# Έλεγχος κίνησης δίποδου ρομπότ με ενισχυτική μάθηση

Νταίζη Χρόνη

# Έλεγχος κίνησης δίποδου ρομπότ με ενισχυτική μάθηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Νταίζη Χρόνη

Α.Μ: 2001010042, Email: daisyxroni@hotmail.com

Εξεταστική Επιτροπή:

Ν. Βλάσσης, Επίκουρος Καθηγητής, Τμήμα Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης (Επιβλέπων)

Η. Κοσματόπουλος, Επίκουρος Καθηγητής, Τμήμα Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης

Μ. Λαγουδάκης, Επίκουρος Καθηγητής, Τμήμα Ηλεκτρονικών Μηχανικών και Μηχανικών Υπολογιστών, Πολυτεχνείο Κρήτης

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΑΡΑΓΩΓΗΣ ΚΑΙ ΔΙΟΙΚΗΣΗΣ

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

© 2008 Νταίζη Χρόνη.

Εικόνα εξωφύλλου: το ρομπότ Nao.

# Έλεγχος κίνησης δίποδου ρομπότ με ενισχυτική μάθηση


**Περίληψη**

Στην παρούσα διατριβή πραγματευόμαστε τον έλεγχο κίνησης δίποδου ρομπότ, χρησιμοποιώντας διάφορους αλγορίθμους ενισχυτικής μάθησης. Το ρομπότ που χρησιμοποιήσαμε είναι το μοντέλο NAO της Aldebaran robotics τα τεχνικά χαρακτηριστικά του οποίου βρίσκονται στο παράρτημα. Το όλο εγχείρημα έγινε χρησιμοποιώντας τον προσομοιωτή webots.

Η κίνηση που επιλέχθηκε είναι μια κλωτσιά. Επιλέχθηκε η συγκεκριμένη κίνηση, ούτως ώστε να συμπεριλάβουμε την στατική ισορροπία του ρομπότ και βασικές έννοιες, όπως το σημείο μηδενικής ροπής και το κέντρο μάζας. Η έννοια του σημείου μηδενικής ροπής, του κέντρου μάζας και του αντίστροφου εκκρεμούς, ειναι βασικές έννοιες για την ευθεία και αντίστροφη κινηματική ανάλυση ενός ρομπότ ούτως ώστε να καταφέρουμε να το ελέγξουμε.

Η ίδια η κίνηση παράχθηκε τελικά, χρησιμοποιώντας αλγορίθμους ενισχυτικής μάθησης. Οι αλγόριθμοι ενισχυτικής μάθησης (Reinforcement learning ( RL )) είναι μια σύγχρονη μέθοδος που αναπτύχθηκε τα τελευταία χρόνια και χρησιμοποιείται ευρέως στην ρομποτική, μιας και η κλασσικές μέθοδοι δυναμικής και στατικής ανάλυσης σε ρομπότ με πολλούς βαθμούς ελευθερίας είναι πολύ δύσκολή έως αδύνατη.

Τελικά καταφέραμε να παρουσιάσουμε μια κλωτσιά προσεγγίζοντας τις βέλτιστες γωνίες και τις επιταγχύνσεις τους ούτως ώστε NAO να μην χάνει την ισορροπία του.

## Ευχαριστίες

Αφιερώνεται στους γονείς μου που όλα τα χρόνια της ζωής μου, πόσο μάλλον της φοιτητικής μου ζωής, με στήριξαν οικονομικά, ηθικά και πάντα βρίσκονταν δίπλα μου, στις χαρές και τις λύπες, και στην αδερφή μου...

Ευχαριστώ τον επιβλέποντα μου, κ. Βλάσση που πρώτα από όλα με ενέπνευσε σαν επιστήμονας και ασχολήθηκα με αυτό το αντικείμενο που πάντα μου προκαλούσε το ενδιαφέρον. Αυτός ήταν που τελικά με μύησε στον κόσμο της ρομποτικής. Τον ευχαριστώ ιδιαίτερα για την υπομονή του, την κατανόηση και την βοήθεια που μου προσέφερε καθόλη την διάρκεια της υλοποίησης της διπλωματικής μου. Ένα μεγάλο ευχαριστώ χρωστάω και στον κ. Λαγουδάκη και τα υπόλοιπα παιδιά της ομάδας, Αλέξανδρο, Γιώργο, Στάθη και Αντρέα, καθώς οι συζητήσεις μας με βοήθησαν πολύ και στην κατανόηση πολλών πραγμάτων πάνω στο ΝΑΟ και όχι μόνο. Κ.Λαγουδάκη, και του χρόνου πρώτοι! Δεν μπορώ να αμελήσω τον Γ. Κόντε που η βοήθειά του πάνω στο RL τόσο σε θεωρητικό επίπεδο, όσο και σε επίπεδο υλοποίησης ήταν πολύτιμη.

Τα όποια ευχαριστώ στον Σπύρο είναι λίγα μιας και οι χρήσιμες συμβουλές του πάνω στην διατριβή και η συμπαράσταση του σε όλα τα επίπεδα απέναντί μου ήταν αμέριστη.

Την όλη μου πορεία σαν φοιτήτρια, δεν μπορώ να την απομονώσω από όλους αυτούς που βρεθήκαμε μαζί στα αμφιθέατρα και στους αγώνες και ελπίζω η αμφισβήτηση και η ελπίδα για έναν καλύτερο κόσμο να συνεχίσει να πλανάται στους διαδρόμους και τα αμφιθέατρα αυτού του ιδρύματος.

# Περιεχόμενα

# Κεφάλαιο 1

# ΕΙΣΑΓΩΓΗ

Στην παρούσα διατριβή πραγματευόμαστε τον έλεγχο κίνησης δίποδου ρομπότ, χρησιμοποιώντας διάφορες μεθόδους ενισχυτικής μάθησης. Το ρομπότ που χρησιμοποιήσαμε είναι το μοντέλο NAO της Aldebaran robotics τα τεχνικά χαρακτηριστικά του οποίου βρίσκονται στο παράρτημα. Το όλο εγχείρημα έγινε χρησιμοποιώντας τον προσομοιωτή webots. Η τελική κίνηση που δημιουργήσαμε είναι μία κλωτσιά με το αριστερό του πόδι. Επιλέγχηκε η κίνηση αυτή για να εισαγάγουμε στο πρόβλημα μας τις συνθήκες ισορροπίας ενός δίποδου ρομπότ. Για την κατανόηση της δυναμικής και στατικής ισσοροπίας του ρομπότ χρειάστηκαν κάποιες βασικές έννοιες που περιγράφονται στο δεύτερο κεφάλαιο.

Αναλυτικότερα παρουσιάζουμε την έννοια του σημείου μηδενικής ροπής, που πρώτος εισήγαγε ο Vucobratovic και χρησιμοποιήθηκε έκτοτε στην ρομποτική, κυρίως στα ανθρωποειδή ρομπότ. Το κέντρο μάζας ενός ρομπότ είναι πολύ σημαντικό για την ευστάθεια του και για τον καθορισμό τελικά της οποιασδήποτε κινησιολογίας του. Το σημείο μηδενικής ροπής (ZMP) σχετίζεται άμεσα με το κέντρο μάζας του. Γνωρίζουμε ότι, στην περίπτωση που οι συνιστώσες, στο επίπεδο του εδάφους, του ZMP βρίσκονται εντός του πολυγώνου στήριξης του ρομπότ, το ρομπότ ισορροπεί. Το πολύγωνο στήριξης ενός δίποδου είναι στην περίπτωση που βρίσκεται σε στάση με τα δύο πόδια στο έδαφος το πολύγωνο που δημιουργείται από τα πέλματά του και την επιφάνεια που περικλύουν και στην περίπτωση που στέκεται στο ένα του πόδι, το πολύγωνο που δημιουργεί το πέλμα του που βρίσκεται σε επαφή με το έδαφος.

Η κίνηση του NAO προσεγγίστηκε χρησιμοποιώντας αλγόριθμους ενισχυτικής μάθησης. Η ενισχυτική μάθηση (Reinforcement learning ( RL )) είναι μια σύγχρονη μέθοδος που αναπτύχθηκε τα τελευταία χρόνια και χρησιμοποιείται ευρέως στην ρομποτική, μιας και η κλασσικές μέθοδοι δυναμικής και στατικής ανάλυσης σε ρομπότ με πολλούς βαθμούς ελευθερίας είναι πολύ δύσκολή έως αδύνατη. Η γενική φιλοσοφία των αλγορίθμων αυτών είναι πως ένα σύστημα περιγράφεται από καταστάσεις και δράσεις:

$$x_{k+1} \sim p\,(x_{k+1}|x_k, u_k)$$

όπου το $u_k \in \Re^M$ είναι η δράση και τα $x_k, x_k + 1 \in \Re^N$ είναι οι τρέχουσες και οι επόμενες καταστάσεις. Υποθέτουμε ότι οι δράσεις του συστήματος προσεγγίζονται από μια πολιτική $u_k \sim \pi_\vartheta(u_k|x_k)$ η οποία παραμετροποιείται από κάποιες παραμέτρους $\vartheta \in \Re^k$. Η ακολουθία των δράσεων και των καταστάσεων διαμορφώνουν μία τροχιά. Κάθε χρονική στιγμή το σύστημα υπολογίζει μία ανταμοιβή $r(x_k, u_k)$ που στην περίπτωσή μας περιλαμβάνει τις συνθήκες ευστάθειας του ρομπότ και τις βέλτιστες τιμές των αρθρώσεων. Σκοπός μας είναι να βελτιστοποιήσουμε τις παραμέτρους του προβλήματος έτσι ούτως ώστε να έχουμε το βέλτιστο αποτέλεσμα και ως προς την τελική κίνηση και ως προς την ευστάθεια .

Ακολουθώντας αυτή την φιλοσοφία στο τέταρτο κεφάλαιο περιγράφουμε την μοντελοποίηση και υλοποίηση του προβλήματος, ενώ στο πέμπτο κεφάλαιο παρουσιάζουμε τα αποτελέσματά μας. Στο έκτο κεφάλαιο παραθέτουμε κάποιους προβληματισμούς ως προς τα αποτελέσματα, την ίδια την αρχική μοντελοποιήση και παροτρύνουμε για περαιτέρω διερεύνηση.

Τέλος, όποιες διευκρινήσεις κρίθηκαν απαραίτητες κατά την συγγραφή του κειμένου βρίσκονται στα παραρτήματα που υπάρχουν και προφανώς στην βιβλιογραφία που χρησιμοποιήθηκε.

# Κεφάλαιο 2

# ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

## 2.1 Zero moment point (Σημείο Μηδενικής Ροπής)

Το 1968 περίπου ο Vucobratovic εισήγαγε την έννοια του Σημείο Μηδενικής Ροπής το οποίο έπαιξε σημαντικό ρόλο στην δυναμική και στατική ανάλυση των δίποδων ρομπότ κατά την κίνηση τους.

Το βασικό χαρακτηριστικό της κίνησης των δίποδων ρομπότ ( ας χρησιμοποιήσουμε ως παράδειγμα το περπάτημα ), είναι ότι κατά το βάδισμα τους εμφανίζονται δύο φάσεις, η στατική φάση διπλής υποστήριξης του δίποδου στα δύο του πόδια και η στατικά ασταθής φάση μονής υποστήριξης όπου το ένα πόδι βρίσκεται στο έδαφος και το άλλο μετακινείται προς τα εμπρός.

Το Σημείο Μηδενικής Ροπής είναι το σημείο εκείνο στο έδαφος, όπου η ροπή που παράγεται από τις δυνάμεις επίγειας αντίδρασης, είναι ίση με μηδέν. Το Σημείο Μηδενικής Ροπής υπολογίζεται αναλυτικά, αν βασιστούμε στην θέση και την επιτάχυνση των αρθρώσεων του ρομπότ κατά την διάρκεια μια κίνησης, ή κατά την οποιαδήποτε θέση του, όταν αυτό είναι ακίνητο.

Το Σημείο Μηδενικής Ροπής είναι ισοδύναμο με το κέντρο της πίεσης (Center of Pressure), που αποτελεί το σημείο εκείνο στο έδαφος που ενεργεί η δύναμη της επίγειας αντίδρασης. Το κέντρο της πίεσης (*CoP*) μπορεί να μετρηθεί άμεσα, αν τοποθετήσουμε αισθητήρες δύναμης (force sensors), συμμετρικά, στο κάθε πέλμα του ρομπότ. Αν και το Σημείο Μηδενικής Ροπής και το *CoP* είναι ισοδύναμα, ο όρος Σημείο Μηδενικής Ροπής χρησιμοποιείται συνήθως όταν αναφερόμαστε στο σημείο που υπολογίσαμε αναλυτικά από τις επιταχύνσεις και τις θέσεις των αρθρώσεων του ρομπότ, ενώ ο όρος *CoP* χρησιμοποιείται συνήθως όταν αναφερόμαστε στο ισοδύναμο σημείο, το οποίο υπολογίσαμε από τους αισθητήρες δύναμης που βρίσκονται στα πέλματα του ρομπότ.

Για τον έλεγχο οποιονδήποτε κινήσεων του ρομπότ ( είτε στατικών , είτε δυναμικών ), η μέθοδος του Σημείο Μηδενικής Ροπής χρησιμοποιείται ευρέως. Σε αυτή τη μέθοδο, η βασική υπόθεση είναι ότι, το πόδι ή τα πόδια εφάπτονται στο έδαφος και το έδαφος είναι επίπεδο. Εφόσον συμβαίνει αυτό, οι τροχιές των αρθρώσεων για μία κίνηση μπορούν να προγραμματιστούν θεωρώντας ότι το ρομπότ είναι σταθερά συνδεδεμένο με το έδαφος. Το πρόβλημα όμως σε αυτή την υπόθεση είναι ότι το πόδι ενός δίποδου δεν είναι γενικά σε τέλεια επαφή με το έδαφος. Το πόδι εύκολα μπορεί να γλιστρήσει.

Για ένα δίποδο, εάν το Σημείο Μηδενικής Ροπής βρίσκεται στην άκρη της βάσης του πολυγώνου στήριξης, το πόδι μπορεί να αρχίσει να κυλά [2]. Η μέθοδος ελέγχου της κίνησης του δίποδου με τη χρήση του Σημείο Μηδενικής Ροπής λύνει αυτό το πρόβλημα, με την εξασφάλιση ότι για τις διάφορες τροχιές των αρθρώσεων που παράγουμε για μία κίνηση, το Σημείο Μηδενικής Ροπής βρίσκεται αυστηρά μέσα στο πολύγωνο υποστήριξης.

Ο υπολογισμός των τροχιών των αρθρώσεων με τη μέθοδο του Σημείο Μηδενικής

Ροπής, έχει ως αποτέλεσμα μια σταθερή και επιτυχημένη κίνηση. Η κίνηση όμως αυτή είναι εύθραυστη λόγω των διαταραχών (εξωτερικών δυνάμεων ) που μπορεί να δεχθεί το ρομπότ. Η απαίτηση του ότι τα πόδια ή το πόδι ( ανάλογα την κίνηση ) πατάνε επίπεδα στο έδαφος, είναι υπερβολικά συντηρητική. Παραδείγματος χάριν, ο άνθρωπος όταν περπατάει δεν υπακούει σε αυτόν τον περιορισμό. Ακολουθώντας αυτή την μέθοδο λοιπόν, η κίνηση του ρομπότ δεν είναι ιδιαίτερα ρεαλιστική ( π.χ. στο περπάτημα περπατάει σαν να έχει πλατυποδία ) και συνεπώς δεν βελτιστοποιείται όσο θα μπορούσε.

Τέλος υπάρχει ένα θεμελιώδες πρόβλημα στη μέθοδο του Σημείο Μηδενικής Ροπής, το οποίο γίνεται κατανοητό αν μελετήσουμε τι πραγματικά εισάγουμε στο πρόβλημα μας για να εξάγουμε το αποτέλεσμα που επιθυμούμε. Ας εξετάσουμε το πρόβλημα του αυτοκινήτου που κινείται με σταθερή ταχύτητα. Για να γίνει αυτό, ο οδηγός ελέγχει την ταχύτητα στο ταχύμετρο ρυθμίζοντας τα πετάλια ( γκάζι, φρένο και ταχύτητες ) αναλόγως. Η θέση των πεταλιών μπορούμε να πούμε ότι αντιπροσωπεύει την επιτάχυνση-δύναμη που εισάγεται στο σύστημα. Αυτό που παράγεται είναι το είδος κάθε φορά της κινήσεως του αυτοκινήτου. Γι' αυτό το παράδειγμα , αλλά και για όλα τα προβλήματα ελέγχου εν γένει, η επιθυμητή συμπεριφορά επιτυγχάνεται από τον έλεγχο αυτού που παράγεται ( δηλαδή του αποτελέσματος της κινήσεως ) και όχι από τον έλεγχο αυτού που εισάγουμε. Κατά συνέπεια, για το συγκεκριμένο πρόβλημα, η επιθυμητή συμπεριφορά είναι η κίνηση και όχι η θέση των πεταλιών. Στην πραγματικότητα είναι πολύ δύσκολο για έναν οδηγό να διατηρήσει σταθερή ταχύτητα στο αυτοκίνητο αν βασιστεί μόνο στο που βρίσκονται τα πετάλια του αυτοκινήτου, χωρίς να συμβουλευτεί το ταχύμετρο.

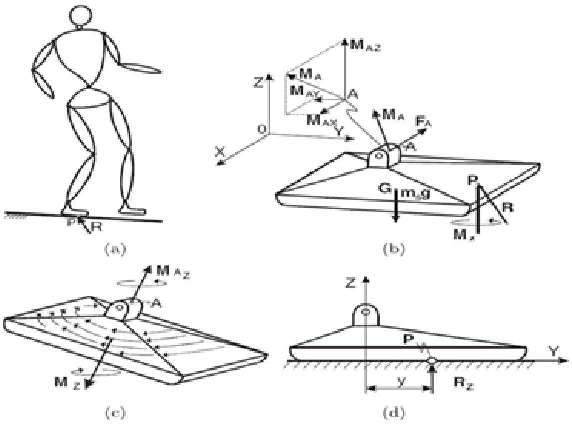
Ο έλεγχος της θέσης των πεταλιών και όχι του ταχύμετρου, δείχνει ακριβώς την προσέγγιση ενός ελέγχου που γίνεται από τον έλεγχο της εισαγωγής των δεδομένων και όχι από τον έλεγχο της παραγωγής (του αποτελέσματος). Γενικά η προσέγγιση αυτή που γίνεται από την εισαγωγή (*input*) όχι και από τον έλεγχο της παραγωγής (*output*), είναι μια ανοιχτού βρόγχου προσέγγιση( $open - loop$ ) και γενικά αποφεύγεται στα περισσότερα συστήματα ελέγχου . Οι ανοιχτού βρόγχου προσεγγίσεις είναι ευαίσθητες στα λάθη και την αστάθεια επειδή δεν ελέγχουν την μεταβλητή που πραγματικά μας ενδιαφέρει.

Η μέθοδος του Σημείου Μηδενικής Ροπής είναι μια προσέγγιση ανοικτού βρόγχου. Το Σημείο Μηδενικής Ροπής σχετίζεται άμεσα με την οριζόντια δύναμη που εφαρμόζεται στο κέντρο μάζας του ρομπότ ( όπως θα εξηγήσουμε παρακάτω ) αποτελώντας έτσι μία επιτάχυνση που εισάγεται στο σύστημα. Το αποτέλεσμα που επιθυμούμε να ελέγξουμε στην περίπτωση της επίτευξης ισορροπίας στο ρομπότ, είναι η θέση του κέντρου μάζας του και η ταχύτητά του. Επειδή η μέθοδος του Σημείο Μηδενικής Ροπής καθορίζεται από το Σημείο Μηδενικής Ροπής, η εν λόγω μέθοδος είναι μια καθαρά ανοικτού βρόγχου προσέγγιση ελέγχου.

Ας εξετάσουμε ένα δίποδο σε δύο φάσεις της κίνησης του για να δούμε πώς υπολογίζεται το Σημείο Μηδενικής Ροπής. Θα μελετήσουμε την περίπτωση όπου το ρομπότ ισορροπεί με το ένα πόδι στο έδαφος.

Γενικά , η συνολική δύναμη αντίδρασης από το έδαφος αποτελείται από τις 3 συνιστώσες της δύναμης $\overrightarrow{R} = (R_x, R_y, R_z)$ και τη ροπή $\overrightarrow{M} = (M_x, M_y, M_z)$. Εφόσον η δύναμη της τριβής επιδρά στο σημείο επαφής του ποδιού με το έδαφος και το πόδι θεωρούμε ότι είναι ακίνητο, τότε οι συνιστώσες της δύναμης $\overrightarrow{R}$ και της ροπής $\overrightarrow{M}$ στο οριζόντιο επίπεδο, είναι ίσες κατά μέτρο με την δύναμη της τριβής.

Επομένως η οριζόντια δύναμη αντίδρασης $(R_x, R_y)$ αντιπροσωπεύει τη δύναμη της τριβής που ισορροπεί την οριζόντια δύναμη $F_A$, ενώ η κάθετη ροπή $M_z$ αντιπροσωπεύει τη ροπή από τις δυνάμεις αντίδρασης της τριβής ( σχήμα *c*). Κατά συνέπεια, αν υποθέσουμε ότι η επαφή του ποδιού με το έδαφος είναι χωρίς ολίσθηση, η στατική τριβή θα αντισταθμίζει τις οριζόντιες συνιστώσες $R_x, R_y$ και την κάθετη ροπή αντίδρασης $M_z$. Η κάθετη δύναμη αντίδρασης $R_z$ αντιπροσωπεύει την επίγεια αντίδραση που



Σχήμα 2.1: ZMP

ισορροπεί τις κάθετες δυνάμεις.

Είναι προφανές ότι η δύναμη επίγειας αντίδρασης που προκαλείται από το πόδι είναι πάντα προσανατολισμένη προς τα πάνω. Οι οριζόντιες συνιστώσες της ροπής λοιπόν, μπορούν να αντισταθμιστούν μόνο από τη μεταβαλλόμενη θέση της δύναμης $\overrightarrow{P}$ που βρίσκεται μέσα στο πολύγωνο υποστήριξης. Επομένως η οριζόντια συνιστώσα της ροπής $M_A$ θα μετατοπίσει την δύναμη αντίδρασης στην αντίστοιχη θέση για να ισορροπήσει το πρόσθετο φορτίο. Αυτό φαίνεται από την εικόνα ( σχήμα *d*), όπου χάριν απλότητας παρουσιάζεται μια απλή περίπτωση στο επίπεδο $Y - Z$. Η συνιστώσα της ροπής $M_A x$ εξισορροπείται αν μετατοπίσουμε το σημείο που ενεργεί η συνιστώσα δύναμη $R_z$, η ένταση της οποίας υπολογίζεται από την εξίσωση ισορροπίας των δυνάμεων που ενεργούν στο πόδι στην απόσταση $y$.

Τονίζουμε ότι κάθε χρονική στιγμή η δύναμη αντίδρασης είναι μέσα στην περιοχή που καλύπτεται από το πόδι ( πολύγωνο στήριξης ). Η οποιαδήποτε αύξηση της ροπής των αστραγάλων αντισταθμίζεται από την αλλαγή της θέσης αυτής της δύναμης και δεν θα δημιουργηθεί καμία συνιστώσα $M_x$ και $M_y$. Γι αυτό το λόγο στην εικόνα ( σχήμα *b*) στο σημείο $P$, εμφανίζεται μόνο η συνιστώσα $M_z$. Παρόλα αυτά αν το πραγματικό πολύγωνο στήριξης δεν είναι αρκετά μεγάλο για να καλύψει τη θέση της δύναμης $\overrightarrow{P}$, θα ενεργήσει στην άκρη του ποδιού και η οριζόντια συνιστώσα της ροπής θα προκαλέσει περιστροφή του μηχανισμού που μπορεί να προκαλέσει και την ανατροπή του. Επομένως μπορούμε να πούμε ότι η απαραίτητη προϋπόθεση για την ισορροπία ενός δίποδου ρομπότ είναι ότι για το σημείο $P$ που ενεργεί η δύναμη επίγειας αντίδρασης

$$\begin{aligned} Mx &= 0, \\ My &= 0 \end{aligned} \tag{2.1}$$

Δεδομένου ότι οι συνιστώσες $M_x$ και $M_y$ είναι μηδενικές στην περίπτωση που βρισκόμαστε σε δυναμική ισορροπία, ονομάζουμε αυτό το σημείο, σημείο μηδενικής ροπής ( Zero Moment Point ). Μια λογική ερώτηση τώρα είναι: ¨*ποια θα έπρεπε να είναι η θέση του Σημείο Μηδενικής Ροπής για να βρίσκεται ένα δίποδο σε δυναμική ισορροπία;*¨ Για να απαντήσουμε σε αυτή την ερώτηση, αρκεί να λύσουμε τις στατικές εξισώσεις ισορροπίας

$$\begin{aligned} \overrightarrow{R} + \overrightarrow{F}_A + m_s\overrightarrow{g} &= 0 \\ \overrightarrow{OP} \times \overrightarrow{R} + \overrightarrow{OG} \times m_s\overrightarrow{g} + \overrightarrow{M}_A + M_Z + \overrightarrow{OA} \times \overrightarrow{F}_A &= 0 \end{aligned} \tag{2.2}$$

Όπου $O$ είναι η αρχή των αξόνων $xyz$ και $P$ είναι το σημείο που ενεργεί η δύναμη

$\overrightarrow{P}$, $\overrightarrow{G}$ είναι το σημείο του κέντρου μάζας του ποδιού και $A$ είναι το σημείο που βρίσκεται η άρθρωση του αστραγάλου. Η μάζα του ποδιού είναι $m_s$.

Αν τοποθετήσουμε την αρχή των αξόνων στο σημείο $P$, τότε:

$$M_z = M_{fr} = -\left(M_A^Z + \left(\overrightarrow{OA} \times \overrightarrow{F}_A\right)^Z\right)$$

Στην γενική αυτή περίπτωση η ροπή είναι διαφορετική του μηδενός, και μπορεί να γίνει μηδέν μόνο μέσω της κατάλληλης δυναμικής όλου του συστήματος. Ωστόσο , στο οριζόντιο επίπεδο έχουμε :

$$\begin{aligned}\left(\overrightarrow{OP} \times \overrightarrow{R}\right)^H &= \begin{bmatrix} \widehat{x} & \widehat{y} & \widehat{z} \\ OP_x & OP_y & OP_z \\ R_x & R_y & R_z \end{bmatrix} = \ldots οριζοντια \quad προβολη \ldots \\ &= \widehat{x}\left(OP_x R_z - R_y OP_z\right) - \widehat{y}\left(OP_x R_z - R_x OP_z\right) \end{aligned} \tag{2.3}$$

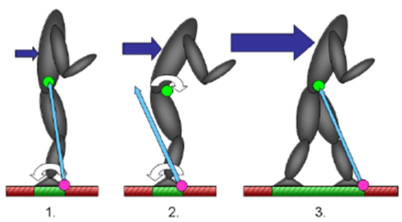
$$\left(\overrightarrow{OP} \times \overrightarrow{R}\right) + \overrightarrow{OG} \times m_s \overrightarrow{g} + \overrightarrow{M}_A^H + \left(\overrightarrow{OA} \times \overrightarrow{F}_A\right)^H = 0 \tag{2.4}$$

Μέσω αυτής της εξίσωσης μπορούμε να υπολογίσουμε το Σημείο Μηδενικής Ροπής.

## 2.2 Κέντρο Μάζας (CoM) και Σημείο Μηδενικής Ροπής (ZMP)

Το κέντρο μάζας του ρομπότ, συναρτήσει των κέντρων μαζών των επιμέρους τμημάτων του σώματος του δίνεται από τον τύπο:

$$CoM = \frac{\sum_{i=1}^{k} m_i r_i}{\sum_{i=1}^{k} m_i} \tag{2.5}$$




Σχήμα 2.2:

Γνωρίζουμε ότι αν η προβολή του κέντρου μάζας του στο οριζόντιο επίπεδο ( ή πιο απλά στο έδαφος ) βρίσκεται μέσα στο πολύγωνο υποστήριξης του ρομπότ, το ρομπότ βρίσκεται σε δυναμική ισορροπία (αυτό για την περίπτωση που το ρομπότ στέκεται ακίνητο, διότι τότε το κέντρο μάζας συμπίπτει με το ZMP). Το πολύγωνο υποστήριξης στην περίπτωση που το NAO πατάει και τα δύο του πόδια στο έδαφος είναι η επιφάνεια που σχηματίζεται από τα δύο του πόδια συμπεριλαμβανομένου και της επιφάνειας που βρίσκεται ανάμεσα από αυτά. Στην περίπτωση που το NAO έχει μόνο το ένα του πόδι στο έδαφος τότε το πολύγωνο στήριξης είναι η κοινή επιφάνεια πέλματος- εδάφους. Ο πρώτος έλεγχος λοιπόν για να δούμε αν το ρομπότ βρίσκεται σε κατάσταση δυναμικής ισορροπίας είναι εάν το ZMP βρίσκεται εντός του πολυγώνου στήριξης. Εάν ισχύει η συνθήκη αυτή τότε βρισκόμαστε σε κατάσταση δυναμικής ισορροπίας. Εάν δεν βρισκόμαστε σε κατάσταση ισορροπίας τότε υπάρχουν τρεις διαφορετικές στρατηγικές για να επέλθουμε σε αυτή.

Στην πρώτη περίπτωση η δύναμη που δέχεται το ρομπότ είναι μικρή. Το *ZMP* (ροζ κουκίδα ) ισορροπεί με το *COM* (πράσινη κουκίδα). Δηλαδή, η συνισταμένη δύναμη που ασκείται από το έδαφος, ¨δείχνει¨ στο *COM*. Σε αυτή την περίπτωση οποιαδήποτε βελτίωση της θέσης του ρομπότ γίνεται με την άσκηση ροπής στον αστράγαλο. Στην δεύτερη περίπτωση η δύναμη που δέχεται το ρομπότ και είναι μεγαλύτερη και για να επέλθει πάλι σε κατάσταση δυναμικής ισορροπίας πρέπει να ασκήσουμε την κατάλληλη ροπή στο *COM*. Στην τρίτη περίπτωση η δύναμη που ασκείται στο ρομπότ είναι τόσο μεγάλη που ο μόνος τρόπος να επανέλθει πάλι σε κατάσταση ισορροπίας είναι να κάνει ένα βήμα, ούτος ώστε να αλλάξει το πολύγωνο στήριξης και η προβολή του *COM* στο οριζόντιο επίπεδο να βρεθεί πάλι μέσα σε αυτό.

Το σημείο μηδενικής ροπής, συναρτήσει των συντεταγμένων $x_c, y_c, z_c$ του κέντρου μάζας, δίνεται από τον τύπο:

$$ZMP_x = x_c - \alpha(t)\ddot{x}_c \tag{2.6}$$

$$ZMP_y = y_c - \alpha(t)\ddot{y}_c \tag{2.7}$$

όπου $\alpha(t) = \frac{z_c - ZMP_z}{\ddot{z}_c + g}$
(βλέπε παράρτημα Γ.)

## 2.3 Το αντίστροφο εκκρεμές

Το αντίστροφο εκκρεμές είναι ένα εκκρεμές που έχει την μάζα του, πάνω από τον άξονα στήριξής του. Εφαρμόζεται συχνά με τον άξονα του στερεωμένο πάνω σε ένα όχημα που μπορεί να κινηθεί οριζόντια. Εκτιμώντας ότι ένα κανονικό εκκρεμές είναι σταθερό όταν κρεμιέται προς τα κάτω, ένα εκκρεμές αντεστραμμένο είναι εγγενώς ασταθές και προκειμένου να παραμείνει όρθιο, πρέπει να ασκείται σε αυτό, είτε κάποια ροπή στον άξονα, είτε με την κίνηση του άξονα οριζόντια, λόγω της κίνησης του οχήματος.

Το αντίστροφο εκκρεμές είναι ένα κλασσικό πρόβλημα στην θεωρία του αυτομάτου ελέγχου και ευρέως χρησιμοποιούμενο για την δοκιμή αλγορίθμων ελέγχου (νευρωνικά δίκτυα , γενετικοί αλγόριθμοι κ.α.). Οι εξισώσεις της κίνησης που διέπουν το αντίστροφο εκκρεμές είναι: Για το σταθερό σημείο του άξονα: Η εξίσωση της κίνησης είναι παρόμοια με αυτήν για το ορθό εκκρεμές με τη διαφορά του ότι η γωνιακή θέση του εκκρεμούς μετράται από την κάθετη ασταθή θέση ισορροπίας:

$$\ddot{\vartheta} - \frac{g}{l}\sin\vartheta = 0 \tag{2.8}$$

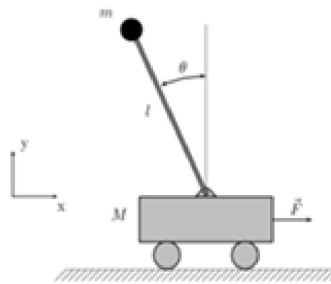
Όταν προστίθεται και στις δύο πλευρές, θα έχει την ίδια γωνιακή επιτάχυνση:

$$\ddot{\vartheta} = \frac{g}{l}\sin\vartheta \tag{2.9}$$

Κατά συνέπεια το εκκρεμές θα επιταχύνει μακριά από την κάθετη ασταθή ισορροπία στην κατεύθυνση που μετατοπίζεται αρχικά και η επιτάχυνσή του θα είναι αντιστρόφως ανάλογη με το μήκος του. Τα ψηλά εκκρεμή πέφτουν πιο αργά από τα πιο κοντά.

### 2.3.1 Το αντίστροφο εκκρεμές πάνω σε όχημα

Οι εξισώσεις κίνησης του συγκεκριμένου προβλήματος μπορούν να παραχθούν εύκολα χρησιμοποιώντας τις εξισώσεις Lagrange, όπως φαίνεται και στην εικόνα, το $x(t)$ είναι η θέση του οχήματος, $\vartheta(t)$ είναι η γωνία του εκκρεμούς που σχηματίζεται με την κάθετη διεύθυνση και οι δυνάμεις που ενεργούν είναι η βαρύτητα και μία εξωτερική δύναμη στην χ- διεύθυνση. Η Λαγκρανζιανή εξίσωση $L = T - V$, όπου το $T_a$ είναι




Σχήμα 2.3: Το αντίστροφο εκκρεμές πάνω σε όχημα

η κινητική ενέργεια στο σύστημα και το $V$ είναι η πιθανή ενέργεια του συστήματος. Έτσι το $L$ είναι:

$$L = \frac{1}{2}Mu_1^2 + \frac{1}{2}mu_2^2 - mgl\cos\partial \tag{2.10}$$

Όπου $u_1$ είναι η ταχύτητα του οχήματος και $u_2$ είναι η ταχύτητα της μάζας του σημείου $m$. Το $u_1$ και $u_2$ μπορεί να εκφραστεί μέσω των $x$ και $\partial$ αν γράψουμε τις ταχύτητες ως τις πρώτες παραγώγους των αντίστοιχων θέσεων.

$$\begin{aligned} u_1^2 &= \dot{x}^2 \\ u_2^2 &= \left(\tfrac{d}{dt}\left(l\cos\partial\right)\right)^2 + \left(\tfrac{d}{dt}\left(x + l\sin\partial\right)\right)^2 \end{aligned} \tag{2.11}$$

Απλοποιώντας την έκφραση της ταχύτητας $u_2$ έχουμε:

$$u_2^2 = \dot{x}^2 + 2\dot{x}l\dot{\partial}\cos\partial + l^2\dot{\partial}^2$$

Η Λαγκρανζιανή δίνεται τώρα ως εξής:

$$L = \frac{1}{2}\left(M + m\right)\dot{x}^2 + ml\dot{x}\dot{\partial}\cos\partial + \frac{1}{2}ml^2\dot{\partial}^2 - mgl\cos\partial \tag{2.12}$$

Και οι εξισώσεις της κίνησης είναι :

$$\begin{aligned} \tfrac{d}{dt}\tfrac{\partial L}{\partial \dot{x}} - \tfrac{\partial L}{\partial x} &= F \\ \tfrac{d}{dt}\tfrac{\partial L}{\partial \dot{\partial}} - \tfrac{\partial L}{\partial \partial} &= 0 \end{aligned} \tag{2.13}$$

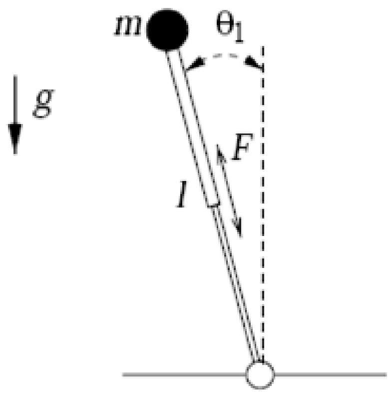
Αντικαθιστώντας τη συνάρτηση $L$ σε αυτές τις εξισώσεις και απλοποιώντας τες, έχουμε τις εξισώσεις που περιγράφουν την κίνηση του εκκρεμούς:

$$\begin{aligned} \left(M + m\right)\ddot{x} + ml\ddot{\partial}\cos\partial - ml^2\dot{\partial}^2\sin\partial &= F \\ ml\left(-g\sin\partial + \ddot{x}\cos\partial + l\ddot{\partial}\right) &= 0 \end{aligned} \tag{2.14}$$

Αυτές οι εξισώσεις είναι μη γραμμικές, αλλά δεδομένου ότι ο στόχος ενός συστήματος ελέγχου θα ήταν να κρατηθεί το εκκρεμές κατακόρυφα, οι εξισώσεις μπορούν να θεωρηθούν γραμμικές για $\partial \approx 0$.

### 2.3.2 Το αντίστροφο εκκρεμές στην ρομποτική

Στην ρομποτική η προσέγγιση του αντίστροφου εκκρεμούς χρησιμοποιήθηκε ευρέως. Η πολυπλοκότητα των ρομπότ με πολλούς βαθμούς ελευθερίας, καθιστά πολύ δύσκολη έως αδύνατη την ευθεία και αντίστροφη κινηματική τους ανάλυση. Έτσι, τα δίποδα ρομπότ από πολλούς ερευνητές, αντιμετωπίστηκαν ως ένα αντίστροφο εκκρεμές, με μάζα στο άκρο του, όσο η μάζα του ρομπότ και αβαρή πόδια. Ένα μοντέλο αντίστροφου εκκρεμούς μου χρησιμοποιείται στην ρομποτική για να προσεγγίσει ένα δίποδο είναι το εξής:




Σχήμα 2.4: Η προσέγγιση με αντίστροφο εκκρεμές

Είναι ένα απλό ανεστραμμένο εκκρεμές μάζας $m$ και μήκους $l$. Η γωνιακή ροπή του εκκρεμούς γύρω από τον άξονά του είναι:

$$H_0 = ml^2\dot{\vartheta}_1 \tag{2.15}$$

Η αλλαγή στην γωνιακή ορμή λόγω της βαρύτητας είναι:

$$\dot{H}_0 = mgl\sin\vartheta_1 \tag{2.16}$$

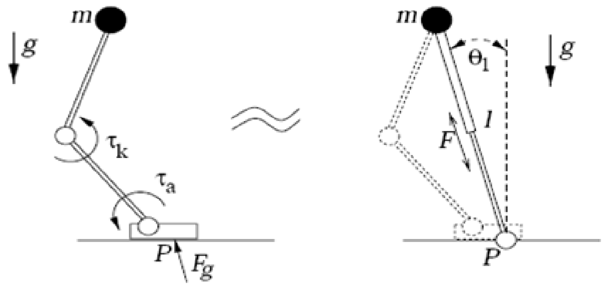
Διαφορίζοντας την $H_0$ και αντικαθιστώντας, έχουμε την δεύτερη παράγωγο της $\vartheta_1$ :

$$\ddot{\vartheta}_1 = \frac{g}{l}\sin\vartheta_1 \tag{2.17}$$

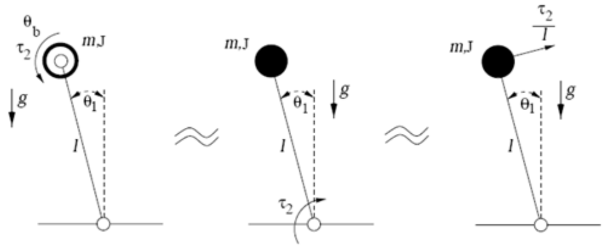
Είναι εύκολο να αντιληφθεί κανείς ότι αν η μάζα στην άκρη του εκκρεμούς είναι στα αριστερά του άξονα, καθώς ανεβαίνει το εκκρεμές, η μάζα επιβραδύνεται. Αυτό συμβαίνει λόγο της επίδρασης της βαρύτητας. Όταν η μάζα περάσει στα δεξιά του άξονα, θα αρχίσει πάλι να επιταχύνεται. Αυτό μαθηματικά φαίνεται στις εξισώσεις κίνησης του εκκρεμούς:

$$\begin{aligned} ml^2\ddot{\vartheta}_1 &= mgl\sin\vartheta_1 - 2ml\dot{l}\dot{\vartheta}_1 \\ m\ddot{l} &= F - mg\cos\vartheta_1 + ml\dot{\vartheta}_1^2 \end{aligned} \tag{2.18}$$

Ακόμα και στην περίπτωση που έχουμε πάνω από μία άρθρωση στο πόδι, όπως φαίνεται στην εικόνα 2.1 που έχουμε δύο περιστροφικές αρθρώσεις, το ανεστραμμένο




Σχήμα 2.5: Το αντίστροφο εκκρεμές




Σχήμα 2.6: Το αντίστροφο εκκρεμές (2)

εκκρεμές μπορεί και πάλι να προσεγγίσει το μοντέλο του ρομπότ. Αυτό είναι αυτονόητο αν η άρθρωση στο γόνατο του ποδιού βρίσκεται σε γωνία $180^o$ με την κνήμη. Αλλά και στην περίπτωση που έχουμε κάποια θέση του ποδιού σε κλίση (βλέπε εικόνα ) και πάλι το ανεστραμμένο εκκρεμές μπορεί να προσεγγίσει το μοντέλο. Για να το αποδείξουμε αυτό αρκεί να δούμε αναλυτικά τις δυναμικές εξισώσεις του συστήματος. Οι εξισώσεις κίνησης του συστήματος στην κατεύθυνση του $l$ και της γωνίας $\vartheta_1$ είναι:

$$
\begin{aligned}
ml^2\ddot{\vartheta}_1 &= mgl\sin\vartheta_1 - 2ml\dot{l}\dot{\vartheta}_1 \\
m\ddot{l} &= f(\tau_2, \tau_3, \vartheta_1, \vartheta_2, \vartheta_3) - mg\cos\vartheta_1 + ml\dot{\vartheta}_1^2
\end{aligned}
\tag{2.19}
$$

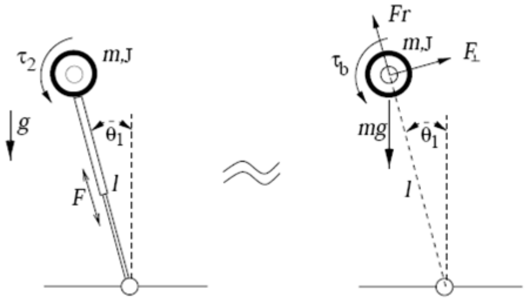
Όπου η $F$ είναι η μη γραμμική δύναμη που προέρχεται από τις ροπές στις αρθρώσεις προς το κέντρο μάζας του εκκρεμούς. Αν συγκρίνουμε τις εξισώσεις αυτές με τις παρατηρούμε ότι είναι ταυτόσημες.

Στην εικόνα φαίνεται ότι το ρομπότ με συγκεντρωμένη την μάζα του στο κέντρο μάζας και με δύο βαθμούς ελευθερίας στο πόδι, προσεγγίζεται με ένα ανεστραμμένο εκκρεμές, μήκους $l$. Σε όλα τα παραπάνω μοντέλα υποθέσαμε ότι η μάζα είναι συγκεντρωμένη σε ένα σημείο. Επομένως, δεν μπορεί να εφαρμοστεί σε αυτή καμία ροπή. Ένα άλλο μοντέλο ανεστραμμένου εκκρεμούς που χρησιμοποιείται στη ρομποτική είναι το ανεστραμμένο εκκρεμές με δυνατότητα άσκησης ροπής στην μάζα του ( flywheel model ).

Οι εξισώσεις κίνησης που προκύπτουν για το μοντέλο ανεστραμμενου εκκρεμούς με σφόνδυλο είναι:

$$
\begin{aligned}
ml^2\ddot{\vartheta}_1 &= mgl\sin\vartheta_1 - \tau_2 \\
J\ddot{\vartheta}_b &= \tau_2
\end{aligned}
\tag{2.20}
$$

Όπου $\vartheta_b$ είναι η γωνία του σφονδύλου που σχηματίζει με το έδαφος, δηλαδή $\vartheta_b = \vartheta_1 + \vartheta_2$. Στην μεσαία εικόνα, υπάρχει ο σφόνδυλος αλλά για την εξισορρόπηση των δυνάμεων αδράνειας έχουμε μια ροπή στον αστράγαλο $\tau_2$. Στην τελευταία εικόνα




Σχήμα 2.7: Το αντίστροφο εκκρεμές

καθορίζουμε και πάλι τον σφόνδυλο, αλλά αντί την χρησιμοποίηση του για την εξισορρόπηση των δυνάμεων, εφαρμόζουμε μια εξωτερική δύναμη $\frac{\tau_2}{l}$ στην μάζα του εκκρεμούς, κάθετα στο πόδι. Η δυναμική του συστήματος και στις τρεις περιπτώσεις περιγράφεται από την εξίσωση

$$ml^2\ddot{\vartheta}_1 = mgl\sin\vartheta_1 - \tau_2 \tag{2.21}$$

Αν αγνοήσουμε την $\vartheta_b$ της εικόνας τότε τα τρία συστήματα είναι δυναμικά ισοδύναμα. Από τα ανωτέρω βλέπουμε πως ένας σφόνδυλος μπορεί να χρησιμοποιηθεί για να ελέγξει την περιστροφική δυναμική ενός εκκρεμούς. Με την προσθήκη ενός γραμμικού ενεργοποιητή ( linear actuator ), (σχήμα από κάτω) και οι τρείς αδρανειακές κινήσεις του εκκρεμούς είναι ελέγξιμες.

Οι εξισώσεις της κίνησης για το μοντέλο στα αριστερά είναι:

$$\begin{aligned} ml^2\ddot{\vartheta}_1 &= mgl\sin\vartheta_1 - 2ml\dot{l}\vartheta_1 - \tau_2 \\ m\ddot{l} &= ml\dot{\vartheta}_1^2 - mg\cos\vartheta_1 + F \\ J\ddot{\vartheta}_b &= \tau_2 \end{aligned} \tag{2.22}$$

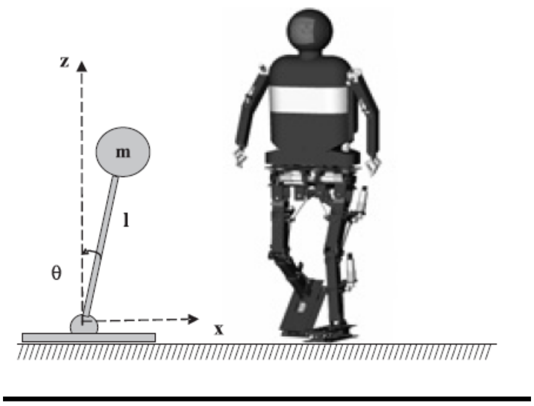
Στο σχήμα στα δεξιά φαίνεται ένα ελεύθερο διάγραμμα ενός εκκρεμούς με τις δυνάμεις και τις ροπές που του ασκούνται. Τα δύο μοντέλα είναι δυναμικά ισοδύναμα αν οι δυνάμεις που ασκούνται στο ελεύθερο διάγραμμα ( εικόνα στα δεξιά ) είναι:

$$\begin{aligned} F_\perp &= \frac{\tau_2}{l} \\ F_r &= F \\ \tau_b &= \tau_2 \end{aligned} \tag{2.23}$$

Είναι απαραίτητο τέλος να τονίσουμε την διαφορά των $\tau_b$ και $\tau_2$ για να γίνει απολύτως κατανοητό. Η $\tau_2$ είναι μία εσωτερική ροπή που εφαρμόζεται μεταξύ του σφονδύλου και του άξονα του εκκρεμούς. Η $\tau_b$ είναι μια εξωτερική ροπή που εφαρμόζεται μεταξύ του σφονδύλου και του κόσμου. Η εσωτερική ροπή $\tau_2$ μεταξύ του σφονδύλου και του εκκρεμούς, είναι δυναμικά ισοδύναμη με την ροπή $\tau_b$ και την δύναμη $F_\perp$ που ασκείται εξωτερικά στον σφόνδυλο. Η δύναμη $F_r$ μπορεί να εφαρμοστεί ανεξάρτητα στο σύστημα, ενώ οι $\tau_b$ και $F_\perp$ είναι αλληλένδετες. Για να κατανοήσουμε τον λόγο που είναι αλληλένδετες, αρκεί να σκεφτούμε ότι η συνολική ροπή γύρω από τον άξονα του εκκρεμούς πρέπει να είναι μηδέν.

$$\tau_b + F_\perp l = 0 \tag{2.24}$$

Δηλαδή η γωνιακή ροπή γύρω από τον άξονα του εκκρεμούς επηρεάζεται μόνο από την βαρύτητα.




Σχήμα 2.8: Προσέγγιση διπόδου με ανεστραμμένο εκκρεμές

## 2.4 Το αντίστροφο εκκρεμές, το ZMP και το CoM

Όπως έχουμε ήδη αναφέρει, το σημείο μηδενικής ροπής, είναι ένα χρήσιμο εργαλείο για να συνθέτουμε κινήσεις των δίποδων ρομπότ. Οι περισσότερες από τις προηγούμενες μελέτες που πραγματεύονται αυτό ακριβώς το θέμα, θα μπορούσαν να χωριστούν σε δύο κατηγορίες. Στις μελέτες που παράγουν κινήσεις σε μη πραγματικό χρόνο (off-line) και σε πραγματικό χρόνο (on-line). Στην πρώτη περίπτωση, οι τροχιές των αρθρώσεων αρθρώσεων του διπόδου, που ακολουθεί για την επίτευξη μιας κίνησης, είναι προϋπολογισμένες και σταθερές. Στην δεύτερη περίπτωση, οι τροχιές των αρθρώσεων του ρομπότ, υπολογίζονται on-line σύμφωνα με την σταθερότητα του σημείου μηδενικής ροπής και τους υπόλοιπους στόχους της εκάστοτε κίνησης.

Εντούτοις, εξαιτίας του γεγονότος, του ότι και οι δύο προσεγγίσεις παραγωγής κινήσεων εξετάζουν τους κινητήρες του ρομπότ σαν ιδανικές γεννήτριες ροπών και τις μηχανικές δομές του, ως ιδανικούς άκαμπτους, συνδέσμους σωμάτων, (προκειμένου να αποφευχθεί το υψηλό υπολογιστικό φορτίο), η δυναμική, όπως οι κοινές σπασμωδικές κινήσεις, ο κορεσμός των ενεργοποιητών, οι διάφοροι τρόποι δονήσεων, οποιοιδήποτε από τους οποίους θα μπορούσαν να επηρρεάσουν την σταθερότητα της κίνησης, δεν χρησιμοποιούνται στην μοντελοποίηση. Επιπλέον, το πλήρες μαθηματικό πρότυπο του σημείου μηδενικής ροπής, είναι πολύ σύνθετο για να χρησιμοποιηθεί στον προγραμματισμό σε πραγματικό χρόνο, των τροχιών των αρθρώσεων. Αυτό οφείλεται στο γεγονός, ότι υπάρχει μια μη γραμμική σχέση μεταξύ της δύναμης επίγειας αντίδρασης, της δυναμικής του ρομπότ και του σημείου μηδενικής ροπής.

Ο μόνος τρόπος να λυθεί το πρόβλημα της πρότυπης πολυπλοκότητας του σημείου μηδενικής ροπής, ειναι να απλοποιήσουμε το σύστημα. Οι απλοποιήσεις βασίζονται στο γεγονός ότι οι περισσότερες ανθρώπινες κινήσεις, υποστηρίζουν την αρχή της συντήρησης της συνολικής γωνιακής ορμής του σώματος, στο κέντρο της μάζας του. Πολλοί ερευνητές προτείνουν την προσέγγιση ενός διπόδου ρομπότ, με ένα ανεστραμμένο εκκρεμές. Με αυτό τον τρόπο προσεγγίζουμε την πλήρη δυναμική ενός διπόδου ρομπότ με τις ακόλουθες εξισώσεις:

Δυναμικές εξισώσεις:

$$m\ddot{x}_c = f_x \tag{2.25}$$

$$m\ddot{y}_c = f_y \tag{2.26}$$

$$m(\ddot{z}_c + g) = f_z \tag{2.27}$$

περιορισμοί του σημείου μηδενικής ροπής:

$$(ZMP_x - x_c)f_z + (z_c - ZMP_z)f_x = 0 \tag{2.28}$$
$$(ZMP_y - y_c)f_z + (z_c - ZMP_z)f_y = 0 \tag{2.29}$$

Όπου το m είναι η συνολική μάζα του ρομπότ,οι $x_c, y_c, z_c$ είναι οι συντεταγμένες του κέντρου μάζας του ρομπότ, οι $f_x, f_y, f_z$ είναι οι συνιστώσες της δύναμης της επίγειας αντίδρασης, τα $ZMP_x, ZMP_y, ZMP_z$ είναι η θέση του σημείου μηδενικής ροπής στο έδαφος και το $g$ είναι η επιτάχυνση της βαρύτητας.

Από τις εξισώσεις αυτές προκύπτει η σχέση του σημείου μηδενικής ροπής με το κέντρο της μάζας:

$$ZMP_x = \frac{x_c m(\ddot{z}_c+g)-(z_c-ZMP_z)m\ddot{x}_c}{m(\ddot{z}_c+g)} = x_c - a(t)\ddot{x}_c \tag{2.30}$$
$$ZMP_y = \frac{y_c m(\ddot{z}_c+g)-(z_c-ZMP_z)m\ddot{y}_c}{m(z_c+g)} = y_c - a(t)\ddot{y}_c \tag{2.31}$$
$$a(t) = \frac{(z_c-ZMP_z)}{(\ddot{z}_c+g)} \tag{2.32}$$

Η τελευταία εξίσωση εκφράζει την φυσική ερμηνεία του $a(t)$. Αν στην εξίσωση αυτή, θεωρήσουμε ότι $\frac{\ddot{z}_c}{g} \approx 0$ φαίνεται ότι η κάθετη συνιστώσα της επιτάχυνσης του ρομπότ είναι πολύ μικρότερη από την επιτάχυνση της βαρυτήτας (στην περίπτωση της στατικής κίνησης), τότε η $a(t)$ μπορεί να θεωρηθεί σταθερή και δίνεται από τον τύπο:

$$a \approx \frac{z_c-ZMP_z}{g} = \frac{1}{\omega^2{}_c} \tag{2.33}$$

όπου $\omega_c$ είναι η φυσική συχνότητα του αντεστραμμένου εκκρεμούς.

# Κεφάλαιο 3

# ΟΙ ΑΛΓΟΡΙΘΜΟΙ ΕΝΙΣΧΥΤΙΚΗΣ ΜΑΘΗΣΗΣ

## 3.1 Γενικά

Οι αλγόριθμοι ενισχυτικής μάθησης (Reinforcement learning ( RL )) είναι μια σύγχρονη μέθοδος που αναπτύχθηκε τα τελευταία χρόνια και χρησιμοποιείται ευρέως στην ρομποτική, μιας και η κλασσικές μέθοδοι δυναμικής και στατικής ανάλυσης σε ρομπότ με πολλούς βαθμούς ελευθερίας είναι πολύ δύσκολή έως αδύνατη. Στην προσπάθεια επεξήγησης του *RL* θα χρησιμοποιήσουμε ένα απλό παράδειγμα. Έστω ένα δίποδο ρομπότ με 6 βαθμούς ελευθερίας στα δύο του πόδια. Θα ονομάσουμε τις αρθρώσεις του $q_i$ (όπου $i = 1, 2, \dots, 6$). Έστω ότι θέλουμε να ¨μάθουμε¨ το ρομπότ, χρησιμοποιώντας τους 6 αυτούς βαθμούς ελευθερίας να κάνει ένα βήμα μπρος. Κατά την κίνηση αυτή η κάθε άρθρωση εκτελεί μια συγκεκριμένη τροχιά συναρτήσει του χρόνου. Παραμετροποιώντας τις τροχιές $q_i(t)$ συναρτήσει κάποιων χαρακτηριστικών του διπόδου που μπορούμε να ελέγξουμε (π.χ. τις γωνίες $\vartheta_i$ ), μπορούμε να προσεγγίσουμε την κίνηση. Δηλαδή, ξεκινώντας από κάποιες τυχαίες γωνίες $\vartheta_1, \vartheta_2, \dots, \vartheta_6$ αρχικά υπολογίζουμε κάποιες τροχιές για τις αρθρώσεις. Αυτή είναι μια αρχική κατάσταση του συστήματος που βρίσκεται στον χώρο των παραμέτρων και στην συνέχεια θα βελτιστοποιηθεί. Η νέες παράμετροι θα υπολογιστούν εκ νέου στο τέλος του επεισοδίου, χρησιμοποιώντας την πληροφορία της κλίσης (gradient) της συνάρτησης $q_i(t)$.

Αρχικά ορίζουμε τις καταστάσεις και δράσεις του συστήματος

$$x_{k+1} \sim p\left(x_{k+1} | x_k, u_k\right)$$

όπου το $u_k \in \Re^M$ δείχνει την τρέχουσα δράση και τα $x_k, x_k + 1 \in \Re^N$ είναι οι τρέχουσες και οι επόμενες καταστάσεις. Υποθέτουμε ότι οι δράσεις του συστήματος προσεγγίζονται από μια πολιτική $u_k \sim \pi_\vartheta(u_k|x_k)$ η οποία παραμετροποιείται από κάποιες παραμέτρους $\vartheta \in \Re^k$. Η ακολουθία των δράσεων και των καταστάσεων διαμορφώνουν μία τροχιά που χαρακτηρίζεται ως $\tau = [x_{0:H}, u_{0:H}]$, όπου το $H$ δείχνει τον ορίζοντα $s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} s_2 \dots s_{H-1} \xrightarrow{a_{H-1}} s_H$. Κάθε χρονική στιγμή το σύστημα υπολογίζει μία ανταμοιβή $r(x_k, u_k)$.

Σκοπός μας είναι να βελτιστοποιήσουμε τις παραμέτρους του προβλήματος έτσι ούτως ώστε να έχουμε το βέλτιστο αποτέλεσμα στην πολιτική μας, $J(\vartheta) = E\{\sum_{K=0}^{H} a_k r_k\}$. Το $a_k$ εκφράζει το χρονικό βήμα ανεξαρτήτως των διαφόρων παραγόντων στάθμισης του προβλήματος και ισούται με $a_k = \gamma^k$ ή $a_k = 1/H$. Επιδιώκουμε οι αλλαγές στις τιμές των παραμέτρων να είναι ομαλές κάτι που μας το εξασφαλίζουν οι διάφορες μέθοδοι αλγορίθμων ενισχυτικής μάθησης που βασίζονται στην χρησιμοποίηση της κλίσης. Με βάση την μέθοδο αυτή, οι καινούριες παράμετροι υπολογίζονται από τον τυπο:

$$\vartheta_{h+1} = \vartheta_h + a_h \nabla_\vartheta J|_{\vartheta=\vartheta_h}, \tag{3.1}$$

**input**: policy parameterization $\theta_h$.
**repeat**
  generate policy variation $\Delta\theta_1$.
  estimate $\hat{J}_j \approx J(\theta_h + \Delta\theta_i) = \left\langle \sum_{k=0}^{H} a_k r_k \right\rangle$ from roll-out.
  estimate $\hat{J}_{\text{ref}}$, e.g., $\hat{J}_{\text{ref}} = J(\theta_h - \Delta\theta_i)$ from roll-out.
  compute $\Delta\hat{J}_j \approx J(\theta_h + \Delta\theta_i) - J_{\text{ref}}$.
  compute gradient $\mathbf{g}_{\text{FD}} = \left(\mathbf{\Delta\Theta}^T \mathbf{\Delta\Theta}\right)^{-1} \mathbf{\Delta\Theta}^T \mathbf{\Delta\hat{J}}$.
**until** gradient estimate $\mathbf{g}_{\text{FD}}$ converged.
**return**: gradient estimate $\mathbf{g}_{\text{FD}}$.

Σχήμα 3.1: Εκτιμητής κλίσης με τη μέθοδο των πεπερασμένων διαφορών

όπου το $a_h \in \Re^+$ συμβολίζει το ποσοστό της μάθησης και το $h \in \{0, 1, 2, \ldots\}$ συμβολίζει τον τρέχοντα αριθμό των επαναλήψεων. Στην περίπτωση που

$$\sum_{h=0}^{\infty} a_h > 0$$

και

$$\sum_{h=0}^{\infty} a_h^2 = \sigma\tau\alpha\vartheta\varepsilon\rho\alpha$$

τότε έχουμε σύγκλιση σε τουλάχιστον ένα τοπικό ελάχιστο.

Προφανώς, είναι αντιληπτό πως ένα βασικό πρόβλημα για την χρησιμοποίηση ενός αλγορίθμου ενισχυτικής μάθησης, είναι η εύρεση ενός καλού εκτιμητή για το $\nabla_{\vartheta} J|_{\vartheta=\vartheta_h}$. Οι παράμετροι εύρεσης ενός καλού εκτιμητή είναι πολλές. Σε πολλές περιπτώσεις δεν είναι δυνατόν να γνωρίζουμε κάθε τεχνική λεπτομέρεια του ρομπότ και του περιβάλλοντός του. Κάποιες άλλες φορές μπορεί να υπάρχει η ανάγκη να βρίσκουμε την κλίση από τα στοιχεία που παράγονται κατά την διάρκεια της εκτέλεσης του αλγορίθμου. Γι αυτό τον λόγο έχουν αναπτυχθεί κατά καιρούς πολλοί τρόποι εύρεσης ενός καλού εκτιμητή.

### 3.1.1 Μέθοδος πεπερασμένων διαφορών

Οι μικρές αυξήσεις των $\triangle\vartheta_i$ και οι παραλλαγές για κάθε παράμετρο της πολιτικής που εκτελείται, παράγουν τις εκτιμήσεις των $\triangle\widehat{J_j} \approx J(\vartheta_h + \triangle\vartheta_i) - J_{ref}$. Υπάρχουν πολλοί τρόποι επιλογής των $J_{ref}$, μπορούμε να θεωρήσουμε $J_{ref} = J(\vartheta_h)$ ή $J_{ref} = J(\vartheta_h - \triangle\vartheta_h)$.

Ένας άλλος τρόπος υπολογισμού της πολιτικής κλίσης είναι και μέσω γραμμικής οπισθοδρόμησης, όπως φαίνεται στον τύπο:

$$g_{FD} = (\triangle\Theta^T \triangle\Theta)^{-1} \triangle\Theta^T \triangle\widehat{J}$$

Πρέπει να σημειώσουμε ότι η $\triangle\vartheta_j$ απαιτεί καλή γνώση του συστήματος.

Σχηματικά οι εκτιμητές κλίσης με την μέθοδο των πεπερασμένων διαφορών φαίνονται στο σχήμα (3.1) .

### 3.1.2 Χρησιμοποιώντας πιθανότητες

Υποθέτουμε ότι οι τροχιές που παράγονται από το σύστημα τις οποίες συμβολίζουμε με $\tau$ είναι $\tau \sim p_{\vartheta(\tau)} = p(\tau|\vartheta)$ με ανταμοιβή $r(\tau) = \sum_{k=0}^{H} a_k r_k$. Σε αυτή την περίπτωση μπορούμε να εκτιμήσουμε την πολιτική κλίσης χρησιμοποιώντας ένα κόλπο:

$$\nabla_{\vartheta} J(\vartheta) = \int_{T} \nabla_{\vartheta} p_{\vartheta}(\tau) r(\tau) d_{\tau} = \mathrm{E}\{\nabla_{\vartheta} \log p_{\vartheta}(\tau) r(\tau)\} \tag{3.2}$$

Για να γίνει κατανοητή η εξίσωση (3.2) αρκεί να κάνουμε την εξής ανάλυση:

$$\begin{aligned} \tfrac{d}{d\vartheta} J(\vartheta) &= \tfrac{d}{d\vartheta} \int_{Y} \pi(u) r(u) du = \\ \int_{Y} \tfrac{d\pi(u)}{d\vartheta} r(u) du &= \int_{Y} \pi(u) \tfrac{1}{\pi(u)} \tfrac{d\pi(u)}{d\vartheta} r(u) du = \\ = \int_{Y} \pi(u) \tfrac{d \log \pi(u)}{d\vartheta} r(u) du &= \mathrm{E}\{\tfrac{d \log \pi(u)}{d\vartheta} r(u)\} \approx \textstyle\sum_{i=1}^{N} \tfrac{d \log \pi(u_i)}{d\vartheta} r(u_i) \end{aligned} \tag{3.3}$$

Εφόσον το

$$\begin{aligned} p_{\vartheta}(\tau) = p(x_0) \textstyle\prod_{k=0}^{H} p(x_{k+1}|x_k, u_k) \pi_{\vartheta}(u_k|x_k) \Longrightarrow \\ \Longrightarrow \nabla_{\vartheta} \log p_{\vartheta}(\tau) = \textstyle\sum_{k=0}^{H} \nabla_{\vartheta} \log \pi_{\vartheta}(u_k|x_k) \end{aligned} \tag{3.4}$$

και το ολοκλήρωμα $\int_T \nabla_{\vartheta} p_{\vartheta}(\tau) d_{\tau} = 0$ μπορούμε να εισάγουμε μία βασική σταθερά στον εκτιμητή της κλίσης. Δηλαδή

$$\nabla_{\vartheta} J(\vartheta) = E\{\nabla_{\vartheta} \log p_{\vartheta}(\tau)(r(\tau) - b)\} \tag{3.5}$$

όπου το $b \in \Re$ μπορεί να είναι αυθαίρετα επιλεγμένο, αλλά συνήθως επιλέγεται έτσι ούτως ώστε να ελαχιστοποιεί την διαφορά του εκτιμητή κλίσης. Τελικά,

$$\mathbf{Y}_{RF} = \langle (\sum_{k=0}^{H} \nabla_{\vartheta} \log \pi_{\vartheta}(u_k|x_k))(\sum_{l=0}^{H} a_l r_l - b) \rangle \tag{3.6}$$

όπου $\langle f(\tau) \rangle = \int_T f(\tau d_{\tau})$ αντιπροσωπεύουν τον μέσο όρο των τροχιών του προβλήματος.

Για να γίνει κατανοητή η χρησιμότητα της σταθεράς $b$ σε έναν αλγόριθμο ενισχυτικής μάθησης αρκεί να κατανοήσουμε ότι αν χρησιμοποιούμε έναν εκτιμητή πολιτικής κλίσης, και η σταθερά $b = 0$, σε ένα σενάριο όπου υπάρχει μόνο μία ανταμοιβή με σταθερή τιμή, δηλαδή $r(x, u) = c \in \Re$ για όλα τα x , u, τότε η διαφορά του εκτιμητή κλίσης θα αυξάνεται τουλάχιστον κυβικά με την πάροδο του χρόνου του προγραμματισμού :

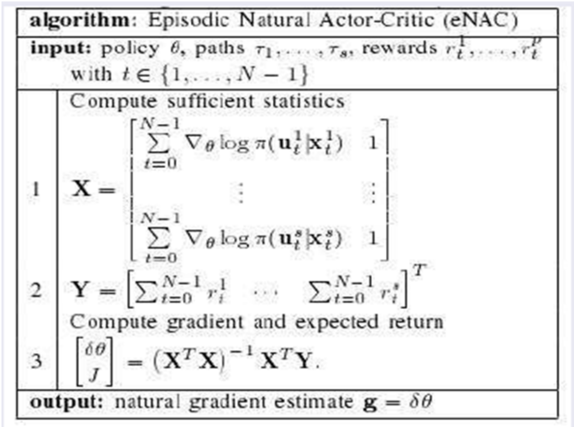
$$Var\{g_{RF}\} = H^2 c^2 \sum_{k=0}^{H} Var\{\nabla_{\vartheta} \log \pi_{\vartheta}(u_k|x_k)\} \tag{3.7}$$

εφόσον το $\{\nabla_{\vartheta} \log \pi_{\vartheta}(u_k|x_k)\} > 0$ για κάθε k. Επιπλέον, θεωρώντας ότι η ανταμοιβή reward παραμένει σταθερή, στην συγκεκριμένη περίπτωση, η μέση τιμή της κλίσης, θα παραμένει και αυτή αμετάβλητη απο το μέγεθος της ανταμοιβής, μιας και αυτή είναι σταθερά (c).

### 3.1.3 Ο φυσικός κριτικός τελεστής eNAC

Με τις μεθόδους που περιγράψαμε, ακόμα και με μία βέλτιστη φυσική γραμμή, η σύγκλιση στην βέλτιστη λύση μπορεί να είναι αργή. Αυτό συμβαίνει γιατί δεν λαμβάνεται υπόψη η πληροφορία που μπορεί να χάνεται κατά την διάρκεια της αναπροσαρμογής της πολιτικής. Κατά συνέπεια, συχνά, αντί να εκμεταλλευόμαστε την γνώση που λαμβάνουμε για την εκτίμηση της κλίσης, ελαχιστοποιούμε την εξερεύνησή της. Η λύση σε αυτό το πρόβλημα είναι η χρησιμοποίηση για την φυσική προσέγγιση της κλίσης, των αλγορίθμων που χρησιμοποιούν έναν φυσικό κριτικό τελεστή.

Η βασική ιδέα αυτού του αλγορίθμου είναι ότι η πληροφορία για τις παραμέτρους της πολιτικής $\vartheta$ που περιλαμβάνονται στις παρατηρηθείσες πορείες $\tau$, δίνονται από την πληροφορία του Fisher $F(\vartheta)$, που προσδιορίζεται ως:



**algorithm**: Episodic Natural Actor-Critic (eNAC)

**input:** policy $\theta$, paths $\tau_1, \ldots, \tau_s$, rewards $r_t^1, \ldots, r_t^p$ with $t \in \{1, \ldots, N-1\}$

| | |
|---|---|
| | Compute sufficient statistics |
| 1 | $\mathbf{X} = \begin{bmatrix} \sum_{t=0}^{N-1} \nabla_\theta \log \pi(\mathbf{u}_t^1 \mid \mathbf{x}_t^1) & 1 \\ \vdots & \vdots \\ \sum_{t=0}^{N-1} \nabla_\theta \log \pi(\mathbf{u}_t^s \mid \mathbf{x}_t^s) & 1 \end{bmatrix}$ |
| 2 | $\mathbf{Y} = \begin{bmatrix} \sum_{t=0}^{N-1} r_t^1 & \cdots & \sum_{t=0}^{N-1} r_t^s \end{bmatrix}^T$ |
| | Compute gradient and expected return |
| 3 | $\begin{bmatrix} \delta\theta \\ J \end{bmatrix} = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{Y}.$ |

**output:** natural gradient estimate $\mathbf{g} = \delta\theta$

Σχήμα 3.2: Αλγόριθμος με φυσικό κριτικό εκτιμητή eNAC

$$F(\vartheta) = E\{\nabla_\vartheta \log p(\tau|\vartheta) \nabla_\vartheta \log p(\tau|\vartheta)^T\} \tag{3.8}$$

όπου είναι η διαφορά των παραγώγων των πορειών $\nabla_\vartheta \log p(\tau|\vartheta)$. Αν αλλάξουμε την πολιτική κατά ένα $\delta\vartheta$, παρατηρούμε μια πληροφορία που χάνεται $l_\vartheta(\delta\vartheta) \approx \delta\vartheta^T F(\vartheta)\delta\vartheta$, που μπορούμε να την χαρακτηρίσουμε και ως το μέγεθος της αλλαγής στην διανομή της πορείας $p(\vartheta|\tau)$. Κατά συνέπεια, αν ερευνούμε για την αλλάγή της πολιτικής $\delta\vartheta$ που μεγιστοποιεί την αναμενόμενη ανταμοιβή $J(\vartheta + \delta\vartheta)$ για ένα σταθερό μέγεθος πληροφορίας που χάνεται $l_\vartheta(\delta\vartheta) \approx \epsilon$, μπορούμε βασικά να ψάξουμε για την μεγαλύτερη τιμή σε μία έλλειψη γύρω από το $\vartheta$.

Με έναν πιο αυστηρό φορμαλισμό, συνειδητοποιούμε ότι η κατεύθυνση της πιο απότομης ανάβασης στην έλλειψη αντιστοιχεί στην:

$$\delta\vartheta = \arg \max_{\delta\vartheta s \cdot t \cdot l(\delta\vartheta) = \epsilon} \delta\vartheta^T \nabla_\vartheta J = F^{-1}(\vartheta) \nabla_\vartheta J \tag{3.9}$$

### 3.1.4 Power RL

Ο αλγόριθμος Power είναι βασισμένος στην φιλοσοφία των αλγορίθμων επεισοδιακής ενισχυτικής μάθησης, με το πλεονέκτημα ότι χρησιμοποιεί την μέγιστη πληροφορία που μπορεί να αντλήσει από μια αλληλουχία επεισοδίων και να την χρησιμοποιήσει για την εύρεση της καλύτερης δυνατής λύσης. Η σύγκλιση του αλγορίθμου σε μία λύση είναι πολύ πιο γρήγορη από οποιονδήποτε άλλο αλγόριθμο. Ο αλγόριθμος υποθέτει ότι η πολιτική παραμετροποιείται στη μορφή:

$$u(t) = (\vartheta + \varepsilon_t)\phi(s, t) \tag{3.10}$$

όπου $\varepsilon_t$ είναι Gaussian θόρυβος που παίζει το ρόλο exploration. Ο πλήρης αλγόριθμος φαίνεται στο παρακάτω σχήμα.

**Algorithm 1** EM Policy learning by Weighting Exploration with the Returns (PoWER)

**Input:** initial policy parameters $\theta_0$
**repeat**
*Sample*: Perform trial(s) using $\mathrm{a} = (\theta + \varepsilon_t)^{\mathrm{T}} \phi(\mathrm{s}, t)$ with $[\varepsilon_t]_{ij} \sim \mathcal{N}(0, \sigma_{ij}^2)$ as stochastic policy and collect all $(t, \mathrm{s}_t, \mathrm{a}_t, \mathrm{s}_{t+1}, \varepsilon_t, r_{t+1})$ for $t = 1, 2, \ldots, T+1$.
*Estimate*: Use unbiased estimate $\hat{Q}^{\pi}(\mathrm{s}, \mathrm{a}, t) = \sum_{\tilde{t}=t}^{T} r(\mathrm{s}_{\tilde{t}}, \mathrm{a}_{\tilde{t}}, \mathrm{s}_{\tilde{t}+1}, \tilde{t})$.
*Reweight*: Compute importance weights and reweight trials, discard low-importance trials.
*Update* policy using $\theta_{k+1} = \theta_k + \left\langle \sum_{t=1}^{T} \varepsilon_t Q^{\pi}(\mathrm{s}, \mathrm{a}, t) \right\rangle_{w(\tau)} \Big/ \left\langle \sum_{t=1}^{T} Q^{\pi}(\mathrm{s}, \mathrm{a}, t) \right\rangle_{w(\tau)}$.
**until** Convergence $\theta_{k+1} \approx \theta_k$

Σχήμα 3.3: Αλγόριθμος Power

# Κεφάλαιο 4

# ΕΦΑΡΜΟΓΗ ΣΤΟ ΡΟΜΠΟΤ NAO

## 4.1 Παραμετροποίηση του συστήματος

Αυτό που μας ενδιαφέρει είναι η εύρεση και βελτίωση των τροχιών που θα διαγράψουν κάποιες αρθρώσεις του *NAO* κατά την εκτέλεση μιας κίνησης. Έτσι λοιπόν κατασκευάζουμε τις τροχιές αυτές, που τις ονομάζουμε, $q_i(t)$ , όπου $i = 0, 1, \ldots, 22$ το πλήθος των αρθρώσεων. Προσέγγιση των τροχιών αυτών μπορεί να γίνει με πολλούς τρόπους, είτε με Splines, είτε με Central Pattern Generators (CPGs). Σε μια πρώτη προσέγγιση χρησιμοποιήσαμε για την παραμετροποίηση βασικές τριγωνομετρικές συναρτήσεις της μορφής:

$$q_i(t) = a_0 + \sum_{k=1}^{K} a_k \cos(2\pi k t) + \sum_{k=1}^{K} b_k \sin(2\pi k t) \tag{4.1}$$

Το $k$ είναι ένας σταθερός αριθμός ( $k = 6$ ). Εμπειρικά υπολογίσαμε ότι μια τιμή του $k$ μεταξύ του 6 και του 10 μπορεί να προσεγγίζει τις περισσότερες τροχιές ικανοποιητικά. Κάποιες αρχικές τιμές των παραμέτρων $a_i$ και $b_i$ τις υπολογίζουμε με αντίστροφο μετασχηματισμό *Fourier*.

Η τελική παραμετροποίηση όμως που χρησιμοποιήσαμε είναι η προσέγγιση των τροχιών με συναρτήσεις Gauss.

$$q_i(t) = a_i \exp\left(-\frac{(t-m)^2}{2\sigma^2}\right) \tag{4.2}$$

Η προσέγγιση αυτή, αποδείχθηκε καλύτερη, γιατί η κλίση της είναι πιο ομαλή, με αποτέλεσμα οι τιμές των γωνιών κατά την διάρκεια της κλωτσιάς άλλαζαν τιμή πιο ομαλά και το αποτέλεσμα ήταν καλύτερο.

## 4.2 Κέντρο μάζας

Χρησιμοποιώντας την εξίσωση (2.5) και χρησιμοποιώντας τα δεδομένα που δίνονται από τους κατασκευαστές του NAO(βλ. παράρτημα A), υπολογίζουμε αρχικά, το κέντρο μάζας του στην περίπτωση που το ρομπότ βρίσκεται στην θέση όπου ολες οι γωνίες του έχουν μηδενική τιμή.

$$x_{CoM} = \frac{\sum_{i=1}^{k} m_i x_i}{\sum_{i=1}^{k} m_i} \Longrightarrow x_{CoM} = 0.0087417m \tag{4.3}$$

$$y_{CoM} = \frac{\sum_{i=1}^{k} m_i y_i}{\sum_{i=1}^{k} m_i} \Longrightarrow y_{CoM} = 0m \tag{4.4}$$

$$z_{CoM} = \frac{\sum_{i=1}^{k} m_i z_i}{\sum_{i=1}^{k} m_i} \Longrightarrow z_{CoM} = -0.062249m \tag{4.5}$$

Αυτό που χρειαζόμαστε όμως, είναι η δυνατότητα υπολογισμού του κέντρου μάζας σε οποιαδήποτε θέση και αν βρίσκεται το ρομπότ. Με άλλα λόγια δηλαδή χρειάζεται να υπολογίσουμε τις συνιστώσες $x$, $y$, $z$ του κέντρου μάζας συναρτήσει των γωνιών των αρθρώσεων του ρομπότ.

Πρέπει να σημειώσουμε ότι το κέντρο μάζας που υπολογίσαμε για την αρχική θέση του ρομπότ, είναι ως προς ένα εγωκεντρικό σύστημα αναφοράς που βρίσκεται στην μέση του. Ο υπολογισμός του κέντρου μάζας του ρομπότ ως προς αυτό το σύστημα αναφοράς δεν μας είναι ιδιαίτερα χρήσιμος για τον έλεγχο της ευστάθειας του, έτσι θα υπολογίσουμε τις συνιστώσες του κέντρου μάζας του ρομπότ μεταφέροντας την αρχή των αξόνων του συστήματος στο δεξί του πέλμα.

Δεδομένης της αρχικής θέσης του ρομπότ, που στέκεται στο δεξί του πόδι, με μια ¨εποπτική¨ προσέγγιση μπορούμε να προσεγγίσουμε το κέντρο μάζας του. Το NAO τώρα μπορούμε να πούμε πως έχει προσεγγιστεί με ένα ανάποδο εκρεμμές που η βάση του είναι στον δεξιό αστράγαλο του με την μάζα του συγκεντρωμένη στο σημείο $x_{com}, y_{com}, z_{com}$. Με αυτή την μετατροπή, ο έλεγχος της ευστάθειας του ρομπότ είναι εύκολος, μιας και όσο οι συνιστώσες του σημείου μηδενικής ροπής του ρομπότ στο επίπεδο του εδάφους βρίσκονται μέσα στα όρια του πολυγώνου στήριξης (βλέπε παράγραφο 2.2) το ρομπότ βρίσκεται σε ισορροπία. Για μία καλύτερη κατανόηση ας ανατρέξουμε στο παράρτημα C.

Το σημείο μηδενικής ροπής του NAO συναρτήσει του κέντρου μάζας του είναι:

$$ZMP_x = x_{com} - a(t)\ddot{x}_{com} \tag{4.6}$$

$$ZMP_y = y_{com} - a(t)\ddot{y}_{com} \tag{4.7}$$

όπου $a(t) = \frac{z_{com} - ZMP_z}{\ddot{z}_{com} + g} \approx 0$

Στην περίπτωση που θεωρήσουμε ότι η κίνηση γίνεται με τέτοιο τρόπο ούτως ώστε όλες οι επιταγχύνσεις του ρομπότ να θεωρούνται σχεδόν μηδέν, το σημείο μηδενικής ροπής του NAO είναι:

$$ZMP_x = x_{com} \tag{4.8}$$

$$ZMP_y = y_{com} \tag{4.9}$$

Δηλαδή το ZMP στο επίπεδο του εδάφους θεωρούμε ότι ταυτίζεται με την προβολή του κέντρου μάζας στο επίπεδο αυτό.

Το κέντρο μάζας τελικά του ρομπότ είναι:

$$x_{COM} = \frac{m_{chest}x_{chest} + m_{head}x_{head}}{\sum m_i} + \frac{m_{upperarm}x_{upperarm} + m_{lowerarm}x_{lowerarm}}{\sum m_i} + \frac{m_{thigh}x_{thigh} + m_{tibia}x_{tibia}}{\sum m_i} + \frac{m_{foot}x_{foot}}{\sum m_i} \tag{4.10}$$

$$y_{COM} = \frac{m_{chest}y_{chest} + m_{head}y_{head}}{\sum m_i} + \frac{m_{upperarm}y_{upperarm} + m_{lowerarm}y_{lowerarm}}{\sum m_i} + \frac{m_{thigh}y_{thigh} + m_{tibia}y_{tibia}}{\sum m_i} + \frac{m_{foot}y_{foot}}{\sum m_i} \tag{4.11}$$

$$z_{COM} = \frac{m_{chest}z_{chest} + m_{head}z_{head}}{\sum m_i} + \frac{m_{upperarm}z_{upperarm} + m_{lowerarm}z_{lowerarm}}{\sum m_i} + \frac{m_{thigh}z_{thigh} + m_{tibia}z_{tibia}}{\sum m_i} + \frac{m_{foot}z_{foot}}{\sum m_i} \tag{4.12}$$

## 4.3 Reinforcement learning

Δοκιμάσαμε κάποιες μεθόδους ενισχυτικής μάθησης για να δούμε που έχουμε καλύτερη σύγκλιση (βλ. κεφάλαιο 3). Ο αλγόριθμός υλοποίησης της μεθόδου έχει συνολικά 3 μέρη. Στην αρχή γίνεται η παραμετροποίηση των τροχιών των 22 αρθρώσεων του *NAO*. Στο 2 μέρος υπολογίζονται οι τροχιές των αρθρώσεων που θα βελτιώσουμε χρησιμοποιώντας κάποιες τυχαίες αρχικές παραμέτρους . Στην συνέχεια εκτελείται η κίνηση. Καθώς οι αρθρώσεις λαμβάνουν τις αντιστοιχες τιμές για την κάθε χρονική στιγμή $t \in [0, 1]$ , για κάθε χρονική στιγμή, υπολογίζεται η ανταμοιβή $r$ ( *rewards* ) της θέσης του ρομπότ. Στο τρίτο μέρος του αλγορίθμου, υπολογίζουμε τις νέες παραμέτρους $a$ και $b$ και με τις νέες αυτές παραμέτρους ξεκινά το νέο επεισόδιο. Τα επεισόδια είναι τόσα όσα χρειάζεται για να έχουμε ικανοποιητικά αποτελέσματα.

Χρησιμοποιήσαμε αρχικά έναν αλγόριθμο πολιτικής κλίσης και επειδή τα αποτελέσματα δεν ήταν ικανοποιητικά, λόγω του σφάλματος που συσσωρεύονταν με την παροδο του χρόνου της προσομοίωσης, χρησιμοποιήσαμε τελικά τον αλγόριθμο Power και είχαμε καλύτερα αποτελέσματα. Η εύρεση καλύτερης λύσης από τον τελευταίο αλγόριθμο, λογικά οφείλεται στο γεγονός του ότι ο υπολογισμός των νέων παραμέτρων γίνεται λαμβάνοντας υπόψη, τον μέσο όρο των rewards για κάποια επεισόδια (δέκα στην περίπτωση μας).

Οι γωνίες που τελικά χρησιμοποιήσαμε για να παράξουμε την κλωτσιά είναι τρεις: η *LKneePitch, ηLAnklePitchκαιηLAnkleRoll.*

### 4.3.1 Η συνάρτηση rewards

Η συνάρτηση της ανταμοιβής καθορίζεται με βάση τήν κίνηση που θέλουμε να υλοποιήσουμε. Στην δικιά μας περίπτωση ξεκινήσαμε το ρομπότ από μία θέση ισορροπίας στο ένα πόδι. Αυτό που τελικά θέλουμε να κάνει το ρομπότ είναι να κλωτσήσει την μπάλα. Προφανώς ένας όρος της συνάρτησης ανταμοιβής θα πρέπει να διασφαλίζει το γεγονός ότι το ρομπότ δεν θα πέσει. Αυτό το επιτυγχάνουμε με τον υπολογισμό σε κάθε βήμα της κίνησης των συνιστωσών στο επίπεδο του εδάφους του ZMP του ρομπότ.

Οι συνισταμένες του ZMP υπολογίζονται ως εξής:

$$ZMP_x = x_{com} - \tfrac{z_{com}-ZMP_z}{\ddot{z}_{com}+g} * \ddot{x}_{com} \tag{4.13}$$

$$ZMP_y = y_{com} - \tfrac{z_{com}-ZMP_z}{\ddot{z}_{com}+g} \ddot{y}_{com} \tag{4.14}$$

Προφανώς, η $\ddot{z}_{com} \approx 0$ μιας και η κίνηση στην κάθετη επιφάνεια θεωρείται αμελητέα Υπολογισμός των πρώτων και δεύτερων παραγώγων του κέντρου μάζας, δεδομένου ότι αλλάζουμε στο ρομπότ 1 γωνία, την $\partial_9$ :

$$x_{com} = c_1 * \cos(c_2 - q_{max} * \exp(-\partial * (t - 0.5)^2)) \tag{4.15}$$

$$\begin{aligned} \dot{x}_{com} = -c_1 * \sin(c_2 - q_{max} * \exp(-\partial * (t - 0.5)^2)) * \\ (2 * q_{max} * \partial * \exp -\partial * (t - 0.5)^2) * (t - 0.5) \end{aligned} \tag{4.16}$$

$$\begin{gathered} \ddot{x}_{com} = c_1 * \sin(c_2 - q_{max} * \exp(-\partial * (t - 0.5)^2)) * \\ 2 * q_{max} * \partial * \exp(-\partial * (t - 0.5)^2) * (2 * \partial * (t - 0.5)^2 - 1) + \\ c_1 * \cos(c_2 - q_{max} * \exp(-\partial * (t - 0.5)^2)) * \\ 2 * q_{max} * \partial * \exp(-\partial * (t - 0.5)^2) * (t - 0.5) \end{gathered} \tag{4.17}$$

Αντικαθιστώντας στις (4.46) και (4.47) τις (4.48),(4.49) και (4.50) έχουμε τις συνισταμένες $ZMP_x$, $ZMP_y$.

Στην περίπτωση που αλλάξουμε μόνο την γωνία $\vartheta_9$ η επιτάγχυνση στην $y$ συνιστώσα του κέντρου μάζας είναι μηδέν, διότι εν εξαρτάται απο την γωνία του αριστερού ποδιού $\vartheta_9$.

Στην περίπτωση που χρησιμοποιήσουμε και την γωνία του αριστερού αστραγάλου $\vartheta_{11}$, που θα την χρησιμοποιήσουμε, έχουμε επιταγχύνσεις κατά την $x$ και κατά την $y$ συνιστώσα του κέντρου μάζας, οι οποίες υπολογίζονται με παρόμοιο τρόπο.

Όσο οι συνιστώσες του ZMP βρίσκονται εντός του πολυγώνου στήριξης του NAO, που στην συγκεκριμένη περίπτωση ταυτίζεται με το πολύγωνο που σχηματίζει η επιφάνεια του ποδιού με το έδαφος,το NAO βρίσκεται σε κατάσταση ισορροπίας. Στην περίπτωση που κάποιες γωνίες του ρομπότ κάποια χρονική στιγμή πάρουν τέτοια τιμή ούτως ώστε η επιτάχυνση στον $x$ και τον $y$ άξονα μεγαλώσει αρκετά, το ρομπότ θα πέσει.

Ενας άλλος όρος της συνάρτησης ανταμοιβής θα είναι η επιθυμητή τιμή που θέλουμε να φτάσει η κάθε μία από τις 2 αρθρώσεις. Συνολικά, η συνάρτηση ανταμοιβής δίνεται από τον ακόλουθο τύπο:

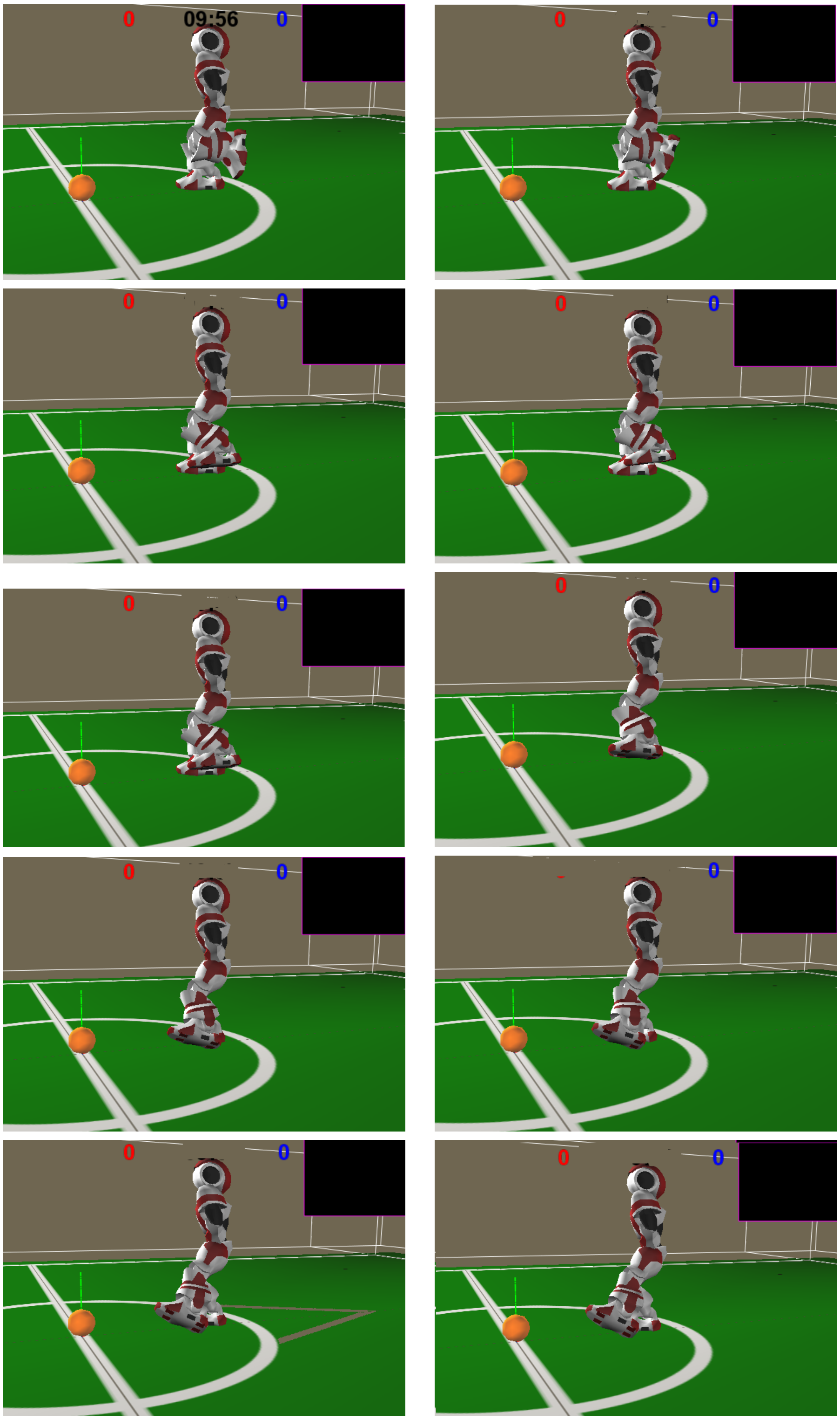
$$r = \frac{1}{3}\exp\Big[-(ZMP_x)^2\Big] + \frac{1}{3}\exp\Big[-(ZMP_y)^2\Big] + \frac{1}{3}\Big[\frac{1}{1+\exp(-\vartheta)}\Big] \tag{4.18}$$

# Κεφάλαιο 5

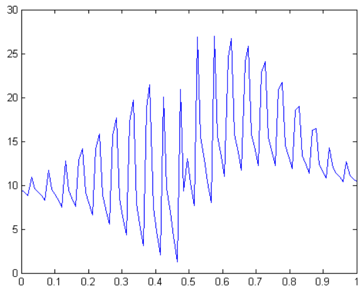
# ΑΠΟΤΕΛΕΣΜΑΤΑ

## 5.1 Τελική κίνηση

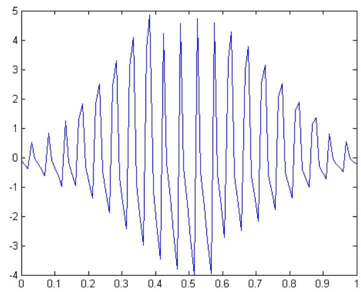
Η κίνηση που επιτύχαμε παραθέτεται στις κάτωθι εικόνες. Η παραμετροποίηση των γωνιών έγινε με συναρτήσεις Gauss της μορφής (εξίσωση 4.2). Οι γωνίες που αλλάξαμε είναι μόνο 3 γωνίες στο αριστερό πόδι. Οι αρχικές και τελικές παράμετροι $a_i$ για τις 3 γωνίες του ρομπότ φαίνονται στον πίνακα. Τα επεισόδια που χρειάστηκαν για την εύρεση της συγκεκριμένης λύσης ήταν 500. Ο αλγόριθμος που χρησιμοποιήσαμε ήταν ο Power RL. Για την εύρεση των τροχιών των αρθρώσεων για την παραγωγή της κλωτσιάς, αρχικά οι παράμετροι που ψάχναμε ήταν κάποιες μέγιστες τιμές των γωνιών χωρίς να πέσει. Στην συνέχεια αφού παράχθηκε η κλωτσιά, θελήσαμε να ελέγξουμε, ποια είναι τα όρια της διασποράς των συναρτήσεων Gauss ούτως ώστε να μην πέφτει το NAO. Η διασπορά σε μία συνάρτηση Gauss όσο αυξάνεται, τόσο η κλίση της τροχιάς μικραίνει, ενώ όσο η διασπορά μειώνεται η κλίση της συνάρτησης γίνεται πιο απότομη. Αυτό σημαίνει πως όσο μικρότερη διασπορά έχουμε, η κίνηση θα γίνεται με μεγαλύτερη επιτάγχυνση. Από τα αποτελέσματα του αλγορίθμου φαίνεται πως η διασπορά αν είναι μικρότερη του $\sigma^2 = 0.064$ το ρομπότ αναπτύσσει τέτοιες ροπές που πέφτει. Η πιο γρήγορη υλοποίηση της κλωτσιάς γίνεται για διασπορά $\sigma^2 = 0.064$ (που αντιστοιχεί σε $\partial = 7.81$. Προφανώς, όσο αυξάνεται η διασπορά, το ρομπότ εκτελεί την κίνηση όλο και πιο αργά.




Σχήμα 5.1: Κίνηση του ρομπότ




Σχήμα 5.2: η συνισταμένη $x$ του ZMP σε $mm$ κατά την διάρκεια ενός επεισοδίου

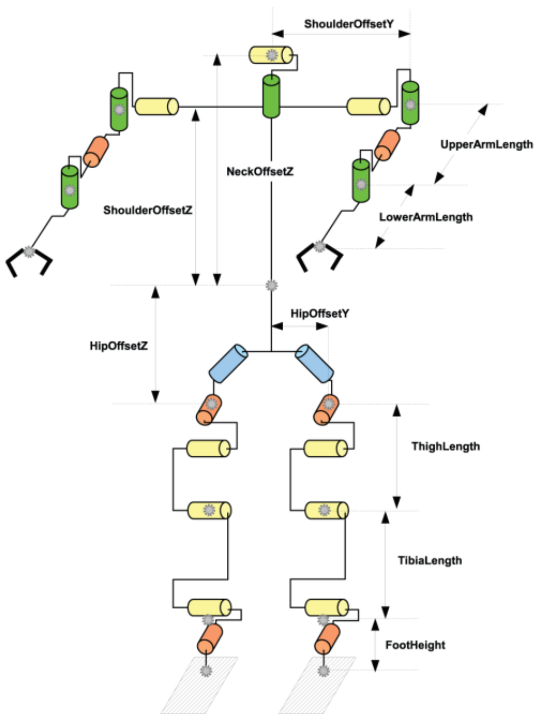



Σχήμα 5.3: η συνισταμένη $y$ του ZMP σε $mm$ κατά την διάρκεια ενός επεισοδίου


Σχήμα 5.4: Ανταμοιβή ως συνάρτηση της παραμέτρου $\vartheta$, όπου $\vartheta = \frac{1}{2\sigma^2}$

# Παράρτημα Α΄

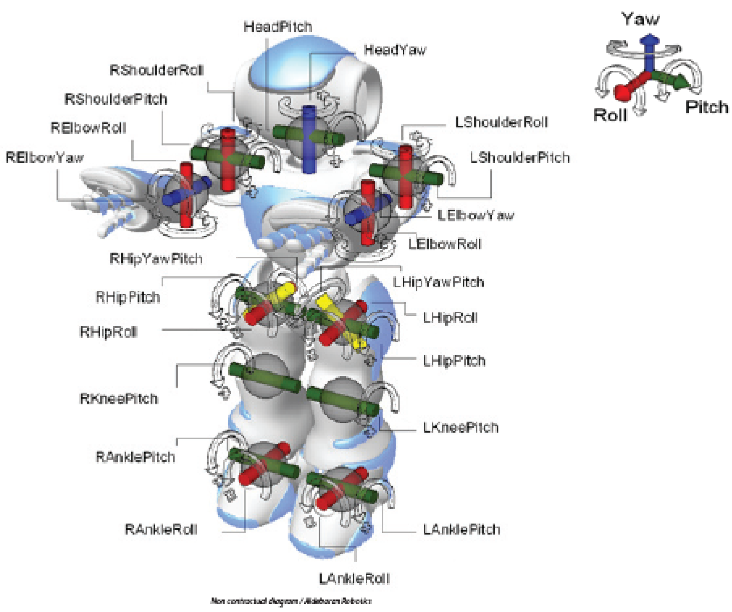
# ΤΕΧΝΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ NAO




Σχήμα Α΄.1: Η σχηματική δομή του NAO

| Table Links length | | | |
|---|---|---|---|
| **Name** | **length (mm)** | **Name** | **length (mm)** |
| **NeckOffsetZ** | 126,5 | **HipOffsetZ** | 85 |
| **ShoulderOffsetY** | 98 | **HipOffsetY** | 50 |
| **UpperArmLength** | 90 | **Thigh Lenght** | 100 |
| **LowerArmLength** | 135 | **Tibia Lenght** | 100 |
| **ShoulderOffsetZ** | 100 | **Foot Height** | 46 |

Σχήμα Α΄.2: Πίνακας με τα μήκη των τμημάτων του ΝΑΟ




Σχήμα Α΄.3: Joints directions

**Motion range**

| PART | JOINT NAME | MOTION | RANGE (degrees) |
|---|---|---|---|
| Head | HeadYaw | Head joint twist (Z) | -120 to 120 |
| | HeadPitch | Head joint front & back (Y) | -45 to 45 |
| Left arm | LShoulderPitch | Left shoulder joint front & back (Y) | -120 to 120 |
| | LShoulderRoll | Left shoulder joint right & left (Z) | 0 to 95 |
| | LElbowRoll | Left shoulder joint twist (X) | -120 to 120 |
| | LElbowYaw | Left elbow joint (Z) | 0 to 90 |
| Left leg | LHipYawPitch | Left hip joint twist (Z45°) | -90 to 0 |
| | LHipPitch | Left hip joint front & back (Y) | -100 to 25 |
| | LHipRoll | Left hip joint right and left (X) | -25 to 45 |
| | LKneePitch | Left knee joint (Y) | 0 to 130 |
| | LAnklePitch | Left ankle joint front & back (Y) | -75 to 45 |
| | LAnkleRoll | Left ankle joint right & left (X) | -45 to 25 |
| Right leg | RHipYawPitch | Right hip joint twist (Z45°) | -90 to 0 |
| | RHipPitch | Right hip joint front and back (Y) | -100 to 25 |
| | RHipRoll | Right hip joint right & left (X) | -45 to 25 |
| | RKneePitch | Right knee joint (Y) | 0 to 130 |
| | RAnklePitch | Right ankle joint front & back (Y) | -75 to 45 |
| | RAnkleRoll | Right ankle right & left (X) | -25 to 45 |
| Right arm | RShoulderPitch | Right shoulder joint front & back (Y) | -120 to 120 |
| | RShoulderRoll | Right shoulder joint right & left (Z) | -95 to 0 |
| | RElbowRoll | Right shoulder joint twist (X) | -120 to 120 |
| | RElbowYaw | Right elbow joint (Z) | -90 to 0 |

Σχήμα Α΄.4: Γωνίες του NAO και τα όρια τους. Αριθμούνται απο το 0 μεχρι το 22

# Παράρτημα Β′

# ΚΕΝΤΡΟ ΜΑΖΑΣ

Οι συνιστώσες $x_i$, $y_i$ του κάθε τμήματος του ρομπότ που βρίσκονται οι επιμέρους μάζες του είναι:

$$y_{Rthigh} = (Tibialenght + thighlenght) * \sin \partial_{17} \tag{B'.1}$$

$$\begin{gathered} x_{Rthigh} = Tibialenght * \cos(\tfrac{\pi}{2} - \partial_{16}) + \\ Thighlengh * \cos(\tfrac{\pi}{2} - \partial_{15}) \end{gathered} \tag{B'.2}$$

$$y_{Rtibia} = tibialenght * \sin \partial_{17} \tag{B'.3}$$

$$x_{Rtibia} = Tibialenght * \cos(\tfrac{\pi}{2} - \partial_{16}) \tag{B'.4}$$

$$\begin{gathered} y_{Lfoot} = (Tibialenght + Thighlenght) * \sin \partial_{17} + \\ 2 * HipoffsetY * \cos(\partial_{17} + \partial_{14}) + \\ (Tibialenght + Thighlenght) * \cos(\tfrac{\pi}{2} - \partial_8 - \partial_{17} - \partial_{14}) + \\ 0.016 \cos(\partial_{11} + \partial_8 + \partial_{17} + \partial_{14} - \tfrac{\pi}{2}) \end{gathered} \tag{B'.5}$$

$$\begin{gathered} x_{Lfoot} = Tibialenght * \cos(\tfrac{\pi}{2} - \partial_{16}) + \\ Thighlengh * \cos(\tfrac{\pi}{2} - \partial_{15} - \partial_{16}) + \\ Thighlengh * \cos(\tfrac{\pi}{2} - \partial_{15} - \partial_{16} - \partial_7 - \partial_{13}) + \\ Tibialenght * \cos(\tfrac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_7 - \partial_{13} - \partial_9) + \\ 0.016 * \cos(\tfrac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_7 - \partial_{13} - \partial_9 - \partial_{10}) \end{gathered} \tag{B'.6}$$

$$x_{Ltibia} = Tibialenght * \cos\left(\frac{\pi}{2} - \partial_{16}\right) + \\ Thighlengh * \cos\left(\frac{\pi}{2} - \partial_{15} - \partial_{16}\right) + \\ Thighlengh * \cos\left(\frac{\pi}{2} - \partial_{15} - \partial_{16} - \partial_{7} - \partial_{13}\right) \tag{Β′.7}$$

$$y_{Ltibia} = (Tibialenght + Thighlenght) * \sin \partial_{17} + \\ 2 * HipoffsetY * \cos\left(\partial_{17} + \partial_{14}\right) + \\ Thighlenght * \cos\left(\frac{\pi}{2} - \partial_{8} - \partial_{17} - \partial_{14}\right) \tag{Β′.8}$$

$$x_{Lthigh} = Tibialenght * \cos\left(\frac{\pi}{2} - \partial_{16}\right) + \\ Thighlengh * \cos\left(\frac{\pi}{2} - \partial_{15} - \partial_{16}\right) \tag{Β′.9}$$

$$y_{Ltibia} = (Tibialenght + Thighlenght) * \sin \partial_{17} + \\ 2 * HipoffsetY * \cos\left(\partial_{17} + \partial_{14}\right) \tag{Β′.10}$$

$$y_{Chest} = (Tibialenght + Thighlenght) * \sin \partial_{17} + \\ HipoffsetY * \cos\left(\partial_{17} + \partial_{14}\right) + \\ Hipoffsetz * \cos\left(\frac{\pi}{2} - \partial_{17} - \partial_{14}\right) \tag{Β′.11}$$

$$x_{Chest} = Tibialenght * \cos\left(\frac{\pi}{2} - \partial_{16}\right) + \\ Thighlengh * \cos\left(\frac{\pi}{2} - \partial_{15} - \partial_{16}\right) + \\ Hipoffsetz * \cos\left(\frac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_{13}\right) \tag{Β′.12}$$

$$y_{Head} = (Tibialenght + Thighlenght) * \sin \partial_{17} + \\ HipoffsetY * \cos\left(\partial_{17} + \partial_{14}\right) + \\ (Hipoffsetz + Neckoffsetz) * \cos\left(\frac{\pi}{2} - \partial_{17} - \partial_{14}\right) \tag{Β′.13}$$

$$x_{Head} = Tibialenght * \cos\left(\frac{\pi}{2} - \partial_{16}\right) + \\ Thighlengh * \cos\left(\frac{\pi}{2} - \partial_{15} - \partial_{16}\right) + \\ (Hipoffsetz + Neckoffsetz) * \cos\left(\frac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_{13}\right) \tag{Β′.14}$$

$$x_{Rlowerarm} = Tibialenght * \cos(\frac{\pi}{2} - \partial_{16}) +$$
$$Thighlengh * \cos(\frac{\pi}{2} - \partial_{15} - \partial_{16}) +$$
$$(Hipoffsetz + shoulderoffsetz) * \cos(\frac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_{13}) +$$
$$(upperarmlengh * \cos(\partial_{19}) + lowerarmlengh * \cos(\partial_{19} + \partial_{21})) *$$
$$\sin(\frac{\pi}{2} - \partial_{16} - theta_{15} - \partial_{13} - \partial_{18}) \tag{B'.15}$$

$$y_{Rlowerarm} = (Tibialenght + thighlengh) * \sin\partial_{17} +$$
$$HipoffsetY * \cos(\partial_{17} + \partial_{14}) +$$
$$shoulderoffsetz * \cos(\frac{\pi}{2} - \partial_{17} - \partial_{14}) + upperarm * \sin\partial_{19} +$$
$$lowerarmlenght * \sin(\partial_{21} + \partial_{19}) \tag{B'.16}$$

$$x_{Rupperarm} = Tibialenght * \cos(\frac{\pi}{2} - \partial_{16}) +$$
$$Thighlengh * \cos(\frac{\pi}{2} - \partial_{15} - \partial_{16}) +$$
$$(Hipoffsetz + shoulderoffsetz) * \cos(\frac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_{13}) \tag{B'.17}$$

$$y_{Rupperarm} = (Tibialenght + thighlengh) * \sin\partial_{17} +$$
$$HipoffsetY * \cos(\partial_{17} + \partial_{14}) +$$
$$shoulderoffsetz * \cos(\frac{\pi}{2} - \partial_{17} - \partial_{14}) +$$
$$upperarm * \sin\partial_{19} \tag{B'.18}$$

$$x_{Lupperarm} = Tibialenght * \cos(\frac{\pi}{2} - \partial_{16}) +$$
$$Thighlengh * \cos(\frac{\pi}{2} - \partial_{15} - \partial_{16}) +$$
$$(Hipoffsetz + shoulderoffsetz) * \cos(\frac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_{13}) \tag{B'.19}$$

$$y_{Lupperarm} = (Tibialenght + thighlengh) * \sin\partial_{17} +$$
$$HipoffsetY * \cos(\partial_{17} + \partial_{14}) +$$
$$shoulderoffsetz * \cos(\frac{\pi}{2} - \partial_{17} - \partial_{14}) +$$
$$upperarm * \sin(\partial_{3}) \tag{B'.20}$$

$$x_{Llowerarm} = Tibialenght * \cos(\frac{\pi}{2} - \partial_{16}) +$$
$$Thighlengh * \cos(\frac{\pi}{2} - \partial_{15} - \partial_{16}) +$$

$$(Hipoffsetz + shoulderoffsetz) * \cos\left(\frac{\pi}{2} - \partial_{16} - \partial_{15} - \partial_{13}\right) +$$
$$(upperarmlengh * \cos\partial_3 +$$
$$lowerarmlengh * \cos(\partial_3 + \partial_5)) *$$
$$\sin\frac{\pi}{2} - \partial_{16} - theta_{15} - \partial_{13} - \partial_2 \tag{Β΄.21}$$

$$y_{Llowerarm} = (Tibialenght + thighlengh) * \sin\partial_{17} +$$
$$HipoffsetY * \cos(\partial_{17} + \partial_{14}) +$$
$$shoulderoffsetz * \cos\left(\frac{\pi}{2} - \partial_{17} - \partial_{14}\right) +$$
$$upperarm * \sin(\partial_3) +$$
$$lowerarmlenght * \sin(\partial_5 + \partial_3) \tag{Β΄.22}$$

όπου οι γωνίες $\partial_0, \partial_1, \ldots, \partial_{21}$ αντιστοιχούν στις γωνίες που αναγράφονται αναλυτικά στο παράρτημα Α.

Οι συνιστώσες $z_i$ υπολογίζονται με αντίστοιχο τρόπο και είναι οι εξής:

$$z_{Lfoot} = \{Tibialenght * \cos(\partial_{16}) +$$
$$Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) +$$
$$2 * HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) +$$
$$Thightlenght * \cos(\partial_{17} + \partial_{14} + \partial_{18}) * \cos(\partial_7 + \partial_{13} + \partial_{15} + \partial_{16}) +$$
$$Tibialengh * \cos(\partial_7 + \partial_{13} + \partial_{15} + \partial_{16} + \partial_9)$$
$$0.016 * \cos(\partial_{17} + \partial_{14} + \partial_8 + \partial_{11}) * \cos(\partial_{15} + \partial_{16} + \partial_9 + \partial_{13} + \partial_7 + \partial_{10}) \tag{Β΄.23}$$

$$z_{Ltibia} = \{Tibialenght * \cos(\partial_{16}) +$$
$$Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) +$$
$$2 * HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) +$$
$$Thightlenght * \cos(\partial_{17} + \partial_{14} + \partial_{18}) * \cos(\partial_7 + \partial_{13} + \partial_{15} + \partial_{16}) \tag{Β΄.24}$$

$$z_{Lthight} = \{Tibialenght * \cos(\partial_{16}) +$$
$$Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) +$$
$$2 * HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) \tag{Β΄.25}$$

$$z_{Rtibia} = Tibialenght * \cos(\partial_{17}) * \cos(\partial_{16}) \tag{Β΄.26}$$

$$z_{Rthigh} = \{Tibialenght * \cos(\partial_{16}) +$$
$$Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) \tag{Β΄.27}$$

$$
\begin{gathered}
z_{Chest} = \{Tibialenght * \cos(\partial_{16}) + \\
Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) + \\
HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) + \\
\sin(\tfrac{\pi}{2} - \partial_2 - \partial_1) * Hipoffsetz
\end{gathered}
\tag{B'.28}
$$

$$
\begin{gathered}
z_{Head} = \{Tibialenght * \cos(\partial_{16}) + \\
Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) + \\
HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) + \\
\sin(\tfrac{\pi}{2} - \partial_2 - \partial_1) * (Hipoffsetz + Neckoffsetz)
\end{gathered}
\tag{B'.29}
$$

$$
\begin{gathered}
z_{Rlowerarm} = \{Tibialenght * \cos(\partial_{16}) + \\
Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) + \\
HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) + \\
\sin(\tfrac{\pi}{2} - \partial_{18} - \partial_1) * (Hipoffsetz + shoulderoffsetz) + \\
\{Upperarm * \cos(\partial_{19}) + Lowerarm * \cos(\partial_{19} + \partial_{21})\} * \cos\partial_{18}
\end{gathered}
\tag{B'.30}
$$

$$
\begin{gathered}
z_{Rupperarm} = \{Tibialenght * \cos(\partial_{16}) + \\
Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) + \\
HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) + \\
\sin(\tfrac{\pi}{2} - \partial_{18} - \partial_1) * (Hipoffsetz + shoulderoffsetz) + \\
Upperarm * \cos(\partial_{19}) * \cos\partial_{18}
\end{gathered}
\tag{B'.31}
$$

$$
\begin{gathered}
z_{Llowerarm} = \{Tibialenght * \cos(\partial_{16}) + \\
Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) + \\
HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) + \\
\sin(\tfrac{\pi}{2} - \partial_2 - \partial_1) * (Hipoffsetz + shoulderoffsetz) + \\
\{Upperarm * \cos(\partial_3) + Lowerarm * \cos(\partial_3 + \partial_5)\} * \cos\partial_2
\end{gathered}
\tag{B'.32}
$$

$$
\begin{gathered}
z_{Lupperarm} = \{Tibialenght * \cos(\partial_{16}) + \\
Thighlengh * \cos(\partial_{15} + \partial_{16})\} * \cos(\partial_{17}) + \\
HipoffsetY * \cos(\partial_{13}) * \cos(\partial_{14} + \partial_{17}) + \\
\sin(\tfrac{\pi}{2} - \partial_2 - \partial_1) * (Hipoffsetz + shoulderoffsetz) + \\
Upperarm * \cos(\partial_3) * \cos\partial_2
\end{gathered}
\tag{B'.33}
$$

# Παράρτημα Γ′

## ΣΧΕΣΗ ZMP ΚΑΙ ΚΕΝΤΡΟΥ ΜΑΖΑΣ

Δυναμικές εξισώσεις συστήματος:

$$f_x = m\ddot{x}_c \tag{Γ.1}$$

$$f_y = m\ddot{y}_c \tag{Γ.2}$$

$$f_z = m(\ddot{z}_c + g) \tag{Γ.3}$$

συνήθως $\ddot{z}_c \approx 0$ μιας και θεωρούμε την κίνηση στην κάθετη επιφάνεια αμελητέα.

$y'y$ άξονας:

$$\begin{aligned} -x_c f_z + ZMP_x R_z - R_x ZMP_z + f_x z_c = 0 \Longrightarrow \\ f_z(-x_c + ZMP_x) + f_x(-ZMP_z + z_c) = 0 \end{aligned} \tag{Γ.4}$$

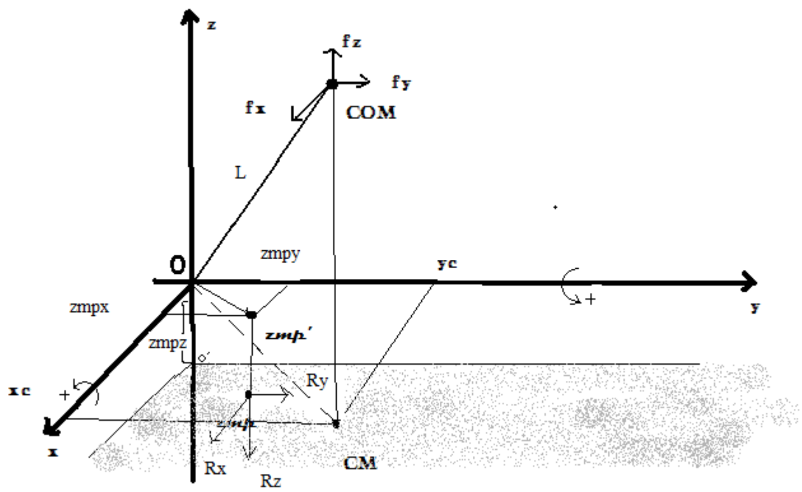
$x'x$ άξονας:

$$\begin{aligned} -f_z y_c + f_z ZMP_y + z_c f_y - R_y ZMP_z = 0 \Longrightarrow \\ (ZMP_y - y_c)f_z + (z_c - ZMP_z)f_y = 0 \end{aligned} \tag{Γ.5}$$

Συνδυάζοντας τις εξισώσεις Γ.1.....Γ.5 έχουμε:

$$ZMP_x = \frac{x_c m(\ddot{z}_c + g) - (z_c - ZMP_z)m\ddot{x}_c}{m(\ddot{z}_c + g)} \tag{Γ.6}$$

$$ZMP_y = \frac{y_c m(\ddot{z}_c + g) - (z_c - ZMP_z)m\ddot{y}_c}{m(\ddot{z}_c + g)} \tag{Γ.7}$$



ή

$$ZMP_x = x_c - \alpha(t)\ddot{x}_c \tag{Γ΄.8}$$
$$ZMP_y = y_c - \alpha(t)\ddot{y}_c \tag{Γ΄.9}$$

όπου $\alpha(t) = \frac{z_c - ZMP_z}{\ddot{z}_c + g}$

# Βιβλιογραφία


[1]. Vucobratovic, M and Borovac, B.,"ZERO- MOMENT POINT - THIRTY FIVE YEARS OF ITS LIFE,"international journal of humanoid robotics 1(1),157-173.

[2]. Goswami,A., 1999, 'Postural stability of biped robots and the foot rotation indicator (FRI) point,' international journal of robotic research 18960,523- 533.

[3]. Hoffman, A.,2006. "Robust Execution of Bipedal walking tasks from Biomechanical Principles"

[4]. Benjamin Stephens, " Humanoid Push Recovery"

[5]. Caballero R, Manuel A., Armada and Pedro Alarcon, "Methodology for Zero-moment Point Experimental modeling in the frequency domain," journal of vibration and control 2006, 12, 1385

[6]. Goswami,A., 2006, " Capture Point: A step toward Humanoid Push Recovery"

[7]. Morimoto,J., Atkenson, C. 2007, " Learning Biped Locomotion"

[8]. Peters,J., 2006, " Policy Gradient Methods for Robotics"

[9]. "*http : //en.wikipedia.org.wiki.Invertedpendulum.*"