
Introduction to Robotics

Julien Carpentier, Stéphane Caron, Yann de Mont-Marin

Class notes by Antoine Groudiev



Last modified 24th November 2024

Contents

1	Introduction	3
2	Position and Orientation	4
2.1	Introduction	4
2.2	Points, frames and transformations	5
2.2.1	Position of a point in space	5
2.2.2	Position and orientation of a body in space	5
2.2.3	Rotation matrices	6
2.3	Rotation representations	6
2.3.1	Euler angles	6
2.3.2	Axis-angle and quaternions	7
2.4	Angular velocity of rotation matrices	7
2.5	Exponential and logarithm map	8
2.6	Rigid Body Transformations	9
2.6.1	Definition	9
2.6.2	Homogeneous transformations	9
2.6.3	Linear and angular velocities	10
3	Forward Kinematics	11
3.1	Articulations and joint speed	11
3.1.1	Joint types and degrees of freedom	11
3.1.2	Joint speed	12
3.2	Direct geometry	12
3.2.1	Transformation matrix	12
3.2.2	Kinematic Jacobian	13
4	Inverse Kinematics	14
4.1	Optimization problem and resolution	14
4.2	Trajectory tracking	15
4.3	Exploiting the redundancy of solutions	16
4.3.1	Relaxed approach	16
4.3.2	Strict approach	16
5	Direct and Inverse Dynamics	17
5.1	Physical equations	17
5.2	Contact forces	18
5.3	Soft contact	18
5.4	Rigid bilateral contacts	19
5.4.1	Optimization problem	19
5.4.2	Solution using the KKT conditions	19
5.4.3	Sparse Cholesky factorization	20
5.4.4	The maximum dissipation principle	21
5.5	Rigid unilateral contacts	21
5.5.1	Unilateral contact as a Nonlinear Complementarity Problem	21
5.5.2	Relaxed contact problem	22
5.6	Inverse dynamics	22
6	Motion Planning	23
6.1	Configuration space	23

6.1.1	Definitions	23
6.1.2	Joints	23
6.1.3	Examples	23
6.2	Obstacles and collisions	24
6.2.1	Representations and modelisation	25
6.2.2	Example: translation without rotation of a rigid body	25
6.3	Motion planning algorithms	26
6.3.1	Algorithm paradigms	26
6.3.2	Combinatorial planning	27
6.3.3	Trajectory optimization (local planning)	28
6.4	Sampling-based motion planning (probabilistic planning)	28
6.4.1	Principle	28
6.4.2	Sampling	28
6.4.3	Collision detection	29
6.4.4	Probabilistic roadmap	29
6.4.5	Tree	29
7	Collision Detection	31
7.1	The broad phase	31
7.1.1	Bounding volumes	31
7.1.2	Example of bounding volumes	31
7.1.3	Dynamic tree representation	32
7.1.4	Sweep and Prune algorithm	32
7.2	The narrow phase	33
7.2.1	Introduction	33
7.2.2	Minkowski-based algorithms	33
7.2.3	Frank-Wolfe algorithm	34
7.2.4	GJK algorithm	34
8	Optimal Control	35

Abstract

This document is Antoine Groudiev's class notes while following the class *Motion planning in robotics and graphical animation* (Planification de mouvement en robotique et en animation graphique) at the Computer Science Department of ENS Ulm. It is freely inspired by the lectures of Justin Carpentier, Stéphane Caron, and Yann de Mont-Marin.

1 Introduction

2 Position and Orientation

Building algorithms for robots requires a good understanding of the position and orientation of the robot in space. This chapter will introduce the mathematical tools to describe these concepts.

2.1 Introduction

Kinematics studies the *movement* of an object – in our case of a robot – without taking into account the *forces* generating it. Instead, it only handles aspects such as position, orientation, speed and momentum of bodies in movement.

Consider for instance a robotic arm. We can design a simplified scheme of the robot and its environment, to create a kinematic pipeline and reference frames associated to each of these objects.

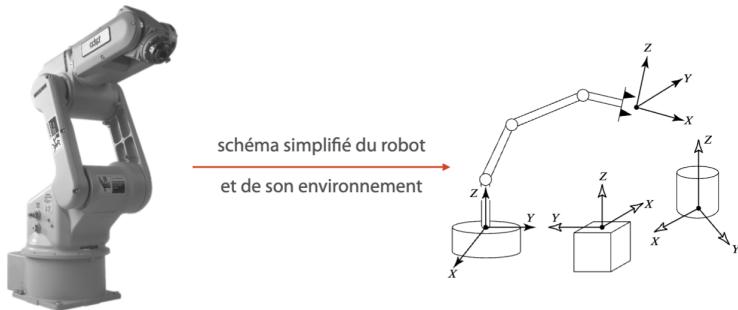
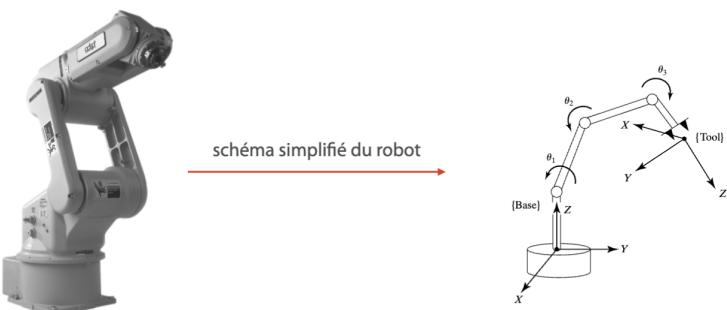


Figure 2.1: Simplified scheme of the robot, the kinematic pipeline and reference frames.

Direct kinematics allows to compute the position and orientation of the terminal organ given, for instance, the angles of the articulations.



Invert kinematics answers the question the other way around: given the position and orientation of a body, how can we compute the values of the articulations angles. Invert kinematics is used for instance for trajectory tracking: given a reference trajectory, how can we compute the speed of the articulations?

2.2 Points, frames and transformations

2.2.1 Position of a point in space

Once that a reference frame $\{A\}$ is defined, we can localize any point of the universe given a *position vector*:

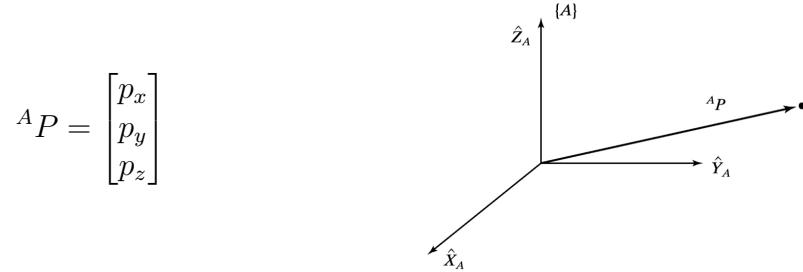


Figure 2.2: Vector and position of the point established in the frame $\{A\}$.

2.2.2 Position and orientation of a body in space

To define the orientation of a body in space, we need to define a reference frame $\{B\}$ attached to this body. The orientation is therefore defined as the expression of this coordinate system in the reference frame $\{A\}$.

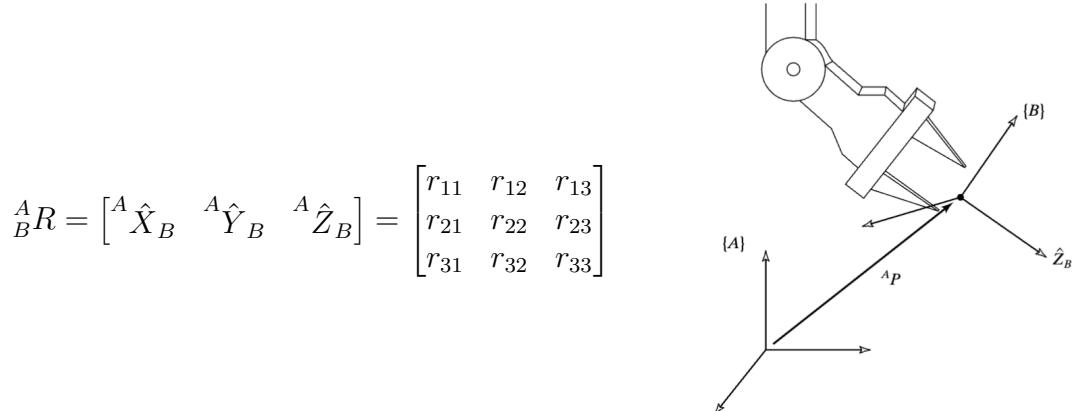


Figure 2.3: Expression of the coordinate system $\{B\}$ in the reference frame $\{A\}$, and the associated position and orientation of the body.

Each element of the matrix ${}^A R$ is the scalar product between the vectors of the two coordinate systems $\{B\}$ and $\{A\}$:

$${}^A R = \begin{bmatrix} {}^A \hat{X}_B & {}^A \hat{Y}_B & {}^A \hat{Z}_B \end{bmatrix} = \begin{bmatrix} \hat{X}_B \cdot \hat{X}_A & \hat{Y}_B \cdot \hat{X}_A & \hat{Z}_B \cdot \hat{X}_A \\ \hat{X}_B \cdot \hat{Y}_A & \hat{Y}_B \cdot \hat{Y}_A & \hat{Z}_B \cdot \hat{Y}_A \\ \hat{X}_B \cdot \hat{Z}_A & \hat{Y}_B \cdot \hat{Z}_A & \hat{Z}_B \cdot \hat{Z}_A \end{bmatrix}$$

The lines correspond to the axes of the frame $\{A\}$ expressed in the frame $\{B\}$. Note that the invert of a rotation matrix is its transpose:

$${}^A R^T = {}^A R^{-1} = {}^B R$$

2.2.3 Rotation matrices

We showed that the orientation of the body in space could be expressed as a 3-dimensional rotation matrix. The group of 3-dimensional rotation matrices is denoted $\text{SO}(3)$, and is composed of all the matrices that are orthonormal, that is orthogonal and with a determinant equal to 1:

$$\text{SO}(3) = \left\{ R \in \mathcal{M}_3(\mathbb{R}) \mid RR^T = I_3 \text{ and } \det(R) = +1 \right\} \quad (2.2.1)$$

If we write:

$$R = [\hat{X} \ \hat{Y} \ \hat{Z}]$$

Then $R \in \text{SO}(3)$ if and only if:

$$\begin{cases} \hat{X} \cdot \hat{Y} = 0 \\ \hat{Y} \cdot \hat{Z} = 0 \\ \hat{Z} \cdot \hat{X} = 0 \end{cases} \quad \text{and} \quad \begin{cases} \hat{X} \cdot \hat{X} = 1 \\ \hat{Y} \cdot \hat{Y} = 1 \\ \hat{Z} \cdot \hat{Z} = 1 \end{cases}$$

Note that we have 9 degrees of freedom and 6 independent constraints, so the group $\text{SO}(3)$ is 3-dimensional.

2.3 Rotation representations

There are several ways to represent a rotation matrix, each with its own advantages and drawbacks. The most common representations are:

- Orthonormal 3 by 3 matrices — 9 components
- Euler angles — 3 components
- Axis-angle representation — 3 components
- Quaternions — 4 components

The number of components used to be an important factor when computers were slow and memory was expensive. Nowadays, the choice of representation is more about the ease of use and the properties of the representation.

2.3.1 Euler angles

Each rotation can be represented by the composition of three elementary rotations around the fixed axes of the frame $\{A\}$.

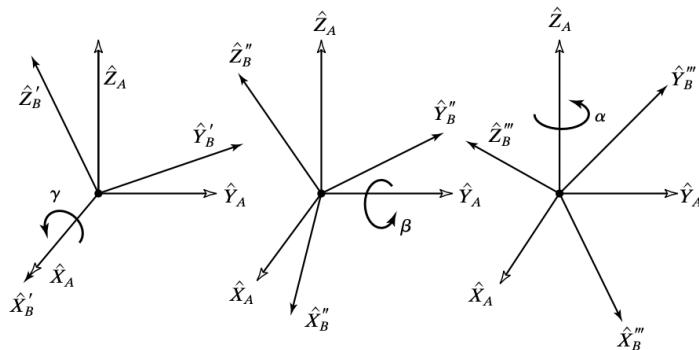


Figure 2.4: Euler angles representation of a rotation.

In terms of rotation matrices, any rotation matrix can be written as ${}^A_B R(\gamma, \beta, \alpha)$, defined by:

$$\begin{aligned} {}^A_B R(\gamma, \beta, \alpha) &= R_Z(\alpha) R_Y(\beta) R_X(\gamma) \\ &= \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \gamma & -\sin \gamma \\ 0 & \sin \gamma & \cos \gamma \end{bmatrix} \end{aligned}$$

The Euler angles are not unique, as the same rotation can be represented by different sets of Euler angles. This is called gimbal lock, and is a major drawback of the Euler angles representation.

We can also express the Euler angles given the rotation matrix:

$${}^A_B R(\gamma, \beta, \alpha) = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \implies \begin{cases} \beta = \text{atan2}(-r_{31}, \sqrt{r_{11}^2 + r_{21}^2}) \\ \alpha = \text{atan2}(r_{21}/\cos \beta, r_{11}/\cos \beta) \\ \gamma = \text{atan2}(r_{32}/\cos \beta, r_{33}/\cos \beta) \end{cases}$$

2.3.2 Axis-angle and quaternions

We are given a vector \vec{k} and an angle θ ; the rotation represented by these two elements is the rotation of angle θ around the axis \vec{k} . We can define:

$$\begin{cases} \varepsilon_1 = k_x \sin \frac{\theta}{2} \\ \varepsilon_2 = k_y \sin \frac{\theta}{2} \\ \varepsilon_3 = k_z \sin \frac{\theta}{2} \\ \varepsilon_4 = \cos \frac{\theta}{2} \end{cases}$$

we have $\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 + \varepsilon_4^2 = 1$, creating a unit quaternion. The rotation matrix associated to the quaternion is:

$$R_\varepsilon = \begin{bmatrix} 1 - 2\varepsilon_2^2 - 2\varepsilon_3^2 & 2(\varepsilon_1\varepsilon_2 - \varepsilon_3\varepsilon_4) & 2(\varepsilon_1\varepsilon_3 + \varepsilon_2\varepsilon_4) \\ 2(\varepsilon_1\varepsilon_2 + \varepsilon_3\varepsilon_4) & 1 - 2\varepsilon_1^2 - 2\varepsilon_3^2 & 2(\varepsilon_2\varepsilon_3 - \varepsilon_1\varepsilon_4) \\ 2(\varepsilon_1\varepsilon_3 - \varepsilon_2\varepsilon_4) & 2(\varepsilon_2\varepsilon_3 + \varepsilon_1\varepsilon_4) & 1 - 2\varepsilon_1^2 - 2\varepsilon_2^2 \end{bmatrix}$$

The invert operation is also simple to compute:

$${}^A_B R(\gamma, \beta, \alpha) = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \implies \begin{cases} \varepsilon_1 = \frac{r_{32} - r_{23}}{4\varepsilon_4} \\ \varepsilon_2 = \frac{r_{13} - r_{31}}{4\varepsilon_4} \\ \varepsilon_3 = \frac{r_{21} - r_{12}}{4\varepsilon_4} \\ \varepsilon_4 = \frac{1}{2}\sqrt{1 + r_{11} + r_{22} + r_{33}} \end{cases}$$

2.4 Angular velocity of rotation matrices

The angular velocity of a body is a vector that describes the rotation of the body in space. It is defined as the derivative of the rotation matrix with respect to time. Angular velocity matrices are skew-symmetric matrices, i.e. matrices of the tangent space of $\text{SO}(3)$, denoted $\text{so}(3)$:

$$\text{so}(3) = \left\{ W \in \mathcal{M}_3(\mathbb{R}) \mid W^\top = -W \right\} = \left\{ \begin{bmatrix} 0 & -w_z & w_y \\ w_z & 0 & -w_x \\ -w_y & w_x & 0 \end{bmatrix} \mid w_x, w_y, w_z \in \mathbb{R} \right\}$$

Indeed, consider the rotation matrix $R(t) \in \text{SO}(3)$, and denote $\dot{R}(t) = \frac{\partial R(t)}{\partial t}$ the derivative of $R(t)$ with respect to time. We have:

$$\begin{aligned} R(t)R(t)^\top &= I_3 \\ \dot{R}R^\top + R\dot{R}^\top &= 0_3 \\ \dot{R}R^\top &= -R\dot{R}^\top \\ \dot{R}R^\top &= -(\dot{R}R^\top)^\top \end{aligned}$$

Hence, by defining $W := \dot{R}R^\top$, we have $W \in \text{so}(3)$.

Note that $R(t)$ is the solution of the differential equation $\dot{R} = WR$, with $R(0) = R_0 \in \text{SO}(3)$. The solution is:

$$R(t) = R_0 \exp(Wt) \quad \text{where} \quad \exp(Wt) = \sum_{n=0}^{+\infty} \frac{(Wt)^n}{n!}$$

This gives us a new parameterization of the group $\text{SO}(3)$, using the angular velocity matrices. Recall that a matrix $W \in \text{so}(3)$ is associated to a vector $w = (w_x, w_y, w_z) \in \mathbb{R}^3$; we denote $\hat{w} = W$. The main question associated to this representation is to compute the infinite sum of the exponential function:

$$\exp(tW) = \sum_{n=0}^{+\infty} \frac{(tW)^n}{n!} \tag{2.4.1}$$

We can use the fact that any skew-symmetric matrix $W \in \text{so}(3)$ is nilpotent, i.e. there exists an integer $n \in \mathbb{N}$ such that $W^n = 0_3$. This allows us to compute the exponential function:

$$\hat{w} = W = \begin{bmatrix} 0 & -w_z & w_y \\ w_z & 0 & -w_x \\ -w_y & w_x & 0 \end{bmatrix} = w \times \cdot$$

Therefore, $\hat{w}^2 = ww^\top - I_3$, and $\hat{w}^3 = -\hat{w}$. We can then regroup the different terms of the exponential function, by transforming the exponents greater than 3:

$$\exp_{\text{so}(3)}(tW) = I_3 + \sin(t)\hat{w} + \frac{1 - \cos(t)}{t}\hat{w}^2 \tag{2.4.2}$$

2.5 Exponential and logarithm map

We saw that the exponential map is an application from the tangent space of $\text{SO}(3)$, $\text{so}(3)$ to the group $\text{SO}(3)$, defined by:

$$\begin{aligned} \exp : \text{so}(3) &\longrightarrow \text{SO}(3) \\ tW &\longmapsto \exp(tW) = I_3 + \sin(t)\hat{w} + \frac{1 - \cos(t)}{t}\hat{w}^2 \end{aligned}$$

This function is surjective and 2π -periodic.

We can define a reciprocal function, the logarithm map, that is the inverse of the exponential map:

$$\begin{aligned} \log_{\text{SO}(3)} \text{SO}(3) &\longrightarrow \text{so}(3) \\ R &\longmapsto w \end{aligned}$$

Note that

$$\|w\| = \cos^{-1} \left(\frac{\text{Tr}(R) - 1}{2} \right)$$

and therefore:

$$w = \frac{\|w\|}{2 \sin(\|w\|)} \begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix}$$

This gives us a way to compute the distance between two rotations:

$$d(R_1, R_2) = \|\log_{SO(3)}(R_1^\top R_2)\|$$

Such a construction comes handy when we want to minimize the distance between two rotations, for instance in the context of trajectory tracking.

2.6 Rigid Body Transformations

2.6.1 Definition

Rigid body transformations are the combination of a rotation and a translation. They are defined as the set of matrices that preserve the vector product $v \times w$ and the scalar product $v \cdot w$; note that the preservation of the scalar product implies the preservation of the norm of the vectors, hence the distances (no deformation).

The group of rigid body transformations is also called the *Special Euclidean group* (denoted $SE(3)$), and defined by:

$$SE(3) = \left\{ g : \mathbb{E}^3 \rightarrow \mathbb{E}^3 \mid \forall x, y \in \mathbb{E}^3, \|g(x) - g(y)\| = \|x - y\| \right\}$$

Note that the rotation matrices ($SO(3)$) form a subgroup of the rigid body transformations group. Therefore, we also denote:

$$SE(3) = \mathbb{R}^3 x \times SO(3) = \left\{ (t, R) \mid t \in \mathbb{R}^3, R \in SO(3) \right\}$$

2.6.2 Homogeneous transformations

A rigid body transformation $M \simeq (t, R) \in SE(3)$ has the following action on a point $p \in \mathbb{E}^3$:

$$M(p) = R(p) + t$$

Therefore, similarly to the expression of $SO(3)$ as orthogonal matrices, one can write $SE(3)$ as a set of 4×4 invertible matrices:

$$M(p) = \begin{bmatrix} R & t \\ 0_3 & 1 \end{bmatrix} \begin{bmatrix} p \\ 1 \end{bmatrix}$$

Formally, one can define a composition law for the representation in homogeneous coordinates:

$$M_1 \circ M_2 = M_1 M_2 = \begin{bmatrix} R_1 & t_1 \\ 0_3 & 1 \end{bmatrix} \begin{bmatrix} R_2 & t_2 \\ 0_3 & 1 \end{bmatrix} = \begin{bmatrix} R_1 R_2 & R_1 t_2 + t_1 \\ 0_3 & 1 \end{bmatrix} \in \mathcal{M}_4(\mathbb{R})$$

The identity of $SE(3)$ is the invertible matrix I_4 . Finally, the invert of M corresponds to:

$$M^{-1} = \begin{bmatrix} R^\top & -R^\top t \\ 0_3 & 1 \end{bmatrix}$$

2.6.3 Linear and angular velocities

Similarly to the case of rotation matrices, one can derive an exponential map for the group $\text{SE}(3)$. To do so, we consider a curve $M(s) \in \text{SE}(3)$; we have:

$$\dot{M}(s)M(s)^{-1} = \begin{bmatrix} \dot{R}(s)R(s)^\top & t(s) - \dot{R}(s)R(s)^\top t(s) \\ 0_3 & 1 \end{bmatrix}$$

We identify the angular velocity $W(s) = \dot{R}(s)R(s)^\top = [w(s)]_\times$ and the linear velocity $v(s) = \dot{t}(s) - w(s) \times t(s)$. Therefore, the pair $(v, w) \in \text{se}(3) \simeq \mathbb{R}^3 \times \text{so}(3)$ corresponds to the tangent vector of the group $\text{SE}(3)$ at the point $M \in \text{SE}(3)$.

This allows us to define the exponential map for the group $\text{SE}(3)$:

$$\begin{aligned} \exp_{\text{se}(3)} : \text{se}(3) &\longrightarrow \text{SE}(3) \\ (v, w) &\longmapsto M \end{aligned}$$

such that:

$$\exp_{\text{se}(3)}(v, w) = \begin{bmatrix} \exp_{\text{so}(3)}(w) & \frac{1}{\|w\|} (ww^\top v + (I_3 - \exp_{\text{so}(3)}(w))w \times v) \\ 0_3 & 1 \end{bmatrix}$$

Since $\exp_{\text{se}(3)}$ is surjective, we can define its reciprocal function, the logarithm map:

$$\begin{aligned} \log_{\text{SE}(3)} : \text{SE}(3) &\longrightarrow \text{se}(3) \\ M &\longmapsto (v, w) \end{aligned}$$

such that:

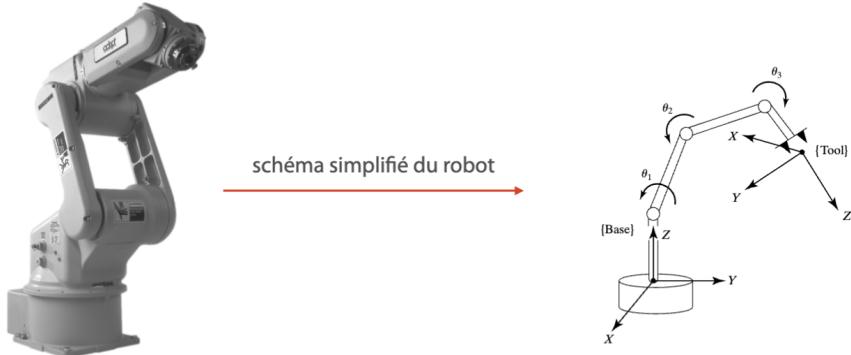
$$w = \log_{\text{SO}(3)}(R) \quad \text{and} \quad v = \|w\|(ww^\top + (I - R)W)^{-1}t$$

This allows once again to define a notion of distance between two rigid body transformations:

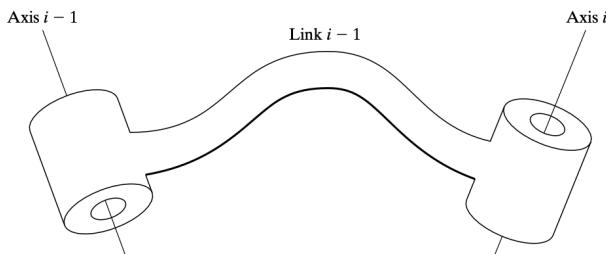
$$d(M_1, M_2) = \|\log_{\text{SE}(3)}(M_1^{-1}M_2)\|$$

3 Forward Kinematics

Forward kinematics allows to compute the position and orientation of the kinematic chain given the joint parameters (e.g. angles, lengths, etc.). For instance, we can compute the position and orientation of the terminal organ given the angles of the articulations.



The kinematic chain is composed of rigid bodies interconnect by joints. The joins define the degrees of freedom of the cinematic chain, which are the parameters that we can control.



3.1 Articulations and joint speed

3.1.1 Joint types and degrees of freedom

The topology of the articulation between two rigid bodies defines the degrees of freedom of the joint.

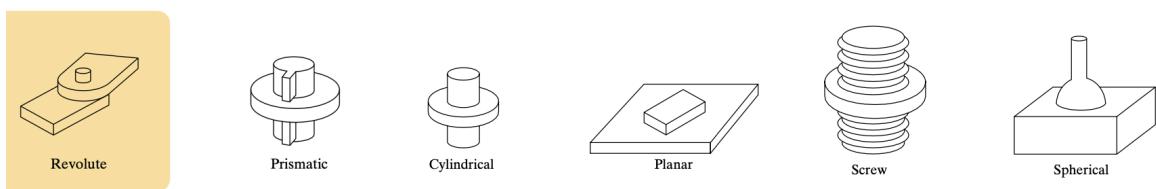


Figure 3.1: Example of possible joints.

Each joint can be represented as a function K_i from a configuration space Q_i to the Special Euclidean Group SE(3):

$$K_i : Q_i \longrightarrow \text{SE}(3)$$

$$q \longmapsto M_i(q_i)$$

Consider for instance the revolute joint, which constrains the motion of two bodies around a fixed axis. It is parameterized by the angle $\theta \in \mathbb{R}$, therefore the configuration space is $Q_i = \mathbb{S}^1 \simeq \mathbb{R}$ (one degree of freedom). The function K_i is then defined as:

$$K_i : \mathbb{S}^1 \simeq \mathbb{R} \longrightarrow \text{SE}(3)$$

$$q \longmapsto \begin{bmatrix} \cos & -\sin q & 0 & 0 \\ \sin q & \cos q & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

3.1.2 Joint speed

The change of configuration of the joint comes from the joint speed, which is the derivative of the joint parameter with respect to time. Consider the relative position of the articulation between the two adjacent frames, given by:

$$K_i : Q_i \longrightarrow \text{SE}(3)$$

$$q_i \longmapsto M_i(q_i)$$

The relative speed of the joint generated by the articulation is given by:

$$k_i : T_{q_i} Q_i \longrightarrow \text{se}(3)$$

$$(q_i, \dot{q}_i) \longmapsto v_i(q_i) = S_i(q_i)\dot{q}_i$$

for some transformation S_i .

In the case of the revolute joint, the speed of the joint is given by:

$$k_i : T_q \mathbb{S}^1 \simeq \mathbb{R}^2 \longrightarrow \text{se}(3)$$

$$(q, \dot{q}) \longmapsto (v, w) = \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \dot{q} \end{bmatrix} \right)$$

Hence, we have:

$$S_i(q) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \in \mathcal{M}_{6,1}(\mathbb{R})$$

which gives us $S_i(q)\dot{q} = (v, w)$ as expected.

3.2 Direct geometry

We aim to compute the position and orientation of the terminal organ given the joint parameters. We can do so by computing the transformation matrix of each body in the kinematic chain, and then multiplying them to get the transformation matrix of the terminal organ.

3.2.1 Transformation matrix

Given an articulation i , we can compute the relative position of the associated frame i with respect to the frame $i - 1$ using the function K_i :

$${}^{i-1}M_i(q_i) = {}^{i-1}P_i K_i(q_i)$$

We can combine the transformations of all the bodies in the kinematic chain to get the transformation matrix of the terminal organ:

$$\begin{aligned} {}^0K_n : \bigtimes_{i=1}^n Q_i &\longrightarrow \text{SE}(3) \\ q = (q_1, \dots, q_n) &\longmapsto \bigtimes_{i=1}^n {}^{i-1}M_i(q_i) \end{aligned}$$

Therefore, the configuration space can be read as the product of the configuration spaces of each joint:

$$Q = \bigtimes_{i=1}^N Q_i \simeq \mathbb{R}^{\sum_{i=1}^N n_i}$$

where N is the number of joints.

3.2.2 Kinematic Jacobian

Our goal is now to link the joint speed to the spatial speed of the bodies in movement. For each articulation, we have a linear mapping of the form:

$$v_i(q_i) = S_i(q_i)\dot{q}_i$$

This corresponds to the temporal derivative of the articular geometry:

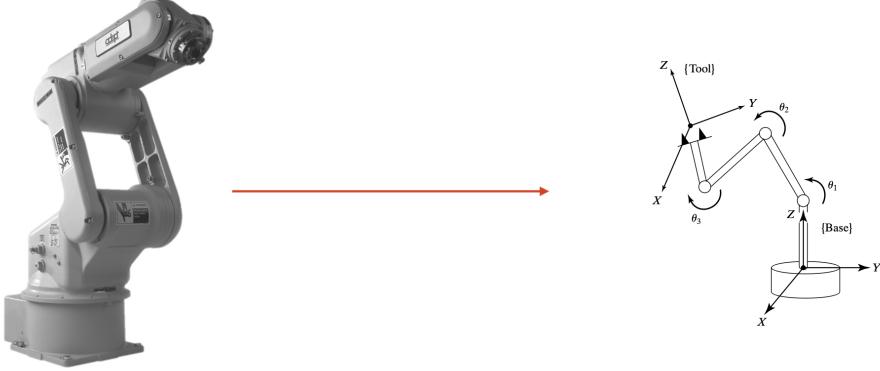
$$\dot{M}_i(q_i)M_i^{-1}(q_i) = S_i(q_i)\dot{q}_i$$

Therefore, one can link the spatial speed of a body to the joint speed using the relation:

$${}^0k_n(q, \dot{q}) = \sum_{i=1}^n k_i(q_i, \dot{q}_i) = \underbrace{\begin{bmatrix} S_1(q_1) & \cdots & S_n(q_n) \end{bmatrix}}_{J(q)} \begin{bmatrix} \dot{q}_1 \\ \vdots \\ \dot{q}_n \end{bmatrix}$$

4 Inverse Kinematics

Inverse Kinematics aims at finding joint parameters that achieve a desired end-effector pose. This is a more complex problem than forward kinematics, as it is not always possible to find a solution, and when it is, there may be multiple solutions.



4.1 Optimization problem and resolution

Given a target position M^* for the end-effector, we aim at solving the following distance minimization problem:

$$\min_{q \in Q} d({}^0K_n(q), M^*)$$

Using the logarithm distance, we can rewrite the problem as:

$$\min_{q \in Q} \|\log({}^0K_n(q)^{-1}M^*)\| \quad (4.1.1)$$

We can cast this initial problem $\min_{q \in Q} \|\log({}^0K_n(q)^{-1}M^*)\|$ as a more general optimization problem:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|f(x)\|_2^2$$

We aim at iteratively solving this optimization problem. A first-order linear approximation of f around x is given by:

$$f(x + p) = f(x) + \underbrace{\frac{\partial f}{\partial x}(x)p}_{J(x)}$$

Hence:

$$\min_{p \in \mathbb{R}^n} \frac{1}{2} \|f(x_0 + p)\|_2^2 = \min_{p \in \mathbb{R}^n} \frac{1}{2} \|f(x) + J(x)p\|_2^2$$

We can now solve the optimization problem by setting the gradient of the objective function to zero:

$$\begin{aligned} & \nabla_p \left(\frac{1}{2} \|f(x) + J(x)p\|_2^2 \right) = 0 \\ \iff & J(x)^\top (f(x) + J(x)p^*) = 0 \\ \iff & p^* = -(J(x)^\top J(x))^{-1} J(x) f(x) \\ \iff & p^* = -J(x)^+ f(x) \end{aligned}$$

where $J(x)^+$ is the *Moore-Penrose pseudoinverse* of $J(x)$. We can then iteratively update:

$$x_{k+1} = x_k + \alpha p_k$$

until we obtain $\|J(x)^\top f(x)\| \leq \varepsilon^*$, where $\varepsilon^* > 0$ is a fixed precision we aim at achieving. This iterative method is known as the *Newton-Raphson* method. If we work on a differential manifold, we can use instead:

$$q_{k+1} = q_l \oplus \alpha p_k$$

4.2 Trajectory tracking

We can extend the previous method to track a trajectory $M(t)$: we want to compute the joint speeds \dot{q} that track the trajectory as closely as possible. Consider a continuous and differentiable time trajectory M^* :

$$\begin{aligned} M^* : [0, T] &\longrightarrow \text{SE}(3) \\ t &\longmapsto M^*(t) \end{aligned}$$

We are also given the corresponding spatial (linear and angular) velocity of the end-effector v^* :

$$\begin{aligned} v^* : [0, T] &\longrightarrow \text{se}(3) \\ t &\longmapsto v^*(t) \end{aligned}$$

At each instant t , the forward kinematics of the robot give us the current end-effector pose $M(t)$ and the corresponding Jacobian $J(t)$:

$$M(q(t)) = \bigtimes_{i=0}^n K_i(q_i(t)) \in \text{SE}(3) \quad \text{and} \quad v(t) = J(q(t))\dot{q}(t) \in \text{se}(3)$$

Using the angular speed of the joints $\dot{q}(t)$, we can control the position $q(t)$ of the joints. Our goal is to minimize the gap between the position and speed of the end-effector $(M(q(t)), v(q(t), \dot{q}(t)))$ and the desired trajectory $(M^*(t), v^*(t))$.

The error can be computed using the following cost function:

$$e(t, q(t)) = M(q(t)) \ominus M^*(t) = \log_{\text{SE}(3)}(M^*(t)^{-1} M(q(t)))$$

The error in velocity space is given by:

$$\dot{e}(t, q(t), \dot{q}(t)) = v(q(t), \dot{q}(t)) - v^*(t) = J(q(t))\dot{q}(t) - v^*(t)$$

We define the error correction profile as:

$$\dot{e}(t, q(t), \dot{q}(t)) = -\lambda e(t, q(t))$$

for some $\lambda > 0$.

We would like to have $\dot{e} + \lambda e = 0$ to respect the error correction profile. Therefore, we want to solve the following optimization problem:

$$\min_{\dot{q}(t)} \frac{1}{2} \|\dot{e} + \lambda e\|_2^2 \tag{4.2.1}$$

By substituting the expressions of e and \dot{e} , we obtain:

$$\min_{\dot{q}(t)} \frac{1}{2} \left\| \underbrace{J(q(t))\dot{q}(t)}_A - \underbrace{v^*(t) + \lambda(M(q) \ominus M^*)}_b \right\|_2^2$$

This is a linear least squares problem, i.e. of the form:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2$$

with $A = J(q(t))$ and $b = v^*(t) + \lambda(M(q) \ominus M^*)$. The solution is given by:

$$x^* = A^+ b = (A^\top A)^{-1} A^\top b$$

We can then update the joint speeds $\dot{q}(t)$ using the computed values x^* .

4.3 Exploiting the redundancy of solutions

In the case of redundant robots, there are multiple solutions to the inverse kinematics problem. We can exploit this redundancy to optimize a given criterion. For instance, we can minimize the energy consumption of the robot by minimizing the norm of the joint speeds.

More generally, we can consider two tasks to be achieved simultaneously, one being strictly prioritized over the other:

$$(A_1, b_1) \gg (A_2, b_2)$$

4.3.1 Relaxed approach

A first approach is to assign each task a coefficient $\alpha_i > 0$, and to choose the cost function to be the sum of the weighted tasks. Given $\alpha_1 \gg \alpha_2 > 0$, we might want to minimize:

$$\min_x \alpha_1 \|A_1 x + b_1\|_2^2 + \alpha_2 \|A_2 x + b_2\|_2^2$$

However, this approach does not guarantee that the first task will be strictly prioritized over the second one. Imagine that the first task is for our robot to stay in a safe position, while the second task is to reach a target position. To ensure that the robot stays in a safe position, we need to strictly prioritize the first task, resulting in a much higher α_1 than α_2 . This would lead to a numerically unstable problem, and the robot would likely not reach the target position.

4.3.2 Strict approach

A more stable approach consists in solving a constrained optimization problem:

$$\min_x \|A_2 x + b_2\|_2^2 \quad \text{s.t. } A_1 x = b$$

This guarantees that the first task is strictly prioritized over the second one, and does not require coefficients. To do so, we can compute the projection of the second task onto the null space of the first task. For any problem of the form:

$$\min_x \|Ax - b\|_2^2$$

a solution is given by $x^* = A^+ b$. If A is not invertible (which is the case when the system is under determined), the solutions are given by:

$$x^* = A^+ b + \underbrace{(I - A^+ A)}_{P_{\ker(A)}} y$$

for some $y \in \mathbb{R}^n$. The general idea of minimizing the second task while respecting the first one is to choose the value of y that minimizes the norm of the second task. This is equivalent to projecting the second task onto the null space of the first task.

Given two tasks $(A_1, b_1) \gg (A_2, b_2)$, we want to choose y such that:

$$\begin{aligned} & A_2(A_1^+ b_1 + (I - A_1^+ A_1)y) = b_2 \\ \iff & \underbrace{A_2(I - A_1^+ A_1)y}_{\tilde{A}_2} = \underbrace{b_2 - A_2 A_1^+ b_1}_{\tilde{b}_2} \\ \iff & y = \tilde{A}_2^+ \tilde{b}_2 \end{aligned}$$

Hence, we obtain the solution to the constrained optimization problem:

$$x^* = A_1^+ b_1 + (I - A_1^+ A_1)\tilde{A}_2^+ \tilde{b}_2$$

5 Direct and Inverse Dynamics

Previously, we studied *kinematics*, which describes the motion of a robot without considering the forces that cause it. We now turn to *dynamics*, which studies the *forces* that cause the motion.

Similarly to kinematics, dynamics can be divided into two problems: the *direct dynamics* problem and the *inverse dynamics* problem. The *direct dynamics* problem consists in finding the motion of a robot given the forces applied to it. Conversely, the goal of *inverse dynamics* is to minimize the gap between the position and speed of the end-effector and the desired trajectory, while taking into account the dynamics of the robot.

5.1 Physical equations

The motion of a robot is governed by the laws of physics. The *Newton-Euler* equations describe the motion of a rigid body in space. They are a set of differential equations that relate the forces and torques applied to the body to its translational and rotational motion. The equations are given by:

$$m_k \ddot{x}_k = f_k \quad (\text{Newton's equation})$$

for the translational motion, and:

$$I_k \dot{\omega}_k + \underline{\omega}_k \times I_k \underline{\omega}_k = \tau_k \quad (\text{Euler's equation})$$

for the rotational motion. Here, m_k is the mass of the body, f_k is the force applied to it, I_k is the inertia matrix, $\underline{\omega}_k$ is the angular velocity, and τ_k is the torque applied to the body. The equations are written in the body frame of the body k , whose origin coincides with the body's center of mass.

Almost all physical problems, from mechanics to relativity, follow the *least-action principle*. This principle states that the actual motion of a system is the one that minimizes a certain quantity, called the *action*. In mechanics, it corresponds to the minimization of the integral of the kinetic-potential energies over time.

We can define \mathcal{D} to be a kinetic metric of least deviation between the actual motion of the robot and the desired motion:

$$\mathcal{D} := \sum_{k=1}^n \frac{1}{2} (\ddot{x}_k - \ddot{\underline{x}}_k)^\top m_k (\ddot{x}_k - \ddot{\underline{x}}_k) + \frac{1}{2} (\dot{\omega}_k - \dot{\underline{\omega}}_k)^\top I_k (\dot{\omega}_k - \dot{\underline{\omega}}_k) \quad (5.1.1)$$

Formally, the least-action principle states that the actual motion of the robot is the one that minimizes \mathcal{D} . We can find the actual motion by solving the following optimization problem:

$$\min_{\ddot{x}_k, \dot{\omega}_k} \mathcal{D} \quad (5.1.2)$$

In the joint space, we can write the equations of motion as:

$$M(q)\ddot{q} + C(q, \dot{q}) - F(q) = 0 \quad (5.1.3)$$

where $M(q)$ is the mass matrix, $C(q, \dot{q})$ is the Coriolis matrix, and $F(q)$ is the vector of forces applied to the joints.

5.2 Contact forces

The poly-articulated system dynamics is driven by the Lagrangian dynamics. If we consider the contact forces, we can write the equations of motion as:

$$\underbrace{M(q)}_{\text{Mass matrix}} \ddot{q} + \underbrace{C(q, \dot{q})}_{\text{Coriolis}} + \underbrace{G(q)}_{\text{Gravity}} = \underbrace{\tau}_{\text{Motor torque}} + \underbrace{J_c^\top(q) \lambda_c}_{\text{External forces}}$$

With this equation in mind, the direct dynamics problem consists in finding the motion of the robot given the forces applied to it. Schematically, the problem can be written as:

$$\ddot{q} = \text{ForwardDynamics}(q, \dot{q}, \tau, \lambda_c)$$

This is used in simulation, and can be solved using the articulated body algorithm.

Conversely, the *inverse dynamics* problem consists in finding the forces applied to the robot given its motion. Schematically, the problem can be written as:

$$\tau = \text{InverseDynamics}(q, \dot{q}, \ddot{q}, \lambda_c)$$

This is used in control, and can also be solved by integrating the equations of motion over time (recursive Newton-Euler algorithm).

Note that naive algorithms for solving the direct dynamics problem have a complexity of $O(n^3)$, which is prohibitive for large systems. However, the articulated body algorithm has a complexity of $O(n)$, which makes it tractable for large systems. Such improvements come from the structure of the mass matrix, which is often sparse since parts of the robots are not connected. Similarly, the Jacobian of the contact forces is often sparse, which can also be exploited to reduce the complexity of the problem.

An important aspect of the direct dynamics problem is the computation of λ_c . Multiple approaches, corresponding to different contact models, can be used:

- Soft contact: using a spring-damper model
- Rigid contact: using a bilateral or unilateral contact model
- Mixed contact: using a relaxed contact model

we will study these models in more detail in the next sections.

5.3 Soft contact

The soft contact model is used when the contact between the robot and the environment is soft, i.e. when the robot can penetrate the environment. This is one of the simplest contact model, which is both very intuitive and straightforward to implement.

The model is based on a spring-damper system, where the contact force is proportional to the penetration depth and the relative velocity between the robot and the environment. The spring will push the robot away from the environment (proportionally to the penetration p), while the damper will slow down the relative motion between the robot and the environment (proportionally to the speed \dot{p}).

We consider the parameter k to be the stiffness of the spring, and b to be the damping coefficient of the damper. The contact force is then given by:

$$\lambda_c = \max(-k \cdot p - b \cdot \dot{p}, 0) \tag{5.3.1}$$

The max function is used to ensure that the contact force is always positive, i.e. that the robot is always pushed away from the environment.

Lower values of k correspond to softer contacts, while higher values of k correspond to harder objects. Higher values of d also mean higher energy dissipation; in the case of a trampoline, for instance, we would have a very low d to allow the robot to bounce back.

Despite being simple and easy, the soft contact model is not relevant for rigid interfaces (when $k \rightarrow +\infty$). Furthermore, it requires a very stable integrator, since the equation is stiff (because of the max function).

5.4 Rigid bilateral contacts

5.4.1 Optimization problem

The rigid contact model is used when the contact between the robot and the environment is rigid, i.e. when the robot cannot penetrate the environment. It uses the least-action principle to write and solve an optimization problem: given q and \dot{q} , we aim at retrieving \ddot{q} and λ_c .

When the contact is symmetric, i.e. when the robot can push the environment and the environment can push the robot, we used the bilateral contact model. To do so, we solve the following optimization problem:

$$\min_{\ddot{q}} \frac{1}{2} \|\ddot{q} - \ddot{q}_f\|_{M(q)}^2 \quad \text{s.t.} \quad c(q) = 0 \quad (5.4.1)$$

where $c(q)$ is the contact function, $J_c(q)$ its Jacobian. The constraint $c(q) = 0$ ensures that the robot is in contact with the floor. Here, we are minimizing the distance between the actual acceleration and the *free acceleration* with respect to the unconstrained acceleration. This distance is measured by the metric induced by the kinetic energy (hence by the mass matrix $M(q)$). The free acceleration is the acceleration that the robot would have if it was not in contact with the floor. It can be expressed as:

$$\ddot{q}_f = M(q)^{-1}(\tau - C(q, \dot{q}) - G(q))$$

Note that from the constraint $c(q) = 0$ we can derive by differentiation the following equation:

$$c(q) = 0 \implies J_c(q)\dot{q} = 0 \implies J_c(q)\ddot{q} + \underbrace{\dot{J}_c(q, \dot{q})\dot{q}}_{\gamma_c(q, \dot{q})} = 0$$

Therefore, we can rewrite Equation 5.4.1 as:

$$\min_{\ddot{q}} \frac{1}{2} \|\ddot{q} - \ddot{q}_f\|_{M(q)}^2 \quad \text{s.t.} \quad J_c(q)\ddot{q} + \gamma_c(q, \dot{q}) = 0 \quad (5.4.2)$$

5.4.2 Solution using the KKT conditions

The solution of this problem can be retrieved by deriving the KKT conditions¹ of the QP problem, via the so-called *Lagrangian*:

$$\mathcal{L}(\ddot{q}, \lambda_c) = \underbrace{\frac{1}{2} \|\ddot{q} - \ddot{q}_f\|_{M(q)}^2}_{\text{cost function}} + \lambda_c^\top \underbrace{(J_c(q)\ddot{q} + \gamma_c(q, \dot{q}))}_{\text{equality constraint}}$$

¹The *Karush-Kuhn-Tucker (KKT) conditions* are conditions for a solution of an optimization problem to be optimal.

where λ_c , the dual variable (or Lagrange multiplier), corresponds to the contact forces. The KKT conditions of the QP problem are given by:

$$\nabla_{\ddot{q}} \mathcal{L} = M(q)(\ddot{q} - \ddot{q}_f) - J_c^\top(q)\lambda_c = 0 \quad (5.4.3)$$

$$\nabla_{\lambda_c} \mathcal{L} = J_c(q)\ddot{q} + \gamma_c(q, \dot{q}) = 0 \quad (5.4.4)$$

where the first equation represents the joint space force propagation, and the second equation translates the contact acceleration constraint.

By rearranging the equations, we obtain:

$$\begin{aligned} M(q)\ddot{q} - J_c^\top(q)\lambda_c &= M(q)\ddot{q}_f \\ J_c(q)\ddot{q} + 0 &= -\gamma_c(q, \dot{q}) \end{aligned}$$

leading to the *KKT dynamics equation*:

$$\underbrace{\begin{bmatrix} M(q) & J_c^\top(q) \\ J_c(q) & 0 \end{bmatrix}}_{K(q)} \begin{bmatrix} \ddot{q} \\ -\lambda_c \end{bmatrix} = \begin{bmatrix} M(q)\ddot{q}_f \\ -\gamma_c(q, \dot{q}) \end{bmatrix} \quad (5.4.5)$$

Note that there might be one, multiple, or no solution to this problem. This depends on whether $J_c(q)$ is full rank, not full rank, or if $\gamma_c(q, \dot{q})$ is not in the range space of $J_c(q)$.

We can analytically inverse the system to obtain the solution in three steps. First, we solve for \ddot{q} by expressing it as a function of \ddot{q}_f and λ_c :

$$\ddot{q} = \ddot{q}_f + M(q)^{-1}J_c(q)^\top\lambda_c$$

Then, we replace \ddot{q} and get an expression depending only on λ_c :

$$\underbrace{J_c(q)M(q)^{-1}J_c(q)^\top}_{G_c(q)}\lambda_c + \underbrace{J_c(q)\ddot{q}_f + \gamma_c(q, \dot{q})}_{a_{c,f}(q, \dot{q}, \ddot{q}_f)} = 0$$

where $G_c(q)$ is called *Delassus' matrix*, and $a_{c,f}(q, \dot{q}, \ddot{q}_f)$ is called the *free contact acceleration*. Finally, we can inverse $G_c(q)$ and find the optimal λ_c :

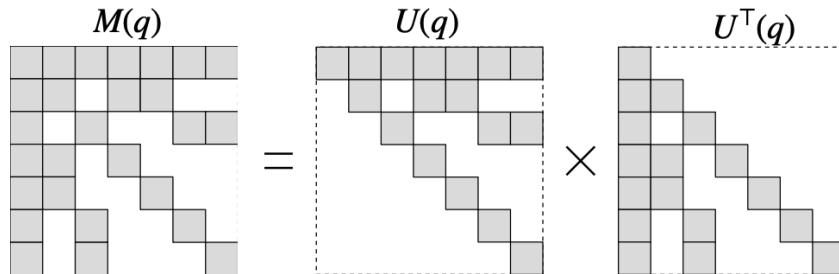
$$\lambda_c = -G_c(q)^{-1}a_{c,f}(q, \dot{q}, \ddot{q}_f)$$

5.4.3 Sparse Cholesky factorization

The bottleneck of this method is the computation and inversion of $G_c(q)$, which has a complexity of $O(n^3)$:

$$G_c(q) := J_c(q)M(q)^{-1}J_c(q)^\top$$

This is prohibitive for large systems, as the inversion of $M(q)$ has a complexity of $O(n^3)$. A solution is to exploit the sparsity in the Cholesky factorization of $M(q)$.



This reduced the complexity to $O(n^2)$ when using a dense Cholesky decomposition.

5.4.4 The maximum dissipation principle

We showed that the contact forces λ_c fulfill the relation:

$$G_c(q)\lambda_c + a_{c,f}(q, \dot{q}, \ddot{q}_f) = 0$$

From an energetic point of view, this solution minimizes:

$$\min_{\lambda_c} \frac{1}{2} \lambda_c^\top G_c(q) \lambda_c + \lambda_c^\top a_{c,f}(q, \dot{q}, \ddot{q}_f)$$

This is equivalent to the *maximum dissipation principle*, which states that the contact forces are chosen to maximize the energy dissipation in the contact. This corresponds to the dual problem of the least-action principle, formally:

$$\max_{\lambda_c} -\frac{1}{2} \underbrace{\left(G_c(q)\lambda_c + 2\lambda_c^\top a_{c,f}(q, \dot{q}, \ddot{q}_f) \right)}_{a_c(q, \dot{q}, \ddot{q}_f)}$$

which is the dual of the least-action principle (primal):

$$\min_{\ddot{q}} \frac{1}{2} \|\ddot{q} - \ddot{q}_f\|_{M(q)}^2 \quad \text{s.t.} \quad J_c(q)\ddot{q} + \gamma_c(q, \dot{q}) = 0$$

5.5 Rigid unilateral contacts

5.5.1 Unilateral contact as a Nonlinear Complementarity Problem

The rigid unilateral contact model is used when the contact between the robot and the environment is rigid, and when the robot can only push the environment. This is the case, for instance, when the robot is standing on the floor.

When dealing with unilateral contact conditions, three conditions must be satisfied:

1. **Maximum dissipation** states that the contact forces should dissipate at most the kinetic energy:

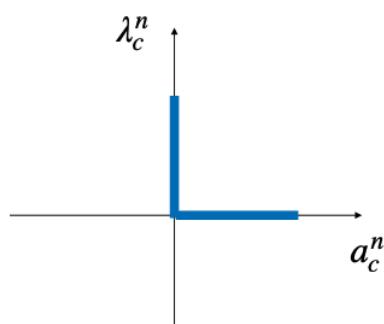
$$\min_{\lambda_c} \frac{1}{2} \lambda_c^\top G_c(q) \lambda_c + \lambda_c^\top a_{c,f}(q, \dot{q}, \ddot{q}_f)$$

2. **Complementary condition** (Signorini's conditions) require that the floor can **only push** (no pulling), and should emit no force when the contact is about to open:

$$0 \leq \lambda_{c,n} \perp a_{c,n} \geq 0$$

3. **Friction cone constraint** (Coulomb law) bounds the lateral forces by the normal force:

$$\sqrt{\lambda_{c,x}^2 + \lambda_{c,y}^2} \leq \mu \lambda_{c,n}$$



Signorini's conditions



Coulomb law

The contact problem then corresponds to a *Nonlinear Complementarity Problem (NCP)*:

$$\begin{cases} \min_{\lambda_c} \frac{1}{2} \lambda_c^\top G_c(q) \lambda_c + \lambda_c^\top a_{c,f}(q, \dot{q}, \ddot{q}_f) \\ 0 \leq \lambda_{c,n} \perp a_{c,n} \geq 0 \\ \sqrt{\lambda_{c,x}^2 + \lambda_{c,y}^2} \leq \mu \lambda_{c,n} \end{cases}$$

which is non-convex, hence hard to solve. More complex solvers can be used to solve this problem in an exact way.

5.5.2 Relaxed contact problem

Another approach to solve the rigid unilateral contact problem is to relax the contact conditions. This is done by removing the complementarity condition, and by regularizing the forces:

$$\begin{cases} \min_{\lambda_c} \frac{1}{2} \lambda_c^\top (G_c(q) + R) \lambda_c + \lambda_c^\top a_{c,f}(q, \dot{q}, \ddot{q}_f) \\ \sqrt{\lambda_{c,x}^2 + \lambda_{c,y}^2} \leq \mu \lambda_{c,n} \\ 0 \leq \lambda_{c,n} \perp a_{c,n} \geq 0 \end{cases}$$

where R is a regularization term. This problem becomes convex, which makes it much easier to solve. Nevertheless, the solution is not exact, which might lead to some physical inconsistencies.

5.6 Inverse dynamics

Similarly to the inverse kinematics problem, we can solve the inverse dynamics problem by minimizing the gap between the actual motion of the robot and the desired motion. Instead of solving for the configuration of the robot, $q \in Q$, we solve for the forces applied to the robot, or equivalently for $\ddot{q}, \tau, \lambda_c$.

Consider the actual position and speed of the end effector, (M, v, a) . To compare it to a reference trajectory (M^*, v^*, a^*) , we define the error as:

$$e := M \ominus M^* = \log_{SE(3)}((M^*)^{-1} M) \quad (5.6.1)$$

The error in velocity space is given by:

$$\dot{e} = v - v^*$$

and the acceleration of the error is:

$$\ddot{e} = a - a^* = J(q)\ddot{q} + \dot{J}(q, \dot{q})\dot{q} - a^*$$

This allows us to define the error correction profile as:

$$\ddot{e} = -K_p e - K_v \dot{e} \quad (5.6.2)$$

for some $K_p, K_v > 0$. We can then solve the following optimization problem:

$$\min_{\ddot{q}, \tau, \lambda_c} \frac{1}{2} \|\ddot{e} + K_p \dot{e} + K_v e\|_2^2 \quad \text{s.t.} \quad M(q)\ddot{q} + C(q, \dot{q}) + G(q) = \tau + J_c^\top(q) \lambda_c$$

with λ_c the contact forces.

6 Motion Planning

Motion planning is the problem of finding a sequence of valid actions that will transform the robot from its initial configuration to a desired configuration. Such actions can be constrained: we might ask the robot to avoid obstacles, to respect its kinematic limits, or to minimize energy consumption.

We will first introduce a mathematical framework to describe the robot's configuration space, obstacles, and collision, then we will present some algorithms to solve the motion planning problem.

6.1 Configuration space

6.1.1 Definitions

The ambient space, also called the *workspace*, is in most cases the Euclidean manifold \mathbb{E}^3 . In this ambient space, a physical system can be described as a set of compact subsets called *bodies*. The vector of the considered physical systems must be in the *state space* \mathcal{S} , the set of all possible sets:

$$(K_1, \dots, K_N) \in \mathcal{S} \subseteq \mathcal{K}(\mathbb{E}^3)^N$$

The state space is parametrized with a transformation q in the *configuration space* \mathcal{C} :

$$q = (\phi_1, \dots, \phi_M) \in \mathcal{C}$$

where each ϕ_i is a *transformation* of a system, that is:

$$\phi_i : \mathcal{K}(\mathbb{E}^3) \rightarrow \mathcal{K}(\mathbb{E}^3)$$

Note that M corresponds to the number of systems of the robot, the other physical systems being obstacles. Therefore, given a configuration q , the space S_q is:

$$S_q = \underbrace{\phi_1(K_1^\circ), \dots, \phi(M^\circ)}_{\text{robot}}, \underbrace{K_{M+1}, \dots, K_N}_{\text{obstacles}}$$

where the K_i° are compacts representing each body, intuitively in their own frames. We will see below that they can be represented numerically with meshes or polynomials.

In general, \mathcal{C} is a manifold, its dimension being the number of degrees of freedom of the system. Note that the numerical dimension used to represent \mathcal{C} is not necessarily the same as the actual dimension of the manifold: for instance, quaternions of dimension 4 can be used to represent rotations of $SO(3)$, a manifold of dimension 3.

6.1.2 Joints

Multiple types of joints can be used to connect the different parts of a robot. Formally, a joint is a mapping from a manifold of dimension $1 \leq p \leq 6$ into the solid placement $SE(3)$. The difference between the joints is the dimension of the manifold, which can also be seen as the number of degrees of freedom of the joint.

6.1.3 Examples

Moving solid Let us consider the Euclidean manifold \mathbb{E}^2 as a workspace, that is 2D objects. A solid object K is described as a compact subset of \mathbb{E}^2 . The configuration space is the set of transformations of this object, $SE(2)$, the special Euclidean group of dimension 2.

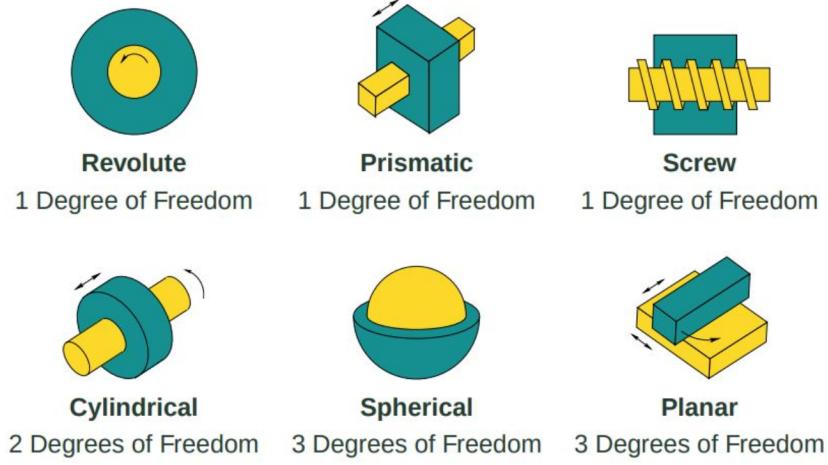


Figure 6.1: Different types of joints.

Double pendulum In the case of a double pendulum, the workspace remains \mathbb{E}^2 . Because of the physical constraints on the structure of the system, the pendulum is parametrized by two angles θ_1 and θ_2 . The configuration space is therefore a torus $T^2 = S^1 \times S^1$, but can also be seen as $[0, 2\pi] \times [0, 2\pi]$.

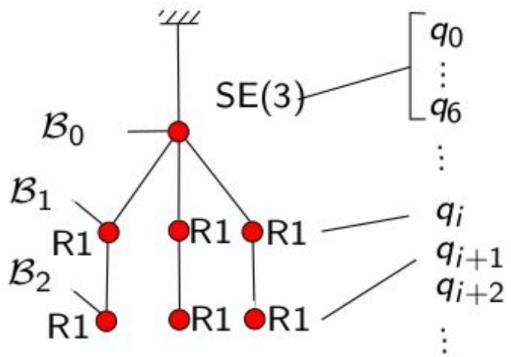
Rigid poly-articulated body The same general framework can be applied to more complex systems, such as humanoid robots. The workspace is \mathbb{E}^3 , and the configuration q of a robot is represented by the concatenation of the parameters of each joint:

$$q \in \underbrace{SE(3)}_{\text{Position}} \times \underbrace{S_1 \times \cdots \times S_1}_{\text{Revolute joints}} \times \underbrace{[0, 1] \times \cdots \times [0, 1]}_{\text{Prismatic & bounded joints}}$$

Recall that the forward kinematic can be used to compute the position of each joint in the global frame, given the configuration q .



A humanoid robot



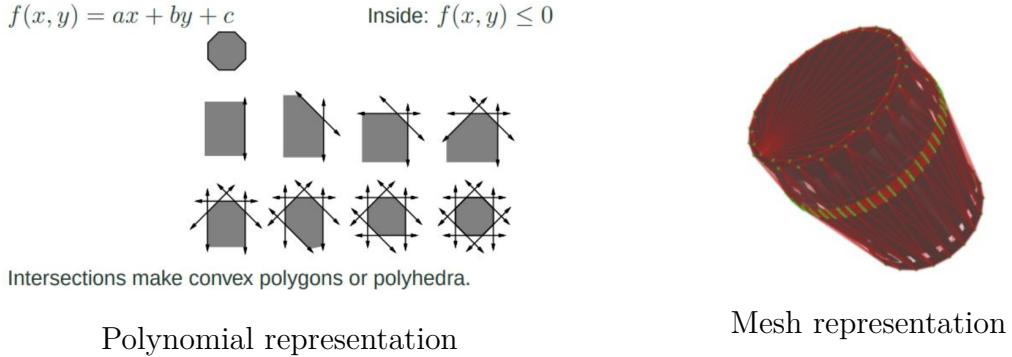
The associated configuration tree

6.2 Obstacles and collisions

Obstacles and collision handling are crucial in motion planning, and are responsible for the complexity of the problem. We will dive deeper into this topic in another chapter, but we introduce here the fundamental notions needed for motion planning.

6.2.1 Representations and modelisation

There are multiple ways to represent obstacles, which are used in different algorithms. The *polynomial representation* consists in representing the obstacles as sets satisfying polynomial inequalities. The *mesh representation* and *primitive representation* define simple sets of points in the workspace that are respectively inside or outside the obstacles.



Given a physical state S_q of the form

$$S_q = \phi_1(K_1^\circ), \dots, \phi(M^\circ), K_{M+1}^\circ, \dots, K_N^\circ$$

we can define the space of configurations \mathcal{C}_{obs} that are in collision with the obstacles:

$$\mathcal{C}_{\text{obs}} = \left\{ q \in \mathcal{C} \mid \underbrace{\exists i \neq j \leq M, \phi_i(K_i^\circ) \cap \phi_j(K_j^\circ) \neq \emptyset}_{\text{Auto-collision}} \text{ or } \underbrace{\exists i \leq M, j > M, \phi_i(K_i^\circ) \cap K_j^\circ \neq \emptyset}_{\text{Environment collision}} \right\}$$

We can then define the *free space* $\mathcal{C}_{\text{free}}$ as the set of configurations that are not in collision with the obstacles, that is:

$$\mathcal{C}_{\text{free}} = \mathcal{C} \setminus \mathcal{C}_{\text{obs}}$$

Therefore, the problem of motion planning is to find a path in the free space $\mathcal{C}_{\text{free}}$ that connects the initial configuration to the final configuration.

6.2.2 Example: translation without rotation of a rigid body

Let us consider the workspace \mathbb{E}^2 and a rigid body K_r that can be translated without rotating. The configuration space is $SE(2)$, and the obstacles are represented by a set of points in the workspace.

It is handy to compute the Minkowski sum of the robot K_r and the obstacles K_o :

$$K_r + K_o = \{ x + y \mid x \in K_r, y \in K_o \}$$

If there is a path outside the Minkowski sum, then it is possible to find a path to move the body without rotating it.



Figure 6.2: Use of the Minkowski sum of the robot and the obstacles.

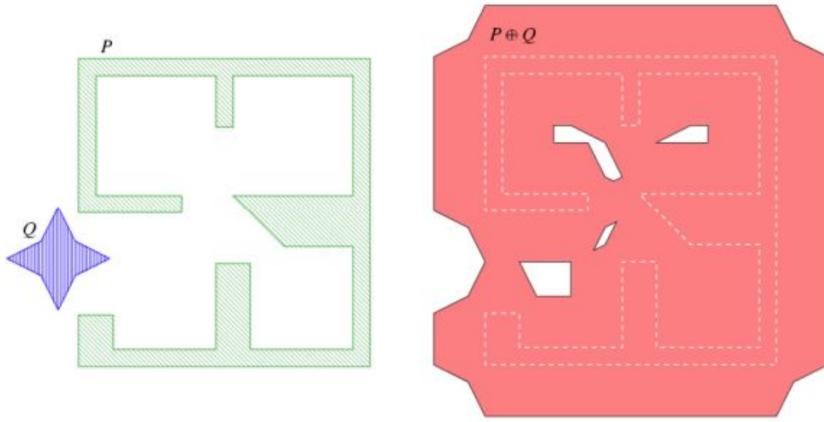


Figure 6.3: Another example, without a feasible path.

Thanks to the configuration space approach, motion planning can be reduced to a punctual trajectory problem. Given $q_i, q_f \in \mathcal{C}_{\text{free}}$, we want to find a continuous path $\gamma \in \mathcal{C}^0([0, 1], \mathcal{C}_{\text{free}})$ such that $\gamma(0) = q_i$ and $\gamma(1) = q_f$.

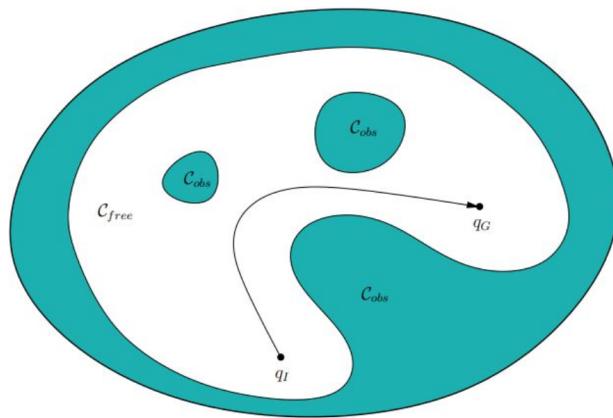


Figure 6.4: Example of a continuous path in the free configuration space.

Nevertheless, this still raises sub-questions. From a mathematical point of view, we can study the existence and optimality of such paths. From a numerical point of view, we need to choose a representation of the configuration space and obstacles, and to design algorithms to compute the path.

6.3 Motion planning algorithms

6.3.1 Algorithm paradigms

Three competing paradigms exist to solve the motion planning problem:

- *Combinatorial planning* consists in finding a global, exact planning. It is “moral” but computationally unrealistic.
- *Trajectory optimization* consists in finding a local planning. It is fast but remains a local approach.
- *Sampling-based motion planning* probabilistically explores the configuration space. It is very efficient but does not guarantee optimality. This is the most used approach in practice.

Such algorithms come with different possible level of completeness:

- A *complete* algorithm returns a solution if one exists, and reports a failure otherwise.
- A *semi-complete* algorithm returns a solution if one exists, but may run forever otherwise.
- For a *probabilistically complete* algorithm, if a solution exists, the probability that it will be found approaches 1 as the number of iterations approaches infinity.

6.3.2 Combinatorial planning

Combinatorial planning was developed in the 80s, and remains close to the mathematical problem studied. It is extremely efficient for low-dimensional problems, and are *complete* and even sometimes *resolution complete*. On the other hand, some are difficult to implement due to numerical issues, and most are intractable even with medium-sized problems.

Principle The main idea is to build finite roadmaps; a *roadmap* is a graph \mathcal{G} such that:

- each vertex is a configuration $q \in \mathcal{C}_{\text{free}}$,
- each edge $e(q, q')$ is a path $\gamma \in \mathcal{C}^0([0, 1], \mathcal{C}_{\text{free}})$ such that $\gamma(0) = q$ and $\gamma(1) = q'$.

Furthermore, a roadmap \mathcal{G} is a *topological graph* for the space $\mathcal{C}_{\text{free}}$ if:

- \mathcal{G} is *accessible*: from anywhere in $\mathcal{C}_{\text{free}}$, it is trivial to compute a path that reaches at least one point along any edge in \mathcal{G}
- \mathcal{G} is *homotopic-preserving*: every close path in $\mathcal{C}_{\text{free}}$ can be homotopically deformed into a path which is the concatenation of edges that form a cycle in \mathcal{G} .

In the following, we will assume that $\mathcal{C}_{\text{free}}$ is piecewise linear, that is, it is the union of a finite number of polytopes. The robot can be either a point, a polygonal, or even a disc, which has the ability to move without rotation (that is, it can be translated).

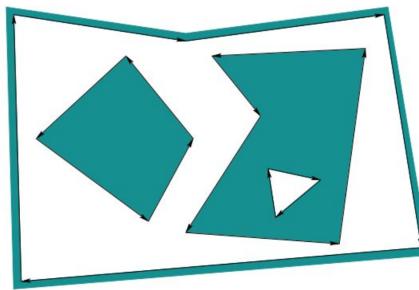


Figure 6.5: Example of piecewise linear obstacles.

Sweep A first approach to build a roadmap is to use a *sweep* algorithm. The idea is to use the plane sweep principle to efficiently determine where the rays terminate on the obstacles. We sort the vertices by abscissa coordinate, and we handle extensions from left to right, while maintaining a vertically sorted list of edges. This leads to a simple to implement, $O(n \log n)$ -time algorithm, where n is the number of vertices of the obstacles.

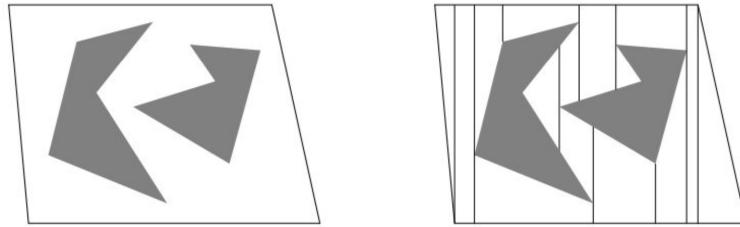


Figure 6.6: Example of sweeping.

6.3.3 Trajectory optimization (local planning)

Trajectory optimization, also known as local planning, use optimal control methods such as DDP or iLQR to find a local solution to the motion planning problem. It has a strong terminal cost, and does not guarantee to explore all the homotopy classes. To solve this issue, we need global optimization methods, that are only viable in small dimensions, or momentum methods, as introduced by El Khadir and Lassere in 2020.

6.4 Sampling-based motion planning (probabilistic planning)

6.4.1 Principle

The idea of sampling-based motion planning is to randomly sample the configuration space, and to connect the samples to build a roadmap. It uses a collision detector to separate planning from input geometry. Single-query algorithms incrementally sample and search the configuration space, while multiple-query algorithms precompute a complete sampling-based roadmap. Such algorithms have been developed since the 90s, with a proof of probabilistic completeness by Latombe in 1999.

Similarly to combinatorial planning, we build approximate roadmaps. The high-level procedure is as follows:

1. Sample a configuration
2. Check whether the configuration is in collision
3. Add the configuration to a graph (roadmap) of the nodes which are free configurations, and the edges of which are collision-free steerable paths

6.4.2 Sampling

The sampling step is fundamental in sampling-based motion planning. The sampling distribution must be chosen carefully, and depend on the specific motion planning task.

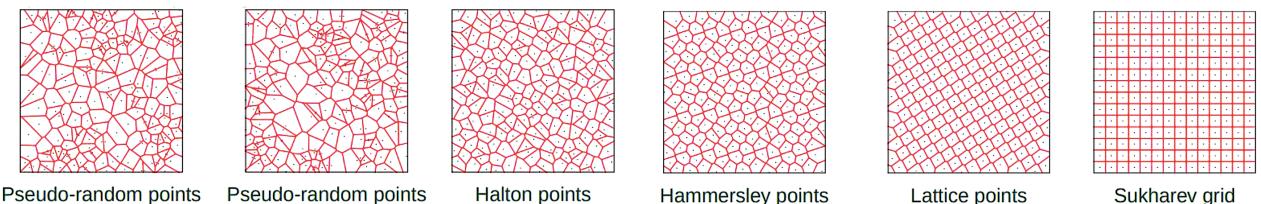


Figure 6.7: Example of different sampling procedures.

6.4.3 Collision detection

The choice of a good collision detector is crucial in sampling-based motion planning. The collision detector must quickly determine whether a configuration is in collision, to be able to efficiently sample valid configurations. The next chapter will dive deeper into this topic.

6.4.4 Probabilistic roadmap

The *Probabilistic Roadmap* (PRM) algorithm was introduced in 1994. It consists in building a roadmap by sampling the configuration space, and connecting the samples with collision-free paths. The roadmap is then used to find a path between the initial and final configurations, using a graph search algorithm such as Dijkstra's.

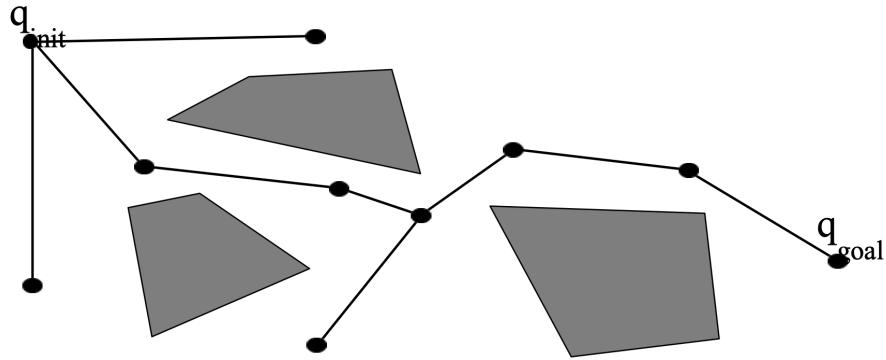


Figure 6.8: Example of the PRM algorithm.

A drawback of PRM is that it samples a lot of useless nodes, which quadratically increases the collision checking time. A solution is to use a concept called *visibility*, as for the art gallery problem. According to Chvátal's art gallery theorem:

"To guard a simple polygon with n vertices,
 $\lfloor n/3 \rfloor$ guards are always sufficient and sometimes necessary."

This lead to the Visibility-PRM algorithm, introduced by Laumond in 2000, which is more efficient than the original PRM.

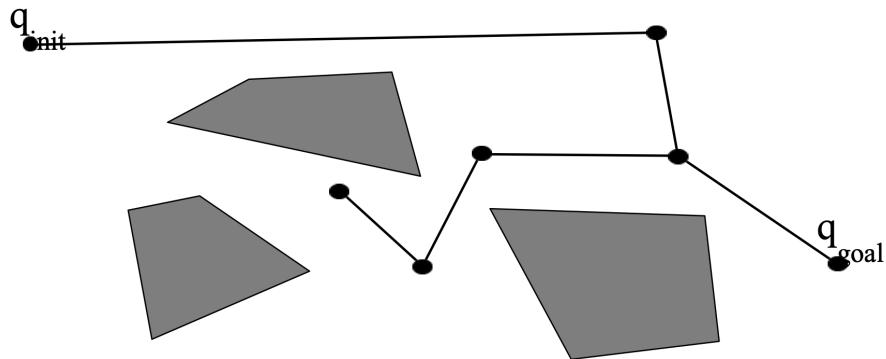


Figure 6.9: Example of the PRM algorithm.

6.4.5 Tree

PRM explores the whole space, which is good for multiple queries. For a single query, we want to explore "from the configuration", which lead to growing a tree from the initial configuration:

1. Pick a random node q of the tree
2. Pick a random direction (an element of $T_q\mathcal{C}$)
3. Steer with 1 unit in this direction

This is the principle of the *Rapidly-exploring Random Tree* (RRT) algorithm:

1. Pick a random node q_{dir} of the configuration space
2. Pick the q in the tree nearest to q_{dir}
3. Steer with 1 unit towards q_{dir}

It improves exploration thanks to Voronoï biases: it statistically explores in the *biggest Voronoï cell* of the tree.

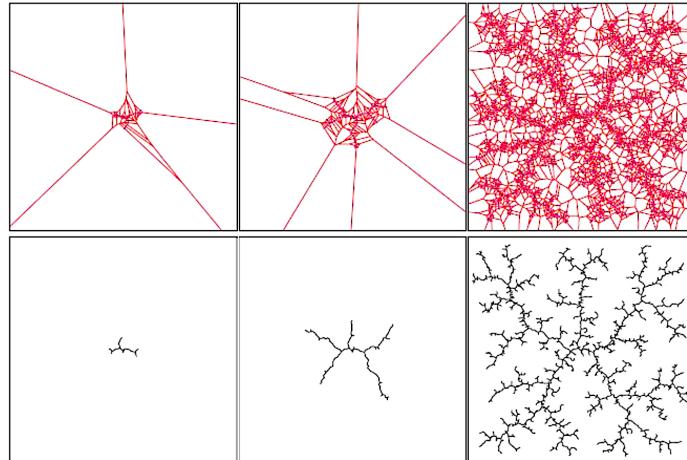


Figure 6.10: Example of Voronoï biases in the RRT algorithm.

Variants include early termination when an obstacle is encountered, bidirectional search, and so on.

7 Collision Detection

Collision detection is a subject at the center of physics simulators. To build a simulation of the robot in its environment, the main loop goes as follows:

1. Collision detection: finding contact points
2. Collision resolution: finding contact forces using physical principles
3. Time integration: update of the quantities of interest (position, velocity, etc.)

It is therefore crucial to have an efficient collision detection algorithm, that is to know whether two objects are in contact or not, and if so, to find the contact points.

Nevertheless, collision detection is a computational bottleneck in physics simulators. Resolving collision detection for one pair of objects takes a significant amount of time, especially for complex shapes, and the number of pairs to check grows quadratically with the number of objects. A general method to optimize such a process is to decompose one collision detection into two phases, the broad phase and the narrow phase. The broad phase uses simple geometric primitives to quickly discard pairs of objects that are far from colliding. The narrow phase then uses more complex geometric primitives to find the exact contact points.

7.1 The broad phase

7.1.1 Bounding volumes

As described previously, the broad phase uses *bounding volumes* (BVs) to prune collisions: we will only check the overlapping BVs, and leave fine-grain detection for the narrow phase. The main goal is to efficiently determine when two objects are far from overlapping, to prune such pairs of objects. Therefore, we will always use an over-approximation of the shape of the object, but never an under-approximation. We accept to mark as a likely collision two objects that are not intersecting, but we do not want to miss any.

The narrow phase will later determine which pairs are actually colliding, and which aren't, using exact representation of the objects. This two-phases approach allows for improved performance without sacrificing the precision of the detector.

7.1.2 Example of bounding volumes

The choice of a bounding volume is always a tradeoff between performance and precision. Larger BVs such as spheres or axis-aligned bounding boxes rely on extremely efficient algorithms, at the cost of a poor representation of the actual volume of the object. On the other hand, complex volumes such as the convex hull represent precisely the objects while requiring a longer time to compute collisions.

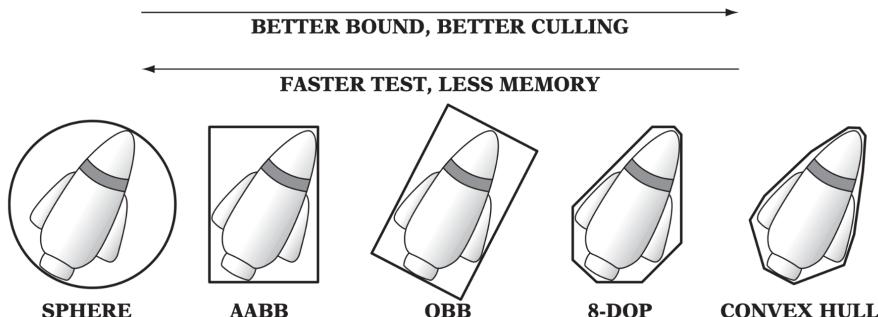
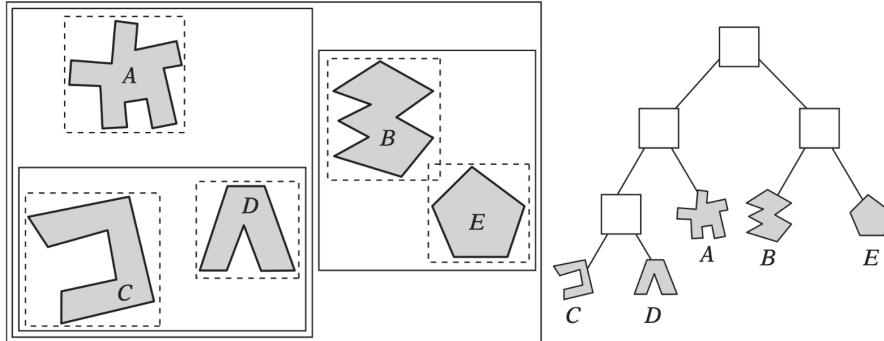


Figure 7.1: Examples of bounding volumes.

It is worth noting two important examples: Axis-Aligned Bounding Boxes (AABB) and Oriented Bounding Boxes (OBB), which can be efficiently computed, and represent accurately most objects we might have to deal with. Note that the choice of the axes is arbitrary.

7.1.3 Dynamic tree representation

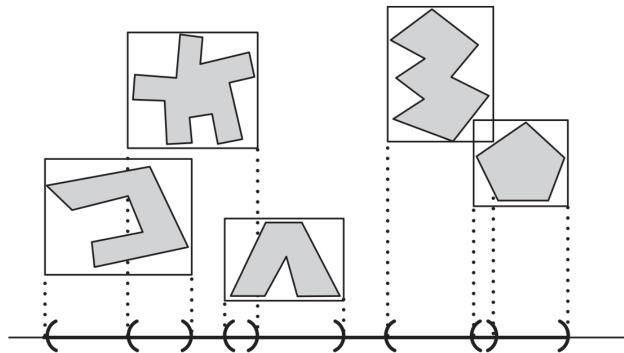
Given any representation of the bounding volumes of the objects of the scene, an efficient way to compute the potential collisions between these objects is to build a *dynamic tree*.



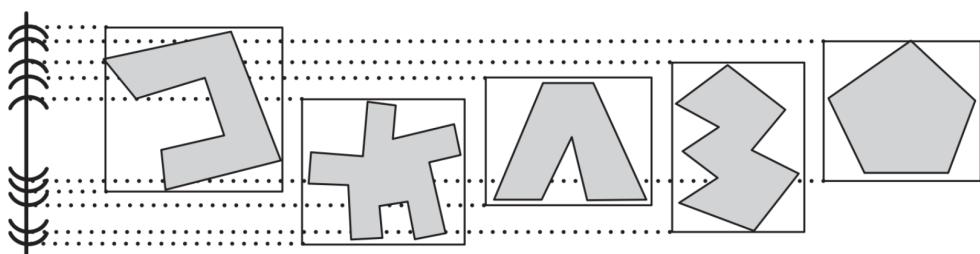
The idea is to split the scene into primitives that are easy to compute, for instance AABBs, such that any object is in exactly one subpart of the scene. Therefore, two objects are colliding only if they are children of the same node of the tree, which drastically reduces the number of collisions to check.

7.1.4 Sweep and Prune algorithm

When using exclusively AABBs, another efficient approach is to use the *Sweep and Prune* (SaP) algorithm. Its idea is to sort the bounding boxes along each axis of their axis. Objects may overlap at their starts or ends; if two objects are overlapping for every axis, one can conclude that their AABBs collide.



Such a method can become problematic if multiple objects are almost aligned on the same line.



7.2 The narrow phase

7.2.1 Introduction

The narrow phase is the second step of the collision detection process. It is used to determine the exact contact points between two objects that have been identified as potentially colliding in the broad phase. The goal is to find the exact contact points, and possibly the contact forces, between the two objects.

Formally, given two objects \mathcal{A}_1 and \mathcal{A}_2 we want to solve the following optimization problem:

$$\min_{x_1 \in \mathcal{A}_1, x_2 \in \mathcal{A}_2} \frac{1}{2} \|x_1 - x_2\|^2$$

If the minimum is zero, then the two objects are in contact. The solution to this problem is the contact point. If the shapes are meshes, we can rewrite the problem as:

$$\min_{x_1, x_2} \frac{1}{2} \|x_1 - x_2\|^2 \quad \text{s.t.} \quad \begin{cases} A_1 x_1 \leq b_1 \\ A_2 x_2 \leq b_2 \end{cases}$$

where the sizes of the matrix A_1 (respectively A_2) is the number of faces of the mesh of \mathcal{A}_1 (respectively \mathcal{A}_2).

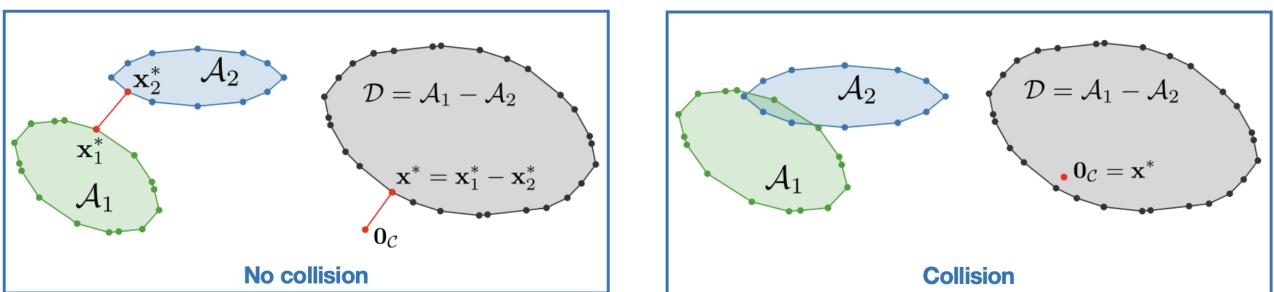
7.2.2 Minkowski-based algorithms

We can introduce the Minkowski difference of our two objects:

$$\mathcal{D} := \mathcal{A}_1 - \mathcal{A}_2 = \{x_1 - x_2 \mid x_1 \in \mathcal{A}_1, x_2 \in \mathcal{A}_2\}$$

Given the Minkowski difference, it is easy to see that the two objects are in contact if and only if the origin is in the Minkowski difference:

$$\mathcal{A}_1 \cap \mathcal{A}_2 \neq \emptyset \iff 0 \in \mathcal{D}$$



Hence, we can formally rewrite the optimization problem as a Minimum Norm Point (MNP) problem:

$$\min_{x_1 \in \mathcal{A}_1, x_2 \in \mathcal{A}_2} \frac{1}{2} \|x_1 - x_2\|^2 = \min_{x \in \mathcal{D}} \|x\|^2$$

In practice, the Minkowski difference is intractable; to solve this problem, we work implicitly with \mathcal{D} using the Frank-Wolfe algorithm.

7.2.3 Frank-Wolfe algorithm

The Frank-Wolfe algorithm is an iterative first-order algorithm for convex optimization under constraints. Generally, it solves a problem of the form:

$$\min_{x \in \mathcal{D}} f(x)$$

where both f and \mathcal{D} are convex. In the case of collision detection, the function f is the squared norm, and the set \mathcal{D} is the Minkowski difference of the two objects:

$$f(x) = \frac{1}{2} \|x\|^2 \quad \text{and} \quad \mathcal{D} = \mathcal{A}_1 - \mathcal{A}_2$$

The algorithm acts as a sort of constrained gradient descent. It goes as follows:

1. Compute the gradient $\nabla f(x_k)$ at current iterate x_k .
2. Compute the point $s_k \in \mathcal{D}$ which is the most in the direction of $-\nabla f(x_k)$, that is:

$$s_k = \arg \min_{y \in \mathcal{D}} \langle y, \nabla f(x_k) \rangle$$

3. Update the iterate x_k towards s_k to get x_{k+1} .

The intuition behind the second step is to minimize the linear approximation of the problem given by the first-order Taylor approximation of f at x_k .

Obviously, the difficulty is to compute the support points:

$$s_k = \arg \min_{y \in \mathcal{D}} \langle y, \nabla f(x_k) \rangle$$

Hopefully, this can be done efficiently in the case of a Minkowski difference. We introduce $S_{\mathcal{A}}(d)$, the support function of a set \mathcal{A} in the direction d , defined as:

$$S_{\mathcal{A}}(d) = \arg \min_{x \in \mathcal{A}} x^\top d$$

Note that:

$$\begin{cases} s_1 \in S_{\mathcal{A}_1}(d) \\ s_2 \in S_{\mathcal{A}_2}(-d) \end{cases} \implies s := s_1 - s_2 \in S_{\mathcal{D}}(d)$$

Hopefully, the support function can be computed efficiently for simple shapes.

7.2.4 GJK algorithm

The Gilbert-Johnson-Keerthi (GJK) algorithm is a popular algorithm to solve the narrow phase problem. It is an acceleration of the Frank-Wolfe algorithm, applied to an MNP.

8 Optimal Control