Journal's Title, Vol. x, 20xx, no. xx, xxx - xxx
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/

An algebraic procedure for reducing the Boltzmann–Hamel equations in nonholonomic systems

Federico Talamucci

DIMAI, Dipartimento di Matematica e Informatica "Ulisse Dini"
University of Florence, Italy
Viale G. B. Morgagni, 67/a, 50134 Florence

Abstract

Nonholonomic systems are increasingly worth considering, because of their application for a wide class of models in mechanical engineering, joint construction, robotics, control for wheels and many other types of mechanisms. A mathematical method suitable for treating such models is based on the definition of quasi–velocities and the formlation of the Boltzmann–Hamel equations.

The paper pursues the aim of discussing the appropriate choice of quasi–velocities: the possibility of simplifying the mathematical problem via the definition of specific quasi–velocities is discussed, mainly focussing on the linear structure of part of the system and developing algebraic procedure.

The technique formulated in the paper is then applied to some models which are exemplars in literature for nonholonomic constrained systems.

Subject Classification: 37J60, 70F25, 70H03

Keywords: nonholonomic constraints, quasi–velocites, Boltzmann–Hamel equations.

1 Introduction

The main purpose of the paper is the mathematical examination of a significant set of equations especially related to the equation of motions of systems subject

to nonholonomic constraints. This kind of restrictions, as it is known, affects directly the velocities of the system and do not confine the possible placements, as the holonomic constraints do.
It is not limiting the assumption that the kinematic resctrictions depend linearly on the velocities: this is commonly true in most of the actual circumstances.
Such a linear dependence gives the possibility on the one hand of defining a new set of kinematic variables (quasi–velocities), which turn out advantageous for the mathematical problem, on the other hand of obtaining a set of equations with the minimum number of variables avoiding the presence of the Lagrange multipliers (Boltzmann–Hamel equations), even though the constraints are not geometrical.
We find it convenient to draw concisely the equations of motion, so that the role of the used variables and of the different terms treated in the analyis will be understandable.
In our analysis the possibility of reducing the mathematical problem of equations of motion concerns the elimination of one or more lagrangian coordinates: the technique is based on algebraic procedures and differs from the differential geometry methods performed in [1], [4] for the case of nonholonomic systems with symmetry.
Since the proposed procedure decisively depends on the particular structure of the studied system (namely the Lagrangian function and the kinematic restrictions), it is worth testing it on nonholonomic systems largely present in literature.
By virtue of these preliminary observations, the text expounds the following points.

- The equations of motion for a system constrained with fixed nonholonomic constrained are introduced.

- The possibility of defining in a suitable way the quasi–velocities in order to simplify the mathematical probelm is discussed.

- A certain number of instances are presented in order to inspect the possibility of implementing the method of reduction and to compare the effects with the techniques commonly performed in literature.

1.1 The equations of motion

We consider a mechanical system subjected to fixed coinstraints, both of geometrical type and of kinematic type. Following the same procedure as in [3], one first exploits the geometrical constraints in order to establish ℓ local lagrangian coordinates $q_1 \dots, q_\ell$ and to write the Lagrangian function

$$\mathcal{L} = T + U, \quad T(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\dot{\mathbf{q}} \cdot A(\mathbf{q})\dot{\mathbf{q}} \tag{1}$$

where T is the kinetic energy and U the potential of the applied forces. The kinematic restrictions are assumed to be expressed by the linear equations

$$\alpha(\mathbf{q})\dot{\mathbf{q}} = \mathbf{0}_\mu, \tag{2}$$

involving the generalized velocities $\dot{\mathbf{q}}$, where α is a μ–by–ℓ matrix, with $\ell > \mu$. The mathematical problem of solving the equations of motion associated with the Lagrangian function $\mathcal{L}$ can be improved via the definition of the quasivelocities

$$\eta_1 = \sum_{j=1}^{\ell} z_{1,j}(\mathbf{q})\dot{q}_j, \quad \dots \quad \eta_\sigma = \sum_{j=1}^{\ell} z_{\sigma,j}(\mathbf{q})\dot{q}_j \qquad \textit{or} \quad \boldsymbol{\eta} = Z(\mathbf{q})\dot{\mathbf{q}} \tag{3}$$

where $z_{i,j}$ are required to guarantee

$$\delta = det \begin{pmatrix} Z \\ \boldsymbol{\alpha} \end{pmatrix} \neq 0 \tag{4}$$

In this way, each set of kinetic variables $\dot{\mathbf{q}}$ is linked to a singular set of quasi–velocities $\boldsymbol{\eta}$, and *vice versa*. More precisely, (3) and (2) give

$$\begin{pmatrix} Z \\ \alpha \end{pmatrix} \dot{\mathbf{q}} = \begin{pmatrix} \boldsymbol{\eta} \\ \mathbf{0}_\mu \end{pmatrix} \quad \text{and} \quad \dot{\mathbf{q}} = \begin{pmatrix} Z \\ \alpha \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\eta} \\ \mathbf{0}_\mu \end{pmatrix} = \begin{pmatrix} \Gamma & \Theta \end{pmatrix} \begin{pmatrix} \boldsymbol{\eta} \\ \mathbf{0}_\mu \end{pmatrix} \tag{5}$$

where $\Gamma(\mathbf{q})$ is a $\ell \times \sigma$ matrix and $\Theta(\mathbf{q})$ is a $\ell \times \mu$ matrix. In order to write the equations of motions, we extract from (5)

$$\dot{\mathbf{q}}(\mathbf{q}, \boldsymbol{\eta}) = \Gamma(\mathbf{q})\boldsymbol{\eta} \tag{6}$$

and define

$$\widetilde{\mathcal{L}}(\mathbf{q}, \boldsymbol{\eta}) = \mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \boldsymbol{\eta})) = \frac{1}{2}\boldsymbol{\eta} \cdot A_\Gamma(\mathbf{q})\boldsymbol{\eta} + U(\mathbf{q}) \tag{7}$$

where

$$A_\Gamma(\mathbf{q}) = \Gamma^T A \Gamma \tag{8}$$

By using the formulae $\nabla_{\dot{\mathbf{q}}}\mathcal{L} = Z^T \nabla_{\boldsymbol{\eta}} \widetilde{\mathcal{L}}$, $\nabla_{\mathbf{q}}\mathcal{L} = \nabla_{\mathbf{q}}\widetilde{\mathcal{L}} + (J^T_{\mathbf{q}}\boldsymbol{\eta})\big|_{\dot{\mathbf{q}}=\Gamma\boldsymbol{\eta}} \nabla_{\boldsymbol{\eta}}\widetilde{\mathcal{L}}$ where $J^T_{\mathbf{q}}\boldsymbol{\eta}(\mathbf{q}, \boldsymbol{\eta})$ is the matrix with elements $\left((J^T_{\mathbf{q}}\boldsymbol{\eta})\big|_{\dot{\mathbf{q}}=\Gamma\boldsymbol{\eta}} \right)_{i,j} = \sum_{k=1}^{\ell}\sum_{s=1}^{\sigma} \gamma_{k,s} \frac{\partial z_{j,k}}{\partial q_i}\eta_s$, $i = 1, \dots, \ell$, $j = 1, \dots, \sigma$ ($\gamma_{k,s}$ are the entries of Γ), we can write

$$\Gamma^T \left(\frac{d}{dt}(\nabla_{\dot{\mathbf{q}}}\mathcal{L}) - \nabla_{\mathbf{q}}\mathcal{L} \right) = \mathbf{0}_\sigma \tag{9}$$

in terms of the demanded variables (we use $Z\Gamma = \mathbb{I}_\sigma$, see (5)):

$$\frac{d}{dt}\left(\nabla_{\boldsymbol{\eta}}\widetilde{\mathcal{L}}\right) - \Gamma^T \nabla_{\mathbf{q}}\widetilde{\mathcal{L}} + \Gamma^T \left(\dot{Z}^T - J_{\mathbf{q}}^T \boldsymbol{\eta}\right) \nabla_{\boldsymbol{\eta}}\widetilde{\mathcal{L}} = \mathbf{0}_\sigma \tag{10}$$

We can identify the holonomic case with $\sigma = \ell$, $\mu = 0$ and $Z = \mathbb{I}_\ell$ so that $\Gamma = \mathbb{I}_\ell$ and (10) are the ordinary Euler–Lagrange equations of motion with $\boldsymbol{\eta} = \dot{\mathbf{q}}$. It is worth mentioning that (10) entails the energy balance $\dfrac{d}{dt}\left(\boldsymbol{\eta}\cdot\nabla_{\boldsymbol{\eta}}\widetilde{\mathcal{L}} - \widetilde{\mathcal{L}}\right) = 0$.
We will need to write (10) more explicitly, sorting the terms in a suitable way. In order to lighten the script, we will hereafter use the following shortening: for a general matrix n–by–m matrix C whose elements are $c_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, m$ it is

$$C_{(i)} = (c_{i,1}, \dots, c_{i,m}) \quad i\text{–th row of } C, \quad C^{(j)} = \begin{pmatrix} c_{1,j} \\ \dots \\ c_{n,j} \end{pmatrix} \quad j\text{–th column of } C$$

(whereas $\cdot_j$ or $(\cdot)_j$ is the j–th element of one vector). By means of the computation $\nabla_{\boldsymbol{\eta}}\widetilde{\mathcal{L}}(\mathbf{q},\boldsymbol{\eta}) = A_\Gamma\boldsymbol{\eta}$, $\nabla_{\mathbf{q}}\widetilde{\mathcal{L}}(\mathbf{q},\boldsymbol{\eta}) = \frac{1}{2}J_{\mathbf{q}}^T(A_\Gamma\boldsymbol{\eta})\boldsymbol{\eta} + \nabla_{\mathbf{q}}U$, (10) takes the structure

$$A_\Gamma(\mathbf{q})\dot{\boldsymbol{\eta}} + Q(\mathbf{q},\boldsymbol{\eta}) - \Gamma^T\nabla_{\mathbf{q}}U = \mathbf{0}_\sigma. \tag{11}$$

with $Q(\mathbf{q},\boldsymbol{\eta}) =$

$$\left\{\sum_{k=1}^{\ell}(\Gamma\boldsymbol{\eta})_k\left[A_\Gamma\left(\frac{\partial Z}{\partial q_k} - J_{\mathbf{q}}Z^{(k)}\right)\Gamma + \frac{\partial A_\Gamma}{\partial q_k}\right]^T - \frac{1}{2}\sum_{r=1}^{\sigma}\eta_r\left((J_{\mathbf{q}}A_\Gamma^{(r)})\Gamma\right)^T\right\}\boldsymbol{\eta} \tag{12}$$

Equations (11) are joined to (6), in order to form a system of $\sigma + \ell$ equations for the $\sigma + \ell$ unknown quantities $\boldsymbol{\eta}$ and $\mathbf{q}$. System (11) + (6) can be written in normal form, since A is a positive–definite square matrix and $rank\,\Gamma = \sigma$, so that even $A_\Gamma = \Gamma^T A\Gamma$ is a positive–definite $\sigma\times\sigma$ symmetrical matrix.
It will turn out useful to rewrite the matrix in braces in (12) as follows:

$$\sum_{r=1}^{\sigma}\left[(A_\Gamma\boldsymbol{\eta})_r\Gamma^T\left(J_{\mathbf{q}}Z_{(r)} - (J_{\mathbf{q}}Z_{(r)})^T\right)\Gamma - \eta_r\left((J_{\mathbf{q}}A_\Gamma^{(r)})\Gamma - \frac{1}{2}\left((J_{\mathbf{q}}A_\Gamma^{(r)})\Gamma\right)^T\right)\right] \tag{13}$$

by virtue of the identities

$$\sum_{k=1}^{\ell}\xi_k\left(\frac{\partial Z}{\partial q_k}\right)^T\mathbf{y} = \sum_{r=1}^{\sigma}y_r\left(J_{\mathbf{q}}Z_{(r)}\right)\boldsymbol{\xi}, \qquad \sum_{k=1}^{\ell}\xi_k\left(J_{\mathbf{q}}Z^{(k)}\right)^T\mathbf{y} = \sum_{r=1}^{\sigma}y_r\left(J_{\mathbf{q}}Z_{(r)}\right)^T\boldsymbol{\xi},$$
$$\sum_{k=1}^{\ell}\xi_k\left(\frac{\partial A_\Gamma}{\partial q_k}\right)^T\mathbf{y} = \sum_{r=1}^{\sigma}y_r\left(J_{\mathbf{q}}A_\Gamma^{(r)}\right)\boldsymbol{\xi}, \quad \boldsymbol{\xi} = (\xi_1,\dots,\xi_\ell), \quad \mathbf{y} = (y_1,\dots,y_\sigma).$$

Index by index, calling $b_{r,s}$, for $r,s=1,\dots,\sigma$, the entries of the matrix A_Γ, we can write the lines of (11) and (6) as

$$\begin{cases} \sum\limits_{r=1}^{\sigma} b_{i,r}\dot\eta_r + \sum\limits_{r,s=1}^{\sigma} \mathcal{Q}^{(i)}_{r,s}\eta_r\eta_s - \sum\limits_{h=1}^{\ell}\gamma_{h,i}\dfrac{\partial U}{\partial q_h} = 0 & i=1,\dots,\sigma \\ \dot q_k = \sum\limits_{j=1}^{\sigma}\gamma_{k,j}\eta_j & k=1,\dots,\ell \end{cases} \tag{14}$$

where $\mathcal{Q}^{(i)}$ is, for each index i, the square matrix of order σ with entries

$$\mathcal{Q}^{(i)}_{r,s}(q_1,\dots,q_\ell) = \sum_{h,k=1}^{\ell}\left(\gamma_{k,i}\gamma_{h,s}\sum_{j=1}^{\sigma} b_{r,j}\left(\frac{\partial z_{j,k}}{\partial q_h}-\frac{\partial z_{j,h}}{\partial q_k}\right) + \gamma_{h,s}\frac{\partial b_{r,i}}{\partial q_h} - \frac{1}{2}\gamma_{h,i}\frac{\partial b_{r,s}}{\partial q_h}\right) \tag{15}$$

Equations (14) can be identified with the Boltzmann–Hamel equations for the Lagrangian function (7).

1.2 Quasi–velocities

Once $Z(\mathbf{q})$ has been established, the quasi–velocities are determined through (3). Owing to the structure of equations (11), it is significant to examine the resulting format of the matrices Γ and A_Γ, once Z has been defined.
Owing to (5), the entries of Γ and Z are related as follows (see also (4)):

$$\gamma_{i,j} = \frac{1}{\delta} det\left(\begin{pmatrix} Z \\ \alpha \end{pmatrix}^{(1)} \dots \begin{pmatrix} Z \\ \alpha \end{pmatrix}^{(i-1)} \mathbf{e}_i \begin{pmatrix} Z \\ \alpha \end{pmatrix}^{(i+1)} \dots \begin{pmatrix} Z \\ \alpha \end{pmatrix}^{(\ell)} \right) \tag{16}$$

for any $i=1,\dots,\ell$, $j=1,\dots,\sigma$, where $\mathbf{e}_i$ is the versor of $\mathbb{R}^\ell$, with 1 at the j-th position and 0 elsewhere.
It is quite recurring the case of the σ variables $\boldsymbol{\eta}$ sorted either as $v\le\sigma$ quasi–velocities $\eta_i = \sum\limits_{j=1}^{\ell} z_{i,j}\dot q_j$, $i=1,\dots,v$ and $\sigma-v$ generalized velocities $\eta_{v+1}=\dot q_{\ell-(\sigma-v)+1}, \dots, \eta_\sigma = \dot q_\ell$, so that (we recall $\ell-\sigma=\mu$)

$$Z = \begin{pmatrix} Z_1 & Z_2 \\ \mathbb{O}_{(\sigma-v)\times(\mu+v)} & \mathbb{I}_{\sigma-v} \end{pmatrix} \tag{17}$$

with Z_1 and Z_2 respectively $v\times(\mu+v)$ and $v\times(\sigma-v)$ matrices. The case $v=\sigma$ corresponds to none of the quasi–velocities coinciding with one of the lagrangian velocities. The null matrix and the identity matrix appear in the blocks of Γ as follows:

$$\Gamma = \begin{pmatrix} \Gamma_1 & \Gamma_2 \\ \mathbb{O}_{(\sigma-v)\times v} & \mathbb{I}_{\sigma-v} \end{pmatrix} \tag{18}$$

where Γ_1 and Γ_2 are matrices of size $(\mu+v)\times v$ and $(\mu+v)\times(\sigma-v)$, respectively. As a consequence of (18), the $\ell\times\ell$ matrix A is splitted in the following blocks:

$$A=\begin{pmatrix} A^{(1,1)} & A^{(1,2)} \\ A^{(2,1)} & A^{(2,2)} \end{pmatrix} \tag{19}$$

where $A^{(1,1)}$ is a square matrix of order $\mu+v$, $A^{(1,2)}$ is a matrix of size $(\mu+v)\times(\sigma-v)$, $A^{(2,1)}=(A^{(1,2)})^T$ and $A^{(2,2)}$ is a square matrix of order $\sigma-v$. The computation of $A_\Gamma=\Gamma^T A\Gamma$ according to the blocks leads to

$$\begin{pmatrix} \Gamma_1^T A^{(1,1)}\Gamma_1 & \Gamma_1^T\left(A^{(1,1)}\Gamma_2+A^{(1,2)}\right) \\ \left[\Gamma_1^T\left(A^{(1,1)}\Gamma_2+A^{(1,2)}\right)\right]^T & \Gamma_2^T\left(A^{(1,1)}\Gamma_2+A^{(1,2)}\right)+\left(\Gamma_2^T A^{(1,2)}\right)^T+A^{(2,2)} \end{pmatrix} \tag{20}$$

where the sizes of the blocks are (left to right, top to down) $v\times v$, $v\times(\sigma-v)$, $(\sigma-v)\times v$, $(\sigma-v)\times(\sigma-v)$.

Remark 1.1 *Choosing one of the quasi–velocities η_i, $1\le i\le\sigma$, as one of the momenta p_j, $1\le j\le\ell$, means $z_{i,1}=a_{j,1}$, …, $z_{i,\ell}=a_{j,\ell}$. In that case, the j–th row of $\Gamma^T A$ is the unit vector of $\mathbb{R}^\sigma$ $(0,\dots,1,\dots,0)$, with 1 at the i–th position. In a more extensive way, the choice of $\eta_1,\dots,\eta_v$ as v of the momenta $p_i=\dfrac{\partial\mathcal{L}}{\partial\dot q_i}$, $i=1,\dots,\ell$, gives $Z=\begin{pmatrix} a_{i_1,1} & \dots & a_{i_1,\mu+v} \\ \dots & \dots & \dots \\ a_{i_v,1} & \dots & a_{i_V,\mu+v}\end{pmatrix}$ for some set of v indices $\{i_1,\dots,i_v\}\subset\{1,\dots,\ell\}$, $i_1<\dots<i_v$. This entails that the submatrix of $\Gamma^T A$ formed by the first v rows and the v columns i_1, …, i_v is $\mathbb{I}_v$ (for instance, if the v rows of Z are chosen as $\dfrac{\partial\mathcal{L}}{\partial\dot q_i}$, $i=\ell-v+1$, …, ℓ, the unitary matrix is the $v\times v$ high–placed block to the right).*

In closing the Section, our attention moves to the v–by–v matrix $\Gamma_1^T A^{(1,1)}\Gamma_1$ of (20). We make use of (16), assuming the structure (17), in order to compute the entries of Γ_1 (see also (25)):

$$\gamma_{i,j}=(-1)^{i+j}\frac{1}{\delta_R}\delta_{j,i}\quad i=1,\dots,\mu+v,\quad j=1,\dots,v \tag{21}$$

where $\delta_{j,i}$ is the determinant of the square matrix of order $\mu+v-1$ obtained by suppressing the j–th row and the i–th column from $\begin{pmatrix} Z_1 \\ \alpha_1\end{pmatrix}$ for $i=1,\dots,\mu+v$, $j=1,\dots,v$. Hence, one can calculate the entries of $\Gamma_1^T A^{(1,1)}\Gamma_1$ of (20) by means of

$$b_{r,s}=\frac{1}{\delta_R^2}\sum_{h,k=1}^{\mu+v}(-1)^{r+s}\delta_{r,h}\widehat{a}_{h,k}\delta_{s,k},\quad \widehat{a}_{h,k}=(-1)^{h+k}a_{h,k},\quad r,s=1,\dots,v. \tag{22}$$

2 Reduction

Our analysis aims at a double target: choose the quasi–velocities $\boldsymbol{\eta}$ in order to

1. make as simple as possible the principal part (20) of the equations of motion (11),

2. eliminate some of the variables $q_1, \ldots, q_\ell$ from the same equations, so that some of the second group of ℓ equations in (14) can be disentangled from the system.

2.1 Reducing the matrix A_Γ

As for the first point, we see that a remarkable simplification in (20) can be obtained whenever Γ_2 vanishes: in terms of choosing the pseudovelocities, such a condition means that $\eta_1, \ldots, \eta_v$ do not depend on $\dot{q}_{\mu+v+1}, \ldots, \dot{q}_\ell$: the next property will concern with the question.

Property 2.1 *Define*

$$\begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} = \alpha \tag{23}$$

the two blocks of α *of size* $\mu \times (\mu + v)$ *and* $\mu \times (\sigma - v)$*, respectively. For a fixed index* j*,* $j = 1, \ldots, \sigma - v$*,* $(\Gamma_2)^{(j)} = \mathbf{0}_{v+\mu}$ *if and only if* $\begin{pmatrix} (Z_2)^{(j)} \\ (\alpha_2)^{(j)} \end{pmatrix} = \mathbf{0}_{v+\mu}$.

Proof. Owing to (5), the blocks of Z, Γ and α fulfill

$$\begin{cases} Z_1\Gamma_2 + Z_2 = \mathbb{O}_{v\times(\sigma-v)} \\ \\ \alpha_1\Gamma_2 + \alpha_2 = \mathbb{O}_{\mu\times(\sigma-v)} \end{cases} \tag{24}$$

Computing each of the $\sigma - v$ columns, it is evident that, if $\gamma_{1,v+j} = \cdots = \gamma_{\mu+v,v+j} = 0$ for some j, $1 \le j \le \sigma - v$, then $z_{1,\mu+v+j} = \cdots = z_{v,\mu+v+j} = \alpha_{1,\mu+v+j} \cdots = \alpha_{\mu,\mu+v+j} = 0$. Conversely, if the latter elements are zero, then the j–th column of (24) is the homogeneous system

$$\begin{pmatrix} Z_1 \\ \alpha_1 \end{pmatrix} (\Gamma_2)^{(j)} = \mathbf{0}_{\mu+v}$$

Since (see (4))

$$\delta_R = \det \begin{pmatrix} Z_1 \\ \alpha_1 \end{pmatrix} = (-1)^{\mu(\sigma-v)}\delta, \tag{25}$$

the homogeneous system has only the null solution. $\square$

Remark 2.1 *The previous property can be also derived from the following result:* $\Gamma^{(h)} = \mathbf{e}_k$ *for some* h *and* k*,* $1 \le h \le \sigma$*,* $1 \le k \le \ell$*, if and only if* $\begin{pmatrix} Z \\ \alpha \end{pmatrix}^{(k)} = \mathbf{e}_h$*, where* $\mathbf{e}_r$ *is the unit vector of* $\mathbb{R}^\ell$ *with* 1 *at the* r*–th position and* 0 *elsewhere.*

As a corollary, we conclude that

$$\Gamma_2 = \mathbb{O}_{(v+\mu)\times(\sigma-v)} \tag{26}$$

if and only if $\begin{pmatrix} Z_2 \\ \alpha_2 \end{pmatrix} = \mathbb{O}_{(v+\mu)\times(\sigma-v)}$.
An additional hint concerning further simplifications in (20) is the following

Property 2.2

$$\Gamma_1^T A^{(1,2)} = \mathbb{O}_{v\times(\sigma-v)} \tag{27}$$

if and only if the column vectors of $A^{(1,2)}$ *are generated by the row vectors of* α_1*.*

Proof. Using one more time the block representation we have from (5) $\alpha_1\Gamma_1 = \mathbb{O}_{\mu\times v}$. Since the v column vectors of Γ_1 are linearly independent, because of (5) and (18), and the μ row vectors of α_1 are also independent, we have, in terms of mutual orthogonal spaces,

$$\mathbb{R}^{\mu+v} = \left\langle (\Gamma_1)^{(1)}, \ldots, (\Gamma_1)^{(v)} \right\rangle \oplus \left\langle (\alpha_1)_{(1)}, \ldots, (\alpha_1)_{(\mu)} \right\rangle \tag{28}$$

where $\langle\ \rangle$ denotes the span space of the contained vectors. Condition (27) geometrically means that each of the column vectors of $A^{(1,2)}$ belongs to the linear space orthogonal to the space generated by the columns of Γ_1 or equivalently, by virtue of (28), that it is a linear combination of the row vectors of α_1. $\square$
Owing to Properties 3.1 and 3.2, if assumptions $(H1)$ and $(H2)$ are met, then the matrix of coefficients (20) is simplified into

$$A_\Gamma = \begin{pmatrix} \Gamma_1^T A^{(1,1)}\Gamma_1 & \mathbb{O}_{v\times(\sigma-v)} \\ & \\ \mathbb{O}_{(\sigma-v)\times v} & A^{(2,2)} \end{pmatrix} \tag{29}$$

2.2 Eliminating some of the coordinates

We are going now to discuss point 2 listed at the beginning of this Section, starting from recalling a customary practise for geometrical constraints.
In the matter of holonomic systems it is well known that the absence of a coordinate q_k, $1 \le k \le \ell$, in the Lagrangian function (cyclic coordinate) leads to the possibility of reducing the problem through the first integral $p_k = \dfrac{\partial \mathcal{L}}{\partial \dot{q}_k}$

and the reduced Lagrangian function. More generally, let $\mathbf{q}_{\ell-\sigma} = (q_{\sigma+1}, \dots, q_\ell)$ be cyclic coordinates and $\mathbf{p}_{\ell-\sigma} = \nabla_{\dot{\mathbf{q}}_{\ell-\sigma}}\mathcal{L}$ the corresponding first integrals of motion, which are the linear relations $\sum\limits_{j=1}^{\ell} a_{i,j}\dot{q}_j - p_i = 0$, $i = \sigma+1, \dots, \ell$. Whenever $det\, A_{\ell-\sigma}(\mathbf{q}_\sigma) \neq 0$, where $A_{\ell-\sigma}$ is the square matrix with the entries $a_{r,s}$, $r, s = \sigma+1\dots,\ell$ and $\mathbf{q}_\sigma = (q_1, \dots, q_\sigma)$, it is possible to deduce from the first integrals

$$\dot{\mathbf{q}}_{\ell-\sigma} = A^{-1}_{\ell-\sigma}(\mathbf{q}_\sigma)\left(\mathbf{p}_{\ell-\sigma} - A_\sigma(\mathbf{q}_\sigma)\dot{\mathbf{q}}_\sigma\right), \tag{30}$$

where A_σ is the $(\ell-\sigma)\times\sigma$ matrix with entries $a_{i,j}$, $i = \sigma+1, \dots, \ell$ and $j = 1, \dots, \sigma$. Setting now $\boldsymbol{\eta} = \dot{\mathbf{q}}_\sigma$ and making use of (30), one can define the reduced Lagrangian $\widetilde{\mathcal{L}}(\mathbf{q}_\sigma, \boldsymbol{\eta}, t) = \mathcal{L}(\mathbf{q}_\sigma, \boldsymbol{\eta}, \dot{\mathbf{q}}_{\ell-\sigma}(\mathbf{q}_\sigma, \boldsymbol{\eta}))$ (without demanding correction terms giving a lagrangian structure to the corresponding equations). The function $\widetilde{\mathcal{L}}$ fulfills the equations of motion

$$\frac{d}{dt}\left(\nabla_{\boldsymbol{\eta}}\widetilde{\mathcal{L}} + A^T_\sigma(\mathbf{q}_\sigma)A^{-1}_{\ell-\sigma}(\mathbf{q}_\sigma)\mathbf{p}\right) = \nabla_{\mathbf{q}_\sigma}\widetilde{\mathcal{L}} + J_{\mathbf{q}}\left[A^{-1}_{\ell-\sigma}(\mathbf{q}_\sigma)\left(\mathbf{p}_{\ell-\sigma} - A_\sigma(\mathbf{q}_\sigma)\boldsymbol{\eta}\right)\right]$$

and they are totally disentangled from the so called reconstruction equations (30), which play in some sense the same role as (2).
The same procedure cannot be employed for nonholonomic systems, because of the kinematics relations: actually, the occurrence of a cyclic coordinate does not entail the constancy of the corresponding momentum, as (9) exhibits.
Nevertheless, it is reasonable to wonder if suitable choices of $\boldsymbol{\eta}$ make some of the $\mathbf{q}$ disappear from (11). A preliminary discrimination among coordinates is the following:

$$\begin{array}{ll} (A) & \text{a coordinate } q_k \text{ is not present neither in } \mathcal{L} \text{ nor in } \alpha, \\[1em] (B) & \text{a coordinate } q_k \text{ appears either in } \mathcal{L} \text{ or in } \alpha. \end{array} \tag{31}$$

In the first case, it is evident that, if q_k is kept away from Z, the same coordinate is definitively missing in (14). If (B) is the case, (11) cannot be solved separately from (6), since all the $\mathbf{q}$ generally appear in it.
Our next analysis will investigate the possibility of giving an appropriate form to Z (i. e. deciding on quasi–velocities) fit for the purpose of eliminating some of the $\mathbf{q}$, say $q_1, \dots, q_p$, $1 \leq p < \ell$, in (11).
Owing to the structure of (14), we start from the following request: it must be possible to split Γ into the two blocks

$$\Gamma = \begin{pmatrix} \Gamma_p \\ \Gamma_{\ell-p} \end{pmatrix} \tag{32}$$

where Γ_p and $\Gamma_{\ell-p}$ are $p\times\sigma$ and $(\ell-p)\times\sigma$ matrices respectively, and

$$\Gamma_{\ell-p} = \Gamma_{\ell-p}\left(q_{p+1}, \dots, q_\ell\right). \tag{33}$$

In that case, the mathematical problem is simplified: only the $\ell - p$ equations

$$\begin{pmatrix} \dot{q}_{p+1} \\ \dots \\ \dot{q}_\ell \end{pmatrix} = \Gamma_{\ell-p}(q_{p+1}, \dots, q_\ell) \begin{pmatrix} \eta_1 \\ \dots \\ \eta_\sigma \end{pmatrix} \tag{34}$$

are needed in order to solve the σ equations (11). The reduced unknown quantities are $\boldsymbol{\eta}$, q_{p+1}, $\dots$, q_ℓ. Once system (11) + (34) has been solved, the motion of the system is completed by integrating $\begin{pmatrix} \dot{q}_1 \\ \dots \\ \dot{q}_p \end{pmatrix} = \Gamma_p \begin{pmatrix} \eta_1 \\ \dots \\ \eta_\sigma \end{pmatrix}$ analogously to (30).

2.3 The main result

Just after we list the assumptions we need, we will state the result concerning the elimination of coordinates.
According to the structure of the (11), we are compelled to require in advance that

$(H0)$ The matrix Γ can be splitted as in (32), (33) and the vector $\Gamma^T \nabla_{\mathbf{q}} U$ of $\mathbb{R}^\sigma$ does not depend on q_1, $\dots$, q_p for some $p < \ell$.

This is, for instance, the circumstance of U not depending on q_1, $\dots$, q_p.
At the same time, we let the system in the situation apt to implement the Properties discussed in Paragraph 2.1: namely, let us assume that Z has the structure (17) for some $v < \ell$ and

$(H1)$ $\begin{pmatrix} Z_2 \\ \alpha_2 \end{pmatrix} = \mathbb{O}_{(v+\mu)\times(\sigma-v)}$ (the blocks are defined in (17) and (23)),

$(H2)$ the column vectors of $A^{(1,2)}$ are generated by the row vectors of α_1,

$(H3)$ the expressions (22) do not contain q_1, $\dots$, q_p.

We remark that, if $p \leq \mu + v$, then the partition (32) is automatic. We also remark that $(H1)$ is only in part linked to the choice of Z, since the condition $\alpha_2 = \mathbb{O}_{\mu\times(\sigma-v)}$ is actually related to the features of the constrained system, as well as $(H2)$. Also $(H3)$ has to be ascribed to the mechanical system, namely to the structure of the kinetic energy (19).
Due to (29), we add the following assumption:

$(H4)$ q_1, $\dots$, q_p are not present in $A^{(2,2)}$.

The latter hypothesis concerns (15), as it will be shown:

$(H5)$ the expressions

$$\sum_{h,k=1}^{\mu+v}\left(\gamma_{h,r}\gamma_{k,s}-\gamma_{h,s}\gamma_{k,r}\right)\frac{\partial z_{i,h}}{\partial q_k},\quad 1\le r<s\le v,$$
$$\sum_{h=1}^{\mu+v}\gamma_{h,r}\frac{\partial z_{i,h}}{\partial q_{\mu+s}},\qquad 1\le r\le v,\ v+1\le s\le\sigma \tag{35}$$

do not contain $(q_1,\dots,q_p)$, for any $i=1,\dots,v$, where $\gamma_{h,s}$ are the entries of Γ_1 (see (18)).

Let us turn now to (11) and state the main result.

Proposition 2.1 *If assumptions from $(H0)$ to $(H5)$ hold, then equations (11) do not include $q_1,\ \dots,\ q_p$.*

Proof. Assumption $(H0)$ makes the last term in (11) free from $q_1,\ \dots,\ q_p$. Moreover, $(H1)$ and $(H2)$ make A_Γ of the form (29), whose blocks are independent of $q_1,\ \dots,\ q_p$ owing to $(H3)$ and $(H4)$: hence the same coordinates do not appear even in the first term of (11). We finally discuss the term (13): according to the blocks (17), (18) one finds

$$\mathcal{M}=\Gamma^T\left(J_{\mathbf{q}}Z_{(r)}\right)\Gamma=\begin{pmatrix}\Gamma_1^T(J_{(q_1,\dots,q_{\mu+v})}(Z_1)_{(r)})\Gamma_1 & \Gamma_1^T(J_{(q_{\mu+v+1},\dots,q_\ell)})(Z_1)_{(r)}\\ \mathbb{O}_{(\sigma-v)\times v} & \mathbb{O}_{(\sigma-v)\times(\sigma-v)}\end{pmatrix}$$

for $r=1,\dots,\sigma$. The first part of (13) consists in $\mathcal{M}-\mathcal{M}^T$, whose entries are independent of $(q_1,\dots,q_p)$ if (35) holds. On the other hand, since A_Γ does not contain $(q_1,\dots,q_p)$, the calculation $(J_{\mathbf{q}}A_\Gamma^{(r)})$, for each $r=1,\sigma$, leads to a $\sigma\times\ell$ matrix where the left block is the $\sigma\times p$ null matrix, owing to $(H3)$ and $(H4)$. Hence, by virtue of (33) even the entries of $J_{\mathbf{q}}A_\Gamma^{(r)}\Gamma$ do not contain $(q_1,\dots,q_p)$, as well as the terms in the second square bracket of (13). $\square$

Remark 2.2 *The additional conditions (35) are independent from the rest of the assumptions: actually, a simple example where A_Γ does not depend on a coordinate q_k but (13) does is the following, for $\ell=3$, $A=diag\,(a_{1,1},a_{2,2},a_{3,3})$ with $a_{i,i}(\mathbf{q})>0$, $\mathbf{q}=(q_1,q_2,q_3)$, $i=1,2,3$, with the constraint $\alpha_{1,3}(\mathbf{q})\dot q_3=0$, $\alpha_{1,3}(\mathbf{q})\neq 0$: choosing $\left\{\begin{array}{l}\eta_1=\sqrt{a_{1,1}}\dot q_1,\\ \eta_2=\sqrt{a_{2,2}}\dot q_2\end{array}\right.$ one has $\Gamma=\begin{pmatrix}1/\sqrt{a_{1,1}} & 0\\ 0 & 1/\sqrt{a_{2,2}}\\ 0 & 0\end{pmatrix}$ and $A_\Gamma=\mathbb{I}_2$. On the other hand, the computation of (13) includes the matrix $\Gamma^T\left(J_{\mathbf{q}}Z_{(1)}-(J_{\mathbf{q}}Z_{(1)})^T\right)\Gamma=\begin{pmatrix}0 & \dfrac{1}{\sqrt{a_{1,1}a_{2,2}}}\dfrac{\partial a_{1,1}}{\partial q_2}\\ -\dfrac{1}{\sqrt{a_{1,1}a_{2,2}}}\dfrac{\partial a_{1,1}}{\partial q_2} & 0\end{pmatrix}$ containing all the $\mathbf{q}$ (the same is true for $Z_{(2)}$).*

3 Some instances

We will go through some typical systems woth nonholonomic constraints (drawn from literature, expecially from [2]) and put them into the context of the examined techniques.

3.1 Knife

The first and well–known example we consider is a homogeneous bar of lenght ℓ_f and mass m_f (see Figure 1, first picture). The midpoint P_f has to move on a horizontal plane Π and, at each time, in the direction that the bar is pointing (knife or blade). In this case $\ell = 1$ and setting $\mathbf{q} = (x_f, y_f, \theta)$, where (x_f, y_f) are the coordinates of P_f with respect to a cartesian system on Π and θ is the angle that the bar forms with the x–axis, the Lagrangian function is $\mathcal{L} = \frac{1}{2}m(\dot{x}_f^2 + \dot{y}_f^2) + \frac{1}{2}I_f\dot{\theta}^2$, $I_f = \frac{1}{2}m_f\ell_f^2$, whenever external forces are absent. We have $\mu = 1$ and the kinematic constraint is $\dot{x}_f \sin\theta - \dot{y}_f \cos\theta = 0$. If one assumes $v = 1$, the matrix in (4) is $\left(\begin{array}{ccc} z_{1,1} & z_{1,2} & 0 \\ 0 & 0 & 1 \\ \sin\theta & -\cos\theta & 0 \end{array}\right)$ where we established $z_{1,3} = 0$ so that assumption $(H1)$ is verified ($Z_2 = \alpha_2 = 0$) and (18) is of type $\Gamma = \left(\begin{array}{cc} \gamma_{1,1} & 0 \\ \gamma_{2,1} & 0 \\ 0 & 1 \end{array}\right)$, by virtue of (18) and of Property 2.1. The entries $z_{1,1}$ and $z_{1,2}$ have to be chosen in order to define the $\sigma = 2$ quasi–velocities η_1 and η_2. Since $A = diag\,(m_f, m_f, I_f)$ and no forces are applied, hypotheses $(H0)$, $(H1)$ and $(H4)$ are clearly fulfilled. The three determinants appearing in (22) are

$$\delta_R = -z_{1,1}\cos\theta - z_{1,2}\sin\theta, \quad \delta_{1,1} = \cos\theta, \quad \delta_{1,2} = -\sin\theta \quad \Longrightarrow \quad b_{1,1} = \frac{1}{\delta_R^2}m_f.$$

By means of (21) one finds $\gamma_{1,1} = -\frac{1}{\delta_R}\cos\theta$ and $\gamma_{2,1} = -\frac{1}{\delta_R}\sin\theta$. Finally, only the second line of (35) has to be checked (the values are $v = 1$, $s = 2$ and $h = 1, 2$):

$$\gamma_{1,1}\frac{\partial z_{1,1}}{\partial\theta} + \gamma_{2,1}\frac{\partial z_{1,2}}{\partial\theta} = \frac{1}{\delta_R}\left(-\cos\theta\sin\theta + \sin\theta\cos\theta\right) = 0.$$

We conclude that, whenever θ is not present in δ_R, it will be missed in the motion equations (11). Furthermore, since x_f and y_f do not appear neither in $\mathcal{L}$ nor in α (type (A) of (31)), it is sufficient to set $z_{1,1} = z_{1,1}(\theta)$, $z_{1,2} = z_{1,2}(\theta)$ in order to make them not appear in (11).
The spontaneous choice which guarantees δ_R free form θ is $z_{1,1} = \cos\theta$, $z_{1,2} = \sin\theta$ so that (3) gives

$$\eta_1 = \dot{x}_f\cos\theta + \dot{y}_f\sin\theta, \quad \eta_2 = \dot{\theta}$$

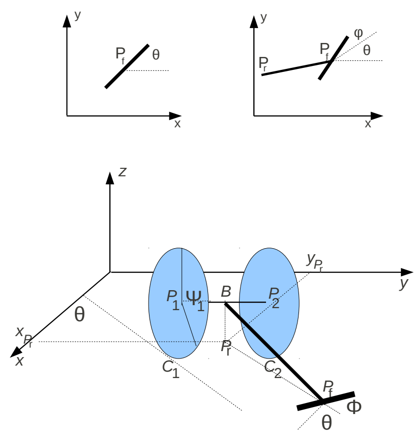

Figure 1: knife, bar driver, tricycle

as adopted in literature. Since $A_\Gamma = diag\,(m_f, I_f)$ and the term (12) vanishes, equations (14) are simply

$$\left\{ \begin{array}{l} m_f \dot\eta_1 = 0, \quad I_f \dot\eta_2 = 0, \\ \dot x_f = \eta_1 \cos\theta, \quad \dot y_f = \eta_1 \sin\theta, \quad \dot\theta = \eta_2 \end{array} \right.$$

In this elementary example the set of first σ equations in (14) is completely disentangled from the second set, which defines the quasi–velocities.

Remark 3.1 *A different situation occurs if Π is a vertical plane: if the y–axis is the vertical direction, we have $U = -m_f g y_f$, so that y_f is not of type (A) in (31). However, since $\Gamma^T \nabla_{\mathbf{q}} U = (-m_f g \gamma_{2,1}, 0)^T$, the procedure we performed eliminates y_f from the equations. On the other hand, θ is not longer removed and even taking $v = 2$ should not produce comfortable conditions in order to eliminate θ, as it can be seen without difficulty.*

3.2 Bar driven by a blade

The previous example can be enhanced in order to describe a sort of tricycle joint ([2]): we add a second bar (lenght ℓ_r, mass m_r) whose one end is pivoted in P_f and the other, say P_r, constrained not to slip sideways. Both of the bars lie on a plane Π (see Figure 1, second picture)). In a basic way, the system models the working principle of a tricycle. Calling θ the angle that $P_f - P_r$ forms with the x–axis and ϕ the angle that the front bar forms with the rear bar, the two nonholonomic constraints write

$$\dot{P}_r \wedge \mathbf{e}_\theta = \mathbf{0}, \quad \dot{P}_f \wedge \mathbf{e}_\beta = \mathbf{0}, \quad \mathbf{e}_\theta = \cos\theta \mathbf{i} + \sin\theta \mathbf{j}, \quad \mathbf{e}_\beta = \cos(\theta+\phi)\mathbf{i} + \sin(\theta+\phi)\mathbf{j} \tag{36}$$

In this case $\ell = 4$ and $\sigma = 2$: choosing $\mathbf{q} = (x_{P_r}, y_{P_r}, \theta, \phi)$, with x_{P_r}, y_{P_r} coordinates of P_r (hence $x_{P_f} = x_{P_r} + \ell_r \cos\theta$, $y_{P_f} = y_{P_r} + \ell_r \sin\theta$) and assuming no external forces, the Lagrangian funcition is the kinetic energy

$$T = \frac{1}{2}(m_f + m_r)\left(\dot{x}^2_{P_r} + \dot{y}^2_{P_r}\right) + \frac{1}{2} m_f I_f \left(\dot{\phi} + \dot{\theta}\right)^2 + \frac{1}{2}\left(\frac{1}{3} m_r \ell_r^2 + m_f \ell_f^2\right)\dot{\theta}^2$$

$$+ \left(m_f \ell_f + \frac{1}{2} m_r \ell_r\right)\left(-\dot{x}_{P_r}\dot{\theta}\sin\theta + \dot{y}_{P_r}\dot{\theta}\cos\theta\right), \qquad I_f = \frac{1}{12} m_f \ell_f^2$$

with matrix $A(\theta) = \begin{pmatrix} m_{tot} & 0 & -\lambda_1 \sin\theta & 0 \\ 0 & m_{tot} & \lambda_1 \cos\theta & 0 \\ -\lambda_1 \sin\theta & \lambda_1 \cos\theta & \lambda_2 & I_f \\ 0 & 0 & I_f & I_f \end{pmatrix}$ where $m_{tot} = m_r + m_f$, $\lambda_1 = m_f \ell_f + \frac{1}{2} m_r \ell_r$, $\lambda_2 = \frac{1}{6} m_r \ell_r^2 + \frac{5}{12} m_f \ell_f^2$. Since $A^{(2,2)} = I_f$, assumption $(H4)$ is obvious; moreover, $(H2)$ is valid for $\phi \neq \pi/2$. The constraints (36) are, in coordinates, $\dot{x}_{P_r}\sin\theta - \dot{y}_{P_r}\cos\theta = 0$, $\ell_r \dot{\theta}\cos\theta - (\dot{x}_{P_r}\cos\theta + \dot{y}_{P_r}\sin\theta)\sin\phi = 0$. Attempting with $v = 1$ and $Z_2 = 0$, the matrix in (4) is $\begin{pmatrix} z_{1,1} & z_{1,2} & z_{1,3} & 0 \\ 0 & 0 & 0 & 1 \\ \sin\theta & -\cos\theta & 0 & 0 \\ -\sin\phi\cos\theta & -\sin\phi\sin\theta & \ell_r\cos\phi & 0 \end{pmatrix}$ so that $\Gamma_2 = 0$ and $\Gamma_1 = \begin{pmatrix} \gamma_{1,1} \\ \gamma_{2,1} \\ \gamma_{3,1} \end{pmatrix}$ in (18). As for $(H4)$, we have $\delta_R = -\ell_r \cos\phi(z_{1,1}\cos\theta + z_{1,2}\sin\theta) - z_{1,3}\sin\phi$ and the only term to be calculated in (22) (for $r = s = 1$, $h, k = 1, 2, 3$) is $b_{1,1} = \frac{1}{\delta_R^2}\left(m_{tot}\ell_r^2\cos^2\phi + \lambda_2 \sin^2\phi\right)$.

On the other hand, (35) (only the second line, $i = 1$, $r = 1$, $\mu + s = 4$, $h = 1, 2, 3$) requires to check only the expression $\gamma_{1,1}\frac{\partial z_{1,1}}{\partial\phi} + \gamma_{2,1}\frac{\partial z_{1,2}}{\partial\phi} + \gamma_{3,1}\frac{\partial z_{1,3}}{\partial\phi}$. Since $\gamma_{3,1} = -\delta_R^{-1}\sin\phi$, a possible and evident choice which allows us to remove θ is $z_{1,1} = \cos\theta$, $z_{1,2} = \sin\theta$, $z_{1,3} = z_{1,3}(\phi)$: the quasi–velocities $\eta_1 = \dot{x}_{P_f}\sin\theta + \dot{y}_{P_r}\cos\theta + z_{1,3}(\phi)\dot{\theta}$, $\eta_2 = \dot{\phi}$ eliminate θ from (11), whose principal matrix (29) is $A_\Gamma = \begin{pmatrix} \dfrac{m_{tot}\ell_r^2\cos^2\phi + \lambda_2\sin^2\phi}{(\ell_r\cos\phi + z_{1,3}(\phi)\sin\phi)^2} & 0 \\ 0 & I_f \end{pmatrix}$.

We finally remark that x_{P_r} and y_{P_r} are of type (A) with respect to (31): the two equations (11) are coupled with only $\dot{\phi} = \eta_2$, while the remaining three equations of (6) will reconstruct the motion of the absent coordinates.

3.3 Adding a pair of wheels

The model can be refined by replacing the rear bar with a pair of actual wheels, whose centres are connected by a transverse axle (see Figure 1, third picture). The anterior device (blade) is the same as before, as well as the second nonholonomic constraint of (36). The rear two wheels are identical (mass m_w and radius R, each) and are required to roll without sliding on the plane Π, remaining orthogonal to it. Calling ψ_1 the pitch angle of one of the disks and θ the angle that $P_f - P_r$ forms with the x–axis (P_r is now the projection of B, midpoint of P_1 and P_2, on Π), the angular velocity of the disk is $\dot{\theta}\mathbf{k} + \dot{\psi}_1\mathbf{e}$, where $\mathbf{k} = \mathbf{i} \wedge \mathbf{j}$ ($\mathbf{i}$, $\mathbf{j}$ versors of the x–axis and y–axis) and $\mathbf{e} = \sin\theta\mathbf{i} - \cos\theta\mathbf{j}$ is the unit vector of $P_1 - P_2$.
Assuming that the plane Π is horizontal, the Lagrangian function is the kinetic energy

$$T = \frac{1}{24}m_f\ell_f^2(\dot{\phi}+\dot{\theta})^2 + \frac{1}{4}m_wR^2\left(\dot{\psi}_1^2 + 2\dot{\psi}_2^2\right) + \frac{1}{2}m_f\dot{P}_f^2 + \frac{1}{2}m_w\left(\dot{P}_1^2 + \dot{P}_2^2\right)$$

As for the nonholonomic conditions, imposing $\dot{C}_1 = \mathbf{0}$ (null velocity of the contact point) produces the kinematic constraints

$$\dot{x}_{P_1} + R\dot{\psi}_1\cos\theta = 0, \quad \dot{y}_{P_1} + R\dot{\psi}_1\sin\theta = 0 \tag{37}$$

At this point, the same condition on the second wheel, that is $\dot{C}_2 = \mathbf{0}$, is a holonomic condition: namely

$$\dot{x}_{P_2} + R\dot{\psi}_2\cos\theta = 0, \quad \dot{y}_{P_2} + R\dot{\psi}_2\sin\theta = 0$$

combined with $P_2 - P_1 = 2a\mathbf{e}$ (a is the half lenght of the axle) gives

$$\dot{x}_{P_1} + 2a\dot{\theta}\cos\theta + R\dot{\psi}_2\cos\theta = 0, \quad \dot{y}_{P_1} + 2a\dot{\theta}\sin\theta + R\dot{\psi}_2\sin\theta = 0$$

which in turn imply, recalling (37), $R(\psi_2 - \psi_1) + 2a\theta = \textit{constant}$ for any θ. Let us opt for the $\ell = 5$ lagrangian coordinates $(x_{P_r}, y_{P_r}, \theta, \psi_1, \phi)$: the matrix of T is $A(\theta) = \begin{pmatrix} 2m_w + m_f & 0 & -2\ell_r\sin\theta & 0 & 0 \\ 0 & 2m_w + m_f & 2\ell_r\cos\theta & 0 & 0 \\ -2\ell_r\sin\theta & 2\ell_r\cos\theta & \bar{I} & -m_waR & I_f \\ 0 & 0 & -m_waR & m_wR^2 & 0 \\ 0 & 0 & I_f & 0 & I_f \end{pmatrix}$ with $I_f = \dfrac{1}{12}m_f\ell_f^2$, $\bar{I} = m_f\ell_r^2 + m_w(4a^2 + R^2/2) + I_f$. Since $P_1 - P_r = -a\mathbf{e} + R\mathbf{k}$, the nonholonomic constraints (36), second condition, and (37) are, in terms of the selected coordinates,

$$\begin{cases} (\dot{x}_{P_r}\cos\theta + \dot{y}_{P_r}\sin\theta)\sin\phi - \ell_r\dot{\theta}\cos\phi = 0 \\ \dot{x}_{P_r} + (-a\dot{\theta} + R\dot{\psi}_1)\cos\theta = 0 \\ \dot{y}_{P_r} + (-a\dot{\theta} + R\dot{\psi}_1)\sin\theta = 0 \end{cases}$$

($\dot{P}_f \wedge \mathbf{e}_\beta = \mathbf{0}$ reduces to the first line by virtue of $\dot{x}_{P_r}\sin\theta - \dot{y}_{P_r}\cos\theta = 0$). In this example $\mu = 3$, $\sigma = 2$ and the purpose of eliminating θ is carried out by setting $v = 1$ and $z_{1,5} = 0$:

$$\begin{pmatrix} Z \\ \alpha \end{pmatrix} = \begin{pmatrix} z_{1,1} & z_{1,2} & z_{1,3} & z_{1,4} & 0 \\ 0 & 0 & 0 & 0 & 1 \\ \cos\theta\sin\phi & \sin\theta\sin\phi & -\ell_r\cos\phi & 0 & 0 \\ 1 & 0 & -a\cos\theta & R\cos\theta & 0 \\ 0 & 1 & -a\sin\theta & R\sin\theta & 0 \end{pmatrix}$$

so that $\Gamma_2 = \mathbb{O}_{4\times 1}$ (assumption $(H1)$). Assuming once more that Π is horizontal, assumption $(H0)$ is automatic. Moreover, $(H2)$ and $(H4)$ are easily checked. As for $(H3)$ and $(H5)$, one finds in this example

$$\delta_R = \ell_r(-Rz_{1,1}\cos\theta - Rz_{1,2}\sin\theta + z_{1,4})\cos\phi - (Rz_{1,3} + az_{1,4})\sin\phi$$

$$\begin{array}{ll} \delta_{1,1} = -R\ell_r\cos\phi\cos\theta, & \delta_{1,2} = R\ell_r\sin\theta\cos\phi, \\ \delta_{1,3} = -R\sin\phi, & \delta_{1,4} = -\ell_r\cos\phi + a\sin\phi \end{array}$$

hence (22) is $b_{1,1} = \dfrac{1}{\delta_R^2}\left[(3m_w + m_f)R^2\ell_r^2\cos^2\phi + (\bar{I} - m_w a^2)R^2\sin^2\phi\right]$, while (35) consists in checking merely

$$-\frac{1}{\delta_R}\ell_r\cos\phi\left(R\frac{\partial z_{1,1}}{\partial\phi}\cos\theta + R\sin\theta\frac{\partial z_{1,2}}{\partial\phi} - \frac{\partial z_{1,4}}{\partial\phi}\right) - \frac{1}{\delta_R}\sin\phi\left(R\frac{\partial z_{1,3}}{\partial\phi} + a\frac{\partial z_{1,4}}{\partial\phi}\right).$$

We can conclude that $z_{1,1} = \cos\theta$, $z_{1,2} = \sin\theta$, $z_{1,3} = z_{1,3}(\phi)$, $z_{1,4} = z_{1,4}(\phi)$ is a choice which removes the coordinate θ from (11). The quasi–velocities (3) turn out to be $\eta_1 = \dot{x}_{P_r}\cos\theta + \dot{y}_{P_r}\sin\theta + z_{1,3}(\phi)\dot{\theta} + z_{1,4}(\phi)\dot{\psi}_1$, $\eta_2 = \dot{\phi}_1$. They allows us to eliminate θ from the two equations (11), which are coupled only with $\dot{\phi} = \eta_2$, since x_{P_r}, y_{P_r} and ψ_1 are missing in $\mathcal{L}$ and α (type (A)).

3.4 Rolling disk with pendulum, or unicycle with rider

The last example is a disk (diameter $2R$ and mass m_d, centre P_0) yawing, rolling and pitching on a horizontal plane Π (ϕ, θ and ψ are respectively the yaw, roll and pitch angles, see Figure 2).
A point P_f (let us say the unicycle "frame") of mass m_f is positioned at distance ρ_1 from the contact point C of the disk with Π, in a way that C (contact point), P_0 and P_f are alligned. In addition, a point P_r (say the unicycle "the rider") of mass m_r is constrained on the plane orthogonal to the disk and containing P_0, C. The model is drawn from [7]). Calling O_1 the point placed at distance ρ_2 from P_f and alligned with C, P_0, the point P_r is at distance ρ from O_1 and oscillates around it.

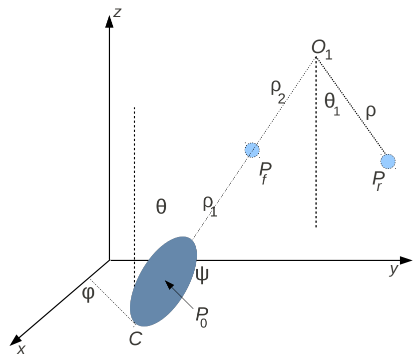

Figure 2: unicycle with rider

In this example $\ell = 6$ and the choice of the Lagrangian coordinates is $\mathbf{q} = (x_C, y_C, \phi, \psi, \theta, \theta_1)$, where θ_1 is the angle that $P_r - O_1$ forms with the downward vertical direction. As discussed in [6], the Lagrangian function includes $U = -\kappa_1 g\cos\theta + m_d g\rho\cos\theta_1$, with $\kappa_1 = m_d R + m_f\rho_1 + m_r(\rho_1+\rho_2)$, and T with matrix A, whose main diagonal and upper triangular part are

$$\begin{pmatrix} m & 0 & -F_1\cos\phi & 0 & -\kappa_1\cos\theta\sin\phi & -m_r\ell\cos\theta_1\sin\phi \\ & m & -F_1\sin\phi & 0 & \kappa_1\cos\theta\cos\phi & m_r\ell\cos\theta_1\cos\phi \\ & & F_2 & -I_d\sin\theta & 0 & 0 \\ & & & I_d & 0 & 0 \\ & & & & \kappa_2 + \frac{1}{2}I_d & \kappa_3\cos(\theta+\theta_1) \\ & & & & & m_r\ell^2 \end{pmatrix}$$

where $m = m_d + m_f + m_r$, $I_d = \dfrac{1}{2}m_d R^2$, $\kappa_2 = m_f\ell_1^2 + m_r(\rho_1+\rho_2)^2 + m_d R$, $\kappa_3 = m_r\rho(\rho_1+\rho_2)$ and $F_1(\theta,\theta_1) = (m_f\rho_1 + m_r(\rho_1+\rho_2) + m_d R)\sin\theta + m_r\rho\sin\theta_1$, $F_2(\theta,\theta_1) = \left(\dfrac{1}{2}I_d + m_d R^2 + m_f\rho_1^2 + m_r(\rho_1+\rho_2)^2\right)\sin^2\theta + m_r\rho^2\sin^2\theta_1 +$

$2m_r\rho(\rho_1+\rho_2)\sin\theta\sin\theta_1$.

The velocity $\dot{C}$ vanishes correspondingly to the kinematic constraints $\dot{x}_C = \dot{\psi}R\cos\phi$, $\dot{y}_C = \dot{\psi}R\sin\phi$ (thus $\mu = 2$, $\sigma = 4$) and the matrix appearing in (4) will be configured with $v = 2$ and $Z_2 = \mathbb{O}_{2\times 2}$ (see (17)), so that $(H1)$ is directly fulfilled: $\begin{pmatrix} Z \\ \alpha \end{pmatrix} = \begin{pmatrix} z_{1,1} & z_{1,2} & z_{1,3} & z_{1,4} & 0 & 0 \\ z_{2,1} & z_{2,2} & z_{2,3} & z_{2,4} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & -R\cos\phi & 0 & 0 \\ 0 & 1 & 0 & -R\sin\phi & 0 & 0 \end{pmatrix}$.

We expect Γ as in (18), with Γ_1 4–by–2 matrix, $\Gamma_2 = \mathbb{O}_{4\times 2}$ (whereas the lower matrices are $\mathbb{O}_{2\times 2}$ and $\mathbb{I}_2$). As for assumption $(H0)$, we see that the term involved in $(H0)$ is $\Gamma^T \nabla_{\mathbf{q}} U = (0,\, 0,\, \kappa_1 g, \sin\theta,\, -m_d \rho g \sin\theta_1)^T$; moreover, x_C, y_C, ψ do not appear neither in $\mathcal{L}$ nor in α (type (A)). Thus, the coordinate which we attempt to eliminate is ϕ (type (B)).

Since $rank \begin{pmatrix} 1 & 0 & 0 & -R\cos\phi \\ 0 & 1 & 0 & -R\sin\phi \\ -\kappa_1\cos\theta\sin\phi & \kappa_1\cos\theta\cos\phi & 0 & 0 \end{pmatrix} = 2$ and the same is true replacing θ_1 with θ, we have that hypothesis $(H2)$ is verified. Furthermore, $(H4)$ is evident, being $A^{(2,2)} = A^{(2,2)}(\theta,\theta_1)$. Moving on to $(H3)$, we calculate $\delta_R = (z_{1,3}z_{2,4} - z_{2,3}z_{1,4}) + R(z_{1,3}z_{2,1} - z_{1,1}z_{2,3})\cos\phi + R(z_{1,3}z_{2,2} - z_{1,2}z_{2,3})\sin\phi$ and (22):

$$\left.\begin{array}{l} \delta_R^2 b_{1,1} = (mR^2 + I_d) z_{2,3}^2 + 2(RF_1 + I_d\sin\theta) z_{2,3}\delta_{1,3} + F_2\delta_{1,3}^2, \\ \delta_R^2 b_{1,2} = (mR^2 + I_d) z_{1,3}z_{2,3} + (RF_1 + I_d\sin\theta)(z_{1,3}\delta_{1,3} + z_{2,3}\delta_{2,3}) + F_2\delta_{1,3}\delta_{2,3}, \\ \delta_R^2 b_{2,2} = (mR^2 + I_d) z_{1,3}^2 + 2(RF_1 + I_d\sin\theta) z_{1,3}\delta_{2,3} + F_2\delta_{2,3}^2 \end{array}\right\} \tag{38}$$

where $\delta_{1,3} = z_{2,4} + Rz_{2,2}\sin\phi + Rz_{2,1}\cos\phi$, $\delta_{2,3} = z_{1,4} + Rz_{1,2}\sin\phi + z_{1,1}R\cos\phi$. Owing to the calculated expressions, in order to eliminate ϕ we are induced to set

$$Z_1 = \begin{pmatrix} \hat{z}_{1,1}(\theta,\theta_1)\cos\phi & \hat{z}_{1,2}(\theta,\theta_1)\sin\phi & z_{1,3}(\theta,\theta_1) & z_{1,4}(\theta,\theta_1) \\ \hat{z}_{2,1}(\theta,\theta_1)\cos\phi & \hat{z}_{2,2}(\theta,\theta_1)\sin\phi & z_{2,3}(\theta,\theta_1) & z_{2,4}(\theta,\theta_1) \end{pmatrix} \tag{39}$$

satisfying the condition

$$\hat{z}_{1,1} = \hat{z}_{1,2}, \quad \hat{z}_{2,1} = \hat{z}_{2,2} \tag{40}$$

so that (38) are free from ϕ, as it can be easily checked. On the other hand, testing $(H5)$ makes us compute (35), first line, for $r=1$, $s=2$, $i=1,2$ and $h,k=1,2,3,4$: the only remaining terms, after the selection (39), (40) are

$$-(\gamma_{1,1}\gamma_{3,2} - \gamma_{1,2}\gamma_{3,1})\hat{z}_{i,1}(\theta,\theta_1)\sin\phi + (\gamma_{2,1}\gamma_{3,2} - \gamma_{2,2}\gamma_{3,1})\hat{z}_{i,2}(\theta,\theta_1)\cos\phi, \quad i=1,2 \tag{41}$$

By means of (21)one finds $\gamma_{1,1}\gamma_{3,2} - \gamma_{1,2}\gamma_{3,1} = \dfrac{1}{\delta_R^2}\mathcal{Y}(\theta,\theta_1)R\cos\phi$, $\gamma_{2,1}\gamma_{3,2} - \gamma_{2,2}\gamma_{3,1} = \dfrac{1}{\delta_R^2}\mathcal{Y}(\theta,\theta_1)R\sin\phi$, with $\mathcal{Y}(\theta,\theta_1) = z_{2,3}(z_{1,4}+R) - z_{1,3}(z_{2,4}+R)$, therefore terms (41) cancel each other out. Finally, the second line of conditions (35) consists in cheking whether the expressions

$$Rz_{h,3}\cos^2\phi\frac{\partial\hat{z}_{i,1}}{\partial q_j} + Rz_{h,3}\sin^2\phi\frac{\partial\hat{z}_{i,2}}{\partial q_j} - \left[z_{h,4} + R(\hat{z}_{h,2}\sin^2\theta + \hat{z}_{h,1}\cos^2\phi)\right]\frac{\partial z_{i,3}}{\partial q_j} + z_{h,3}\frac{\partial\hat{z}_{i,4}}{\partial q_j},$$

are free from the coordinate ϕ, with $h, i = 1, 2$, $j = 5, 6$ ($q_5 = \theta$, $q_6 = \theta_1$). It is evident that, under the specifications (39) and (38), such requirement is fulfilled. Through (17) and (39), we finally write the obtained quasi–velocities

$$\begin{aligned}
&\eta_1 = \hat{z}_{1,1}(\theta, \theta_1)[\dot{x}_C \cos\phi + \dot{y}_C \sin\phi] + z_{1,3}(\theta, \theta_1)\dot{\phi} + z_{1,4}(\theta, \theta_1)\dot{\psi},\\
&\eta_2 = \hat{z}_{2,1}(\theta, \theta_1)[\dot{x}_C \cos\phi + \dot{y}_C \sin\phi] + z_{2,3}(\theta, \theta_1)\dot{\phi} + z_{2,4}(\theta, \theta_1)\dot{\psi},\\
&\eta_3 = \dot{\theta}, \qquad \eta_4 = \dot{\theta}_1,
\end{aligned}$$

whose employement allows us to consider the equations of motion (14), first line, coupled with only 2 of the 6 equations (14), second line, namely for $k = 5, 6$. The rest of the system (4 equations) will form the reconstruction equations for x_C, y_C, ϕ, ψ, trough (18) with $\delta_R = z_{1,3}z_{2,4} - z_{2,3}z_{1,4} - R(z_{2,3} - z_{1,3})$ and $\Gamma_1 = \frac{1}{\delta_R}\begin{pmatrix} -Rz_{2,3}\cos\phi & rz_{1,3}\cos\phi \\ -Rz_{2,3}\sin\phi & Rz_{1,3}\sin\phi \\ z_{2,4} + \hat{z}_{2,1} & -z_{2,4} - R\hat{z}_{1,1} \\ -z_{2,3} & z_{1,3} \end{pmatrix}$.

Remark 3.2 *The choice of η_1, η_2 (performed, for instance, in [1]) as the conjugate momenta $\eta_1 = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = F_2\dot{\phi} - I_d\dot{\psi}\sin\theta - F_1(\dot{x}_C\cos\phi + \dot{y}_C\cos\phi)$, $\eta_2 = \frac{\partial \mathcal{L}}{\partial \dot{\psi}} = I_d\dot{\psi} - I_d\dot{\phi}\sin\theta$, $\eta_3 = \dot{\theta}$, $\eta_4 = \dot{\theta}_1$ matches our conclusions, by carrying out the choice $\hat{z}_{1,1} = \hat{z}_{1,2} = F_1(\theta, \theta_1)$, $\hat{z}_{1,3} = F_2(\theta, \theta_1), z_{1,4} = I_d\sin\theta, \hat{z}_{2,1} = \hat{z}_{2,2} = 0$, $z_{2,3} = -I_d\sin\theta$, $z_{2,4} = I_d$.*

4 Conclusion

The opportunity of confining the resolution of (14) to a reduced number of equations is a certain advantage from the mathematical point of view: in the study of the stability od the system, for instance, linear approximation and eigenvalues computation are simplified.
The elimination of one or more coordinates is carried out either by verifying particular features of the mechanical system and by searching for a set of suitable quasi–velocities, fulfilling specific and explicit conditions.
If, one the one hand, the assumptions listed in Par. 2.3 may appear somehow restrictive, on the other hand they reflect usual situations, as, for instance, the dependence of the kinematic conditions only on a low number of lagrangian velocities, or the absence of many coordinates in the applied forces (actually, the motion of systems containing disks or wheels and subject only to gravity is largely studied in literature: the example in Par. 3.4, if one neglects the two isolated masses, is the well–known model of rolling falling disk). The main task of the examples listed in Section 3 is precisely to check the pertinence of the procedure for common instances.
Nevertheless, the result can be generalised to systems not fulfilling all the listed assumptions, by following the same procedure and adjusting the requests.
The last point is one of the topics of forthcoming investigations, together to the following questions, come to light in preparing the present paper:

- investigate whether a link exists between the choice suggested by our method and the one motivated by the presence of simmetries,

- discover the exact role of choosing quasi–velocities as conjugate momenta with respect to the Lagrangian function of the system,

- take advantage of the described procedure in order to handle complex mechanical systems constrained by kinematic conditions, as the bicycle, a rough model of which was sketched in [5], supposing that the mathematical model falls within the typology contempled here.

References

[1] Bloch, A. ,M. , Mardsen, J. E. , Zenkov, D. V. , Quasivelocities and symmetries in non–holonomic systems, *Dynamical Systems*, **24**(2), 187–222 (2009)

[2] Cameron, J. . M. , J. M. , Book, J. W. , Modeling mechanisms with nonholonomic joints using the Boltzmann-Hamel equations, *Journal International Journal of Robotics Research*, **16**(1), 47–59 (1997)

[3] Gantmacher F. R. , Lectures in Analytical Mechanics, MIR 1975

[4] Maruskin, J. M. , Bloch, A. M. , The Boltzmann–Hamel equations for the optimal control of mechanical systems with nonholonomic constraints, *Int. J. Robust. Nonlinear Control* **21**, 373–386 (2011)

[5] Talamucci F. , The Lagrangian method for a basic bicycle, *Journal of Applied Mathematics and Physics*, **2** (4), 46–60 (2014)

[6] Talamucci, F. , Motion of Nonholonomous Rheonomous Systems in the Lagrangian Formalism, *Journal of Applied Mathematics and Physics*, **3**, 295–309 (2015)

[7] Zenkov, V. , Bloch, A. M. , Mardsen, J. E., Stabilization of the Unicycle with Rider, *Systems and Control Letters* **46**, 293–302 (2002)

Received: Month xx, 20xx