\right)B^{\top}P-Q$$

$$LL(\Delta,P(t))=$$

$$2W\Lambda\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\left(A\left(K_a,K_b\right)\Delta+B\left(-R^{-1}B^{\top}P(t)\Delta\right)+B\frac{B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial\Delta}\right]}{\left\|B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial\Delta}\right]\right\|}\right)\sigma^{\top}\left(\Delta,\Theta\right)-$$

$$2W(t)\Lambda\left(\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\right)^{\top}BR^{-2}B^{\top}P(t)\Delta\sigma^{\top}\left(\Delta,\Theta\right)-$$

$$4W\Lambda\left(\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\right)^{\top}BR^{-2}B^{\top}\left(\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\right)W^{\top}\Lambda W\sigma^{\top}\left(\Delta,\Theta\right)$$

Notice here that if

$$-\frac{d}{dt}P(t)+Ric(R,Q)=0 \tag{4.3}$$

and if

$$-2\frac{d}{dt}W(t)\Lambda\sigma\left(\Delta,\Theta\right)\sigma^{\top}\left(\Delta,\Theta\right)+LL(\Delta,P(t))=0 \tag{4.4}$$

And in view of the boundedness for the variable $\tilde{V}$, hence the HJB is satisfied considering the approximate model. In particular for the learning law, notice that

$$\frac{d}{dt}W(t)=$$

$$W\Lambda\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\left(A\left(K_a,K_b\right)\Delta+B\left(-R^{-1}B^{\top}P(t)\Delta\right)+\frac{BB^{\top}\left[\frac{\partial V(\Delta,t)}{\partial\Delta}\right]}{\left\|B^{\top}\left[\frac{\partial V(\Delta,t)}{\partial\Delta}\right]\right\|}\right)\frac{\sigma^{\top}\left(\Delta,\Theta\right)}{\left\|\sigma\left(\Delta,\Theta\right)\right\|}\Lambda^{-1}-$$

$$W(t)\Lambda\left(\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\right)^{\top}BR^{-2}B^{\top}P(t)\Delta\frac{\sigma^{\top}\left(\Delta,\Theta\right)}{\left\|\sigma\left(\Delta,\Theta\right)\right\|}\Lambda^{-1}-$$

$$2W\Lambda\left(\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\right)^{\top}BR^{-2}B^{\top}\left(\nabla_{\Delta}\sigma\left(\Delta,\Theta\right)\right)W^{\top}\Lambda W\frac{\sigma^{\top}\left(\Delta,\Theta\right)}{\left\|\sigma\left(\Delta,\Theta\right)\right\|}\Lambda^{-1} \tag{4.5}$$

4.5 Reinforcement Learning based on recurrent-differential approximate models by neural networks

According to the results presented in the previous section, once we have designed a controller u such that $u=u(\Theta)$ with Θ representing all the fixed parameters at the beginning of the recurrent

cycle, hence the optimization after each cycle can be presented as follows

$$u = u(\Theta) \in U_{adm}$$

$$u^{*,k}(\cdot,\Theta) = \min_{u \in U_{adm}} \max_{\eta(t,x_a,x_b,u) \in \Omega_\eta} J^k(u(\cdot),\Theta)$$

$$\Theta^{*,k+1} = F\left(\Theta^{*,k}, J^k(u(\cdot,\Theta^{*,k}))\right)$$

such that

$$J^k(u(\cdot,\Theta)) = h_0(kT) + \int_{t=(k-1)T}^{kT} h(x(t),u(t,\Theta))dt$$

Here, the time T is actually what it is defined as the evaluation window period for the RL algorithm. Thus, the total period of the reinforcement algorithm would be defined as

$$T_r = N * T$$

where N can be a fixed number of evaluation rounds or a number that can be changed depending on the expected value J^*, such that for that given N, it holds that

$$J(u^N(t,\Theta)) - J^* < 0$$

Under these definitions, the reinforcement process implies that

$$\Theta(kT) = h((k-1)T, J(u^k(t,\Theta))$$

Note that the function h is a recurrent function that depends on the variation of $J(u^k(t,\Theta)$ with respect to $J(u^{k-1}(t,\Theta)$. Although the control may have a quasi-linear form, its expression is considered using dynamic programming techniques, given by

$$u = u(\Delta,\Theta(kT)) = u(\Delta,\nabla\mathcal{V}^k)$$

where the function $\mathcal{V}^k$ is the value function derived from the HJB equation evaluated at the time window with time $t = T$. However, it is known that the solution of the HJB equation has an analytical form only in the case of linear systems with constant parameters or in the case of time optimization. If the system is nonlinear or even uncertain, the HJB equation has an analytical solution only in very particular cases. For this reason, approximate solutions are proposed for the HJB equation, that is,

for the value function, and in the case that the system model is uncertain, we use neural networks. According to the previous arguments, this reinforcement law is working only at the discrete times kT.

Hence, taking into consideration the recurrent laws , one may consider different strategies for adjusting the parameters that define the operation of the controller within each period where the continuous approximation presented in the previous section holds. Among the availabel options, we may consider

1. Q-learning
2. Policy iteration
3. Value iteration

The DDPG algorithm proposed by [48] is based on the Deterministic Policy Gradients that at the same time, it is learning a Q-function, that is it learns a policy, i.e., uses off-policy data and the approximated Bellman equation learns the Q-function, which turns to learn the policy.

This off-policy algorithm is based on the Q-learning algorithm [133] that learns the value of an action given a particular state, reaching this by maximizing the expected value of the total reward along all the updating steps. Indeed, it is designed to deal with continuous action spaces.

Considering the Bellman equation for the optimal action-value, $Q_{\pi^*}(s_t, a_t)$ (1.4) [1]. This equation is the key to learning an approximation of Q_*. Assuming that this approximation is a Neural Network (NN) whose parameters are φ, and there exists a set $\mathcal{T}$ of transitions (s_t, a_t, r_t, s_{t+1}) sampled from the environment.

The off-policy method employed by this algorithm looks for the minimization of the φ-parametrized mean-squared Bellman error [134, 135] as:

$$L_\varphi(\mathcal{T}) = \mathrm{E}\left[\left(Q_{\varphi(s_t,a_t)} - \left(r + \gamma \max_{a_{t+1}} Q_\varphi\left(s_{t+a}, a_{t+1}\right)\right)\right)^2\right] \tag{4.6}$$

The Q-learning algorithms used as function approximations are mainly based on minimizing the (4.6) function.

All these laws work on some fixed parameters within at the beginning of each time-window,

$$K_a((k-1)T), K_b((k-1)T), W((k-1)T), P((k-1)T), \Theta((k-1)T)$$

The proposed laws are giving the rule such that

$$K_a(kT) \quad = \quad \Omega_{K_a}(J((k-1)T, kT, v^*))$$

[1] In order to shorten the notation, the equation will be written as Q_* from this point in the document.

$$
\begin{aligned}
K_b(kT) &= \Omega_{K_b}(J((k-1)T, kT, v^*)) \\
W(kT) &= \Omega_W(J((k-1)T, kT, v^*)) \\
P(kT) &= \Omega_P = P((k-1)T) \\
\Theta(kT) &= \Omega_\Theta = \Theta((k-1)T)
\end{aligned}
$$

4.6 Virtual platform

For the evaluation of the proposed RL algorithm, it was used a virtual version of the biped robot (Figure 4.1). This virtual model was proposed in SolidWorks® considering mechanical properties as material, weight and dynamics. The bipedal robotic model was exported to Matlab® applying Simescape Multibody®.

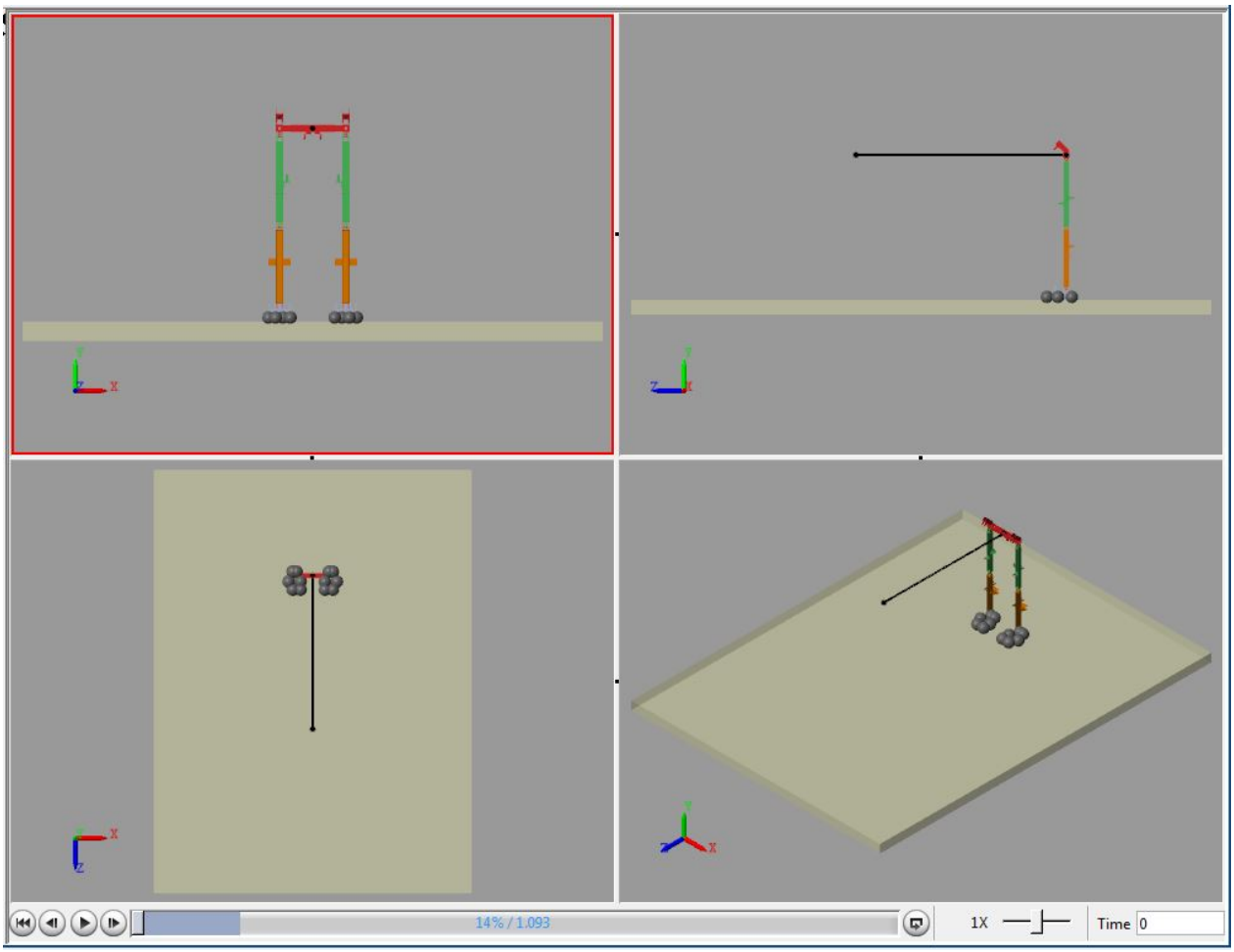


Figure 4.1: Virtual robot to prove the RL scheme

In Figure 4.1 is shown the BR from four points of view As can be seen in the image, along

the foot were implemented virtual sensors (seen as spheres) to know the forces generated by the interaction between the foot and the ground.

4.7 Design

To attain the objectives of design of trajectories while minimizing the energetic consumption of the BR, it was proposed a MARL scheme of two Agents. The first agent modified the proposed human-based trajectories to fit properly to the BR movements. The second Agent intended to find the best gains for the controller.

A deeper description of the RL construction is given in the next sections.

4.7.1 Neural network sketch

The training was made considering the interaction of two MARL agents each of them consisting on an actor and critic networks. Both actor networks have the same structure, consisting on nine layers. The inputs have a dimension of 53 each, followed by two 100-units fully connected layers, and 8-units fully connected layer. Also, two ReLU layers, one LSTM, one Tanh, and a output scaling layer are part of the structure (see 4.1). The outline can be seen in Figures 4.2a and 4.3a.

Table 4.1: Actor-network structure

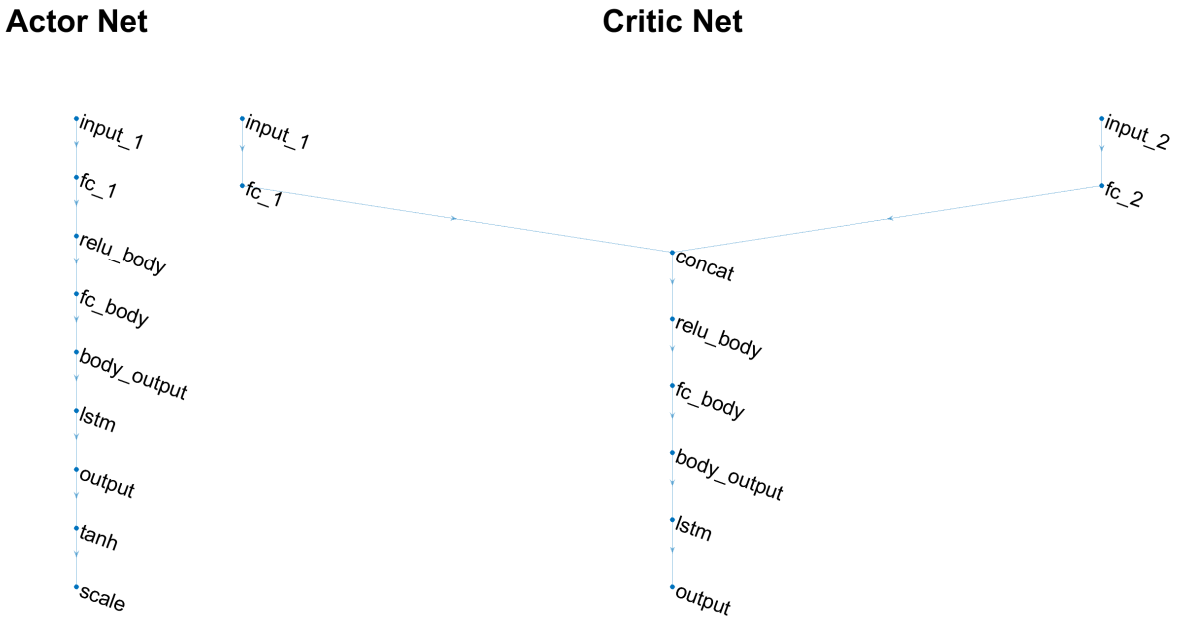
	Reference	Name	Description
1	'input_1'	Sequence Input	Sequence input with 53 dimensions
2	'fc_1'	Fully Connected	100 fully connected layer
3	'relu_body'	ReLU	ReLU
4	'fc_body'	Fully Connected	100 fully connected layer
5	'body_output'	ReLU	ReLU
6	'lstm'	LSTM	LSTM with 100 hidden units
7	'output'	Fully Connected	8 fully connected layer
8	'tanh'	Tanh	Hyperbolic tangent
9	'scale'	ScalingLayer	Scaling layer

The critic network is made out of 10 layers (see Table 4.2, having two different input layers, the first one with dimension of 53, and the second one being 8-dimensioned. Following the inputs, each of the hidden fully connected layer has 100 neurons. Afterwards, both branches get concatenated. The rest of the structure consist on two ReLU layers, two fully connected layers with 100 neurons

each, and a LSTM layer (also 100 hidden neurons). The structure is shown in Figures 4.2b and 4.3b.

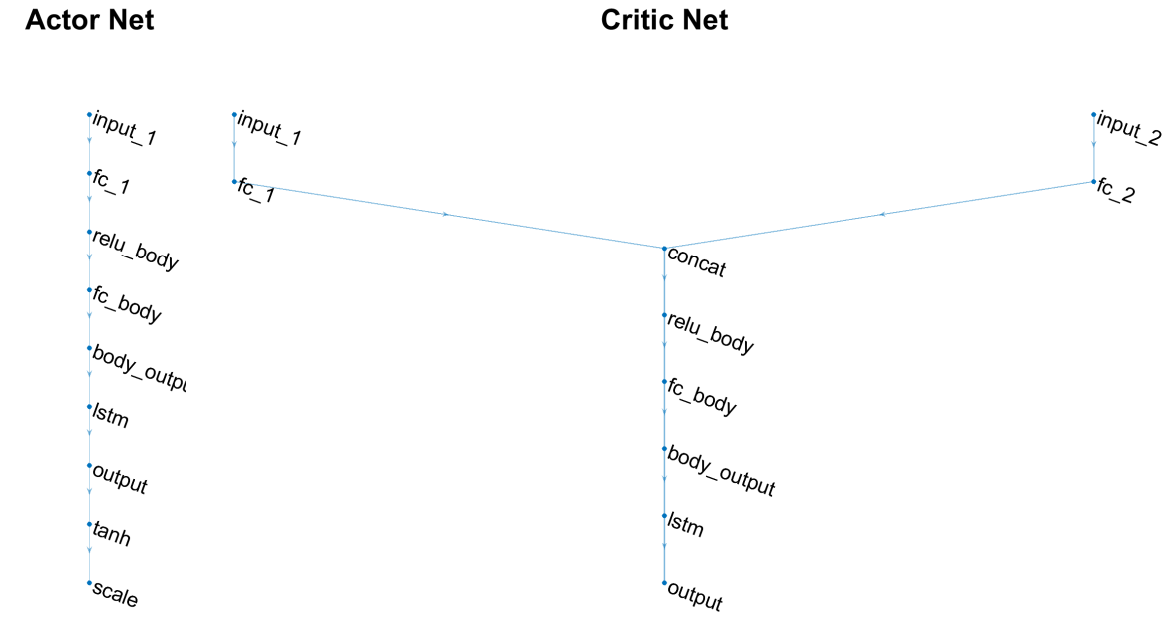
Table 4.2: Critic-network structure

	Reference	Name	Description
1	'concat'	Concatenation	Concatenation of 2 inputs along dimension 1
2	'relu_body'	ReLU	ReLU
3	'fc_body'	Fully Connected	Fully Connected
4	'body_output'	ReLU	ReLU
5	'lstm'	LSTM	LSTM
6	'input_1'	Sequence Input	Sequence Input
7	'fc_1'	Fully Connected	Fully Connected
8	'input_2'	Sequence Input	Sequence Input
9	'fc_2'	Fully Connected	Fully Connected
10	'output'	Fully Connected	Fully Connected



(a) Actor network *(b) Critic network*

Figure 4.2: Critic and actor networks used during the RL from the trajectories agent



(a) Actor network *(b) Critic network*

Figure 4.3: Critic and actor networks used during the RL from the gains agent

4.7.2 Reward design

Assuring that the robot retrieve enough information from its environment and used it to attain a consistent walk, the reward function was designed considering five main parameters: velocity along the Z axis, norm of tracking errors, the interaction of the feet with the ground, the distance covered during the walk and the previous actions taken by both agents. The rewards fed both agents equally.

As it was described in 1.4, for the functioning of every RL algorithm is necessary to declare those variables which will provide information from the environment to the system. For the reward, it was implemented the following structure:

$$\boldsymbol{R} = 7V_z - 0.3\sum_{i=1}^{8} e_i^2 + pos_Z + 5F_N + 3d_{LF} + 3d_{RF} + pos_Z + trays - 0.005\sum_{i=1}^{8} u_A^2 \qquad (4.7)$$

here, $\boldsymbol{R}$ represents the reward, e means the error obtained by the difference between the reference trajectory and the actual position of each joint, V_z stands for the velocity along Z axis, meanwhile, pos_Z expresses the distance from the origin to the CoM of the hip.

F_N is constructed considering that if neither foot touches the ground (the BR is not walking) the normal force of both feet will be zero leading to a penalty of -1. However, if the normal force is different from 0, the reward is $+1$.

d_{LF} and d_{RF} represents the distance from each foot to the ground, and both are constructed the same: $1 - \exp(d - M_{dist})$, d represents the measured distance to the floor and M_{dist} is a constant in meters, considered as the maximum limit. $trays$ is made taking every $(t-1)$ action verifying if they are inside the boundaries (explained further in Table 5.1), giving a reward of $+1$ for each joint if inside the range and 0 otherwise, then, the output vector is dot-multiplied with itself. u_A^2 is computed as the summation of the squared action at $(t-1)$.

4.8 Numerical evaluation

For training the selected artificial neural network that are used to develop the RL strategy, it was considered a maximum number of episodes 1000 with $\frac{Tf}{Ts}$ maximum steps per episode. The evolution of the reward function through the training is shown in Figure 4.4

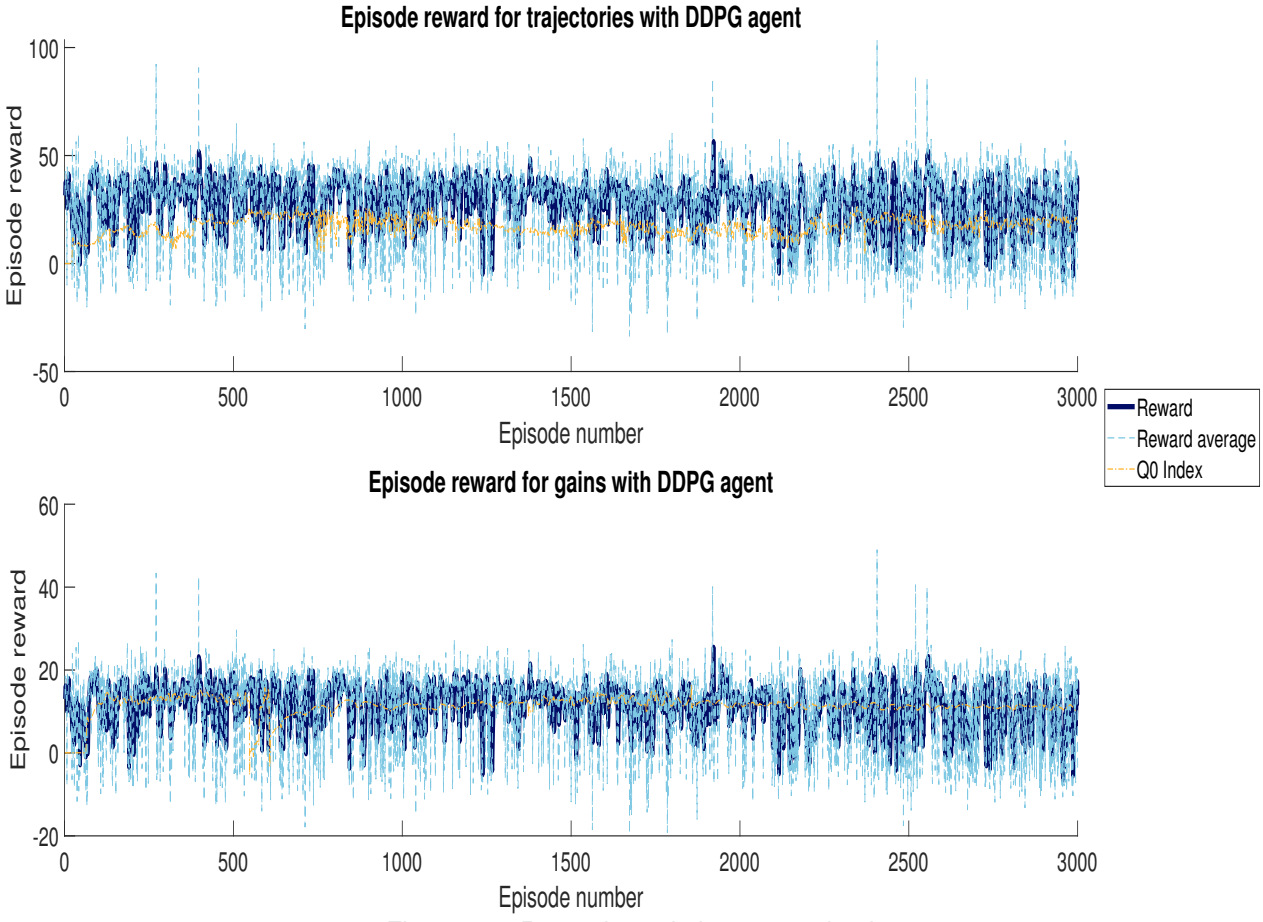


Figure 4.4: Rewards evolution over episode

Each simulated episode was two seconds long. The results are shown in the following figures,

being grouped in right and left limb. The numerical simulation settings were established as follows: fixed integration step at $10\mu s$, integration method Runge-Kutta.

With the aim of comparing the performance of the suggested controller, a classical state feedback control was also evaluated considering a decentralized structure as in the case of the proposed controller. For this comparative controller, the Sliding Mode-based super-twisting algorithm (STA) was used as a method to obtain the on-line estimation of the time-derivative for the tracking errors [89, 128]. The graphs showing the comparative results share the same legend: green line refers to the results attained by the PID tuned by the RL scheme, and the red line is used to highlight the results of the classical state feedback control.

Considering the RL design considered in this thesis, two agents were implemented to attain the data driven based optimization during the automatic walking: minimize the energy used by the controller introducing adaptations for the control gains, while modifying the reference trajectories for the robot joints, considering their motion limitations, and enforcing the bioinspired gait cycle. The Gains Agent added (algebraically) values to a previously tuned PID controller (used for comparison), which were reintroduced as the modified K_P, K_I, and K_D gains. The second Agent, modified reference trajectories (based on human walking patterns [26]), by adding or subtracting values from the reference vectors.

4.8.1 Right limb

The robot's right limb was considered as a reference during the gait analysis, in accordance with what was presented in Section 1.2. The detailed results concerning this limb with its respective joints and movements are discussed in the next paragraphs.

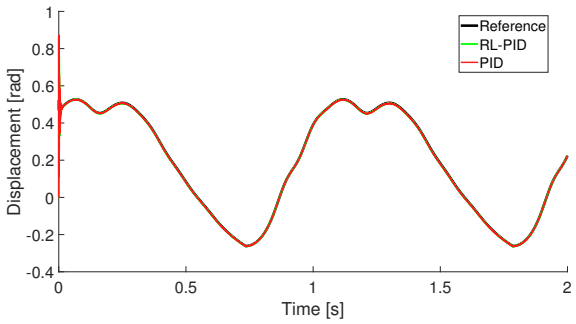
Hip

Hip is the articulation that connects the pelvis-equivalent channel with the limb. This joint mobilizes the entire limb in two movements: Flexion-Extension (FE) in the sagittal plane, and Abduction-Adduction (Abd-Add) in the frontal plane.

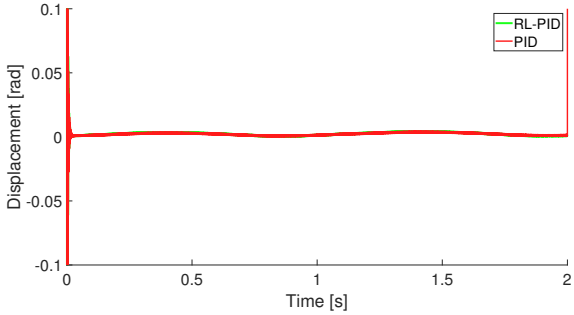
Flexion - Extension

The results for the FE movement of the right hip are presented in Figure 4.5. All the graphs keep the same color code: black for the reference trajectory, green the Reinforcement Learning-based PID (RL-PID) gains, and the red line for the state feedback controller (PID).

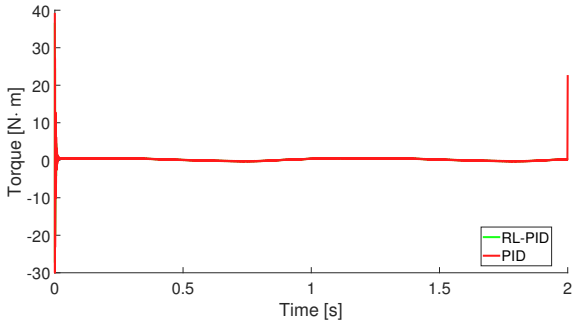
Figure 4.5a shows the outcomes of the tracking simulation of the hip both reach the reference trajectory in approximately $15ms$. Once stabilized, the maximum amplitude of the error for the RL-PID in the FE hip joint is $4.5 \times 10^{-3} rad$ and $0.3009 rad$ for the Proportional–Integral-Derivative (PID) controller (see Figure 4.5b). With regard to the control signals generated by the use of both schemes, shown in Figure 4.5c, the maxima are $0.6512 N \cdot m$ and $22.7090 N \cdot m$ for RL-PID and PID controllers, respectively.



(a) Right hip joint (flexion - extension) tracking



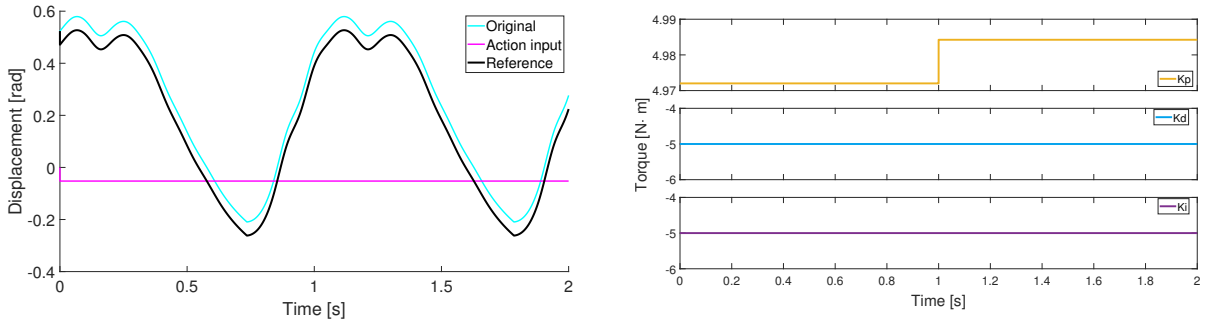
(b) Tracking error



(c) Control trajectories

Figure 4.5: Evolution of the right hip joint (flexion - extension): tracking, error, and control signals

The actions performed by the agents in the reinforcement strategy can be seen in Figure 4.6. The adjustment of the reference trajectory for this joint is presented in Figure 4.6a, where the initial reference trajectory is presented in cyan, the action generated by the Agent is depicted in magenta, and black represents the final trajectory reference. It can be seen that at $50 \mu s$ the trajectory changed $-0.052 rad$ from the original. Together with the actions of the first agent, the actions exerted by the second agent are presented in Figure 4.6b, where the action taken on each of the gains (K_P in orange, K_I in blue, and purple for K_D) can be observed. The K_P gain was modified adding 4.972 to its original value (75), while the K_I and K_D gains were subtracted 5 units from both, being originally 0 and 40, respectively. These modifications are attained after the sequence of gains modifications.

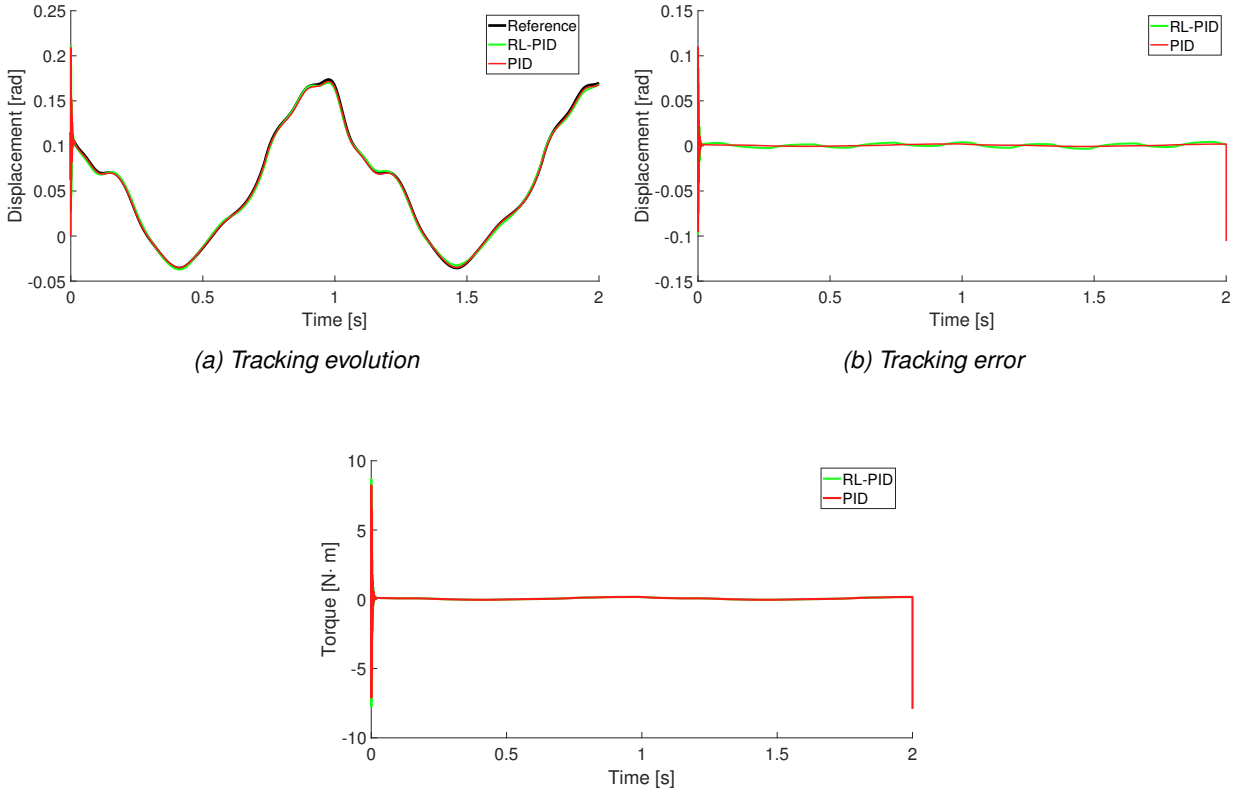


(a) Trajectories *(b) Evolution of the gains*

Figure 4.6: Actions of the Agents over the right hip joint (flexion - extension): correction of the reference trajectory, and evolution of the gains

Abduction - Adduction

The results for the Abd-Add movement of the right limb are presented in Figure 4.7.



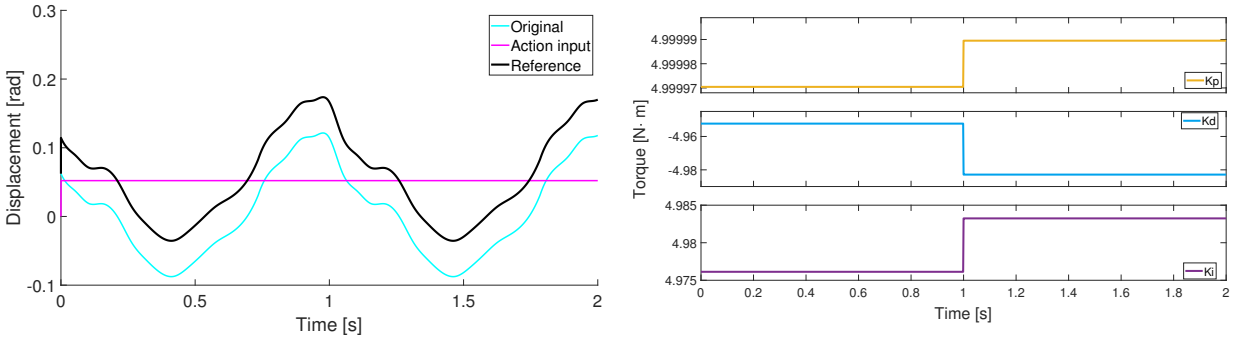
(a) Tracking evolution *(b) Tracking error*

(c) Control trajectories

Figure 4.7: Evolution of the right hip (abduction - adduction) joint: tracking, error, and control signals

Figure 4.7a shows the tracking outcomes corresponding to the tracking simulation of the hip that reaches the reference trajectory in approximately $10ms$. Once stabilized, the maximum amplitude

of the error for the RL-PID in the Abd-Add hip joint is $4.5 \times 10^{-3} rad$ and $0.1057 rad$ for the PID controller (see Figure 4.7b). With regards to the control signals generated after the stabilization point, (Figure 4.7c), the maximum value is $0.3215 N \cdot m$ for the RL-PID controller, and $22.7090 N \cdot m$ from the PID controller.



(a) Trajectories *(b) Control adjustments*

Figure 4.8: Actions of the Agents over the right hip joint (Abduction - Adduction): correction of the reference trajectory, and evolution of the gains

The actions performed by the agents of the RL can be seen in Figure 4.8. The adjustment of the reference trajectory for this joint is presented in Figure 4.8a, the color code is cyan for the original reference trajectory, the agent's action in magenta, and in black the final reference trajectory. The adjustment, occurring at $50 \mu s$, corresponds to $0.05214 rad$ from the original, and it is kept at that value during the walking.

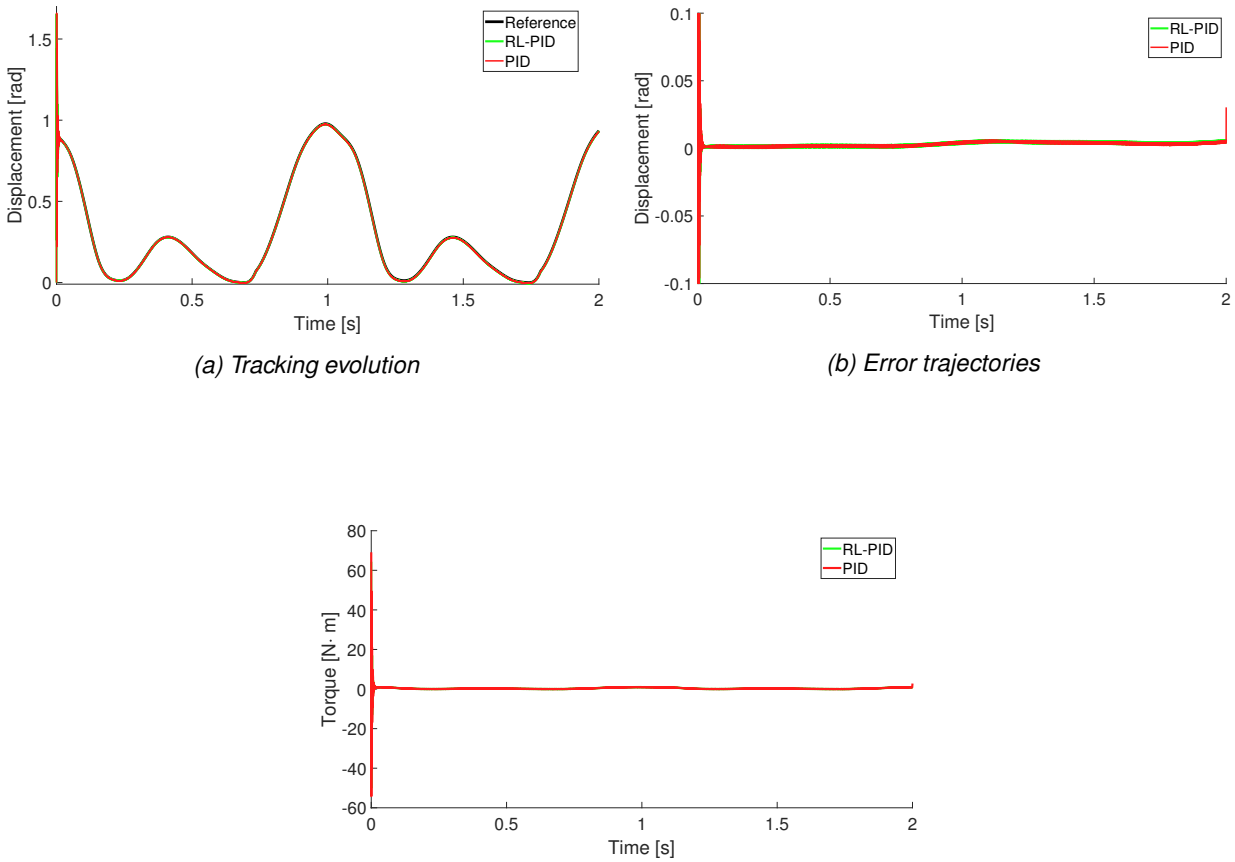
Alongside the actions of the trajectories agent, the gains agent evolution is depicted in Figure 4.8b. This result confirms that gains are modified since the beginning 4.99997 units to the original 75 units for K_P (depicted in orange color). After 1 second this gain was readjusted to 4.99999 units. The original K_D gain (0 units) was reduced 4.9523 units at the beginning. Then, after 1 second, it was readjusted to -4.98285 units (blue line). Regarding K_I gain, in purple and originally 0, it was set from the beginning to 4.9761 which, after 1 second, was readjusted to 4.9834.

Knee

Knee is the articulation that mobilizes the leg with respect to the thigh along the sagittal plane. The results of the simulation can be seen in Figure 4.9.

Figure 4.9a shows the outcomes of the tracking evolution of the knee, which reaches stability in approximately $12 ms$. The RL-PID controller led to a maximum error amplitude of $13.1 \times 10^{-3} rad$, and the PID $30.3 \times 10^{-3} rad$, depicted in Figure 4.9b. Regarding the control signals generated, shown in Figure 4.9c, the maximum of the RL-PID is $1.4155 N \cdot m$, while for the PID controller is

$2.7836N \cdot m$.



(a) Tracking evolution

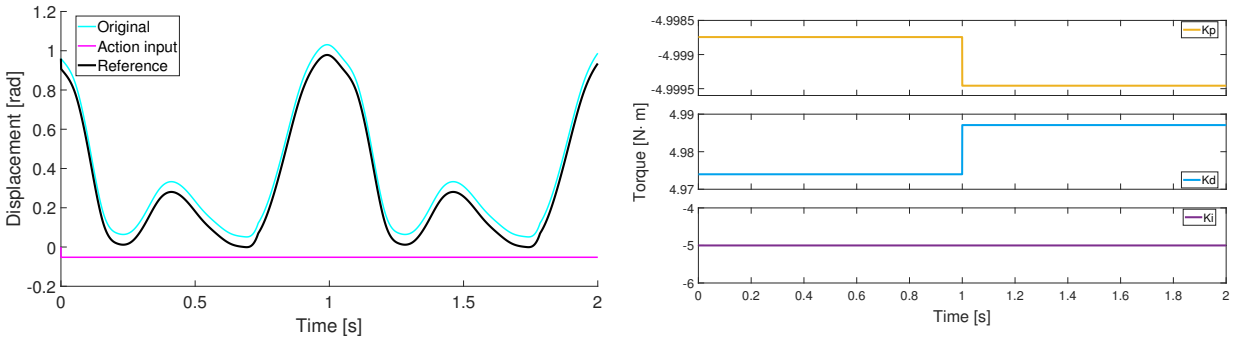
(b) Error trajectories

(c) Control signals

Figure 4.9: Evolution of the right knee joint: tracking, error, and control signals

The actions performed by the agents can be seen in Figure 4.10. The adjustment of the reference trajectory for this joint is presented in Figure 4.10a, where the original reference trajectory is presented in cyan, the action generated by the Agent is depicted in magenta, and black represents the final trajectory reference. It can be seen that at $50\mu s$ the trajectory is changed $-0.0522rad$ from the original, and it is kept during the walking.

Together with the actions of the first agent, the actions exerted by the second agent are presented in Figure 4.10b, where the action taken on each of the gains (K_P in orange, K_I in purple, and blue for K_D) can be seen. The K_P gain was modified by subtracting -4.99875 to its original value (72) since the beginning, and after 1 second the gain was readjusted -4.9994 units. The derivative gain K_D, preset to 45 units, was modified by the addition of 4.974 units since the beginning, modified after 4.9871. Finally, from K_I gain were subtracted 5 units.



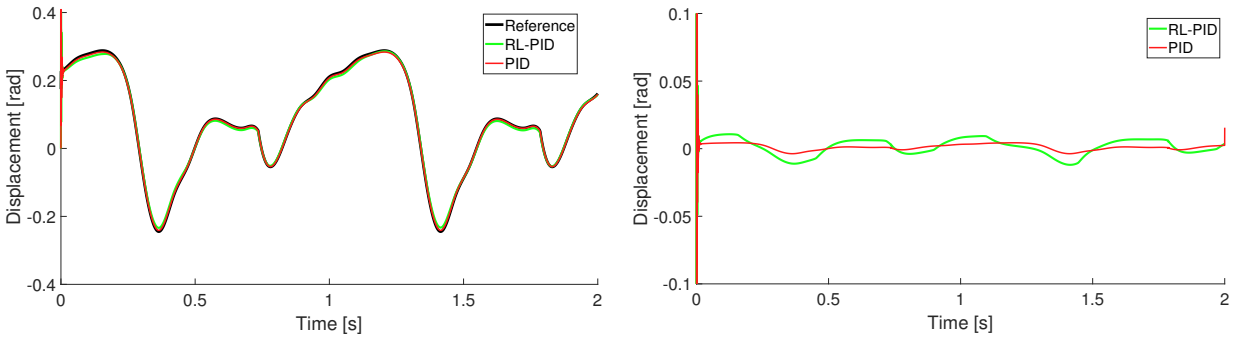
(a) Trajectories *(b) Control adjustments*

Figure 4.10: Actions of the Agents over the right knee: correction of the reference trajectory, and evolution of the gains

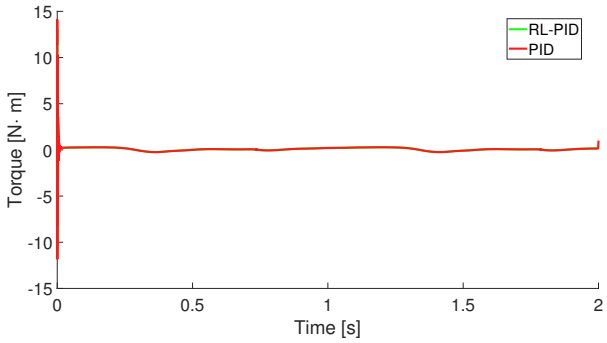
Ankle

Ankle is the articulation where the foot and the leg meet. The results of the simulation for this joint are presented in Figure 4.11.

Figure 4.11a shows the outcomes of the tracking simulation of the hip reaches the reference trajectory with both control schemes in approximately $8ms$. The tracking error evolution is shown in Figure 4.11b, after the tracking is stabilized, the maximum amplitude of the error for the RL-PID is $11.9 \times 10^{-3} rad$ and $15.5 \times 10^{-3} rad$ for the PID controller. The controller signals can be seen in the Figure 4.11c, which after being stable, the maximum amplitudes are $0.4370 N \cdot m$ and $1.0093 N \cdot m$ for RL-PID and PID controllers, respectively.



(a) Tracking evolution *(b) Error trajectories*

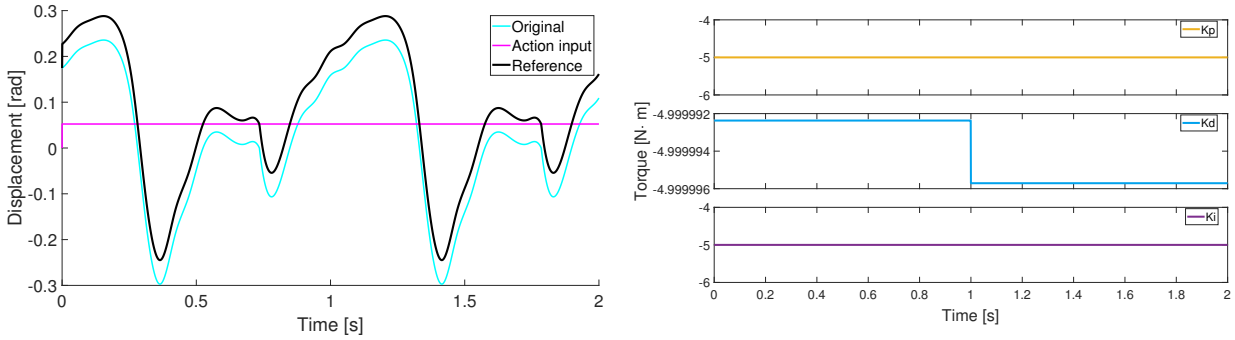


(c) Control signals

Figure 4.11: Evolution of the right ankle joint: tracking, error, and control signals

The actions performed by the agents can be seen in Figure 4.12. The adjustment of the reference trajectory for this joint is presented in Figure 4.12a, where the original reference trajectory is presented in cyan, the action generated by the Agent is depicted in magenta, and black represents the final trajectory reference. It can be seen that at $50\mu s$ the trajectory is changed $0.05235rad$ from the original, and it is kept during the walking.

Together with the actions of the first agent, the actions exerted by the second agent are presented in Figure 4.12b, where the action taken on K_P in orange, and K_I in purple, were constant, adding 5 units to each of them. The derivative gain K_D in blue, however, presented an adjustment of -4.99999 since the beginning, and another small one at 1 second of 3.33786 units. The preset gains were K_P 65, K_D 0, and K_I 0.



(a) Trajectories *(b) Control adjustments*

Figure 4.12: Actions of the Agents over the ankle knee: correction of the reference trajectory, and evolution of the gains

4.8.2 Left limb

The results obtained from the performance of both controllers for this leg are presented below.

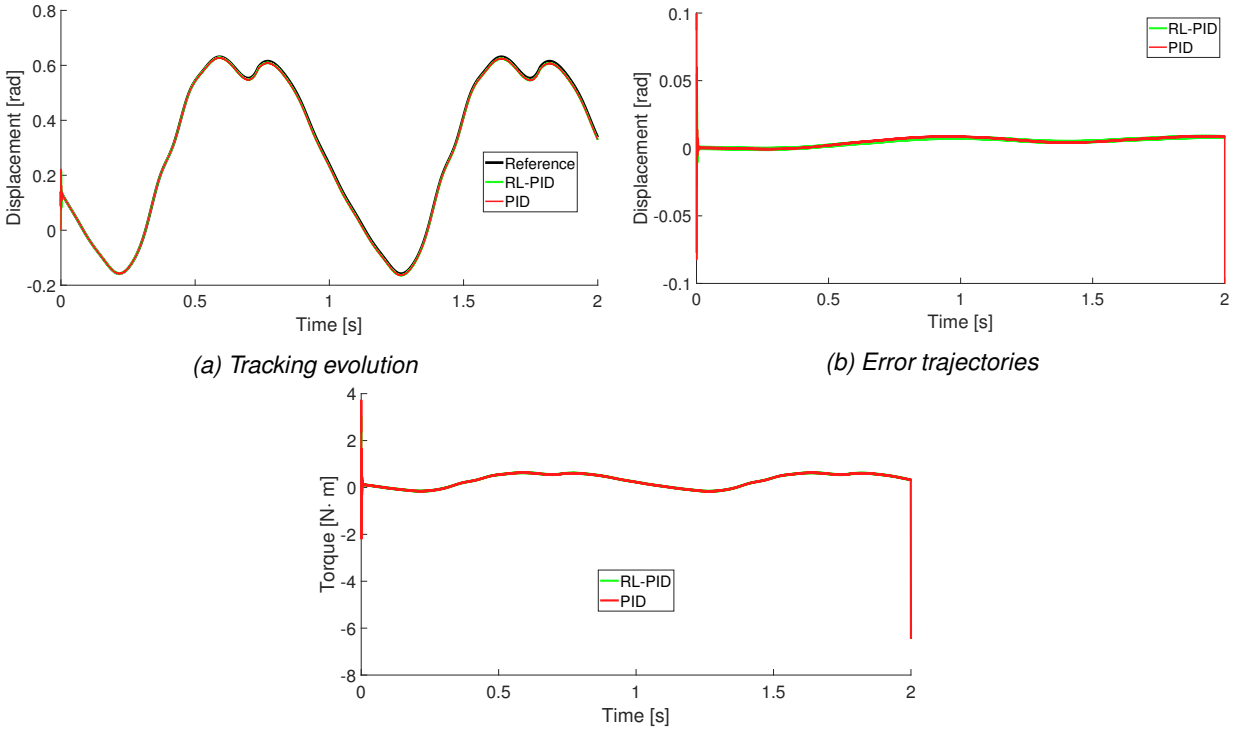
Hip

Hip is the articulation which connects the pelvis-equivalent channel with the limb. This joint mobilizes the entire limb in two movements: FE in the sagittal plane, and Abd-Add in the frontal plane.

Flexion - Extension

The results for the FE movement of the right hip are presented in Figure 4.13.

Figure 4.13a shows the outcomes of the tracking simulation of the hip reaches the reference trajectory with RL-PID control in approximately $15ms$, while PID controlled did it in $12ms$. The tracking error is depicted in Figure 4.13b, which once the tracking was stable, the RL-PID generated a maximum amplitude of $9.3 \times 10^{-3} rad$, while for PID it was $0.2425 rad$. With regard to the control signals generated by the use of both schemes, shown in Figure 4.13c, the maxima are $0.6477 N \cdot m$ and $6.4628 N \cdot m$ for RL-PID and PID controllers, respectively.



(a) Tracking evolution

(b) Error trajectories

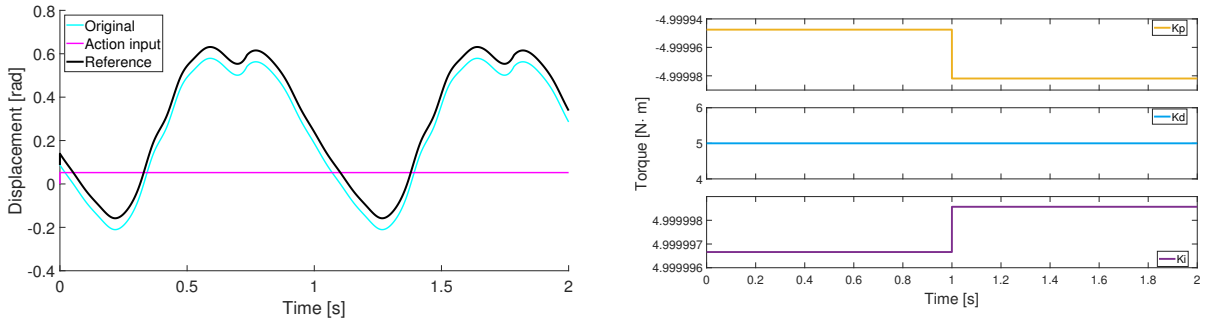
(c) Control signals

Figure 4.13: Evolution of the left hip joint (flexion - extension): tracking, error, and control signals

The actions performed by the agents can be seen in Figure 4.14. The adjustment of the reference trajectory for this joint is presented in Figure 4.14a, where the original reference trajectory is presented in cyan, the action generated by the Agent is depicted in magenta, and black represents

the final trajectory reference. It can be seen that at $50\mu s$ the trajectory is changed $-0.05236rad$ from the original, and it is kept during the walking.

In cooperation with the actions of the first agent, the actions exerted by the second agent are presented in Figure 4.14b, where the action taken on each of the gains (K_P in orange, K_I in blue, and purple for K_D) can be seen. The K_P gain was modified by subtracting -4.99994 to its original value (27) and at 1 second another 34.332×10^{-6}units, while the K_D gain were adjusted $+5$ units from its original settings at 23, and K_I varied from 17 (preset) to 4.999998, and at 1 sec to 4.999996.



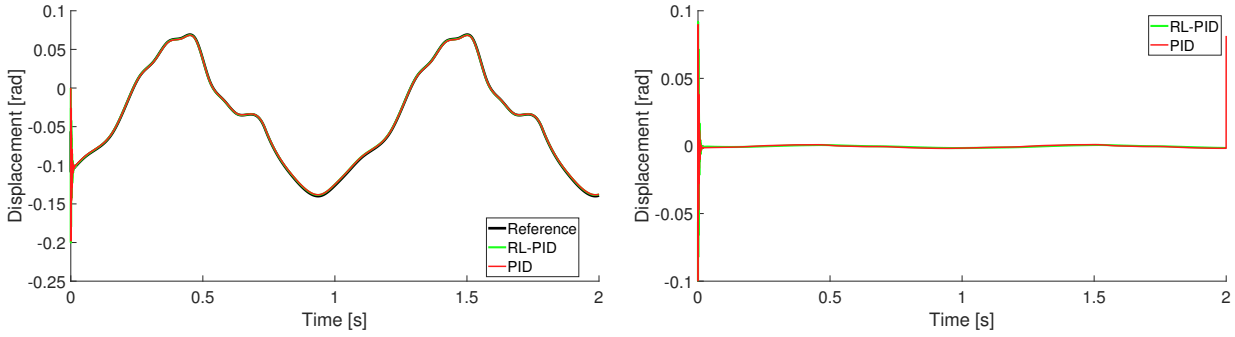
(a) Trajectories *(b) Evolution of the gains*

Figure 4.14: Actions of the Agents over the left hip joint (flexion - extension): correction of the reference trajectory, and evolution of the gains

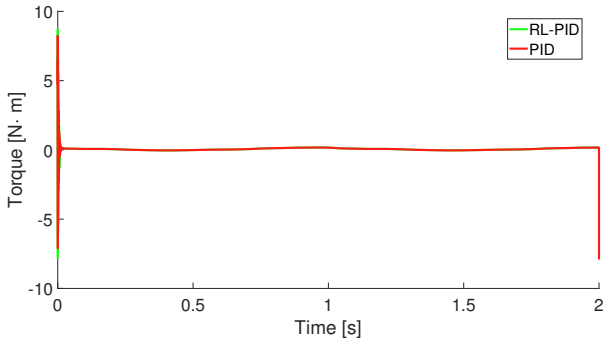
Abduction - Adduction

The results for the Abd-Add movement of the right limb are presented in Figure 4.15.

Figure 4.15a shows the outcomes of the tracking simulation of the hip where both reach the reference trajectory in approximately $15ms$, showing similar performance on both; however, once stabilized, the maximum amplitude of the error for the RL-PID controller is $16 \times 10^{-3}rad$ while the PID controller has a $0.0814rad$ amplitude (Figure 4.15b). With regard to the control signals generated after the stabilization point, shown in Figure 4.15c, the maxima are $0.154N \cdot m$ for the RL-PID controller, and $6.096N \cdot m$ from the PID controller.



(a) Tracking evolution *(b) Error trajectories*

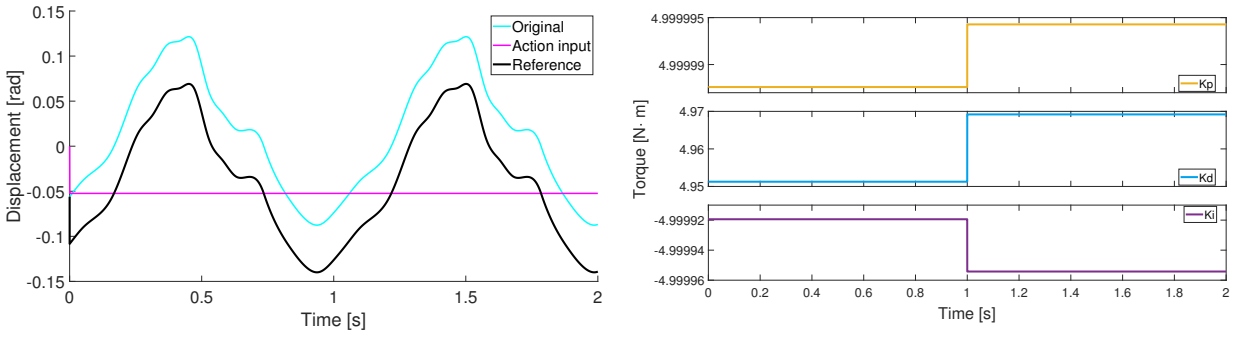


(c) Control signals

Figure 4.15: Evolution of the left hip joint (abduction - adduction): tracking, error, and control signals

The adjustments made by the agents to the tracking trajectory and gains are displayed in Figure 4.16. For the first set of actions, i.e., trajectory modification is presented the Figure 4.6a, the action exerted by the Agent (magenta) on the original trajectory (cyan) generated the final trajectory (black). This action is the constant subtraction of $0.0523rad$ It can be seen that at $50\mu s$ the trajectory is changed $-0.0522rad$ from the original.

Together with the actions of the first agent, the actions exerted by the second agent are presented in Figure 4.10b, where the action taken on each of the gains (K_P in orange, K_I in blue, and purple for K_D) can be seen. The K_P gain was modified by adding $4.999994units$ to its original value (75) since the beginning, and after 1 second the gain was readjusted to 4.99998 units. The derivative gain K_D, preset to 0 units, was modified by the addition of 4.951 units since the beginning, and at 1 s were added 0.0179. Finally, from K_I gain (preset at 10) were subtracted $4.99991units$ since the beginning to finalize in -4.99991.



(a) Trajectories

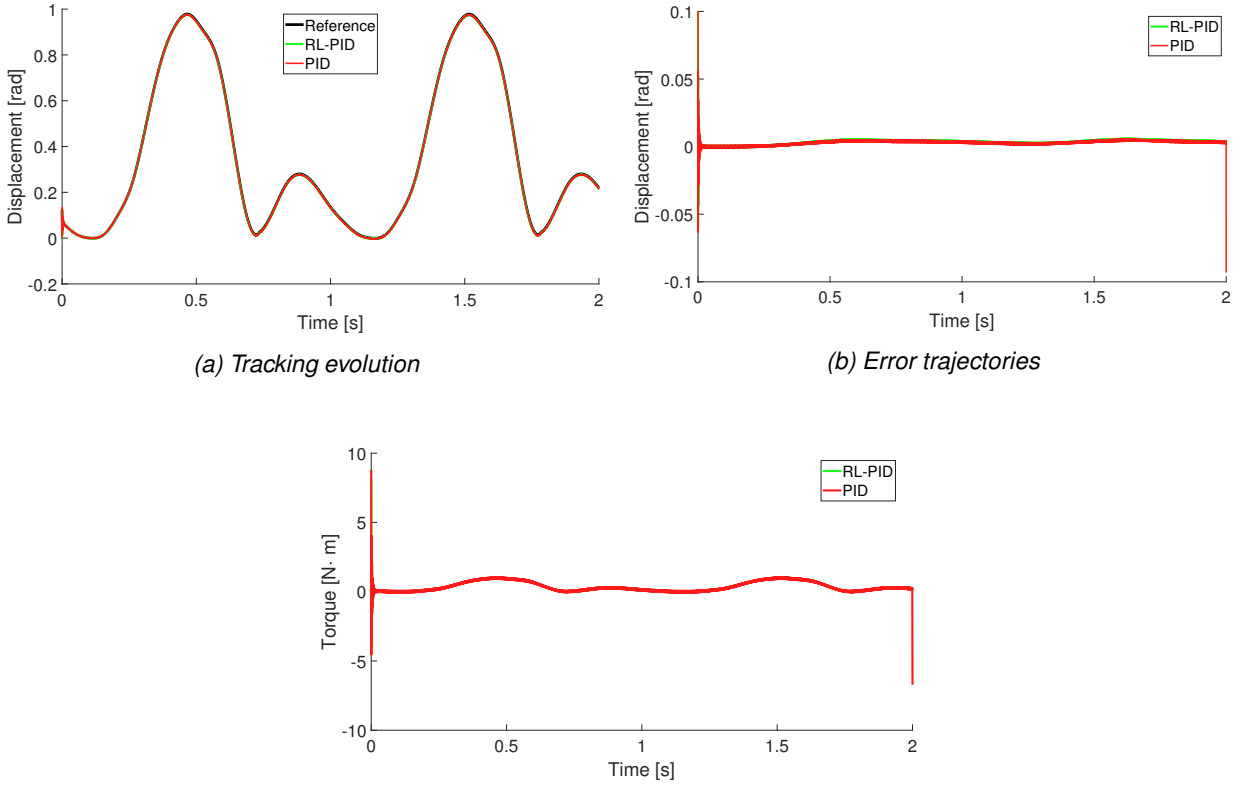
(b) Control adjustments

Figure 4.16: Actions of the Agents over the left hip joint (Abduction - Adduction): correction of the reference trajectory, and evolution of the gains

Knee

Knee is the articulation that mobilizes the leg with respect to the thigh along the sagittal plane. The results of the simulation can be seen in Figure 4.17.

Figure 4.17a shows the outcomes of the tracking evolution of the knee, which reaches stability in approximately $9ms$. The RL-PID controller led to a maximum error amplitude of $5.94 \times 10^{-3} rad$, and the PID $93.12 \times 10^{-3} rad$, shown in Figure 4.17b. Regarding the control signals generated at the given time, shown in Figure 4.17c, the maximum of the RL-PID is $1.01933 N \cdot m$, while for the PID controller is $6.71597 N \cdot m$.



(a) Tracking evolution

(b) Error trajectories

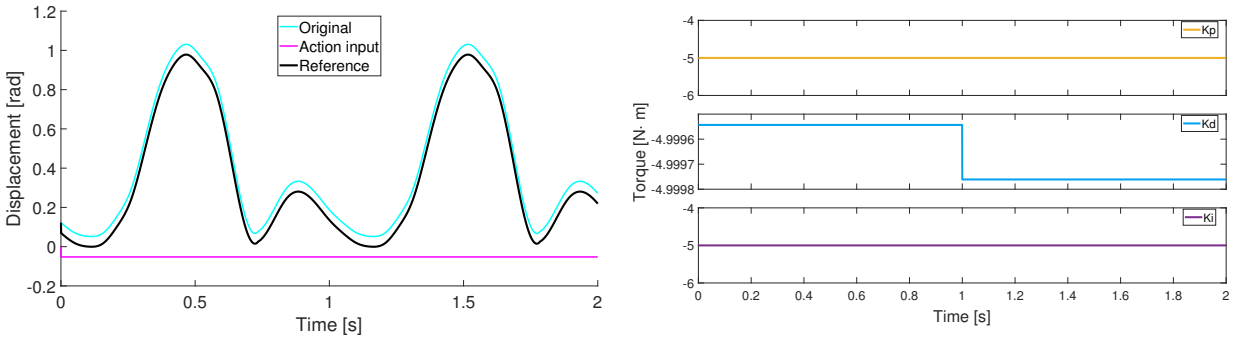
(c) Control signals

Figure 4.17: Evolution of the left knee joint: tracking, error, and control signals

The actions performed by the agents can be seen in Figure 4.18. The adjustment of the reference trajectory for this joint is presented in Figure 4.18a, where the original reference trajectory is presented in cyan, the action generated by the Agent is depicted in magenta, and black represents the final trajectory reference. It can be seen that at $50 \mu s$ the trajectory is changed $-0.05235 rad$ from the original, and it is kept steady during the walking.

The second agent's adjustments are presented in Figure 4.18b, where the action taken on K_P (orange, set at 72), and K_I (blue, set at 30), was the constant subtraction of $5 units$. In purple is

presented the evolution of K_D was modified by subtracting -4.99954 to its original value (45) since the beginning, and after 1 second the gain was readjusted -4.99976 units.

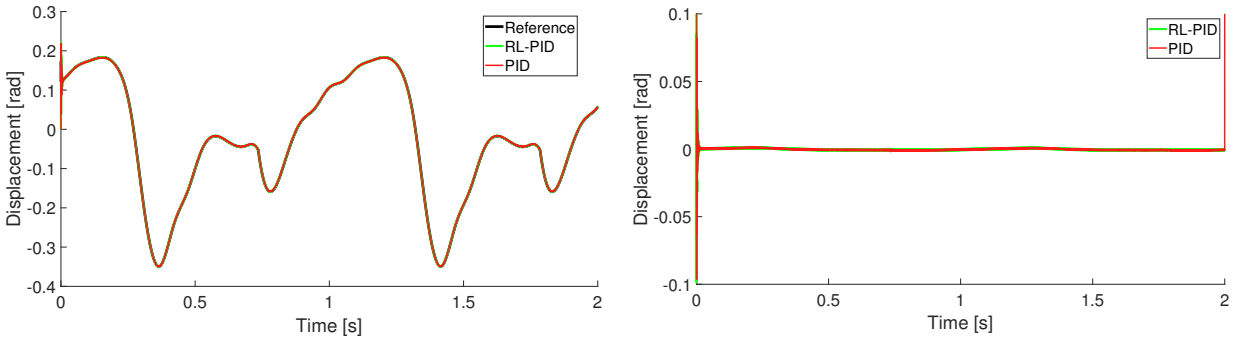


(a) Trajectories *(b) Control adjustments*

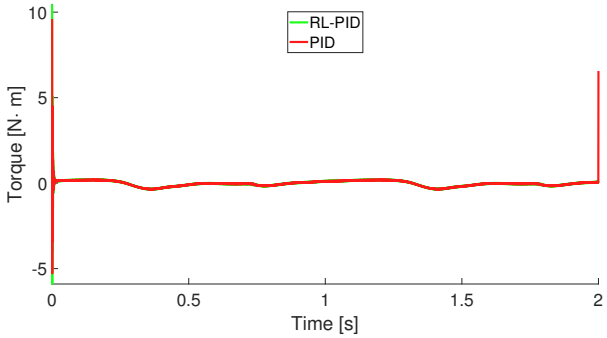
Figure 4.18: Actions of the Agents over the left knee: correction of the reference trajectory, and evolution of the gains

Ankle

The ankle is the joint where the foot and leg meet and for which the simulated results are shown in Figure 4.19. Figure 4.19a displays the development of the tracking simulation, it stabilizes in approximately $10ms$. The tracking error evolution is shown in Figure 4.19b, after the tracking is stabilized, the maximum amplitude of the error is $2.6 \times 10^{-3} rad$ for the RL-PID control, and $117.48 \times 10^{-3} rad$ for the PID controller. The control signals, displayed in Figure 4.19 agree with the results obtained in the evolution of the errors since the maximum amplitude of the RL-PID controller is barely $250.87 N \cdot m$, while for the PID is $6.55892 N \cdot m$, being at its maximum point 25.14407 times greater than the proposed one.



(a) Tracking evolution *(b) Error trajectories*

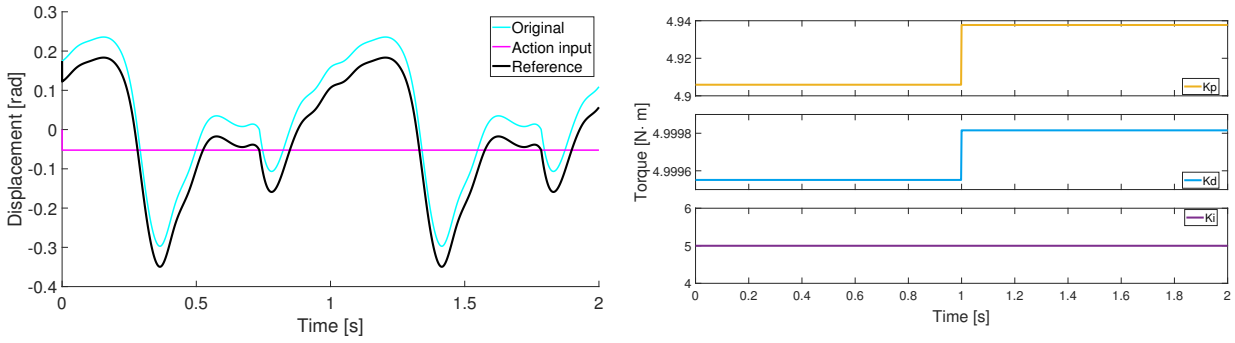


(c) Control signals

Figure 4.19: Evolution of the left ankle joint: tracking, error, and control signals

The actions performed by the agents on this articulation can be seen in Figure 4.20. The adjustment of the reference trajectory for this joint is presented in Figure 4.20a, where the original reference trajectory is presented in cyan, the action generated by the Agent is depicted in magenta, and black represents the final trajectory reference. It can be seen that at $50\mu s$ the trajectory is changed $-52.35 \times 10^{-3} rad$ from the original, and it is kept during the walking.

As to the adaptation of the gains, presented in Figure 4.20b, the initial K_P gain (orange line) was set to 55 units, being modified two times by the addition of 4.9059 units, at the beginning, and 0.03182 at 1 sec. The initial proportional gain K_D (set at 31, blue line) was also modified twice: at the beginning by the addition of 4.99955, and at 1 sec to finalize adding 4.99981. K_I, in purple, was a constant addition of 5 to the originally set gain of 17.



(a) Trajectories *(b) Control adjustments*

Figure 4.20: Actions of the Agents over the ankle: correction of the reference trajectory, and evolution of the gains

4.8.3 Overall performance

The evaluation of the performance indexes J_c and J_e in Table 4.3 were calculated as $J_c = \int_0^{t_f} \left|u_i^2\right|$, with u as the control vector made up of the signals obtained from the eight joints with both RL-PID, and PID controllers, and $J_e = \int_0^{t_f} \left|e_i^2\right|$, with e being the vector of tracking errors measures in the whole structure, for both controllers. Numerical integration was performed by the trapezoidal method, which approximates the integration over an interval by breaking the area down into trapezoids with more easily computable areas. The norm J_e for the RL-PID is $.4\%$ greater than the PID, However, the L_2 norm signal for the control signals is 16.73% smaller for the RL-PID than the one obtained using the PID. This implies that the benefit of reducing the energy consumption to mobilize the robot is sufficient for its choice over a classical feedback controller.

Table 4.3: Value of the two-norm corresponding to the tracking error and control actions of the robotic joints using RL.

Parameter	Controller	
	RL-PID	PID
J_c	92.57198×10^3	108.0661×10^3
J_e	18.07623	17.99555

Chapter 5

Experimental setup for the biped robot

The design and assembly of the BR was based on a linear methodology with feedback, it means that each stage was developed sequentially then, once those stages were finished, the prototype was assembled and proved each section individually and as a whole.

The mechanical structure was assembled using prefabricated aluminium structures, which was electronically instrumented in order to measure the angular movements at each DOF using linear actuators with embedded angular resistors as an indirect measurement tool and then, attain an input trajectory-position. Those linear actuators were set such that the trajectory tracking at each joint could be reached.

The stages that were followed to finish the structure are described in this chapter.

5.1 Introduction

The proposed controller was tested on a real biped robot that was specifically designed for this study. The structural components of the biped robot were developed in a manner that integrates the mechanics well with the electrical components required for actuation or measurement of the robot's action.

Linear actuators were used to provide individual motion at each joint for the bipedal robotic device. The positions of these actuators were chosen to induce maximum angular displacements at the hip, ankle, and knee. The actuators used were 2, 4, and 8 inches in stroke length, with 115 lb Thrust Heavy-Duty Linear Actuation devices HDA2-50, HDA4-50, and HDA8-50, respectively (ServoCity, USA). These devices feature a 12.0 volts direct current motor with a maximum power of 200 Watts. Each actuator has a linear variable resistor that provides indirect measurement of the linear actuator position. The voltage is acquired with the analog-to-digital converter module (12

bits, sampling period of 0.5 ms, and voltage range of 3.3 volts) included in the TIVA1294 C-Series board (Texas Instruments, USA), which has a 120 MHz 32-bit ARM Cortex-M4 CPU processor, a dual 12-bit 2 MSPS analog-to-digital converter (ADC), 1 MB Flash memory, 256 KB SRAM memory, and 6 KB EEPROM memory.

To mobilize each linear actuator at each articulation, a direct-to-analog power converter (ST's™IC VNH5019 AC-DC two-branch power converter in Pololu's™carrier board, USA) is utilized. The power converter is augmented with an optoisolation to provide power separation between the digital board (TIVA1294) where the controller is implemented and the linear actuator, ensuring the safety of the digital device. The implementation of the controller, as given in eq. (3.13), uses an explicit Euler discretization based on a first-order approximation (Euler type, ODE1), and is programmed into the TIVA1294 C-Series (Texas Instruments, USA) board with a sampling time operation of 0.5 ms.

The bipedal robot was designed to satisfy the angular values detected in a typical gait cycle [26], which include step length, cadence sequence, and walking velocity of 72 cm/s, 1.87 steps/s, and 1.37 m/s, respectively, corresponding to an individual with a height of 180 cm. The mechanical structure of the bipedal robot can be divided into two parts: (1) the hip section that is connected to both lower limbs, with a length of 33.81 cm, and (2) the bilaterally symmetrical lower sections separated by the knee joint, comprising a 45.72 cm thigh and a 41.19 cm leg.

To ensure that the angular variations of the bipedal robotic device align with a natural walking cycle, it incorporates four degrees of freedom (DOF) along the sagittal and coronal anatomical planes. The simplified structure features a single joint associated with the ankle, which enables the plantarflexion/dorsiflexion movement.

The articulation arrangement of the bipedal robot converts the linear displacement of the actuator to angular displacement, making it a significant contribution to this study. This configuration is considered a simplified option for the instrumentation of newly developed bipedal robots. The design aims to introduce a novel, robust mechanical section to the robot's arrangement.

The proposed instrumentation shown in Figure 5.1 comprises an electrical section for measuring joint displacement. To actuate the articulation joints, technical elements were selected that correspond to the linear actuation device. The mechanical settings considered when selecting the actuators include maximum load displacement of $3.5\ cm/s$, dynamic thrust of 11.34 KgF, static load of 226.8 KgF, and a single threaded spindle with a $3mm$ pitch length.

The actuation engines have an electrical voltage active range of $6-12V$ and can handle an electrical current of up to $3.8A$ at maximum weight load. Each linear actuator contains a $10K\Omega$ variable resistor device for position sensing, with a tolerance of $\pm 5\%$ of the value. The angular variation of the joint is linearly dependent on the shank movement. The electrical section operates on a voltage range of $5.5-24V$, delivering $12A$ continuously and working with logical levels of

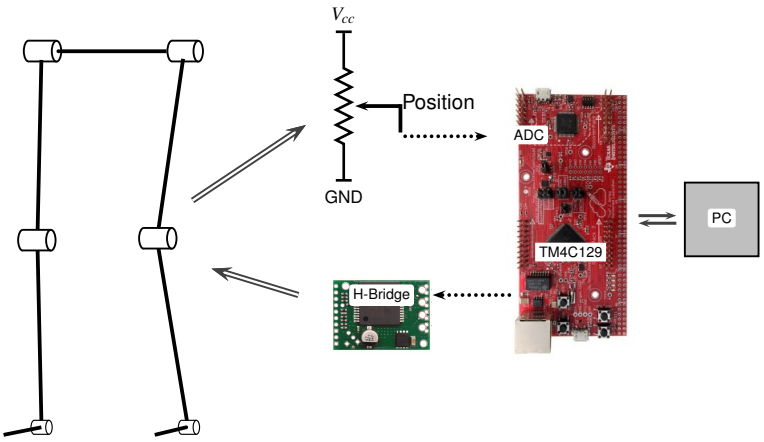


Figure 5.1: Simplified version of the electronic instrumentation system used for each articulation in the biped robot.

$2.5 - 5V$. The device supports square pulsed width modulated signals at frequencies greater than $20KHz$. Additionally, the device has built-in electrical protection against reversal-voltage, overload-voltage, under-voltage, over-temperature, and over-current.

The proposed electrical instrumentation configuration justifies the design of the suggested control system by taking into account the constraints of the actuators used in the bipedal robot's operation. Moreover, the embedded electrical instrumentation is capable of processing the control action that has been designed. These electrical components are assumed to work with the digital section, which implements numerical representations of the control design.

5.2 Design approach

Even the fact that the human gait is an individual characteristic, there have been many efforts to standardize not only the movements made during each phase of the gait cycle, but also to find *normal* values that can describe the human walk.

The gait descriptors listed in [26] were obtained averaging results of three independent investigations carried out in New York, Philadelphia and Amsterdam, in which 2300 adult pedestrians, male and females, that did not know that they were watched while they were walking were measured.

The BR design was based on the spatial and temporal so-called gait descriptors, they are: stride length, which is the distance between two successive heel contacts of the same foot, in this case, it was considered 72 cm; cadence, that is the number of steps per minute, for the BR was 1.87 steps/s; the walking speed, that is the distance covered in a given amount of time, it was considered

1.37 m/s.

Besides the previous parameters, it is necessary to take into account the maximum movement amplitude at each considered articulation for the correct movement of the BR. These movements must be do at least in two anatomical planes. Generally, sagittal and frontal planes [136] are enough for a human-like gait. In the Table 5.1, it is observed the previously mentioned amplitude movements.

Table 5.1: Minima angle movements at each joint considered for the execution of the gait by the BR [137]

	Sagittal Plane		Frontal Plane	
	Maximum	Minimum	Maximum	Minimum
Hip	35°	-10°	7°	-5°
Knee	60°	5°	-	-
Ankle	15°	-20°	6°	-2°

The angles content in the Table 5.1 led to designing a 12 DOF BR that are distributed as follows (see Figure 5.2):

- DOF for hip Abd-Add movement (θ_1)
- DOF for hip FE movement (θ_1)
- DOF for knee FE movement (θ_2)
- DOF for ankle FE movement (θ_3)

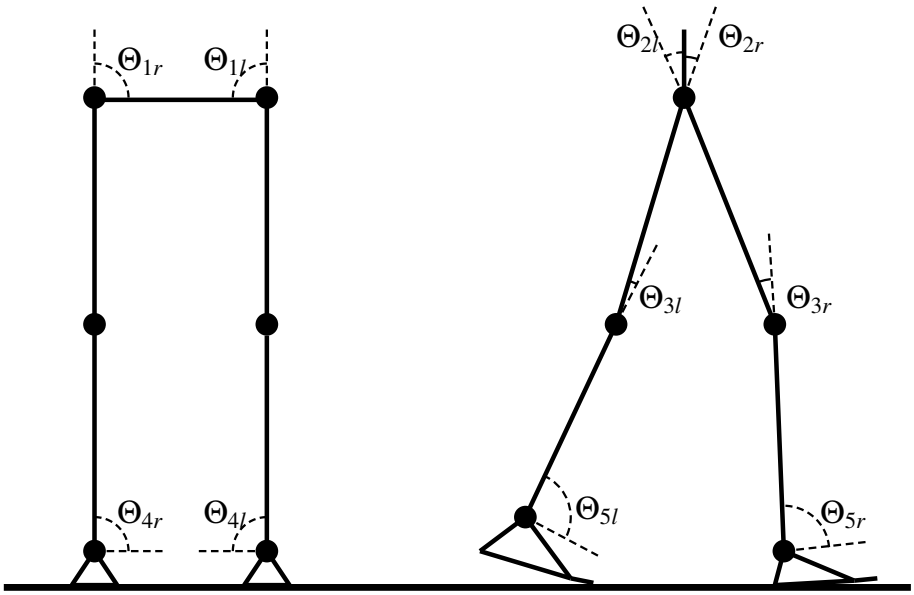


Figure 5.2: Angular movements in the BR

This DOF arrangement corresponds to movements made on sagital and frontal planes, both planes let the execution of a three-dimensional gait [138, 139].

The range of motion presented on the Table 5.1 must be considered for designing reference trajectories through Artificial Neural Network (ANN) meanwhile, at the same time, must be reachable by the controller. For the implementation of the closed-loop controller it is necessary a complete electronic scheme that allows not only measure the sensors but also drive each joint actuator. All of them managed in a couple of embedded development boards connected between them. Those boards must have one acquisition module for each sensor, two digital ports for each actuator and a processing speed that supports the necessary clocking frequency for designing reference trajectories and their tracking by the controller.

5.3 Mechanical Design

As it was previously mentioned, the mechanical structure of the BR was assembly using Servocity ®'s prefabricated aluminum structures. They were chosen because of its lightness and the hole pattern makes easy assemble another structural elements.

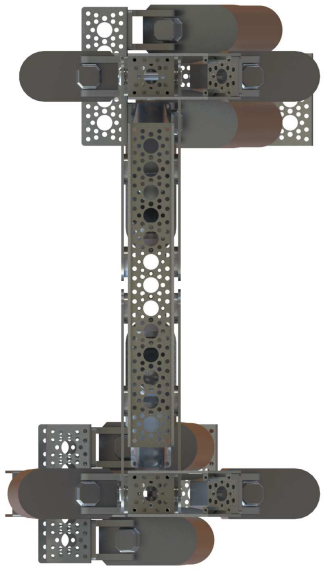
The election of lengths was made using as reference the anthropomorphic proportions of an 180 cm height person, which correspond to the characteristics of the gait analysis study of [26]. The anatomical movements made by each DOF are condensed in the table 5.2.

Table 5.2: Movements made at each DOF by the BR

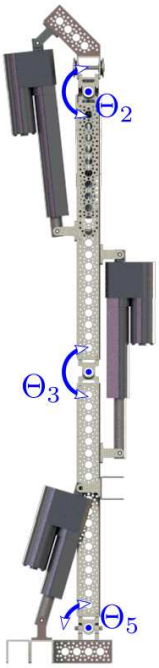
	Sagittal Plane	Frontal plane
Hip	Flexion-extension Θ_{r2}, Θ_{l2}	Adduction-abduction Θ_{r1}, Θ_{l1}
Knee	Flexion-extension Θ_{r3}, Θ_{l3}	-
Ankle	Plantarflexion-dorsiflexion Θ_{r5}, Θ_{l5}	Inversion-eversion Θ_{r4}, Θ_{l4}

Structurally, the joints are based on pillow block bearings attached to u-channels, this mechanical arrangement is adequate to turn the linear movement made by the actuator into angular movement through a rotation axis. Each articulation set-up described is bilaterally symmetric for both limbs.

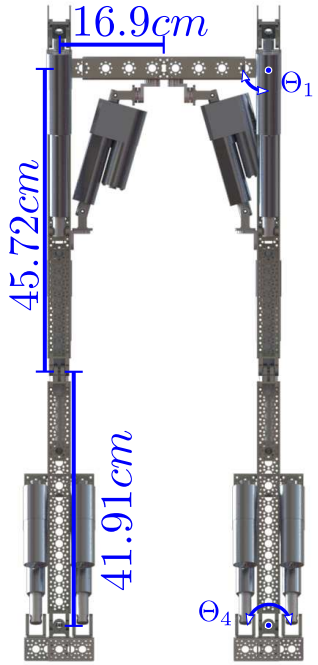
(a) Isometric



(b) Superior



(c) Left



(d) Frontal

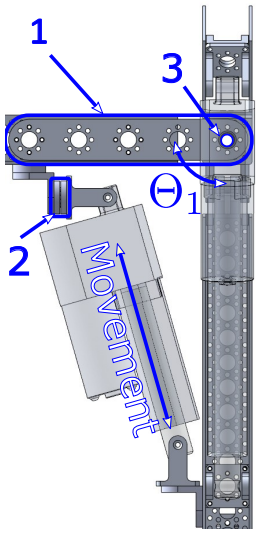
Figure 5.3: Schematic views of the proposed biped robot detailing dimensions and angles taken into account for the analysis of the mechanical structure for the Biped Robot

5.3.1 Linear Actuators

There exists many way to address how the joints are mobilized in this kind of robots and it must be taken into account variables as speed and position. A wide variety of actuators are used in many humanoid robots, including gearboxes (harmonic drive or planetary gears), hydraulic actuators, electric motors, series elastic actuator [10, 140, 141] and LA [142–146]. The latter mainly consist on an electrical DC motor attached to a ball-screw driving system whose characteristics include resizeability, high torque, robustness against external perturbations, among others; this kind of actuators also could be used as a mechanical stop for the structure.

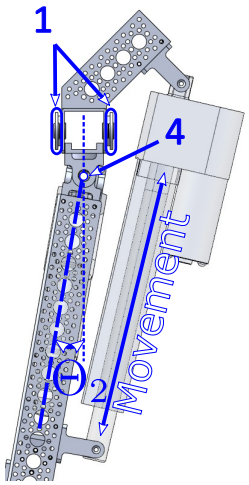
With all this advantages, the mechanical arrangement was designed to embed linear actuators enabling the complete range of movements referred in 5.1.

In the Figure 5.3 are displayed the dimensions and angles in the BR: in 5.3a and 5.3b from an isometric and top views, respectively, the position of each LA is also displayed, meanwhile in the 5.3c is displayed the position of the three LA used for make movements in the sagittal plane (Θ_2, Θ_3 and Θ_5). Figure 5.3d displays lengths corresponding to the distance between rotation axes and the movements along the coronal plane (Θ_1 and Θ_4). A thorough description of each articulation is detailed below.



(a) Mechanical configuration for the abduction-aduction movement

5.3.2 Hip articulation



(b) Mechanical configuration for the flexion-extension movement

Figure 5.4: Structure conforming the hip. This configuration allows the hip movements in two anatomical planes

The hip articulation can be seen in Figure 5.4 and is designed as a spherical joint enabling two-plane movement and is considered as two intersected axes. On the first hand, the movements are made in the coronal plane, its structure consists of a couple of flat brackets, as shown in the Figure 5.4a (labeled with number 1), front and back working as connecting elements between hip and limbs structures. This movement is equivalent to Abd-Add and is guided by an 2 in - length linear actuator nudging the entire limb structure (including the FE structure) when rotating around a ball bearing embedded in the flat brackets, labeled 3. The number 2 element allows a simultaneous movement of the abduction-adduction lineal actuator with the FE displacement. The FE rotational axis is marked with the number 4.

On the other hand, the FE movements (5.4b) is made in the sagital plane using a couple of pillow blocks bearings attained to a square u-channel rotating around the axis numbered 4. This structure transfers the movement made by the 6 in - length linear actuator to the leg.

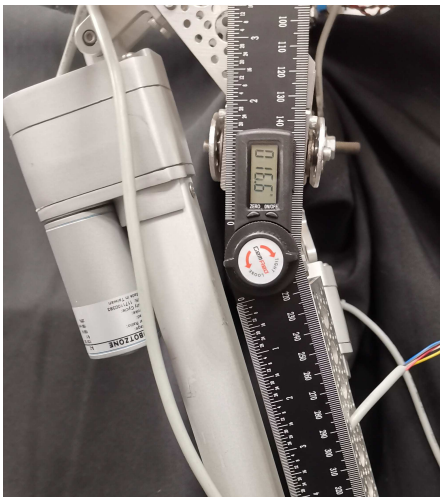
The final configuration of the hip joint can be seen in the Figure 5.5. The linear displacement of the $1in$ actuator is seen in the Figures 5.5a and 5.5b, according to the mechanical arrangement, resulting in the Abd-Add movement. In the same way, Figures 5.5c and 5.5dshow the movement of the $6in$ actuator rod that allows the FE movement.

The movement range reached by the Abd-Add actuator spans from $-3.9°$to $19.8°$and for the FE actuator the range goes from -$16.6°$to $46.4°$, for all of them, it is signed the zero reference as

the angle where the legs are parallel to the coronal axis of the robot. These measures warrant the minima angle movements presented in 5.1.

(a) Adduction movement in the BR

(b) Abduction movement in the BR



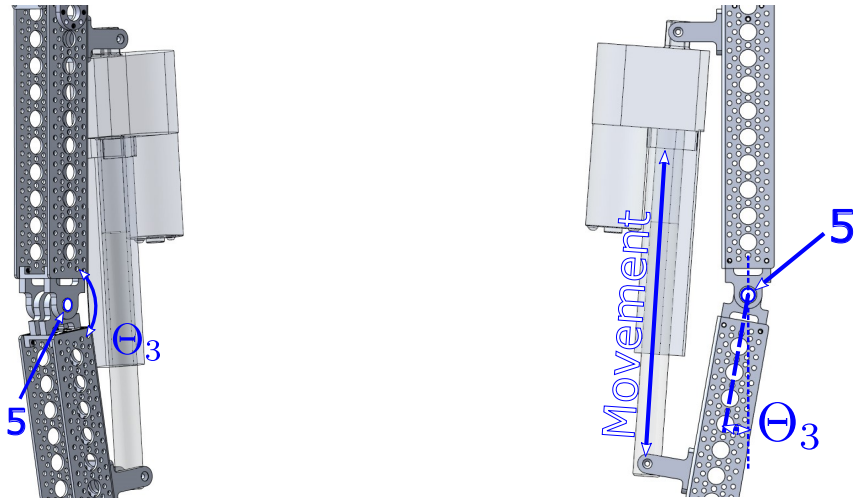
(c) Maximum angle reached by the hip for the extension movement

(d) Maximum angle reached by the hip for the flexion movement

Figure 5.5: Final assembly of the hip

5.3.3 Knee articulation

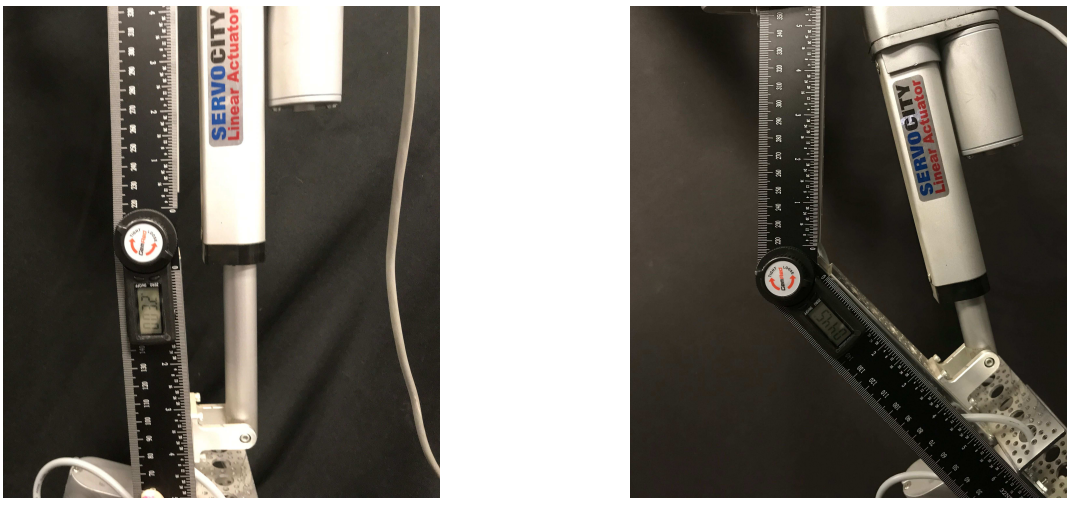
The knee articulation was designed as hinge joint, i. e., the poli-centricity of the knee articulation was not taken into account, it consists in a 4 in - length linear actuator attained to the u-channel equivalent to the thigh The linear actuator rod displacement is marked in the figure 5.6, this linear motion is turned into rotation by the axis.



(a) Isometric view *(b) Side view*

Figure 5.6: Knee articulation

The final configuration of the knee joint is shown at Figure 5.7, from a lateral perspective. It shows its structure hinge type and the axis of rotation of the articulation (joint). It can be seen as well the movement of rotation executed from the linear movement of the actuator, this angular movement is corresponding to the FE of the articulation. The range of movement for the knee joint goes from $3.2°$to $44.5°$, using as reference to measure it in the DOF the previous link to the joint, considering it as $0°$ when the subsequent link is parallel to the axis of the preceding link.

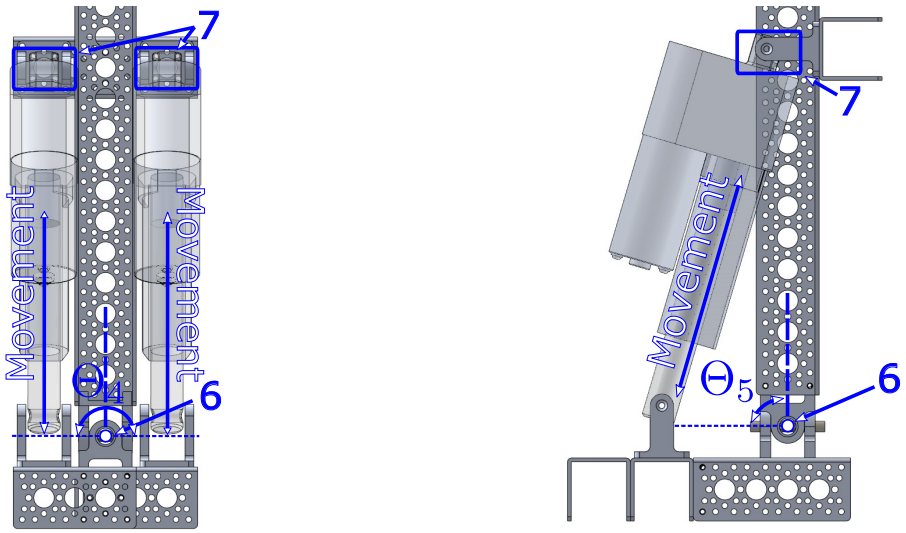


(a) Maximum angle for extension movement *(b) Maximum angle for flexion movement*

Figure 5.7: Final assembly of the knee

5.3.4 Ankle Articulation

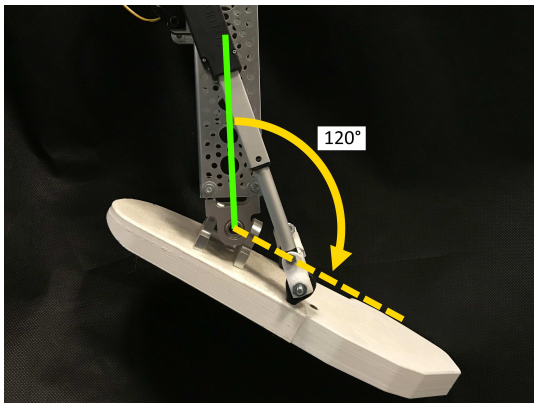
For the ankle, the joint was also designed as a two-dimension movement based on and universal joint, i.e., a pair of pillow block bearings oriented at $90°$to each other connected by a cross shaft 5.8, number 6. The movement of the articulation was made by a couple 2 in length LA, attached to the shank-like channel (marked as 7), moving coordinately: when both move at the same time and direction the foot makes a dorsi-plantar flexion movement; whereas if the movement of both LA is not coordinated the movement made is inversion and eversion of the foot.



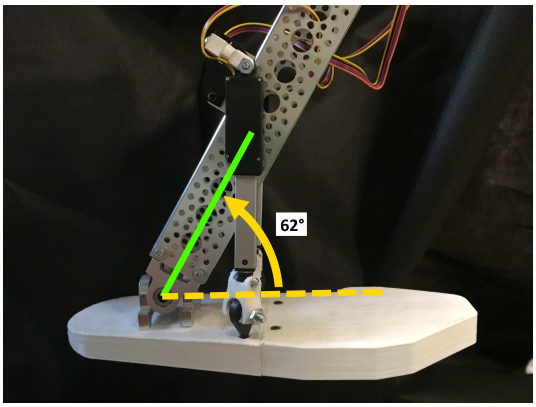
(a) Fontal view *(b) Lateral view*

Figure 5.8: Ankle articulation

The final structure of the ankle can be seen in 5.9. It is shown the universal joint articulation that supports the 2D movement, as well as the base of the LA. The range of movement for the dorsi-plantar flexion, i.e., when both LA move synchronously covers from $62°$5.9a to $120°$5.9b, also considering that 0 when the foot link is parallel to the axis of the shank-like link. Alternatively, the inversion-eversion movement reaches $72°$during the eversion movement and $108°$during for the inversion of the foot.



(a) Lateral view. Maximum plantarflexion angle



(b) Lateral view. Maximum dorsiflexion angle

Figure 5.9: Maxima angular displacements done by the ankle articulation

5.4 Electronic Design

The integrated system fromed by the proposed controller and the mechanical structure was driven by an electronics arrangement based on an development board reading the sensors embedded into the LA, then, once calculated the controller, the movement was driven by H-Bridges through Pulse Wwidth Modulation signals. In the following paragraphs, a detailed description of each component is given.

5.4.1 Development board

The Texas Instrument's EK-TM4C129XL TIVA C Series LaunchPad™ evaluation electronic board was used as main electronic component letting communication between the sensors and actuators through the implemented controller.

The main characteristics considered in the selection of the electronic board are sown in Table 5.3.

Three GPIOs were used for each Linear Actuator: a couple of for drive the H-Bridge and one for the position sensor. Since handling the LA implies the utilization of high voltages and currents, the following isolation system was needed to protect the TIVA.

Table 5.3: Main characteristics of EK-TM4C129XL TIVA

Feature	Quantity	Description
Processor	-	ARM Cortex-M4F
Core	-	120 MHz; 150 DMIPS

Feature	Quantity	Description
External Peripheral Interface	-	8/16/32 bit dedicated interface
GPTM	8	16/32-bit
Flash	-	1024 KB
EEPROM		6KB
SRAM		256 KB single-cycle System
UART	8	-
GPIO	15	-
PWM	8	-
ADC	2	12-bit, 2 MS/s

5.4.2 Optical isolation

In order to separate the logic electronic part from the power voltages, an 4N25 Photocoupler was selected. Its election was leaded considering the response (up to $500K\Omega$, ease of installment and low energy consumption. In the Figure 5.10 [147] is displayed the schematic internal structure.

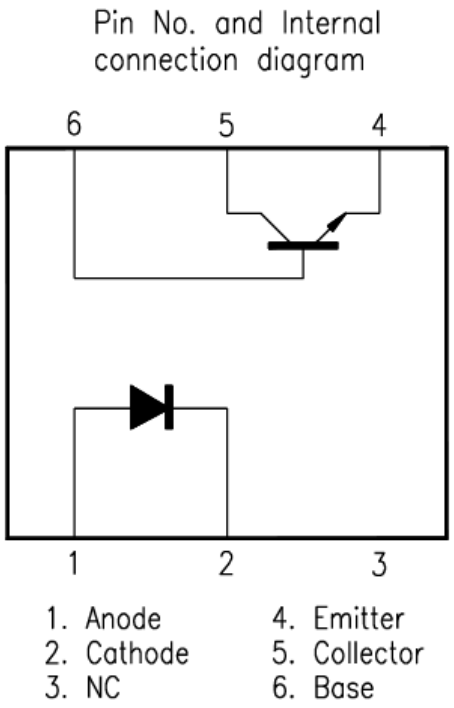


Figure 5.10: Schematic internal structure of the 4N25 Optocoupler

The internal functioning of the IC is based on a LED and a transistor working in switched-mode.

The first part of the design consisted on the election of proper resistors to let the LED to its maximum bright.

Having the forward current (i_F) to reach the maximum bright of LED equals to $10mA$ and taking into account the output voltage as 3.3 V (delivered by the TIVA™), using Ohm's Law to set the appropriate resistance, we have :

$$R = \frac{V}{i_R} = \frac{3.3V}{10 \times 10^{-3}A} = 330\Omega \tag{5.1}$$

Once it was assured the maximum brightness, the second part of the design is work with the transistor. The cutoff region of its operation occurs if the base lead is open, then the current I_{CEO} is extremely small thus can be neglected so that $V_{CE} = V_{CC}$, in our case:

$$V_{CE} = V_{CC} = 5V \tag{5.2}$$

For the saturation region, and isolating R_C we have:

$$R_C = \frac{V_{CC} - V_{CE_{sat}}}{I_{C_{sat}}} = 2.45K\Omega \approx 2.7K\Omega \tag{5.3}$$

Once selected the values of the resistors, the circuit shown in Figure 5.11 was assembled, where R1 stands for the output resistor ($2.7K\Omega$) and R2 notes the input resistor (330Ω). The base of the transistor was no connected. The label marked as TIVA notes that the signal comes from the development board. PWM label means that the output go to the H-Bridge to mobilize the LA. It is also shown that each section of the optocoupler (input and output) has its own ground to prevent reverse current, short circuit, electrical surges or spikes, among others circumstances that could damage the electronics.

For each limb of the BR one electronic board was assembled. Each pair of integrated circuit is used for one LAs, thus allowing forward and backward movement of each joint, the final assembly is displayed in Figure 5.12. In the design each ground (input and output) is shared among all the others. Number one indicates the connection to the $5V$ voltage source, supplying energy to the output part of the optocoupler. Number two signs the ground for the input part. Number three tags the input, as well as number four, does it for the output. Both resistors are labeled, R2 for the input one and R1 for the output.

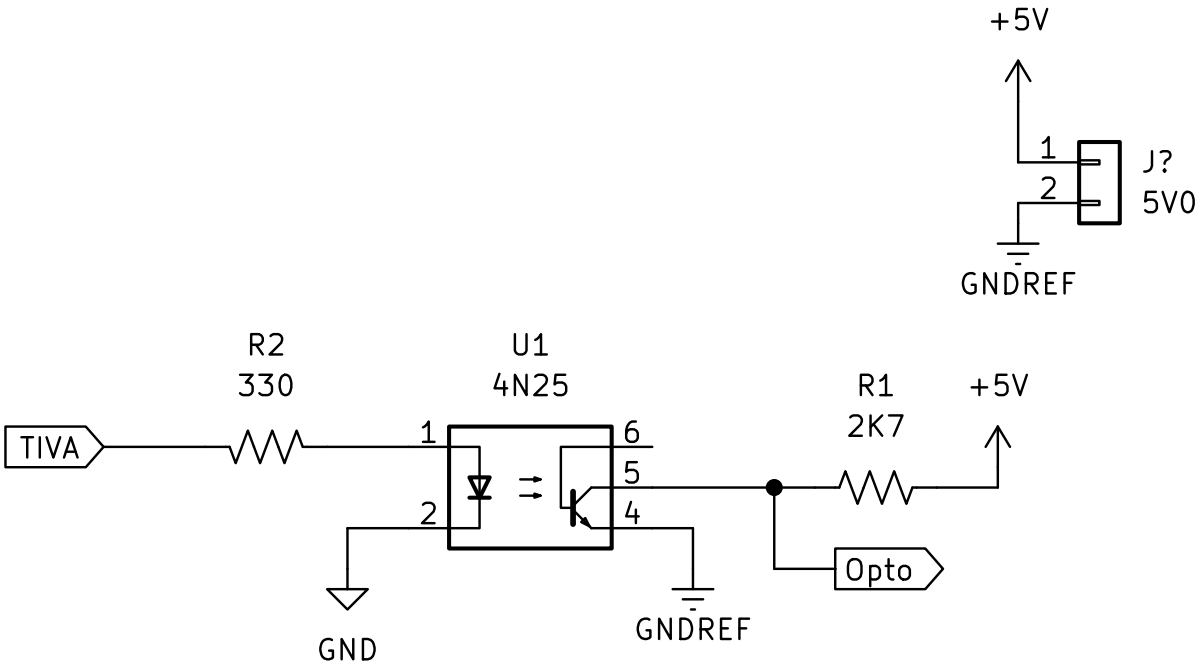
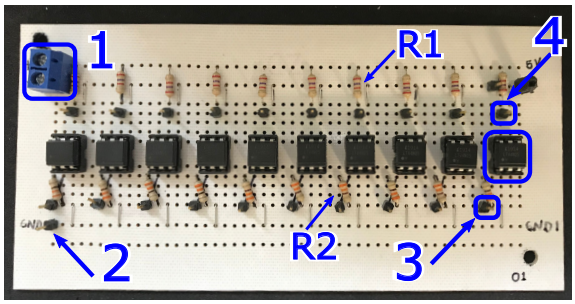
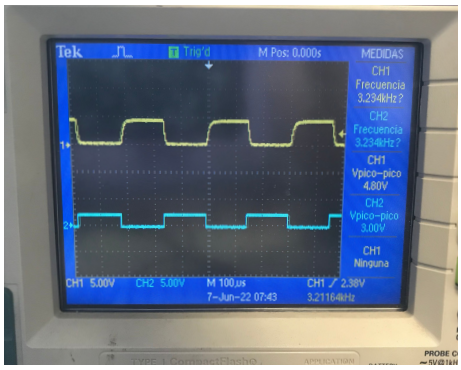


Figure 5.11: Optocoupler circuit electronic diagram

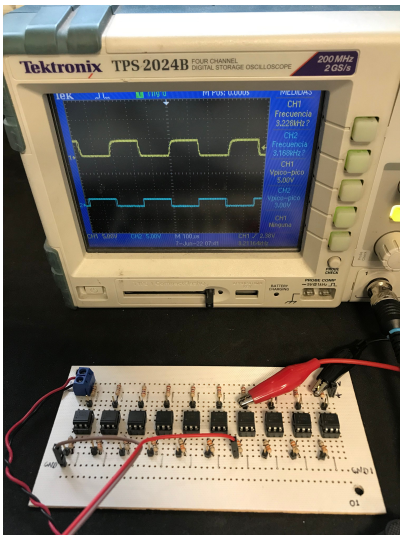
To prove their functioning, a $3kHz$ square signal was transmitted through both circuits (see Figures 5.12c). In Figure 5.12b is shown the input signal (blue) which is $3V$ amplitude and the output (yellow line) is $5V$ and $4.8V$ for the left and right circuits, respectively.



(a) Final layout of the circuit used for the left limb



(b) Response of the left limb optocoupler circuit to a square signal



(c) Proofs of functioning for the left limb optocoupler circuit

Figure 5.12: Final assembly and response of the designed optocoupler circuit

5.4.3 Motor driver

The carrier board for ST®VNH5019 general-purpuse motor driver shown in 5.13 is a fully integrated H-bridge that works from 5.5 to 24 V and can deliver continuous 12 A, working for bidirectional speed control. It works with 2.5 to 5 V logic levels. In combination with embedded protection against reverse-voltage, over-voltage, under-voltage, over-temperature, and over-current. To set up the speed of operation, it must be introduced to the board via an external input signal.

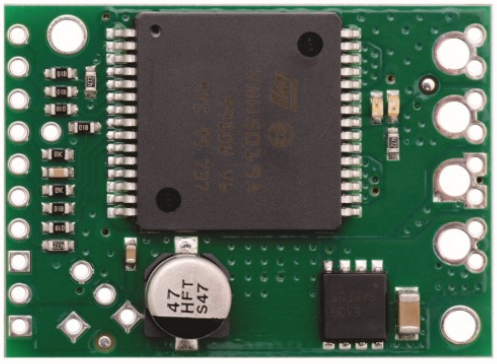


Figure 5.13: VNH5019 Motor Driver Carrier

The main features of the H-bridge are summarized in the Table 5.4

Table 5.4: Main characteristics of VNH5019 Motor driver carrier

Feature	Description
Operating voltage	$5.5 - 24V$
Output current	$12A$
PWM operation	Up to $20kHz$
Protection	
Reverse voltage	$-16V$
Extreme input voltage	$41V$
Undervoltage and overvoltage shutdown	$24V$
High-side and low-side thermal shutdown	
Short-to-ground and short-to-Vcc protection	

The connections are shown in the Figure 5.14, this was the same scheme used for each motor.

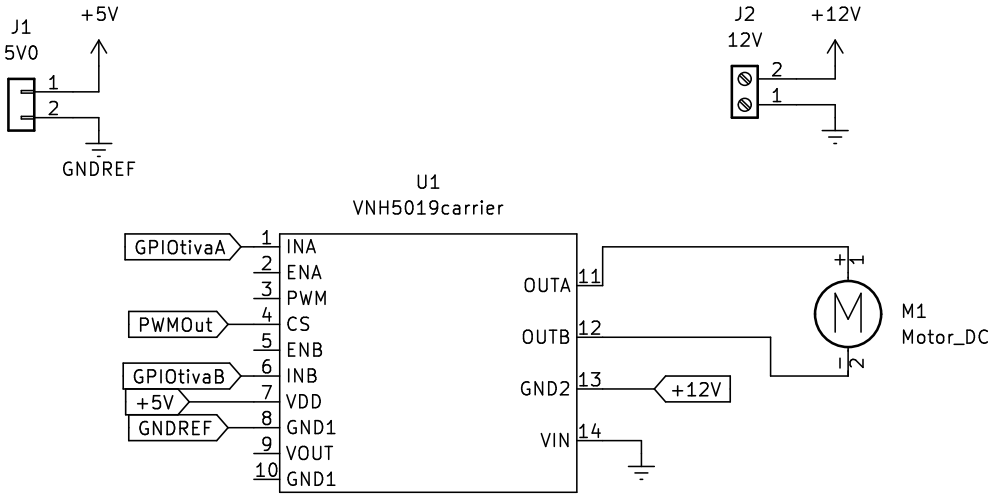


Figure 5.14: VNH5019 Motor Driver carrier connection schematic

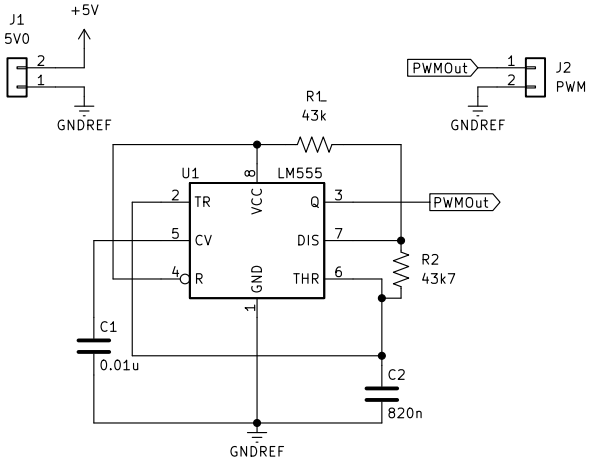
5.4.4 Oscillator

To set the operating frequency of the motors, it was necessary to implement a 555 timer IC in astable configuration. It was chosen as oscillation frequency 19.75 KHz, with duty cycle of 66.5% and a external capacitor of $560pF$. Thus, the resistances were obtained as follows [148]:

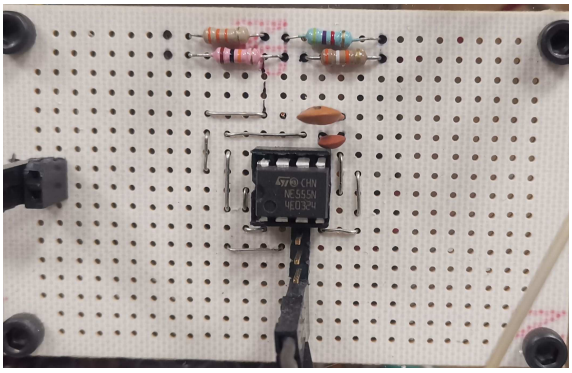
$$R_2 = \frac{17 \times 10^{-6}s}{0.694\,(560 \times 10^{-12}F)} = 43.7k\Omega$$

$$R_1 = \frac{17 \times 33.7^{-6}s}{0.694\,(560 \times 10^{-12}F)} - 43700\Omega = 43k\Omega$$

The final schematic configuration of the oscillator can be seen in 5.15. In 5.15a are shown the connections used to obtain the 19.75 KHz square signal (referred in the figure as *PWMOut*, this signal was used to establish the working frequency in the motor driver (see 5.4.3). There were chosen appropriated resistor values (E12 values), resulting in $R_1 = 39K\Omega + 4.7K\Omega$ and $R_2 = 33K\Omega + 10K\Omega$, which can be seen in 5.15b.



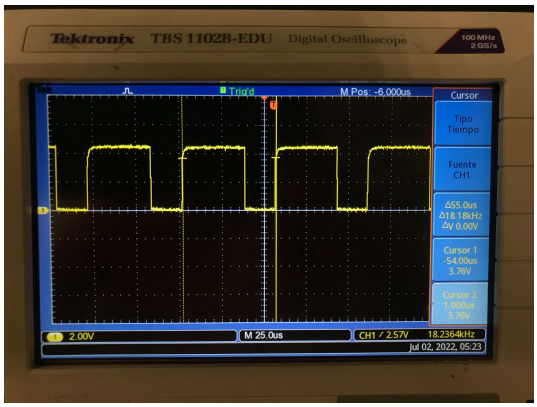
(a) NE555 Oscilator connection schematic



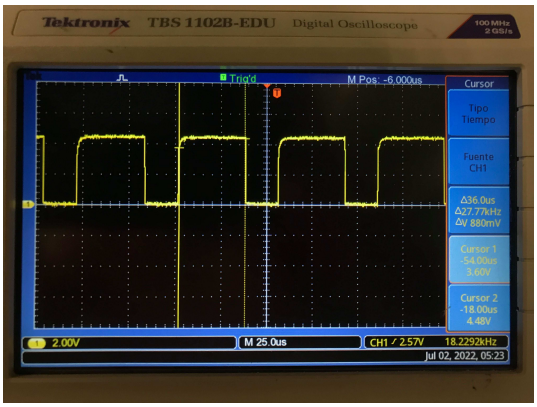
(b) NE555 Oscilator final layout

Figure 5.15: Oscillator final configuration

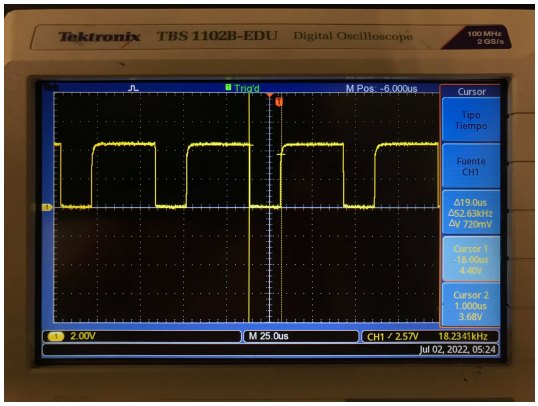
The output signal given by the previous circuit is shown in 5.16. Here we can see that the final output frequency is 18.18 KHz (Figure 5.16a), which means 8.63% error, generated by the variations between the theoretical and real values both resistors and capacitor. In 5.16b is shown that the t_H is $36\mu s$, representing 6.82% of error. For the t_L it is seen in Figure 5.16c that it last $19\mu s$, representing 11% of error.



(a) Final oscillation frequency



(b) t_H final output



(c) t_L final output

Figure 5.16: Output signal obtained from the 555 IC

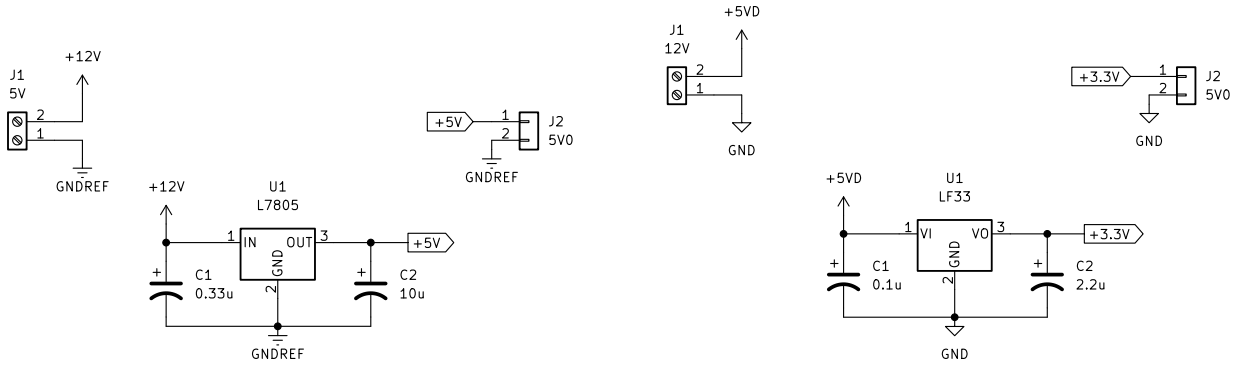
5.4.5 Fixed voltage source

A couple of stable voltage sources were designed to maintain the measurements accurate. Both circuits were assembled accordingly to its datasheet proof of concept circuit, for L7808 IC as shown in Figure 5.17a and LF33 IC displayed in Figure 5.17b . On the first hand, the $5V$ stable voltage was used for the output part in the optical isolation section of the electronics (5.4.2), to power the motor driver carrier 5.4.3 and the oscillator 5.4.4, this IC is energized by a $+12V$ voltage source.

On the other hand, the LF33 IC that delivers 3.3V meets the working voltage of the TIVA (which works with 3.3-V-tolerant in input). This sets the input voltage for the indirect-position sensor, measured by the potentiometer embedded in the LA. The stability in the input voltage guarantees the consistency in the measurement for the tracking trajectories at each linear actuator, to energized it, it was used an independent $5V$ source connected directly to the general line.

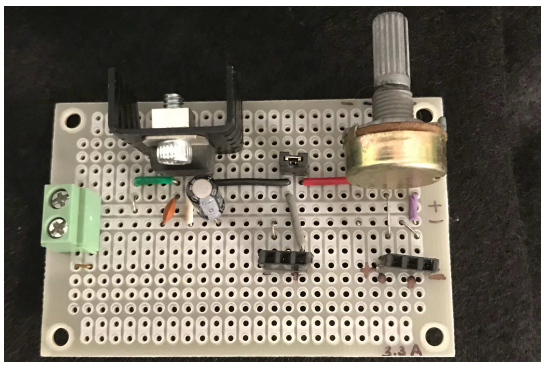
The final outcome of both fixed voltage ICs is displayed in Figure 5.18. The final assembly of both *7805* and LF33 ICs can be seen in Figures 5.18b and 5.18a, respectively. To verify that the

voltages were as needed, the output was verified through oscilloscope signal, as can be seen in 5.18c for the $3.3V$ IC where the delivered output is $3.38V$, representing 2.42% error. For the $5V$ expected outcome, Figure 5.18d shows that the voltage is $5.08V$, meaning an error of 1.6%.

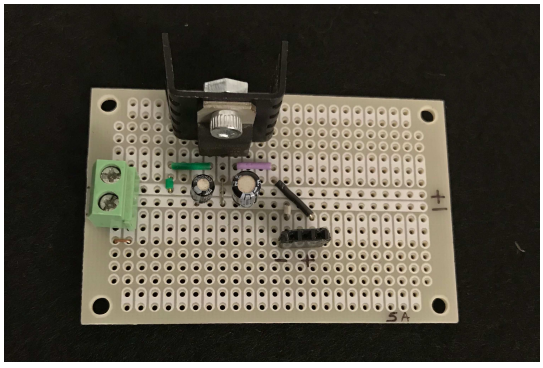


(a) 7805 Electronic diagram (b) LF33 Electronic diagram

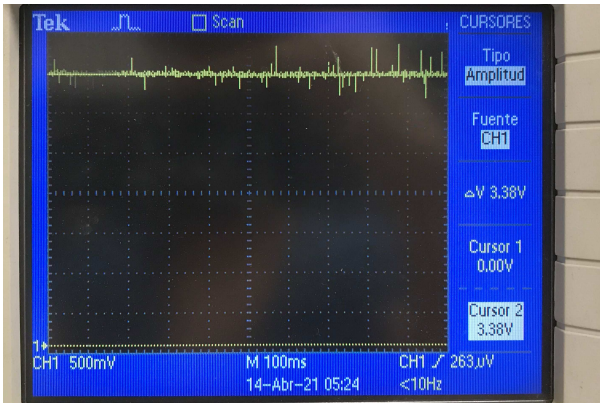
Figure 5.17: Diagram of stable voltage sources



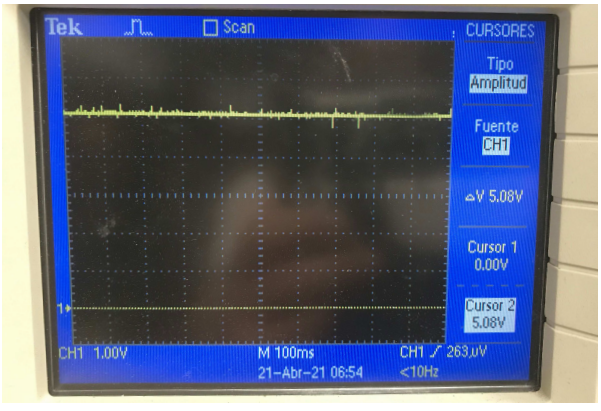
(a) Final arragement of LF33 IC



(b) Final arragement of 7805 IC



(c) LF33 IC output signal



(d) 7805 IC output signal

Figure 5.18: Final results of the fixed voltage sources

5.4.6 Linear actuators

The actuators used to drive all the joints in the BR were Servocity®linear actuators. Its characteristics are shown in the Table 5.5. They were chosen due to its Velocity, stroke and dynamic thrust that can reach the characteristics of a normal Human Gait (HG)

Table 5.5: General characteristics of the linear actuators used in the hip and knee

Actuators					
Actuator	Stroke (cm)	Mass (Kg)	Velocity (cm/s)	Dynamic thrust (Kg)	Current draw (A)
2”	4.9784	1.0829	3.4798	11.3398	3.8
4”	9.9822	1.1396	3.4798	11.3398	3.8
6”	14.986	1.2275	3.4798	11.3398	3.8

All the linear actuators have an embedded $10K\Omega$ linear pot position feedback. Since the pot wiper delivers the position signal to the TIVA, it is necessary to ensure that the voltage does not exceed the $3.3V$ and that it is constant trough the whole rod displacement.

5.5 Software Implementation

Once all the previous sections were evaluated and implemented, the controller was programmed in the TIVA following the scheme depicted in Figure 5.19. Each step is described below:

- **(a)** Here, the GPIOs are set as write or read. there was necessary two for moving backwards and forwards each linear actuator and one for the position sensor.
- **(b)** Each reference trajectory was presented as a vector of *251* positions, each position is in terms of degrees. The initial gains, constants and variable initialization were defined here.
- **(c)** A counter was initialized, letting advance the program along the trajectory vector. If the counter is still in the vector the program continues to **(d)**. Otherwise the counter is reinitialized.
- **(d)** The error variables (measured and estimated) are updated for each position. This allows to implement numerical differentiation and integration over the vector.
- **(e)** A value is read from each LA.
- **(f)** The value read is adjusted to a relationship between the value read, the maximum voltage $(3.3V)$ and the resolution of the ADC (12-bits). This, in turn, is used to obtain the ratio between the voltage and the angular position of the joint.
- **(g)** This point allows getting the current discrepancy between the reference trajectory and the current position.

(h) A subroutine to define the sign of the error is accomplish.

(i) The controller structure proposed in 3.13 was implemented. Here, once defined the magnitude of the control signal, the supertwisting algorithm is implemented following the structure shown in 3.29. Having defined this parameter, each matrix was defined component by component and the terms $\Gamma\left(\xi^E(t)\right)$, $z\left(\delta^E\right)$ and were also mince.

(j) The following step consists in writing through the proper outputs (considering the direction of the rotational displacement) the magnitude of the controller in term of values between zero and 255, leaving a delay of $10\mu s$.

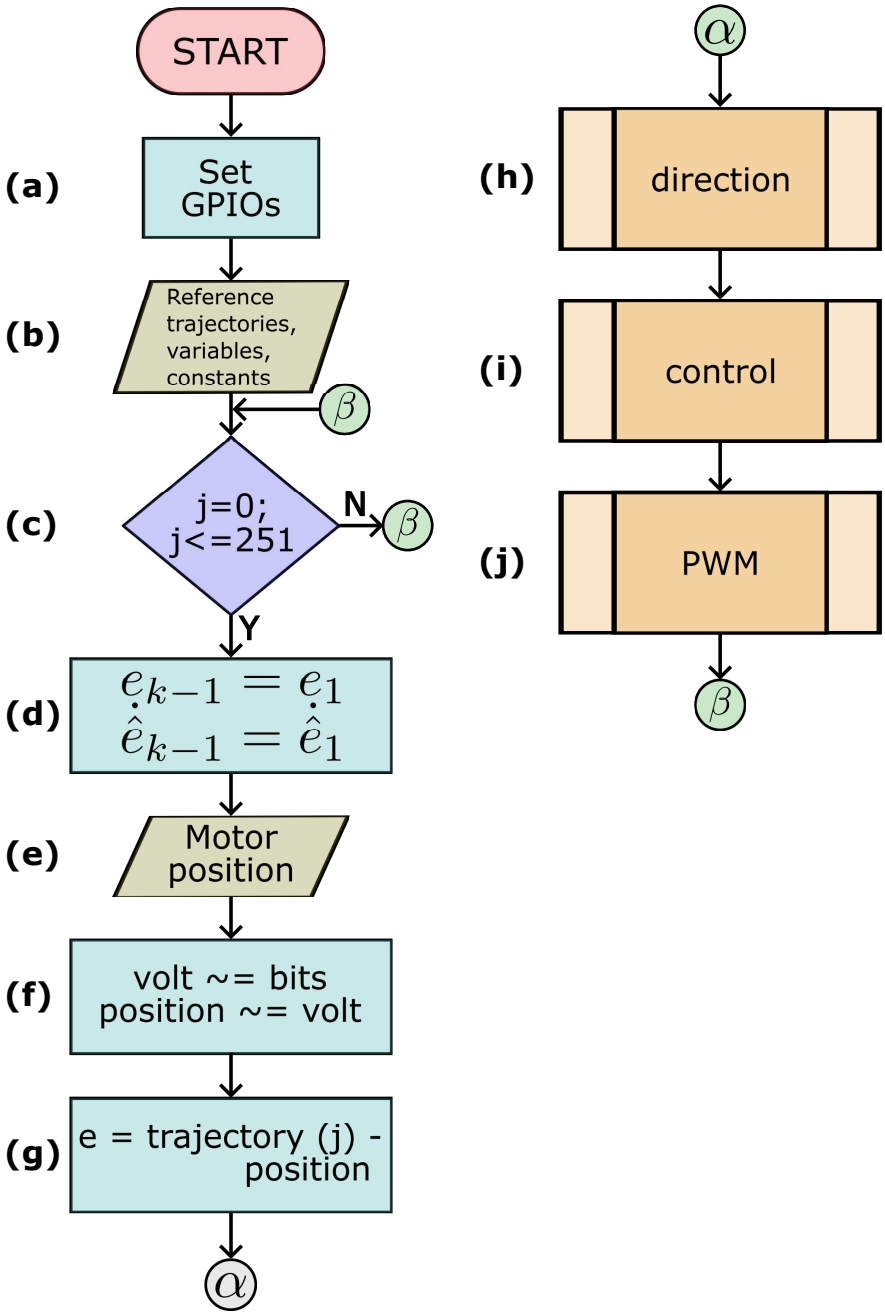


Figure 5.19: Control implementation flowchart

5.6 Experimental evaluations

Figure 5.20 shows the temporal evolution of the proposed controller over a period of 26 seconds. The sequence of photos demonstrates the relative placement of linear actuators with respect to the mechanical structure of the biped robot, which is a contribution of this study. The photos confirm the

smooth movement of both legs for the suspended biped robotic device, demonstrating the effectiveness and potential application of the proposed controller. It is important to note that the slow gait cycle depicted in this figure was intentionally selected to improve the visualization of the controller's effect on the biped robot configuration over time.

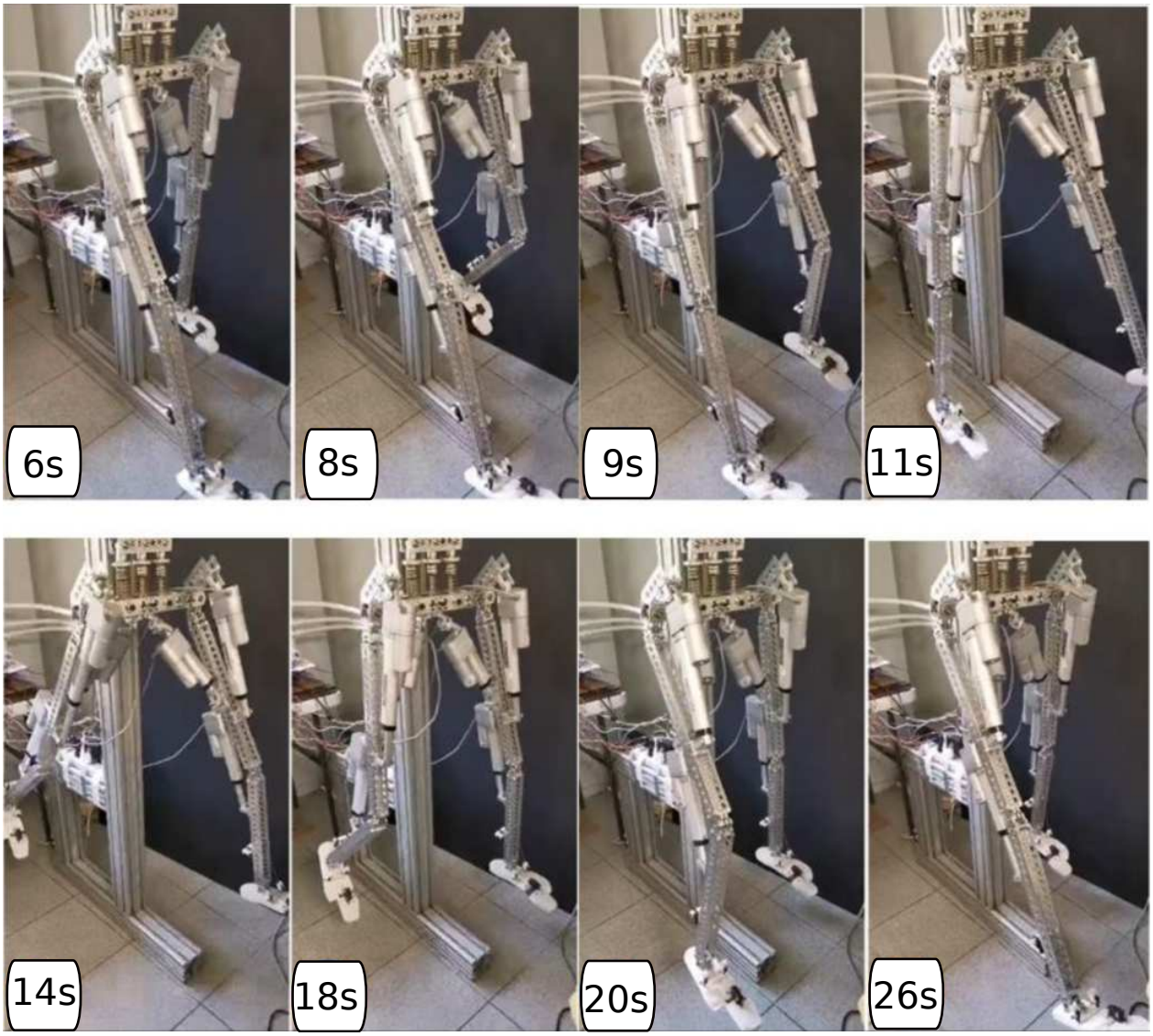


Figure 5.20: Frames of the behavior of the biped structure using the proposed controller

Chapter 6

Conclusions

This thesis presents several contributions in the field of bipedal robotics. Two modeling techniques are proposed to overcome the dynamical differences arising from using linear actuators as mobilizing joint devices, where the joints act as closed-kinematic chains. The dynamical model of the linear actuator is developed, considering its velocity limitations and restricted screw movement range. Moreover, an adaptive controller is proposed for a class of simplified biped robots with state restrictions, which are handled using state-dependent gains in a state feedback controller. The proposed controller is inspired by a traditional state feedback formulation that contains gains that are adjusted using the online evaluation of the tracking error and its derivative. A robust and exact differentiator based on the STA is applied to simplify the implementation of the proposed controller, making a contribution to the automatic control of biped robots. The proposed approach yields a class of output-based controller with the gains depending on the estimated derivative of the tracking error. The stability analysis allows obtaining the time-dependent gain adjusting law, which guarantees the tracking of reference trajectories.

Experimental results confirm that the proposed controller provides better tracking of the reference states than the traditional state feedback control realization, guaranteeing the satisfaction of state constraints. The proposed controller appears to be a significant basis for developing finite time controllers for state-restricted systems. Additionally, a mechanical structure is proposed that allows for the three-dimensional movement necessary for bipedal gait, enabling hip abduction and adduction, and flexion and extension movements to be performed safely. The electronic design enables the coordinated human-based reference trajectory tracking of the complete biped robot, as well as providing protection and sensing of the robot's positions during the execution of the proposed activities. Furthermore, the use of MARL allows the generation and control of suitable trajectories and control signals, respectively, resulting in a decrease of more than 15% in energy consumption

compared to a classical feedback controller.

Future work includes adjusting the gains in the proposed bLf controller through the MARL, which would lead to the inclusion of barriers during the simulated walk. Moreover, the integration of the RL algorithm using appropriate hardware adjustments to the use of simulated MARL-bLfin the biped robotic structure would allow implementing biped walking using the linear actuators-based structure.

Outcomes

Journal Publications

1. **Karla Rincon**, Isaac Chairez, and Wen Yu. “Finite-Time Output Feedback Robust Controller Based on Tangent Barrier Lyapunov Function for Restricted State Space for Biped Robot”. In: *IEEE Trans. Syst. Man Cybern, Syst.* (2021), pp. 1–14
2. **Karla Rincon-Martinez**, Isaac Chairez, and Wen-Yu Liu. “Mathematical Modeling and Robust Control of a Restricted State Suspended Biped Robot Implementing Linear Actuators for Articulation Mobilization”. In: *Applied Sciences* 12.17 (Jan. 2022), p. 8831

Conference Publications

3. **Karla Rincon**, Isaac Chairez, and Wen Yu. “Fixed-time robust output feedback control of a restricted state biped robot based on a tangent barrier Lyapunov function”. In: *2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*. 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Sept. 2019, pp. 1–6
4. **Karla Rincon**, Wen Yu Liu, and Isaac Chairez. “Adaptive control of a biped robot mobilized by linear actuators considering articular restrictions”. In: *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*. vol. 1. May 2022, pp. 457–462

Bibliography

[1] Gabe Nelson, Aaron Saunders, and Robert Playter. “The PETMAN and Atlas Robots at Boston Dynamics”. In: *Humanoid Robotics: A Reference*. Ed. by Ambarish Goswami and Prahlad Vadakkepat. Dordrecht: Springer Netherlands, 2019, pp. 169–186. ISBN: 978-94-007-6046-2. DOI: 10.1007/978-94-007-6046-2_15. URL: https://doi.org/10.1007/978-94-007-6046-2_15.

[2] Ruixiang Zhang and Prahlad Vadakkepat. “Motion Planning of Biped Robot Climbing Stairs”. In: (), p. 6.

[3] Scott Kuindersma et al. “Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot”. In: *Autonomous Robots* 40.3 (Mar. 2016), pp. 429–455. ISSN: 0929-5593, 1573-7527. DOI: 10.1007/s10514-015-9479-3.

[4] Koh Hosoda et al. “Biped robot design powered by antagonistic pneumatic actuators for multi-modal locomotion”. In: *Robotics and Autonomous Systems* 56.1 (Jan. 2008), pp. 46–53. ISSN: 09218890. DOI: 10.1016/j.robot.2007.09.010. (Visited on 07/06/2022).

[5] Satoshi Ito et al. “Design and adaptive balance control of a biped robot with fewer actuators for slope walking”. en. In: *Mechatronics* 49 (Feb. 2018), pp. 56–66. ISSN: 09574158. DOI: 10.1016/j.mechatronics.2017.11.007. (Visited on 07/06/2022).

[6] Baek-Kyu Cho and Jung-Yup Kim. “Dynamic Posture Stabilization of a Biped Robot SUBO-1 on Slope-Changing Grounds”. In: *International Journal of Precision Engineering and Manufacturing* 19.7 (July 2018), pp. 1003–1009. ISSN: 2234-7593, 2005-4602.

[7] “Valkyrie: NASA’s First Bipedal Humanoid Robot”. en. In: *Journal of Field Robotics* 32.3 (2015), pp. 397–419. ISSN: 1556-4967. DOI: 10.1002/rob.21560. (Visited on 07/06/2022).

[8] Michael A. Hopkins et al. “Design of a compliant bipedal walking controller for the DARPA Robotics Challenge”. In: *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. Nov. 2015, pp. 831–837. DOI: 10.1109/HUMANOIDS.2015.7363450.

[9] Matthew Spenko, Stephen Buerger, and Karl Iagnemma, eds. *The DARPA Robotics Challenge Finals: Humanoid Robots To The Rescue*. Vol. 121. Springer Tracts in Advanced Robotics. Springer International Publishing, 2018. ISBN: 978-3-319-74665-4.

[10] Saeed Saeedvand et al. "A comprehensive survey on humanoid robot development". In: *The Knowledge Engineering Review* 34 (2019), e20. DOI: 10.1017/S0269888919000158.

[11] FIRA. *FIRA RoboWorld Cup official website*. en-US. URL: https://firaworldcup.org/ (visited on 07/06/2022).

[12] Jacky Baltes et al. "HuroCup: competition for multi-event humanoid robot athletes". In: *The Knowledge Engineering Review* 32 (2017).

[13] *RoboCup Humanoid League — Official website of the RoboCup Humanoid League*. URL: https://humanoid.robocup.org/ (visited on 07/06/2022).

[14] S.H. Collins and A. Ruina. "A Bipedal Walking Robot with Efficient and Human-Like Gait". In: *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*. 2005 IEEE International Conference on Robotics and Automation. Barcelona, Spain: IEEE, 2005, pp. 1983–1988. ISBN: 978-0-7803-8914-4. DOI: 10.1109/ROBOT.2005.1570404. (Visited on 07/06/2022).

[15] J.E. Pratt and R. Tedrake. "Velocity-Based Stability Margins for Fast Bipedal Walking". In: *Fast Motions in Biomechanics and Robotics: Optimization and Feedback Control*. Ed. by Moritz Diehl and Katja Mombaur. Lecture Notes in Control and Information Sciences. Berlin, Heidelberg: Springer, 2006, pp. 299–324. ISBN: 978-3-540-36119-0.

[16] Anna Jadlovska, Slavka Jadlovska, and Lukas^{v} Koska. "Modeling, Analysis and Control of the Compass Gait Biped Robot and Extensions: A Review". In: *Acta Electrotechnica et Informatica* 21.4 (2021), p. 9.

[17] Aaron D. Ames et al. "Control Barrier Functions: Theory and Applications". In: *2019 18th European Control Conference (ECC)*. 2019 18th European Control Conference (ECC). June 2019, pp. 3420–3431. DOI: 10.23919/ECC.2019.8796030.

[18] Quan Nguyen et al. "3D dynamic walking on stepping stones with control barrier functions". In: *2016 IEEE 55th Conference on Decision and Control (CDC)*. 2016 IEEE 55th Conference on Decision and Control (CDC). Dec. 2016, pp. 827–834.

[19] Shao-Chen Hsu, Xiangru Xu, and Aaron D. Ames. “Control barrier function based quadratic programs with application to bipedal robotic walking”. In: *2015 American Control Conference (ACC)*. 2015 American Control Conference (ACC). July 2015, pp. 4542–4548. DOI: 10.1109/ACC.2015.7172044.

[20] Quan Nguyen et al. “Dynamic Walking on Stepping Stones with Gait Library and Control Barrier Functions”. In: *Algorithmic Foundations of Robotics XII*. Ed. by Ken Goldberg et al. Vol. 13. Series Title: Springer Proceedings in Advanced Robotics. Cham: Springer International Publishing, 2020, pp. 384–399. DOI: 10.1007/978-3-030-43089-4_25.

[21] Ayush Agrawal and Koushil Sreenath. “Discrete Control Barrier Functions for Safety-Critical Control of Discrete Systems with Application to Bipedal Robot Navigation”. In: *Robotics: Science and Systems XIII*. Robotics: Science and Systems 2017. Robotics: Science and Systems Foundation, July 12, 2017. DOI: 10.15607/RSS.2017.XIII.073.

[22] Peter Wieland and Frank Allgöwer. “CONSTRUCTIVE SAFETY USING CONTROL BARRIER FUNCTIONS”. In: *IFAC Proceedings Volumes* 40.12 (2007), pp. 462–467. (Visited on 07/07/2022).

[23] William EH Harcourt-Smith and Leslie C Aiello. “Fossils, feet and the evolution of human bipedal locomotion”. In: *Journal of anatomy* 204.5 (2004), pp. 403–416.

[24] William EH Harcourt-Smith. “The origins of bipedal locomotion”. In: *Handbook of paleoanthropology* 3 (2007), pp. 1483–1518.

[25] Chandra Prakash, Rajesh Kumar, and Namita Mittal. “Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges”. In: *Artificial Intelligence Review* 49.1 (2018), pp. 1–40.

[26] Donald A. Neumann. *Kinesiology of the musculoskeletal system-e-book: foundations for rehabilitation*. Elsevier Health Sciences, 2013.

[27] J. Perry and B. Schoneberger. *Gait Analysis: Normal and Pathological Function*. SLACK, 1992. ISBN: 9781556421921.

[28] Donald A. Neumann. *Kinesiology of the Musculoskeletal System 3rd Edition*. 3rd ed. ELSEVIER, 2018. ISBN: 9780323287531,2016032304.

[29] J. R. Leigh. *Control theory*. In collab. with Institution of Electrical Engineers. 2nd ed. IEE control engineering series 64. OCLC: ocm53460840. London: Institution of Electrical Engineers, 2004. 297 pp. ISBN: 978-0-86341-332-2 978-0-86341-339-1.

[30] Tom M Mitchell and Tom M Mitchell. *Machine learning*. Vol. 1. 9. McGraw-hill New York, 1997.

[31] Muddasar Naeem, Syed Tahir Hussain Rizvi, and Antonio Coronato. “A Gentle Introduction to Reinforcement Learning and its Application in Different Fields”. In: *IEEE Access* 8 (2020), pp. 209320–209344. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2020.3038605. URL: https://ieeexplore.ieee.org/document/9261348/ (visited on 07/10/2022).

[32] Richard S. Sutton and Andrew G Barto. *Reinforcement Learning: An Introduction*. Second. Cambridge, MA: The MIT Press, 2018. ISBN: 9780262039246.

[33] Forest Agostinelli et al. “From Reinforcement Learning to Deep Reinforcement Learning: An Overview”. In: *Braverman Readings in Machine Learning. Key Ideas from Inception to Current State*. Ed. by Lev Rozonoer, Boris Mirkin, and Ilya Muchnik. Vol. 11100. Series Title: Lecture Notes in Computer Science. Springer International Publishing, 2018, pp. 298–328. ISBN: 978-3-319-99491-8. DOI: 10.1007/978-3-319-99492-5_13. URL: http://link.springer.com/10.1007/978-3-319-99492-5_13 (visited on 07/10/2022).

[34] Stephen E. Glickman and Bernard B. Schiff. “A biological theory of reinforcement.” In: *Psychological Review* 74.2 (1967), pp. 81–109. DOI: 10.1037/h0024290. (Visited on 07/08/2022).

[35] Edward Lee Thorndike. *Animal Intelligence*. New York: The Macmillan company, 1911, p. 328.

[36] Frank L. Lewis and Draguna Vrabie. “Adaptive dynamic programming for feedback control”. In: *Proceedings of 2009 7th Asian Control Conference, ASCC 2009* (2009), pp. 1402–1409.

[37] L. P. Kaelbling, M. L. Littman, and A. W. Moore. “Reinforcement Learning: A Survey”. In: *Journal of Artificial Intelligence Research* 4 (May 1, 1996), pp. 237–285. ISSN: 1076-9757. DOI: 10.1613/jair.301.

[38] *What Is Reinforcement Learning? - MATLAB & Simulink*. The MathWorks, Inc. URL: https://www.mathworks.com/help/reinforcement-learning/ug/what-is-reinforcement-%20learning.html (visited on 07/10/2022).

[39] Ismini Lourentzou. “Markov Games and Reinforcement Learning”. In: ().

[40] Frank L. Lewis, Draguna Vrabie, and Kyriakos G. Vamvoudakis. “Reinforcement Learning and Feedback Control: Using Natural Decision Methods to Design Optimal Adaptive Controllers”. In: *IEEE Control Systems* 32.6 (Dec. 2012), pp. 76–105. ISSN: 1066-033X, 1941-000X. DOI: 10.1109/MCS.2012.2214134. URL: https://ieeexplore.ieee.org/document/6315769/ (visited on 07/11/2022).

[41] Shiyu Chen and Yanjie Li. “An Overview of Robust Reinforcement Learning”. In: *2020 IEEE International Conference on Networking, Sensing and Control (ICNSC)*. Oct. 2020, pp. 1–6. DOI: 10.1109/ICNSC48988.2020.9238129.

[42] John Schulman. “Optimizing expectations: From deep reinforcement learning to stochastic computation graphs”. PhD thesis. UC Berkeley, 2016.

[43] *Part 2: Kinds of RL Algorithms — Spinning Up documentation*. 2018. URL: spinningup.openai.com/en/latest/spinningup/rl_intro2.html (visited on 07/04/2022).

[44] Volodymyr Mnih et al. “Asynchronous methods for deep reinforcement learning”. In: *International conference on machine learning*. PMLR. 2016, pp. 1928–1937.

[45] John Schulman et al. “Proximal policy optimization algorithms”. In: *arXiv preprint* (2017).

[46] Volodymyr Mnih et al. “Playing atari with deep reinforcement learning”. In: *arXiv preprint arXiv:1312.5602* (2013).

[47] Marc G Bellemare, Will Dabney, and Rémi Munos. “A distributional perspective on reinforcement learning”. In: *International Conference on Machine Learning*. PMLR. 2017, pp. 449–458.

[48] Timothy P Lillicrap et al. “Continuous control with deep reinforcement learning”. In: *arXiv preprint arXiv:1509.02971* (2015).

[49] Fadi AlMahamid and Katarina Grolinger. “Reinforcement Learning Algorithms: An Overview and Classification”. In: *2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*. IEEE. 2021, pp. 1–7.

[50] Jean-Jacques E Slotine, Weiping Li, et al. *Applied nonlinear control*. Vol. 199. 1. Prentice hall Englewood Cliffs, NJ, 1991.

[51] H.K. Khalil. *Nonlinear Systems*. Upper Saddle River, NJ: Prentice Hall, 2012.

[52] James B Rawlings. “Tutorial overview of model predictive control”. In: *IEEE control systems magazine* 20.3 (2000), pp. 38–52.

[53] Eric C Kerrigan and Jan M Maciejowski. “Designing model predictive controllers with prioritised constraints and objectives”. In: *Proceedings. IEEE International Symposium on Computer Aided Control System Design*. IEEE. 2002, pp. 33–38.

[54] Kim Peter Wabersich and Melanie N Zeilinger. “A predictive safety filter for learning-based control of constrained nonlinear dynamical systems”. In: *Automatica* 129 (2021), p. 109597.

[55] Wei Sun et al. “Adaptive fuzzy tracking control of flexible-joint robots with full-state constraints”. In: *IEEE transactions on systems, man, and cybernetics: systems* 49.11 (2018), pp. 2201–2209.

[56] M Navabi, Hamidreza Mirzaei, and Ali Davoodi. “Optimum design of robust adaptive controller with actuator constraints”. In: *International Journal of Control, Automation and Systems* 18.11 (2020), pp. 2734–2741.

[57] K.P. Tee, S.S. Ge, and E.H. Tay. “Barrier Lyapunov Functions for the control of output-constrained nonlinear systems”. In: *Automatica* 45.5 (2009), pp. 918–917.

[58] Waseda University. *Development of Waseda robot*. URL: http://www.humanoid.waseda.ac.jp/booklet/katobook.html (visited on 06/24/2020).

[59] Joachim von Zitzewitz et al. “Quantifying the Human Likeness of a Humanoid Robot”. In: *International Journal of Social Robotics* 5.2 (2013), pp. 263–276. ISSN: 18754791.

[60] Ann Nowè and Tim Brys. “A gentle introduction to reinforcement learning”. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9858 LNAI (2016), pp. 18–32. ISSN: 16113349.

[61] Duško Katić and Miomir Vukobratović. “Survey of Intelligent Control Techniques for Humanoid Robots”. In: *Journal of Intelligent and Robotic Systems* 37.2 (June 1, 2003), pp. 117–141. DOI: 10.1023/A:1024172417914. (Visited on 07/06/2022).

[62] **Karla Rincon-Martinez**, Isaac Chairez, and Wen-Yu Liu. “Mathematical Modeling and Robust Control of a Restricted State Suspended Biped Robot Implementing Linear Actuators for Articulation Mobilization”. In: *Applied Sciences* 12.17 (Jan. 2022), p. 8831.

[63] F. Negrello et al. “A modular compliant actuator for emerging high performance and fall-resilient humanoids”. In: (2015), pp. 414–420. DOI: 10.1109/HUMANOIDS.2015.7363567.

[64] K. Kaneko et al. “Humanoid robot HRP-2Kai — Improvement of HRP-2 towards disaster response tasks”. In: (2015), pp. 132–139. DOI: 10.1109/HUMANOIDS.2015.7363526.

[65] Kenji Hashimoto. “Mechanics of humanoid robot”. In: *Advanced Robotics* 34.21-22 (Nov. 2020), pp. 1390–1397. ISSN: 0169-1864. DOI: 10.1080/01691864.2020.1813624.

[66] M. Langard et al. “Design and Optimization of a Planar Biped Leg Based on Direct Drive Linear Actuators”. In: *Mathematical Problems in Engineering* (2022).

[67] Pavel I. Philatenkov, Elena V. Morozova, and Tatyana S. Morozova. “Linear Actuator Modeling”. In: *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. Jan. 2020, pp. 807–809.

[68] S. Lohmeier et al. “Leg Design for a Humanoid Walking Robot”. In: (2006), pp. 536–541. DOI: 10.1109/ICHR.2006.321325.

[69] Hun-ok Lim and Atsuo Takanishi. “Biped walking robots created at Waseda University: WL and WABIAN family”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365.1850 (2007), pp. 49–64.

[70] Y. Sugahara et al. “Development of a Biped Locomotor with the Double Stage Linear Actuator”. In: (2007), pp. 1850–1855. DOI: 10.1109/ROBOT.2007.363591.

[71] M. Gienger, K. Loffler, and F. Pfeiffer. “Towards the design of a biped jogging robot”. In: 4 (2001), 4140–4145 vol.4. DOI: 10.1109/ROBOT.2001.933265.

[72] Matthew Haywood and Ferat Sahin. “A novel 3D printed leg design for a biped robot”. In: *2017 12th System of Systems Engineering Conference (SoSE)*. IEEE, June 2017, pp. 1–6. ISBN: 978-1-5090-5945-4. DOI: 10.1109/SYSOSE.2017.7994978.

[73] Zhang Qingxin, Zhang Hongmei, and Wang Fengxiang. “The mathematical model of a novel linear actuator and its control strategy”. In: *2007 Chinese Control Conference*. IEEE. 2007, pp. 165–168.

[74] Hayder F. N. Al-Shuka et al. “Modeling, stability and walking pattern generators of biped robots: a review”. In: *Robotica* 32 (2013), pp. 907–934.

[75] TH Lee et al. “Intelligent control of precision linear actuators”. In: *Engineering Applications of Artificial Intelligence* 13.6 (2000), pp. 671–684.

[76] Jiru Wang et al. “Development and analysis of a dynamic model for parasitic motion principle piezoelectric actuator”. In: *Mechanical Systems and Signal Processing* 147 (2021), p. 107079.

[77] Abdelmalek Zahaf et al. “Robust fault tolerant optimal predictive control of hybrid actuators with time-varying delay for industrial robot arm”. In: *Asian Journal of Control* 24.1 (2022), pp. 1–15.

[78] Arun Manohar Gollapudi, Vasu Velagapudi, and Srikanth Korla. “Modeling and simulation of a high-redundancy direct-driven linear electromechanical actuator for fault-tolerance under various fault conditions”. In: *Engineering Science and Technology, an International Journal* (2020).

[79] Mitio Nagumo. “Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen”. In: 1942.

[80] Walter Murray and Margaret H. Wright. “Projected Lagrangian Methods Based on the Trajectories of Penalty and Barrier Functions.” In: 1978.

[81] Aubrey B. Poore and Qassem M. Al-Hassan. “The expanded Lagrangian system for constrained optimization problems”. In: *Siam Journal on Control and Optimization* 26 (1988), pp. 417–427.

[82] Roman A. Polyak. “Modified barrier functions (theory and methods)”. In: *Mathematical Programming* 54 (1992), pp. 177–222.

[83] Zongyu Zuo and Lin Tie. “Hyperbolic tangent function based adaptive trajectory tracking control for quadrotors”. In: *2013 25th Chinese Control and Decision Conference (CCDC)* (2013), pp. 3721–3726.

[84] Shao-Chen Hsu, Xiangru Xu, and A. Ames. “Control barrier function based quadratic programs with application to bipedal robotic walking”. In: *2015 American Control Conference (ACC)* (2015), pp. 4542–4548.

[85] Aaron D Ames. “Human-inspired control of bipedal walking robots”. In: *IEEE Transactions on Automatic Control* 59.5 (2014), pp. 1115–1130.

[86] Nadhynee Martínez-Fonseca et al. “Robust disturbance rejection control of a biped robotic system using high-order extended state observer”. In: *ISA transactions* 62 (2016), pp. 276–286.

[87] Alexander S. Poznyak. “6 - Dynamic Lagrange Equations”. In: *Classical and Analytical Mechanics*. Ed. by Alexander S. Poznyak. Elsevier, Jan. 1, 2021, pp. 189–219. ISBN: 978-0-323-89816-4. DOI: 10.1016/B978-0-32-389816-4.00017-X. (Visited on 04/03/2023).

[88] Arie Levant. “Higher-order sliding modes, differentiation and output-feedback control”. In: *International journal of Control* 76.9-10 (2003), pp. 924–941.

[89] Jaime A Moreno and Marisol Osorio. “Strict Lyapunov functions for the super-twisting algorithm”. In: *IEEE transactions on automatic control* 57.4 (2012), pp. 1035–1040.

[90] A. Poznyak, A. Polyakov, and V. Azhmyakov. *Attractive Ellipsoids in Robust Control*. Springer, 2014.

[91] **Karla Rincon**, Isaac Chairez, and Wen Yu. “Finite-Time Output Feedback Robust Controller Based on Tangent Barrier Lyapunov Function for Restricted State Space for Biped Robot”. In: *IEEE Trans. Syst. Man Cybern, Syst.* (2021), pp. 1–14.

[92] M. Vukobratovic and D. Juricic. “Contribution to the synthesis of biped gait”. In: *IEEE Transactions on Biomedical Engineering* BME-16.1 (Jan. 1969), pp. 1–6. ISSN: 0018-9294.

[93] Jonathan R Gosyne et al. “Bipedial Locomotion Up Sandy Slopes: Systematic Experiments Using Zero Moment Point Methods”. In: *2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2018, pp. 994–1001.

[94] Matthew J Powell, Ayonga Hereid, and Aaron D Ames. “Speed regulation in 3D robotic walking through motion transitions between human-inspired partial hybrid zero dynamics”. In: *2013 IEEE international conference on robotics and automation*. IEEE. 2013, pp. 4803–4810.

[95] Jessy W Grizzle et al. “3D bipedal robotic walking: models, feedback control, and open problems”. In: *IFAC Proceedings Volumes* 43.14 (2010), pp. 505–532.

[96] Ting Wang and Christine Chevallereau. “A new control law for a 3D biped robot based on regulation of the zero moment point and joint path”. In: *2010 10th IEEE-RAS International Conference on Humanoid Robots* (2010), pp. 27–32.

[97] Ting Wang, Christine Chevallereau, and David Tlalolini. “Stable walking control of a 3D biped robot with foot rotation”. In: *Robotica* 32 (2013), pp. 551–570.

[98] Christine Chevallereau, Jessy W. Grizzle, and Ching-long Shih. “Asymptotically Stable Walking of a Five-Link Underactuated 3-D Bipedal Robot”. In: *IEEE Transactions on Robotics* 25 (2009), pp. 37–50.

[99] Ting Wang, Christine Chevallereau, and Carlos Felipe Rengifo. “Walking and steering control for a 3D biped robot considering ground contact and stability”. In: *Robotics Auton. Syst.* 60 (2012), pp. 962–977.

[100] Dalila Djoudi, Christine Chevallereau, and Jessy W. Grizzle. “A Path-Following Approach to Stable Bipedal Walking and Zero Moment Point Regulation”. In: *Proceedings 2007 IEEE International Conference on Robotics and Automation* (2007), pp. 3597–3602.

[101] Kyu-Cheon Choi, Hyun-Jeong Lee, and Min Cheol Lee. “Fuzzy posture control for biped walking robot based on force sensor for ZMP”. In: *2006 SICE-ICASE International Joint Conference*. IEEE. 2006, pp. 1185–1189.

[102] Pavan K Vempaty et al. “Model Reference Adaptive Control of Biped Robot Actuators for Mimicking Human Gait”. In: *Engineering Letters* 18.2 (2010), p. 165.

[103] Hussein Obeid et al. “Barrier function-based adaptive sliding mode control”. In: *Automatica* 93 (2018), pp. 540–544.

[104] M Sanchez-Magos et al. “Terminal Sliding-Mode Control of Virtual Humanoid Robot with Joint Restrictions Walking on stepping objects”. In: *Cybernetics and Systems* (2020), pp. 1–24.

[105] Shihong Ding, Ju H Park, and Chih-Chiang Chen. “Second-order sliding mode controller design with output constraint”. In: *Automatica* 112 (2020), p. 108704.

[106] Andrey Polyakov and Leonid Fridman. “Stability notions and Lyapunov functions for sliding mode control systems”. In: *Journal of the Franklin Institute* 351.4 (2014), pp. 1831–1865.

[107] Arie Levant. “On fixed and finite time stability in sliding mode control”. In: *52nd IEEE Conference on Decision and Control*. IEEE. 2013, pp. 4260–4265.

[108] Emmanuel Moulay and Wilfrid Perruquetti. “Finite time stability of differential inclusions”. In: *IMA Journal of Mathematical Control and Information* 22.4 (2005), pp. 465–475.

[109] V.I. Utkin. *Sliding Modes in Control and Optimization*. Springer-Verlag, 1992.

[110] Francisco Bejarano and Leonid Fridman. “High order sliding mode observer for linear systems with unbounded unknown inputs”. In: *International Journal of Control* 83.9 (2010), pp. 1920–1929.

[111] Yuri Shtessel et al. *Sliding Mode Control and Observation*. Springer New York, 2014.

[112] Yuri B Shtessel, Leonid Fridman, and Alan Zinober. “Higher order sliding modes”. In: *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal* 18.4-5 (2008), pp. 381–384.

[113] Vadim Utkin. “Discussion aspects of high-order sliding mode control”. In: *IEEE Transactions on Automatic Control* 61.3 (2015), pp. 829–833.

[114] Wen-Jun Cao and Jian-Xin Xu. “Nonlinear integral-type sliding surface for both matched and unmatched uncertain systems”. In: *IEEE Transactions on Automatic Control* 49.8 (2004), pp. 1355–1360.

[115] Bijnan Bandyopadhyay and Deepak Fulwani. “High-performance tracking controller for discrete plant using nonlinear sliding surface”. In: *IEEE Transactions on Industrial Electronics* 56.9 (2009), pp. 3628–3637.

[116] Zongyu Zuo. “Non-singular fixed-time terminal sliding mode control of non-linear systems”. In: *IET control theory & applications* 9.4 (2014), pp. 545–552.

[117] Andrey Polyakov, Dennis Efimov, and Wilfrid Perruquetti. “Finite-time and fixed-time stabilization: Implicit Lyapunov function approach”. In: *Automatica* 51 (2015), pp. 332–340.

[118] Xuan Xiao and Fumihiko Asano. “Analysis of steady and target walking speeds in limit cycle walking”. In: *International Journal of Dynamics and Control* 5.3 (2017), pp. 454–465.

[119] M Mera and I Salgado. “Robust control of linear systems under input saturation using Barrier Lyapunov functions”. In: *International Journal of Dynamics and Control* 6.3 (2018), pp. 1231–1238.

[120] Keng Peng Tee and Shuzhi Sam Ge. “Control of state-constrained nonlinear systems using integral barrier Lyapunov functionals”. In: *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE. 2012, pp. 3239–3244.

[121] Junxing Zhang et al. “Novel prescribed performance-tangent barrier Lyapunov function for neural adaptive control of the chaotic PMSM system by backstepping”. In: *International Journal of Electrical Power & Energy Systems* 121 (2020), p. 105991.

[122] Yongming Li, Kewen Li, and Shaocheng Tong. “Finite-time adaptive fuzzy output feedback dynamic surface control for MIMO nonstrict feedback systems”. In: *IEEE Transactions on Fuzzy Systems* 27.1 (2018), pp. 96–110.

[123] Yongming Li, Tingting Yang, and Shaocheng Tong. “Adaptive neural networks finite-time optimal control for a class of nonlinear systems”. In: *IEEE Transactions on Neural Networks and Learning Systems* (2019).

[124] Shaocheng Tong, Xiao Min, and Yuanxin Li. “Observer-based adaptive fuzzy tracking control for strict-feedback nonlinear systems with unknown control gain functions”. In: *IEEE Transactions on Cybernetics* (2020).

[125] Jianyong Yao and Wenxiang Deng. “Active disturbance rejection adaptive control of uncertain nonlinear systems: theory and application”. In: *Nonlinear Dynamics* 89 (Aug. 2017).

[126] Huann-Keng Chiang, Jason Sh Tsai, and York-Yih Sun. “Extended Ackermann formula for multivariable control systems”. In: *International Journal of Systems Science* 21.11 (1990), pp. 2113–2127.

[127] Andrey Polyakov. “Nonlinear feedback design for fixed-time stabilization of linear control systems”. In: *IEEE Transactions on Automatic Control* 57.8 (2011), pp. 2106–2110.

[128] Marco Tulio Angulo, Jaime A Moreno, and Leonid Fridman. “Robust exact uniformly convergent arbitrary order differentiator”. In: *Automatica* 49.8 (2013), pp. 2489–2495.

[129] Ivan Salgado et al. “Super-twisting sliding mode differentiation for improving PD controllers performance of second order systems”. In: *ISA transactions* 53.4 (2014), pp. 1096–1106.

[130] Carlos Aguilar-Ibañez et al. “Output-feedback stabilization of the PVTOL aircraft system based on an exact differentiator”. In: *Journal of Intelligent & Robotic Systems* 90.3-4 (2018), pp. 443–454.

[131] Michael J. Coleman, Anindya Chatterjee, and Andy Ruina. “Motions of a Rimless Spoked Wheel: a Simple 3D System with Impacts”. In: *Dynamics and Stability of Systems* 12 (1997), pp. 139–159.

[132] Tony Owen. “Scientific Fundamentals of Robotics 7-Biped Locomotion by M. Vukobratovic B. Borovac, D. Surla and D. Stokic Springer-Verlag, Berlin 1990, 349 pages with index (DM 158.00; c. £57.00)”. In: *Robotica* 8 (1990), pp. 325–325.

[133] Christopher J. C. H. Watkins and Peter Dayan. “Q-learning”. In: *Machine Learning* 8.3 (May 1992), pp. 279–292. ISSN: 1573-0565. DOI: 10.1007/BF00992698. URL: https://doi.org/10.1007/BF00992698 (visited on 07/14/2022).

[134] Thomas Degris, Martha White, and Richard S. Sutton. *Off-Policy Actor-Critic*. June 20, 2013. DOI: 10.48550/arXiv.1205.4839. arXiv: 1205.4839[cs]. URL: http://arxiv.org/abs/1205.4839 (visited on 10/28/2022).

[135] Qi Cai et al. *Neural Temporal-Difference and Q-Learning Provably Converge to Global Optima*. Apr. 15, 2020. DOI: 10.48550/arXiv.1905.10027. arXiv: 1905.10027. URL: http://arxiv.org/abs/1905.10027 (visited on 10/28/2022).

[136] Nurfarahin Onn et al. “Human Gait Modelling Considerations of Biped Locomotion for Lower Limb Exoskeleton Designs”. In: *13th International Conference on Robotics, Control and Manufacturing Technology (ROCOM’13)* (2013), pp. 59–64.

[137] D.A. Neumann. *Fundamentos de rehabilitación física*. Primera. Medicina. Badalona, España: Paidotribo, 2007. ISBN: 9788480198134.

[138] N A Borghese, L Bianchi, and F Lacquaniti. “Kinematic determinants of human locomotion.” In: *The Journal of Physiology* 494.3 (Aug. 1996), pp. 863–879. ISSN: 00223751.

[139] Seungsuk Ha, Youngjoon Han, and Hernsoo Hahn. “Adaptive gait pattern generation of biped robot based on human’s gait pattern analysis”. In: *International Journal of Electrical and Computer Engineering* 1.10 (2007), pp. 1460–1465.

[140] Francesca Negrello et al. “A modular compliant actuator for emerging high performance and fall-resilient humanoids”. In: *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE. 2015, pp. 414–420.

[141] Gill A Pratt and Matthew M Williamson. “Series elastic actuators”. In: *Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*. Vol. 1. IEEE. 1995, pp. 399–406.

[142] Hun-ok Lim and Atsuo Takanishi. “Biped walking robots created at Waseda University: WL and WABIAN family”. In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 365.1850 (2007), pp. 49–64.

[143] Yusuke Sugahara et al. “Realization of dynamic human-carrying walking by a biped locomotor”. In: *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA’04. 2004*. Vol. 3. IEEE. 2004, pp. 3055–3060.

[144] Sebastian Lohmeier et al. “Leg design for a humanoid walking robot”. In: *2006 6th IEEE-RAS International Conference on Humanoid Robots*. IEEE. 2006, pp. 536–541.

[145] Yusuke Sugahara et al. “Walking up and down stairs carrying a human by a biped locomotor with parallel mechanism”. In: *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2005, pp. 1489–1494.

[146] Junghoon Choo and Jong Hyeon Park. “Increasing payload capacity of wearable robots using linear actuators”. In: *IEEE/ASME Transactions on Mechatronics* 22.4 (2017), pp. 1663–1673.

[147] *General Purpose Type Photocoupler*. LTV-4N25 Series. Lite-On Corp.

[148] Thomas L. Floyd. *Electronic devices*. Conventional current version, tenth global. Pearson, 2018. ISBN: 9781292222998.

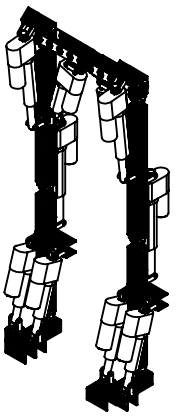
[149] **Karla Rincon**, Isaac Chairez, and Wen Yu. “Fixed-time robust output feedback control of a restricted state biped robot based on a tangent barrier Lyapunov function”. In: *2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*. 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Sept. 2019, pp. 1–6.

[150] **Karla Rincon**, Wen Yu Liu, and Isaac Chairez. “Adaptive control of a biped robot mobilized by linear actuators considering articular restrictions”. In: *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*. Vol. 1. May 2022, pp. 457–462.

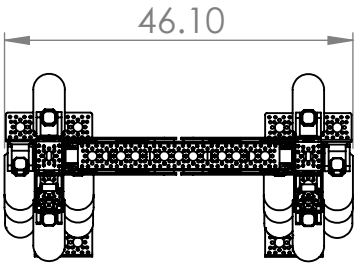
Appendix A

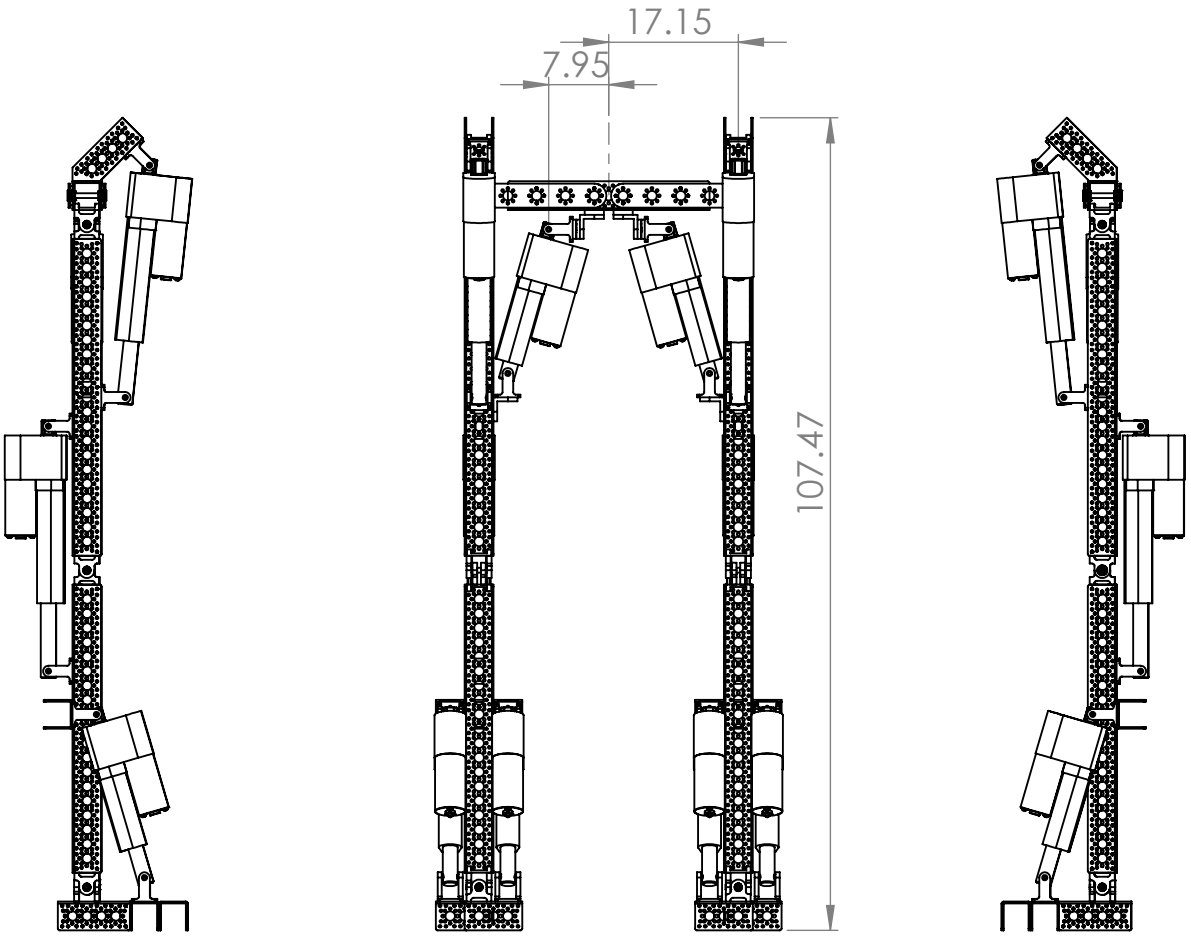
Mechanical drawings of the proposed structure

The robotic device consists of three main parts: the hip and the left and right limbs, which in turn have their own segments. The following images show the dimensions and hinges associated with each main section.

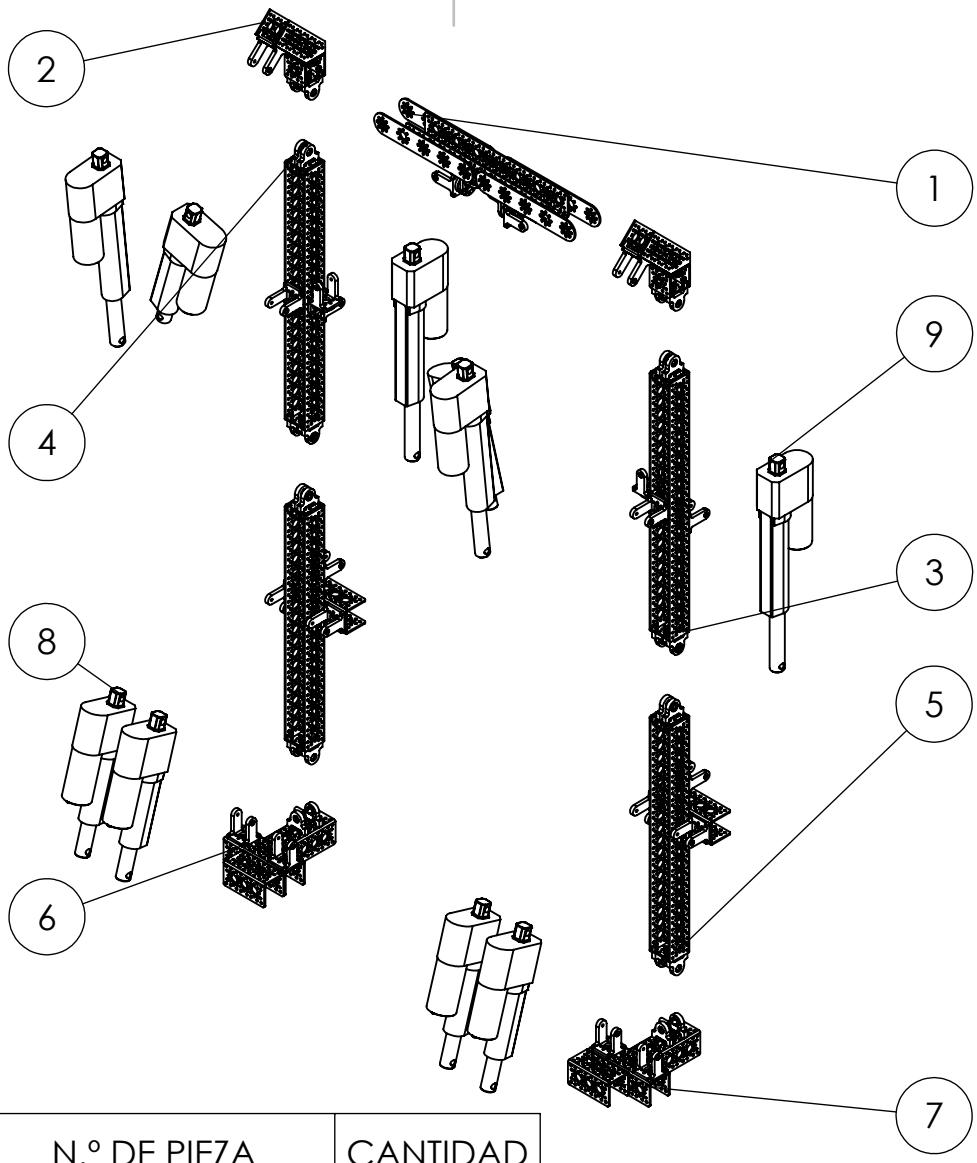


ESCALA 1:20





CINVESTAV	DCA
Biped	
TÍTULO: Complete robot assembly	
N.º DE DIBUJO FullRobot	A4
	KRM
ESCALA:1:10	HOJA 1 DE 2



N.º DE ELEMENTO	N.º DE PIEZA	CANTIDAD
1	Hip	1
2	MobHip	2
3	legLeft	1
4	legRight	1
5	calf	2
6	rightFoot	1
7	leftFoot	1
8	HDA2	6
9	HDA4	4

CINVESTAV | DCA

Biped

TÍTULO:

Robot exploded and components

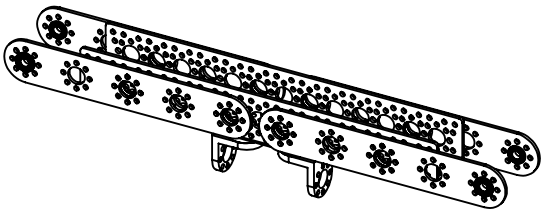
N.º DE DIBUJO

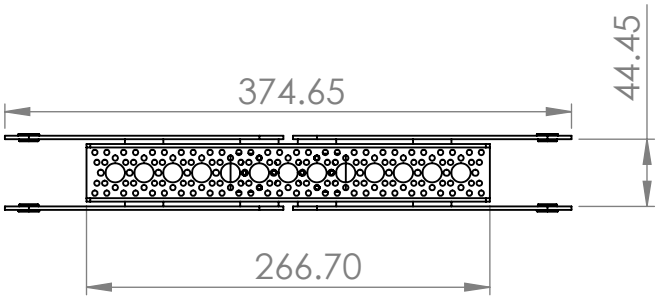
FullRobot

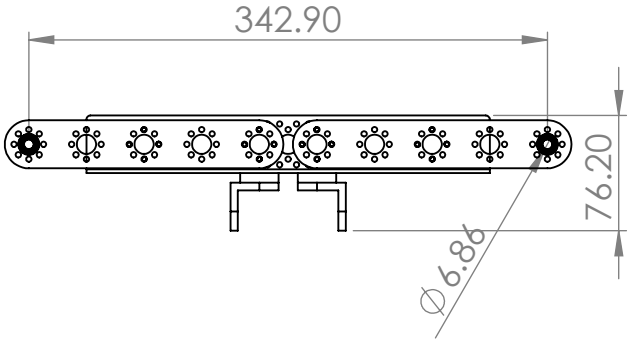
A4

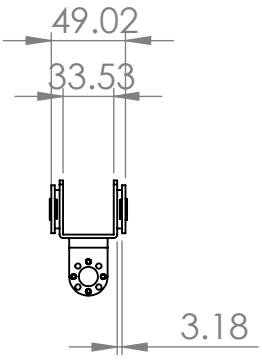
KRM

ESCALA:1:10 | HOJA 2 DE 2

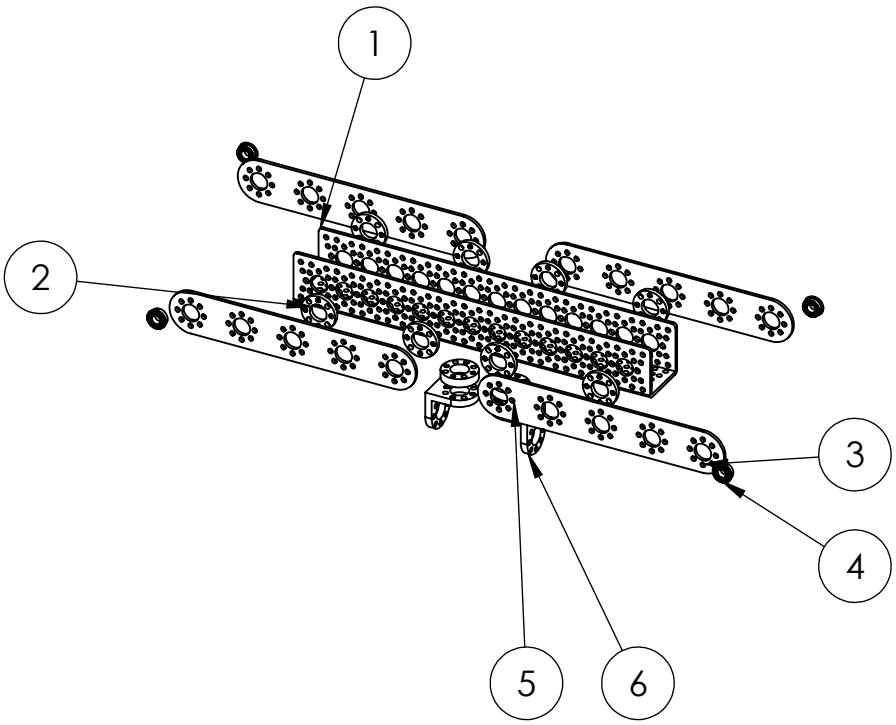






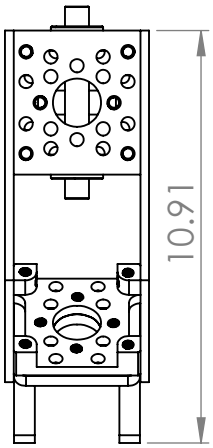


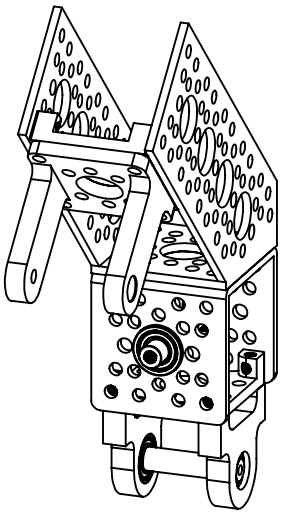
CINVESTAV	DCA
Biped	
TÍTULO: Hip assembly	
N.º DE DIBUJO HipFixed	A4
	KRM
ESCALA:1:5	HOJA 1 DE 2

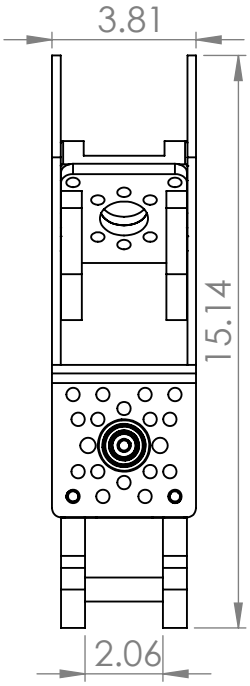


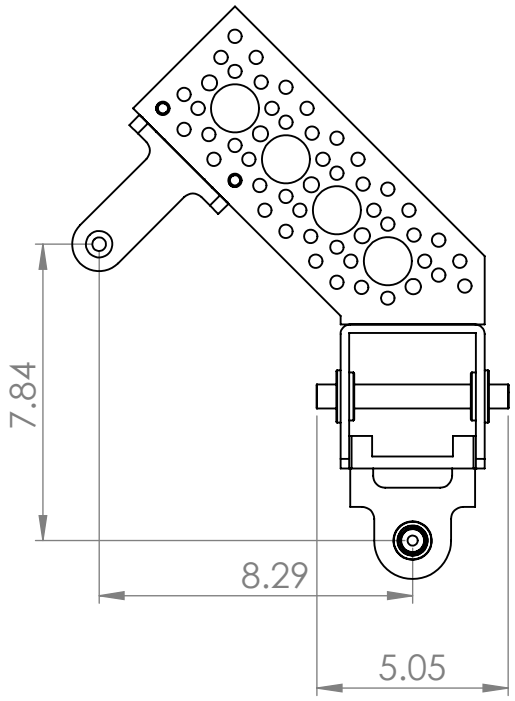
N.º DE ELEMENTO	N.º DE PIEZA	CANTIDAD
1	Canal 10_5''	1
2	espaciador_125	8
3	bracketF	4
4	rodamiento .25ID.5OD	4
5	espaciador_25	2
6	montaje 90o	2

CINVESTAV	DCA
Biped	
TÍTULO: Hip exploded and components	
N.º DE DIBUJO HipFixed	A4
	KRM
ESCALA:1:5	HOJA2 DE 2

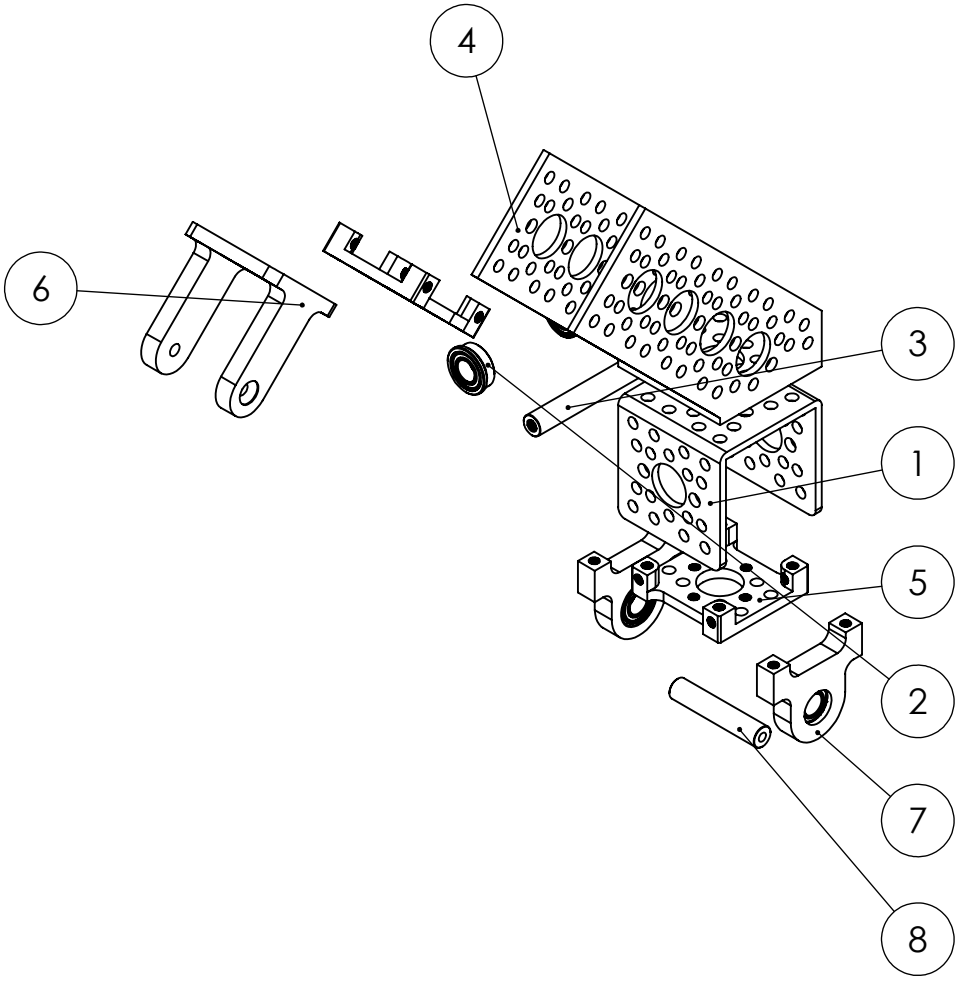






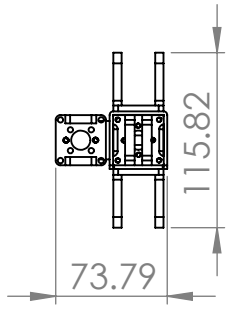


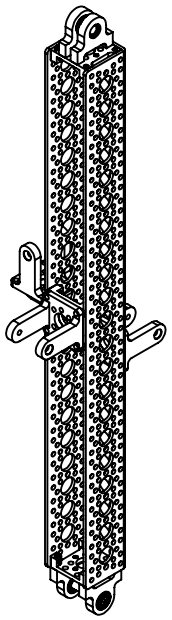
CINVESTAV	DCA
Biped	
TÍTULO: Mobile hip assembly	
N.º DE DIBUJO MobHip	A4
	KRM
ESCALA:1:2	HOJA 1 DE 2

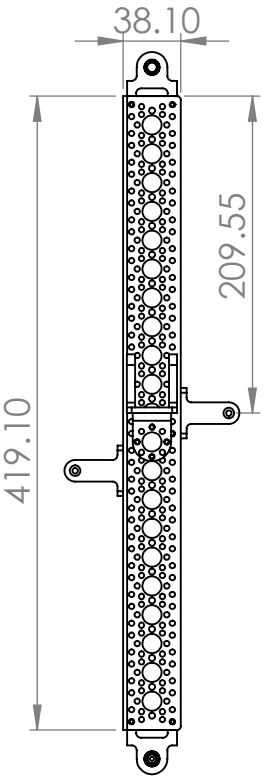


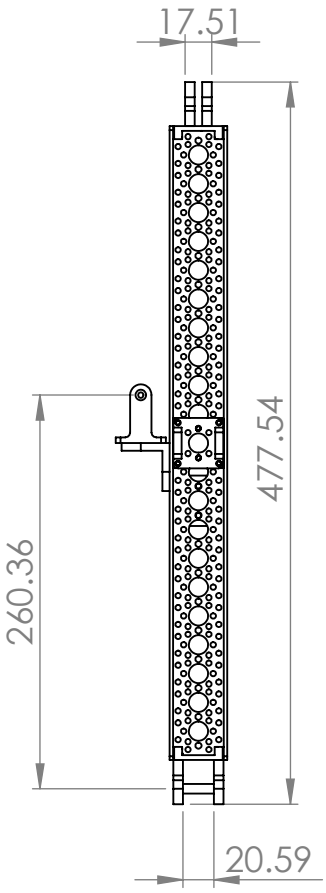
N.º DE ELEMENTO	N.º DE PIEZA	CANTIDAD
1	Canal 1_5"	1
2	rodamiento .25ID.5OD	4
3	standoff 2in	1
4	pattern bracket	1
5	Base	2
6	BaseActuador	1
7	TappedPillow	2
8	standoff 125in	1

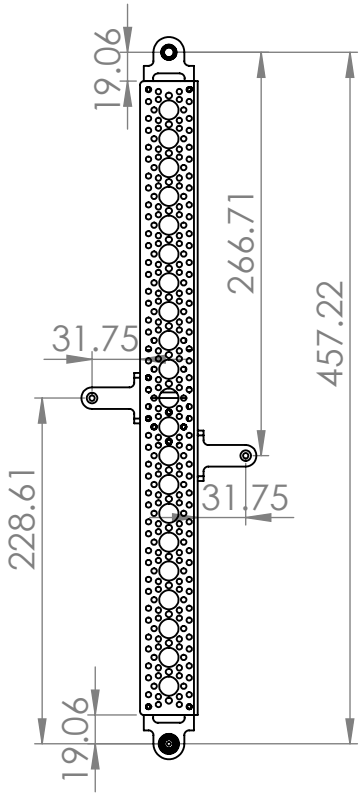
CINVESTAV	DCA
Biped	
TÍTULO: Mob hip exploded and components	
N.º DE DIBUJO MobHip	A4
	KRM
ESCALA:1:2	HOJA 2 DE 2



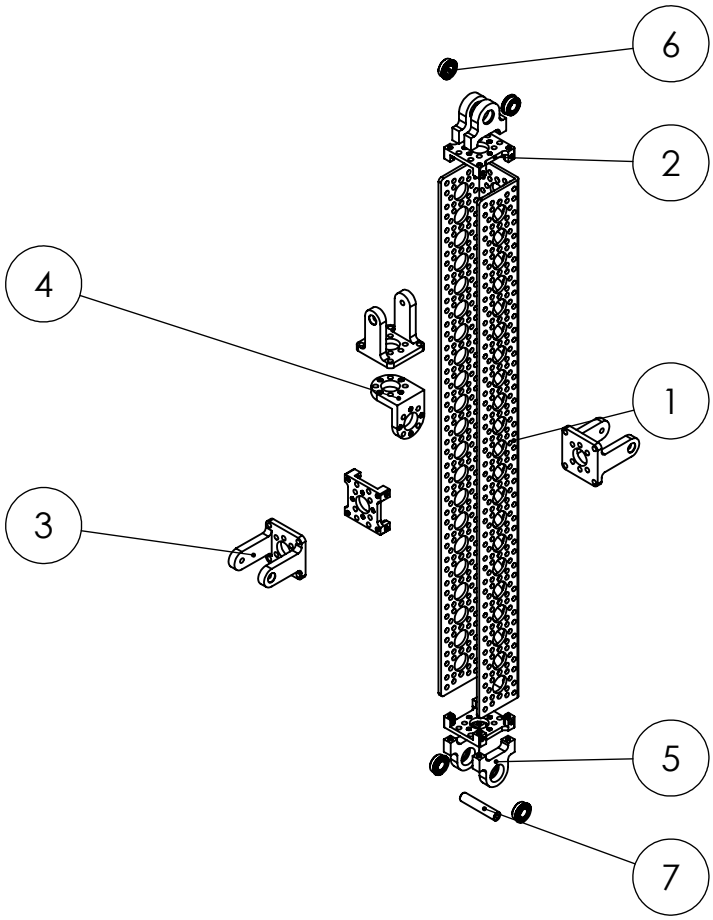






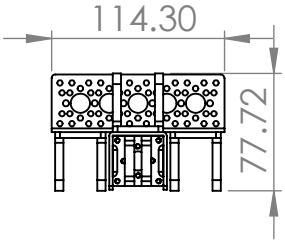


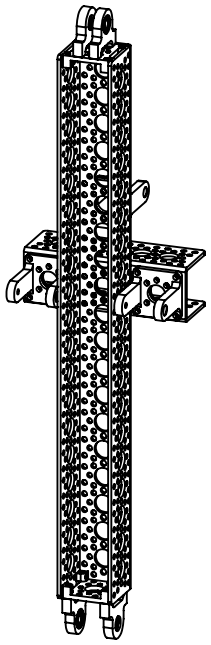
CINVESTAV	DCA
Biped	
TÍTULO: Left leg assembly	
N.º DE DIBUJO LeftLeg	A4
	KRM
ESCALA:1:5	HOJA 1 DE 2

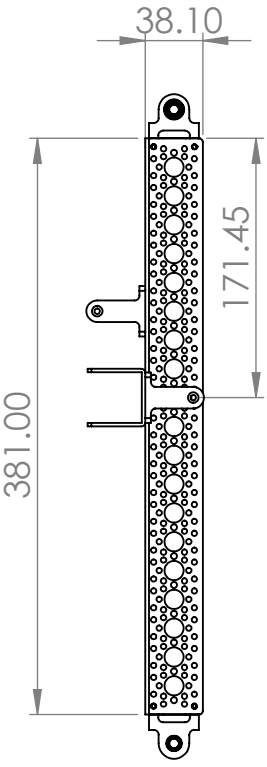


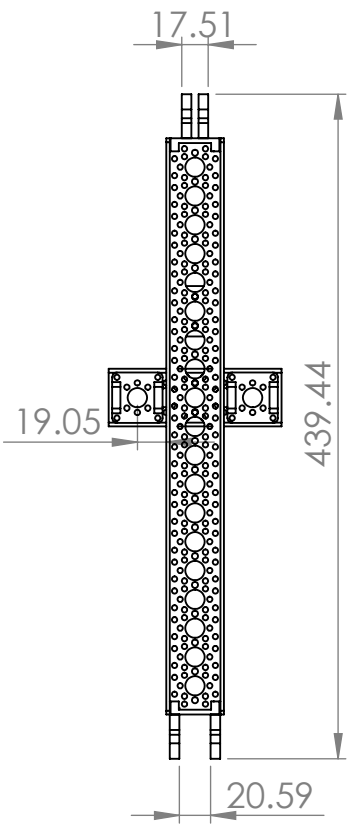
N.º DE ELEMENTO	N.º DE PIEZA	CANTIDAD
1	Canal 16_5"	1
2	BaseDoctoradoVersion	3
3	BaseActuador	3
4	montaje 90o	1
5	TappedPillow	4
6	rodamiento .25ID.5OD	4
7	standoff 125in	1

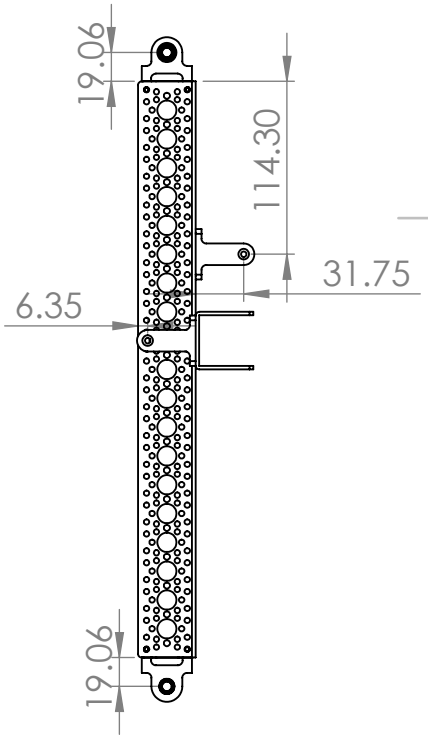
CINVESTAV	DCA
Biped	
TÍTULO: Left leg exploded and components	
N.º DE DIBUJO LeftLeg	A4
	KRM
ESCALA:1:5	HOJA 2 DE 2



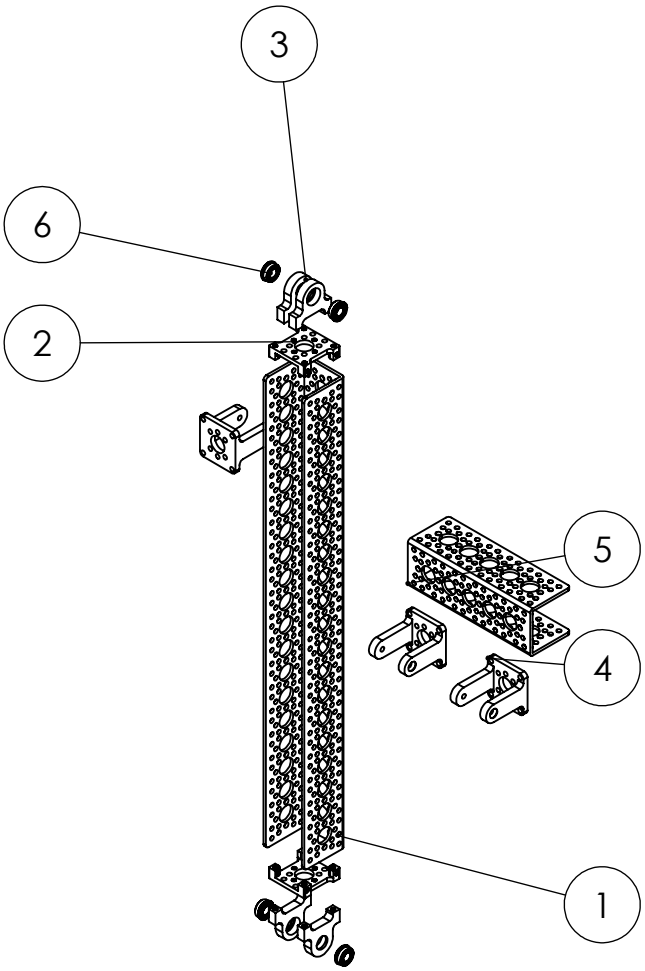






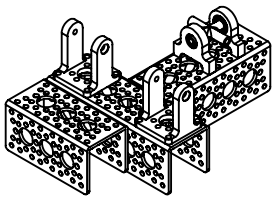


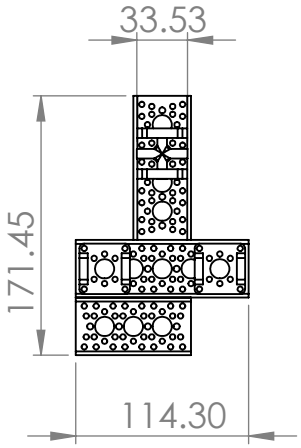
CINVESTAV	DCA
Biped	
TÍTULO: Calf assembly	
N.º DE DIBUJO Calf	A4
	KRM
ESCALA:1:5	HOJA 1 DE 2

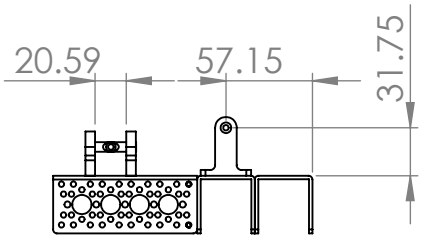


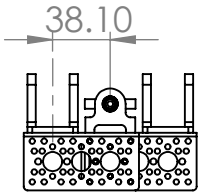
N.º DE ELEMENTO	N.º DE PIEZA	CANTIDAD
1	Canal 15''	1
2	Base	2
3	TappedPillow	4
4	BaseActuador	3
5	Canal 4_5''	1
6	rodamiento .25ID.5OD	4

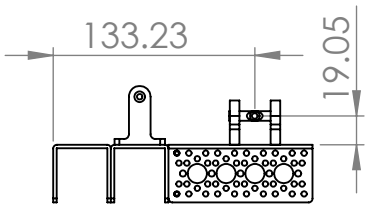
CINVESTAV	DCA
Biped	
TÍTULO: Calf exploded and components	
N.º DE DIBUJO Calf	A4
	KRM
ESCALA:1:5	HOJA 1 DE 1



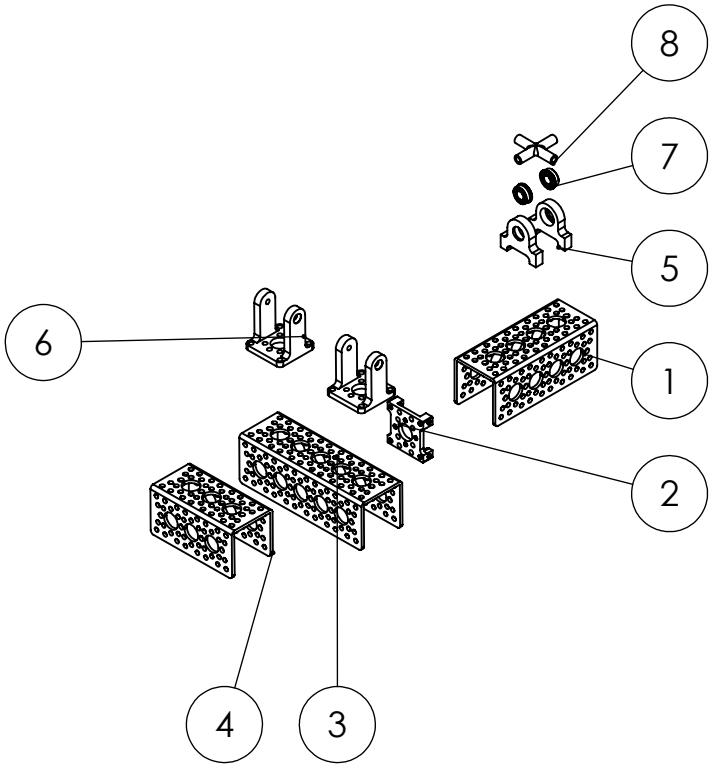








CINVESTAV	DCA
Biped	
TÍTULO: Left foot assembly	
N.º DE DIBUJO leftFoot	A4
	KRM
ESCALA:1:5	HOJA 1 DE 2



N.º DE ELEMENTO	N.º DE PIEZA	CANTIDAD
1	Canal 3_75"	1
2	Base	1
3	Canal 4_5"	1
4	Canal 3"	1
5	TappedPillow	2
6	BaseActuador	2
7	rodamiento .25ID.5OD	2
8	cruceta	1

CINVESTAV	DCA
Biped	
TÍTULO: Left foot exploded and components	
N.º DE DIBUJO leftFoot	A4
	KRM
ESCALA:1:5	HOJA 2 DE 2