CHAPTER 3

Constraints

There are various kinds of constraints a material object may have to satisfy while in motion. If a constraint only concerns position in space, the constraint is called *holonomic*; if it concerns velocities, it is called *non-holonomic*. The effect of these constraints may be represented in various ways, that we are going to explore in this chapter. A fundamental notion, in this respect, consists of the infinitesimal motions that, in various ways, are allowed by the constraints. In particular, we are going to introduce the *virtual* displacements and velocities, that are the allowed infinitesimal displacements and velocities, at a given time t. The term "virtual" indicates the fact that such motions, while allowed by the constraints, are not necessarily going to take place. Some examples will clarify the matter.

Ex1: take a material point P on a ring as represented by Figure 1, whose equation in cartesian coordinates is given by:

$$x^2 + y^2 = R^2, \quad z = 0 \tag{3.1}$$

Condition (3.1), that the material point position belongs to the ring, represents a holonomic constraint, which is most easily implemented in polar coordinates. Indeed, adopting such coordinates, the motion on the ring, *P(t)=(x(t),y(t),z(t))*, takes the following simple form:

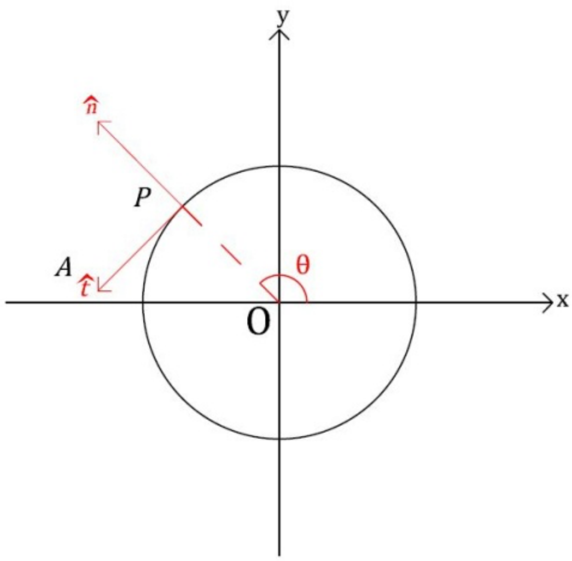

Figure 3.1. *Constrained motion of point P on the ring*

$$\begin{cases} x(t) = R\cos\vartheta(t) \\ y(t) = R\sin\vartheta(t) \\ z(t) = 0 \end{cases}$$

which implies $\begin{cases} \dot{x} = -R\dot{\vartheta}\sin\vartheta \\ \dot{y} = R\dot{\vartheta}\cos\vartheta \\ \dot{z} = 0 \end{cases}$ and $\begin{cases} \ddot{x} = -R\dot{\vartheta}^2\cos\vartheta - R\ddot{\vartheta}\sin\vartheta \\ \ddot{y} = -R\dot{\vartheta}^2\sin\vartheta + R\ddot{\vartheta}\cos\vartheta \\ \ddot{z} = 0 \end{cases}$

One may also like to introduce a mobile frame in which P is at rest, with axis $\hat{t}$ oriented like the tangent and axis $\hat{n}$ oriented like the normal of the ring at P, and with origin in P. One may then write:

$$\begin{cases} \hat{n} = \cos\vartheta\,\hat{\imath}_1 + \sin\vartheta\hat{\imath}_2 \\ \hat{t} = -\sin\vartheta\,\hat{\imath}_1 + \cos\vartheta\,\hat{\imath}_2 \end{cases}$$

if $\hat{\imath}_1, \hat{\imath}_2$ and $\hat{\imath}_3$ are the unit vectors of the reference frame in which the ring, not the point, is at rest. Then, the velocity and acceleration of P in the frame of the ring may also be expressed by:

$$\begin{cases} \vec{v} = R\dot{\vartheta}\hat{t} \\ \vec{a} = -R\dot{\vartheta}^2\hat{n} + R\ddot{\vartheta}\,\hat{t} \\ \vec{v}\cdot\hat{n} = 0 \end{cases}$$

Given the radius R, the condition $\vec{v} \cdot \hat{n} = 0$ identifies the motion on a circle. Therefore, the holonomic constraint (3.1) may also be expressed by a non-holonomic constraint.

Ex2: consider the mechanism represented in Figure 3.2, in which a hinge *A* is bound to move on a straight line, while holding a rigid bar of center *G* that is allowed to rotate. The constraint on *A* may be represented by the expression $(A - O) \cdot \hat{\imath}_2 = 0$.

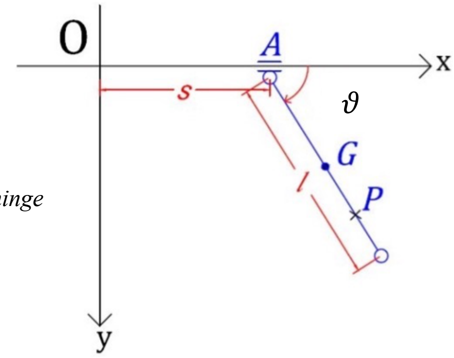

Figure 3.2. *bar attached to a sliding hinge*

With this choice, the coordinates (s, ϑ) automatically take into account the action of the constraint, and one can write:

$$(A - O) = (A(s, \vartheta) - 0) = s\hat{\imath}_1 \tag{3.2}$$

and

$$(G - O) = (G(s, \vartheta) - 0) = \left(s + \frac{l}{2}\cos\vartheta\right)\hat{\imath}_1 + \frac{l}{2}\sin\vartheta\,\hat{\imath}_2 \tag{3.3}$$

Recalling that the velocity of *A* can be expressed as

$$\vec{v}_A = \frac{\partial(A - O)}{\partial s}\frac{ds}{dt} + \frac{\partial(A - O)}{\partial\vartheta}\frac{d\vartheta}{dt} = \dot{s}\hat{\imath}_1 \tag{3.4}$$

while the velocity of *G* is given by:

$$\vec{v}_G = \frac{\partial(G - O)}{\partial s}\frac{ds}{dt} + \frac{\partial(G - O)}{\partial\vartheta}\frac{d\vartheta}{dt} = \dot{s}\hat{\imath}_1 + \dot{\vartheta}\frac{l}{2}(-\sin\vartheta\,\hat{\imath}_1 + \cos\vartheta\,\hat{\imath}_2)$$

$$= \left(\dot{s} - \frac{l}{2}\dot{\vartheta}\sin\vartheta\right)\hat{\imath}_1 + \frac{l}{2}\dot{\vartheta}\cos\vartheta\,\hat{\imath}_2 \tag{3.5}$$

Given a point *P* of coordinates *(q₁, q₂)*, in a two-dimensional reference frame of origin *O*, an elementary displacement, is in general expressed as:

$$\boxed{dP := \frac{\partial(P - O)}{\partial q_1}dq_1 + \frac{\partial(P - O)}{\partial q_2}dq_2} \tag{3.6}$$

In example Ex2, $q_1 = s$ and $q_2 = \vartheta$ are the coordinates which automatically take into account the action of the constraints. Then, one may easily consider instantaneous motions consistent with the constraints, in their configuration at a given time t. This amounts to neglecting the previous as well as the future evolution of the system, and to focus on the constraints, as if they were frozen in the configuration realized at time t. The virtual velocities, *i.e.* all those allowed in this situation, are expressed by:

$$\vec{v}_p' := \sum_{k=1}^{n} \frac{\partial P}{\partial q_k} \frac{\delta q_k}{\delta t} \tag{3.7}$$

where n, the number of coordinates that are *necessary and sufficient* to describe the motion of the system, is known as number of *degrees of freedom*; the prime denotes a *virtual* quantity and δq_k stands for *virtual* (*i.e.* merely possible) *coordinate variation,* not for its real variation, that is instead denoted by dq_k. Combining all virtual variations of all coordinates, we can express the virtual displacemente of point P as:

$$\delta P := \sum_{k=1}^{n} \frac{\partial P}{\partial q_k} \delta q_k \tag{3.8}$$

The vector δP differs from the actual displacement, because it is required to agree with the constraint only at time t, rather than at all times. For instance, δP could be constrained to be tangent to a given surface, but there are many different ways to be so, and δP does not discriminate among them. Moreover, δP may obey the constraints at time t, but if constraints change in time, it would not necessarily obey them at later times.

Ex3: consider a point that slides along a rotating rigid beam. This constitutes a mobile constraint, that is expressed by

$$(P - O) \wedge \hat{u}(t) = 0 \tag{3.9}$$

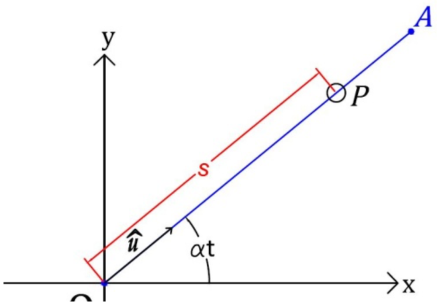

Figure 3.3. *Point mass sliding on rotating guide*

where $\hat{u}(t) = (\cos \alpha t\,, \sin \alpha t)$, and

$(P - O)(s, \vartheta) = (P - O)(s, t) = s \cos \alpha t\, \hat{\imath}_1 + s \sin \alpha t\, \hat{\imath}_2$

since

$$(P - O) \wedge \hat{u} = \begin{vmatrix} \hat{\imath} & \hat{\jmath} & \hat{k} \\ x & y & 0 \\ \cos \alpha t & \sin\alpha\, t & 0 \end{vmatrix} = \hat{k}(x \sin \alpha t - y \cos \alpha t) \tag{3.10}$$

Hence, the constraint means:

$$y \cos \alpha t - x \sin \alpha t = 0 \tag{3.11}$$

Consequently, the true velocity of the point is expressed by:

$$\vec{v}_P = \frac{\partial (P - O)}{\partial s} \dot{s} + \frac{\partial (P - O)}{\partial t} \frac{dt}{dt} = \dot{s}[\cos \alpha t\, \hat{\imath}_1 + \sin \alpha t\ \hat{\imath}_2] + s\alpha[\cos \alpha t\, \hat{\imath}_2 - \sin \alpha t\ \hat{\imath}_1] \tag{3.12}$$

How many degrees of freedom do we have? One point in the plane has two, but equation (3.11) reduces them by one, so we only have one degree of freedom. This is understood considering that the motion perpendicular to the bar is not a freedom of P , this point must stay with the bar. Then, in the reference frame of the bar, only one displacement is allowed: δs. Let us now fix s and t, and let us consider all motions allowed by the constraint at this time t. We note that the corresponding $\vec{v}_P'$ must be tangent to the bar, that makes the angle αt, so one can write:

$$\vec{v}_p' = \sum_{k=1}^{n} \frac{\partial P}{\partial s} \frac{\delta s}{\delta t} \tag{3.13}$$

which represents all velocities tangent to the bar at time t. The corresponding virtual displacement, parallel to the bar, is given by:

$$\delta P = \sum_{k=1}^{n} \frac{\partial P}{\partial s} \delta s \qquad (3.14)$$

We conclude that the real displacement is not included in the set of virtual displacements, when the constraint depends on time. This is a general fact.

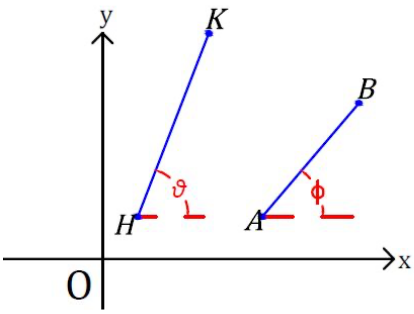

Figure 3.4. *disconnected bars*

Ex4: On a plane, two rigid bars are represented only by (A, φ) and (H, ϑ), which makes 3+3=6 coordinates, because in the plane, A and H have two coordinates and the configuration of a rigid bar is fully determined by one of its points together with one angle that the bar makes with respect to the reference frame. Let us add the constraints that the points K and A are joined by a hinge, that H rotates about a hinge in the origin of the axes, and that B slides on the x-axis. These constraints are mathematically expressed by: $H=0$, $K=A$ and $B=(x_B,0)$. In this case, we have two external constraints and one internal constraint, that reduce the degrees of freedom from 6 to 1, hence the mechanism is fully described by a single coordinate. We have indeed imposed five conditions:

$$\begin{cases} (B-O)\cdot \hat{\imath}_2 = 0 & \text{says that } (B-O) \text{ remains parallel to } \hat{\imath}_1 \\ x_H = 0 = y_H & \text{says that } H \text{ is anchored at O} \\ x_K = x_A & \text{says that } A \text{ and } K \text{ have same horizontal coordinate} \\ y_K = y_A & \text{says that } A \text{ and } K \text{ have same vertical coordinate} \\ y_B = 0 & \text{says that } B \text{ lies in the horizontal axis} \end{cases} \qquad (3.15)$$

reducing the possible independent variables from six to one. In this situation, we can now write:

$$x_A - L\cos\vartheta = 0,\ y_A - L\sin\vartheta = 0,\ y_A + l\sin\varphi = 0$$

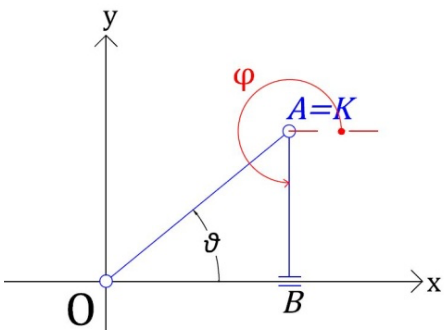

Figure 3.5. *connected bars*

The second and the third equations of (16) yield

$$L\sin\vartheta = -l\sin\varphi$$

which expresses ϑ as a function of φ. Then, in the case $L > l$, the angle φ can vary over the whole range *[0,2π],* while ϑ cannot reach $\pm\pi/2$.

Ex5: Consider a rigid bar hanging on a point A, that is allowed to slide on the x-axis. Let the x-axis also act as a wall that does not allow the bar to move in the upper vertical half plane, identified by negative y coordinates. This is expressed imposing:

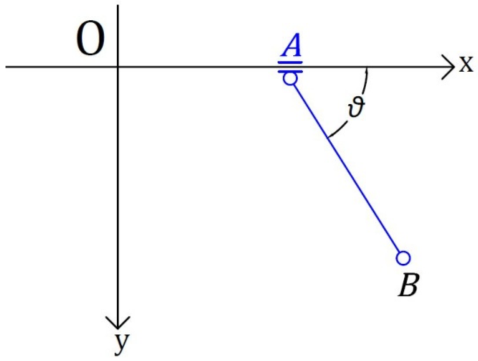

Figure 3.6. *bar attached to a sliding hinge*

$$(A-O)\cdot\hat{\imath}_2=0 \qquad (3.16a)$$

and the *unilateral constraint*

$$(B-A)\cdot\hat{\imath}_2\geq 0 \qquad (3.16b)$$

which means that the bar stays below the *x* axis, or that ϑ is bound to stay in the interval $[0,\pi]$. This kind of constraints acts only when the border of possible configurations is reached, *i.e.* for $\vartheta=0,\pi$. In these cases, the virtual displacements are only directed downwards, and this is the reason why the constraint is called *unilateral*.

Their opposite displacements are not allowed, hence the term "*unilateral*". Differently, we call *reversible* a virtual displacement if its opposite is also a virtual displacement.

If all virtual displacements or velocities are reversible, the constraint is called *bilateral.*

NOTE: the virtual displacement obeys the standard expression

$$\delta P=\frac{\partial(P-O)}{\partial s}\delta s+\frac{\partial(P-O)}{\partial\vartheta}\delta\vartheta$$

even in the case of the unilateral constraints. Only at $\vartheta=0$ or π, $\delta\vartheta$ is affected by the condition (3.16b).

The characteristic equation (2.51) of the instantaneous motion of a rigid body or rigid system $\mathcal{B}$ is:

$$\vec{v}_P=\vec{v}_Q+\vec{\omega}\wedge(P-Q)$$

where P and Q are two points of $\mathcal{B}$, while $\vec{\omega}$ is the angular velocity with respect to a fixed reference frame. The rigidity property is thus a constraint that can be expressed by:

$$\vec{v}'_P=\vec{v}'_Q+\vec{\omega}'\wedge(P-Q) \qquad (3.17)$$

where $\vec{v}_Q{}'$ and $\vec{\omega}'$ are arbitrary vectors, since the virtual quantities have to obey only the (rigidity) constraint. If we further add the constraint that a point $C\in\mathcal{B}$ is fixed, virtual motions have to obey:

$$\vec{v}'_P=\vec{\omega}'\wedge(P-C) \qquad (3.18)$$

where now only $\vec{\omega}'$ is an arbitrary vector.

Ex6: Consider a rigid bar with one end hinged at a fixed point O, and with the other end constrained to slide on a vertical rail. The rotation of the bar that appears to be possible is only virtual, but indeed a virtual rotation is allowed by the constraint in the configuration illustrated by Figure 3.7, because the constraint requires:

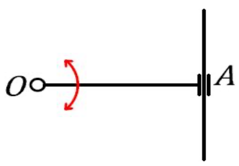

Figure 3.7. *constrains allow virtual rotation*

$$\vec{v}'_P=\vec{\omega}'\wedge(A-O) \quad \text{with} \quad \vec{\omega}'=\omega'\hat{\imath}_3$$

which has a solution if the rotation occurs in the plane containing the rail and the point O. No velocity is real, because the bar should stretch to allow a rotation and, being rigid, it cannot do so. However, from the viewpoint of displacements, the initial elongation of the bar would be second order infinitesimal, hence negligible to the first order infinitesimal displacement, as expressed by the following first order relations:

$$dP = dQ + \vec{\varepsilon} \wedge (P - Q), \quad \vec{\varepsilon} = \vec{\omega} dt \tag{3.19}$$

and

$$\delta P = \delta Q + \vec{\varepsilon}' \wedge (P - Q) \tag{3.20}$$

where $\vec{\varepsilon}$ is the real infinitesimal rotation vector and $\vec{\varepsilon}'$ any angular velocity parallel to $\hat{\imath}_3$. The situation represented by Figure 3.8 is different, because the rail implies that the distance between O and A has to undergo a first order variation, under an infinitesimal rotation of the bar. Such an elongation is not negligible with respect to the infinitesimal displacements, hence it is not allowed by the constraints.

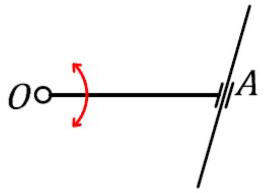

Figure 3.8. *constraints allow no virtual displacement, hence no virtual work*

DEFINITION 3.1: *A system of "<u>free or generalized coordinates</u>" for a set of configurations $\mathbb{C}$ is a map $m: \mathbb{R}^N \to \mathbb{C}$ that associates any point $(q_1, \cdots, q_N) \in U \subset \mathbb{R}^N$ with a configuration in $\mathbb{C}$, where U is open and connected, and may have boundaries.*

For systems that move in time, one requires the function $P(q_1(t), \cdots, q_N(t); t)$, denoting a point of the object of interest at time *t*, to be smooth (no less than twice differentiable, because accelerations are second derivatives of positions), and the different coordinates to be independent.

Depending on the kind and number of constraints, one has different situations, that are briefly illustrated by Figure 3.9. In the case in which the active constraints are not enough to block all degrees of freedom, the system of interest is called *unstable*; in this case, motion is possible, although it may be limited. If the system is moved away from its configuration, it cannot spontaneously restore its original state. When the active constraints are the minimum necessary to block all degrees of freedom, motion cannot take place, and the system is called *determinate*, or *isostatic*, which basically means just right to be static. If the active constraints are more numerous than those strictly necessary to block all degrees of freedom, the system is called *indeterminate*, or *hyperstatic*. The term "determined" refers to the fact that in this case one may directly proceed to the evaluation of the constraining forces, also called constraint reactions. In the "indeterminate" case, direct evaluation leads instead to infinitely many solutions for the values of the constraint reactions, and special procedures and further information must be introduced in order to overcome this indeterminacy.

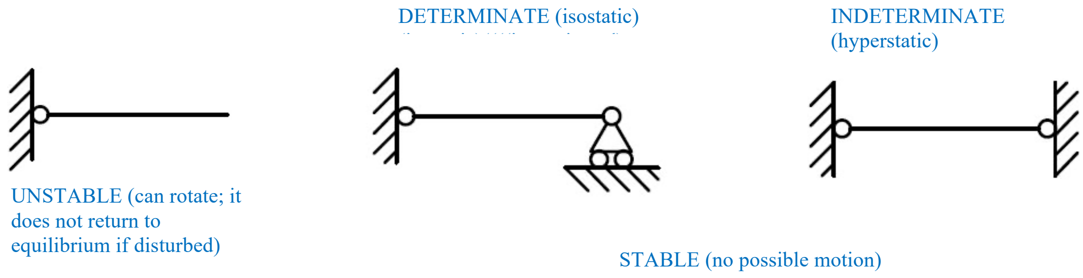

Figure 3.9. *possible static configurations.*

HOLONOMIC BILATERAL CONSTRAINTS: This kind of constraints is defined by the fact that the generalized coordinates obey a relation like:

$$f(q_1, \cdots, q_N; t) = 0 \tag{3.21}$$

where t represents time, and the other variables are coordinates characterizing the configuration of the system. Holonomic means that f does not depend on velocities, while bilateral means that all virtual displacements and velocities are reversible. The consequence of these facts is that such a constraint is expressed by an equation, rather than by an inequality. If t does not appear in Eq.(3.21), the constraint is called *time independent.*

NONHOLONOMIC CONSTRAINTS: in this case velocities or infinitesimal displacements are constrained.

The above is best understood considering explicit examples.

Ex7: Rigid body: this is the constraint for which any two points P_i and P_j of the body preserve their distance in time. Mathematically this is expressed by

$$|(P_i - P_j)| = C_{ij} = constant$$

which is a bilateral holonomic constraint.

Ex8: Constant speed. Mathematically this constraint is expressed by

$$|\vec{v}| = constant$$

which is a non-holonomic constraint.

Ex9: the condition that the velocity be orthogonal to a given direction, *e.g.*

$$\vec{v} \perp \hat{\imath}_h$$

Is another non-holonomic constraint.

Ex10: The motion of a car has constrained velocities: for instance, it cannot move perpendicularly to itself. Therefore, it is affected by *nonholonomic* constraints.

Ex11: Rolling without slipping of one surface on another is nonholonomic. Here, at a given time t, the

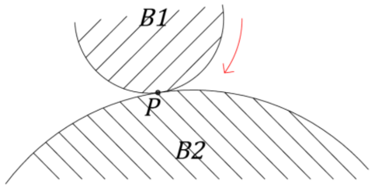

Figure 3.10. *surfaces rolling on each other*

contact point P coincides with one point of the body B_1, P_1 say, and with one point P_2 of the body B_2. As time goes P, P_1, P_2 change, and the constraint is expressed by the equality of the velocities of P_1 and P_2, which is to say: $\vec{v}_1 = \vec{v}_2$.

Ex12: If contact is preserved, but the two surfaces may slip, the constraint is mathematically expressed by $\vec{v}_1 \cdot \hat{n} = \vec{v}_2 \cdot \hat{n}$, where $\hat{n}$ is normal to both surfaces in P.

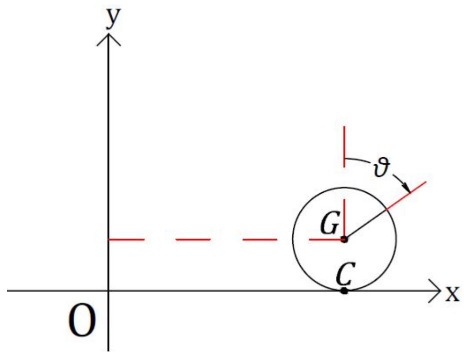

Figure 3.11. *disk rolling on a horizontal rail*

Ex13: Consider a rigid disc; its configuration is determined by the coordinates of its center G and by ϑ, the angle between the normal to the *x* axis at the contact point C and one radius of the disk. The constraint, represented by

$$y_G - R = 0$$

is holonomous. If the *x* axis is time independent and the disk does not slip on it, we have the non-holonomous constraint $\vec{v}_C = \mathbf{0}$, where it is to be noted that the contact point is not a given point of the disk, but it changes in time. In any event, C is the only point of the disk that is at rest in the reference frame, hence it is the instantaneous center of rotation. One further has:

$$\begin{cases} \vec{v}_G = \dot{x}_G\,\hat{\imath}_1 \\ \vec{\omega} = -\dot{\vartheta}\,\hat{\imath}_3 \\ (C - G) = -R\,\hat{\imath}_2 \end{cases} \tag{3.22}$$

which yields

$$\vec{v}_C = \vec{v}_G + \vec{\omega} \wedge (C - G) = \dot{x}\hat{\imath}_1 + \left(-\dot{\vartheta}\hat{\imath}_3\right) \wedge (-R\hat{\imath}_2) = \left(\dot{x}_G - R\dot{\vartheta}\right)\hat{\imath}_1 \tag{3.23}$$

Imposing the non-holonomous constraint $\vec{v}_C = \mathbf{0}$, we get $\dot{x} = R\dot{\vartheta}$ *i.e.* $x = R\vartheta + constant$, where the $constant$ is determined by the initial conditions. One may as well take $x = R\vartheta + k$ as a constraint, which is holonomous. This possibility of expressing one constraint as holonomous and as non-holonomous is not always allowed, especially in 3D. For instance, a constraint on velocities may pose no constraint on positions, as in the case of cars motion.

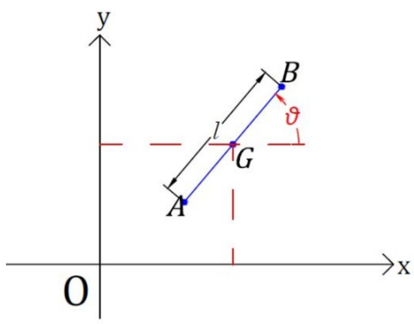

Figure 3.12. *representation of skate*

Ex14: Consider a skate, idealized by a straight-line segment of length l, whose center is G. The forward motion of the skate on ice enjoys the property that $\vec{v}_G$ is parallel to the line segment. This is mathematically expressed by:

$$\vec{v}_G = \dot{x}_G\hat{\imath}_1 + \dot{y}_G\hat{\imath}_2, \qquad \vec{v}_G \wedge (B - A) = \mathbf{0}$$

where

$$\begin{aligned}(B-A) &= l\cos\vartheta\,\hat{\imath}_1 + l\sin\vartheta\hat{\imath}_2 = (B-G)+(G-A)\\ &= \left[\left(x+\frac{l}{2}\cos\vartheta\right)\hat{\imath}_1 + \left(y+\frac{l}{2}\sin\vartheta\right)\hat{\imath}_2 - x\hat{\imath}_1 - y\hat{\imath}_2\right] + x\hat{\imath}_1 + y\hat{\imath}_2\\ &\quad - \left[\left(x-\frac{l}{2}\cos\vartheta\right)\hat{\imath}_1 + \left(y-\frac{l}{2}\sin\vartheta\right)\hat{\imath}_2\right]\end{aligned}$$

that can also be expressed as

$$(x\hat{\imath}_1 + y\hat{\imath}_2)\wedge(l\cos\vartheta\,\hat{\imath}_1 + l\sin\vartheta\hat{\imath}_2) = l\sin\vartheta\,\dot{x}\,\hat{\imath}_3 - l\cos\vartheta\,\dot{y}\,\hat{\imath}_3 = \mathbf{0}$$

i.e.

$$\dot{x}\sin\vartheta - \dot{y}\cos\vartheta = 0 \tag{3.24}$$

Could this ever be written as a holonomous constraint? Could this ever be obtained as the time derivative of a constraint for coordinates only? Something like $f(x,y,\vartheta)=0$? One can prove that this is impossible.

The number of generalized coordinates for a determined system that is subjected to holonomic constraints, so that it preserves in time its state of motion or rest, is the number of *degrees of freedom.* This number equals the number of possible virtual displacements.

Note: **Ex3** reveals the general fact that the real displacements under time dependent constraints are not included among the virtual displacements.

Note: **Ex6** highlights the technical (not common language) nature of the terminology used in this course. For instance, “real” and “virtual” are referred to the model, not to the “true nature” of things. This is necessary, because we can only understand the properties of the models we develop to describe things and phenomena. Mechanical models happen to be among the absolutely most effective models one can think of. Nevertheless, it is important to keep an eye on what we understand of things, in order to continually test and improve such models. In the case of **Ex6**, the virtual displacements better represent, in a sense, reality: a real (in the sense of natural language) bar constrained to rotate with one end on a vertical rail will in fact move a little, if loaded. The motion will stop when the strain in it balances the load. In the model, this displacement is possible because the elongation of the bar is infinitesimal of second order, hence negligible. In the real world, the higher the rigidity of the object, the smaller the true displacements will be. Then, “real” in the model and real in the world match when the rigidity is so high that no load we can produce moves the bar. Mathematically this situation is properly represented by the infinite rigidity described by our definition of rigid object.

In Italian:

Time independent = scleronomo
Time dependent = reonomo