Journal of

Nonlinear Science

Hamilton's Principle as Variational Inequality for Mechanical Systems with Impact

R.I. Leine · U. Aeberhard · C. Glocker

Abstract The classical form of Hamilton's principle holds for conservative systems with perfect bilateral constraints. Several attempts have been made in literature to generalise Hamilton's principle for mechanical systems with perfect unilateral constraints involving impulsive motion. This has led to a number of different variants of Hamilton's principle, some expressed as variational inequalities. Up to now, the connection between these different principles has been missing. The aim of this paper is to put these different principles of Hamilton in a unified framework by using the concept of weak and strong extrema. The difference between weak and strong variations of the motion is explained in detail. Each type of variation leads to a variant of the principle of Hamilton in the form of a variational inequality. The conclusion of the paper is that each type of variation leads to different necessary and sufficient conditions on the impact law. The principle of Hamilton with strong variations is valid for perfect unilateral constraints with a completely elastic impact law, whereas the weak form of Hamilton's principle only requires perfect unilateral constraints and no condition on the energy.

Communicated by G. Stepan.

R.I. Leine (✉) · U. Aeberhard · C. Glocker
Institute of Mechanical Systems, Department of Mechanical and Process Engineering, ETH Zurich, 8092 Zurich, Switzerland
e-mail: remco.leine@imes.mavt.ethz.ch

1 Introduction

Classical analytical mechanics is concerned with mechanical systems with perfect bilateral (mostly holonomic) constraints (Papastavridis 2002) and is closely related with the calculus of variations as many principles of statics and dynamics are formulated in terms of variational problems, e.g. the principle of virtual work and Hamilton's principle (see, for instance, Lanczos 1962). Unilateral constraints, which are basically inequality constraints, are completely ignored in classical analytical mechanics because inequalities are not discussed by the classical calculus of variations and, also, because unilateral constraints in dynamics lead to shocks with discontinuities in the velocity. The mathematical tools to handle unilateral problems in statics, and later also in dynamics, have only been developed since the last 4 decades. The field of non-smooth dynamics is now rapidly developing. We refer the reader to the available textbooks on this subject (Leine and Nijmeijer 2004; Leine and van de Wouw 2008; Glocker 2001; Brogliato 1999).

Variational problems involving inequality constraints can be described by variational inequalities and were first introduced by Hartman and Stampacchia (1966) to study partial differential equations. The applicability of the theory has since been expanded to include problems from mechanics, finance, optimisation, and game theory. References on variational inequalities can be found in the standard reference (Kinderlehrer and Stampacchia 1980) or in the more recent books (Goeleven et al. 2003a, 2003b; Kravchuk and Neittaanmäki 2007). Variational problems with non-convex inequality constraints are described by hemi-variational inequalities and are discussed in the seminal work of Panagiotopoulos (1993). Moreover, (hemi-)variational inequalities are related to (non-)convex super-potentials through the sub-derivative known from non-smooth analysis; see the work of Moreau (1968) and Panagiotopoulos (1981). In the following, we will simply use the term 'variational inequalities' to denote (hemi-)variational inequalities.

Various authors study the classical principle of Hamilton, which is stated in the form of a variational equality, in the context of mechanical systems with impact. In Kozlov and Treshchëv (1991), Fetecau et al. (2003) and Cirak and West (2005), it is stated that the principle of Hamilton in its classical form as variational equality is also valid for motion with completely elastic impact (see also the review of Brogliato 1999). In the proof of Kozlov and Treshchëv (1991), it becomes clear that in this case only comparison curves are considered which have an impact, i.e. the variations of the position functions are not completely free. The classical principle of Hamilton is used in Fetecau et al. (2003) and Cirak and West (2005) to develop a symplectic-momentum preserving numerical scheme for systems with impact. The principle of Hamilton in equality form used in Kozlov and Treshchëv (1991), Fetecau et al. (2003) and Cirak and West (2005) does not hold for a system with persistent contact, i.e. a contact being closed during a time lapse.

The extension of classical analytical mechanics to perfect unilateral constraints, with possible persistent contact, asks for a reformulation of the variational principles of mechanics in terms of variational inequalities. The principle of d'Alembert–Lagrange in the form of a variational inequality has been discussed in Panagiotopoulos and Glocker (1998, 2000), Goeleven et al. (1997, 1999) and May (1984a, 1984b).

Various forms of the principle of Hamilton as variational inequality can be derived from the principle of d'Alembert–Lagrange as variational inequality (May 1984a, 1984b; Panagiotopoulos and Glocker 1998, 2000). The principle of Hamilton in May (1984a, 1984b) is valid for frictionless unilateral constraints (i.e. perfect unilateral constraints) and non-impulsive motion. Hence, impacts which naturally arise due to collisions if the motion is unilaterally constrained are not considered in May (1984a, 1984b). In Panagiotopoulos and Glocker (1998), a principle of Hamilton is presented (Theorem 3) which is valid for perfect unilateral constraints with impulsive motion. Another principle of Hamilton is presented in Theorem 4 of Panagiotopoulos and Glocker (1998), which is basically identical to Proposition 4 in Panagiotopoulos and Glocker (2000), and which is valid for perfect unilateral constraints with a completely elastic impact law.

A very different approach has been taken by Tornambè (1999) and Percivale (1985, 1991) in which the action integral is modified with a Lagrangian multiplier term and the principle of Hamilton is stated as variational equality. In Percivale (1985, 1991), it is proven that a trajectory with a completely elastic impact law is a stationary point of the modified action integral. In Tornambè (1999), the inequality constraints are transformed into equality constraints by making use of so-called Valentine variables. The equality constraint on the Valentine variables is subsequently incorporated in the action integral with a Lagrangian multiplier and the stationarity of the resulting functional is shown to lead to solutions of the completely elastic impact problem. The principles of Hamilton in Tornambè (1999) and Percivale (1985, 1991) are variational equalities using a modified action integral.

The great difference between the approach used by May (1984a, 1984b) and Panagiotopoulos and Glocker (1998, 2000) and the approach used by Tornambè (1999) and Percivale (1985, 1991) is that the latter approach does not derive the principle of Hamilton from the principle of d'Alembert–Lagrange and the principle of virtual work. In the current paper, we will derive various forms of the principle of Hamilton starting from the principle of virtual work. We will therefore follow the first approach which leads to variational inequalities.

The mathematical literature on the calculus of variations distinguishes between so-called weak and strong local extrema of a functional $J(y) = \int_I f(x, y, y')\,dx$ (Cesar 1984; Cesari 1983; Troutman 1996). The strong norm measures the maximal function value $\max(|y(x)|)$ on the interval I, whereas the weak norm also takes the derivative y' into account. A weak/strong local extremal is extremal with respect to comparison functions within a neighbourhood defined by the weak/strong norm. It is known that a weak local extremal fulfills the Euler–Lagrange equations almost everywhere and fulfills the first Weierstrass–Erdmann corner condition, i.e. the left and right limit of $f_{y'}$ agree on its corner points. A strong local extremal fulfills in addition the second Weierstrass–Erdmann corner condition, i.e. the left and right limit of $f - y' f_{y'}$ agree on its corner points. There is some confusion about this in literature. Some authors incorrectly state that a weak local extremal necessarily fulfills the second Weierstrass–Erdmann corner condition (see Cesar 1984 for a discussion and counter examples). Moreover, we note that the first Weierstrass–Erdmann corner condition generalises to an inclusion if it is applied to a variational inequality. Furthermore, in our knowledge, Hamilton's principle has not been put within the concepts of weak and strong local extrema.

The current paper derives weak and strong forms of the principle of Hamilton as variational inequality using the concepts of weak and strong extrema (Cesar 1984; Troutman 1996) and concepts from measure and integration theory. It therefore becomes much more clear how the existing forms of the principle of Hamilton, which have been derived in Panagiotopoulos and Glocker (1998, 2000), Tornambè (1999) and Percivale (1985, 1991), have to be understood. Moreover, we demonstrate how the concept of virtual displacements and the principle of virtual work are to be used in a non-smooth setting and that we have in fact to distinguish between weak and strong virtual displacements. The general variation of the action integral leads to a strong form of Hamilton's principle and to assumptions which are generalisations of the classical Weierstrass–Erdmann corner conditions.

The principle of Hamilton occupies a central role in classical analytical mechanics as it relates dynamics to the calculus of variations. From the practical point of view, Hamilton's principle is used in mechanics to derive (numerical) approximation methods, such as the Ritz method. The advantages of the principle of Hamilton are its concise formulation and the fact that approximations which maintain the symmetries of the Hamiltonian automatically preserve the corresponding conservation laws (Salomon 1983). Similarly, the generalisations of Hamilton's principle, discussed in this paper, can be used to derive difference schemes which can form the basis for dedicated numerical integration schemes for impulsive mechanical systems.

The paper is organised as follows. In Sect. 2, we give the reader an introduction to cones from convex analysis in functional spaces and set the notation. Variational inequalities are discussed in Sect. 3 and special attention is paid to weak and strong local extrema. Subsequently, we discuss in Sect. 4 the principle of d'Alembert–Lagrange in inequality form, which constitutes a variational inequality. The principle of Hamilton in inequality form is derived from the principle of d'Alembert–Lagrange in Sect. 5 for non-impulsive motion under the supposition of perfect geometric unilateral constraints. In Sect. 6, the general variation of the action integral is used to derive necessary and sufficient conditions for the validity of a strong form of Hamilton's principle in inequality form for impulsive motion. Concepts of measure and integration theory are used in Sect. 7 to derive a weak form of the principle of Hamilton. It turns out that only the supposition of perfect geometric unilateral constraints is necessary and sufficient for the validity of a weak form of Hamilton's principle as variational inequality. Finally, the difference between the two forms of the principle of Hamilton is illustrated in Sect. 8 by applying a Ritz-type of method on the action integral for an example system. Conclusions are drawn in Sect. 9 and the paper is closed with a discussion of the usefulness of the results and insight gained in this paper.

2 Cones

The analysis of inequalities requires appropriate geometric objects (Aubin and Frankowska 1990). The usual geometric objects in classical dynamics with equality constraints are manifolds and tangent spaces to manifolds. The study of inequalities,

Fig. 1 Contingent, tangent, and normal cone on a non-convex set $K \subset \mathbb{R}^2$
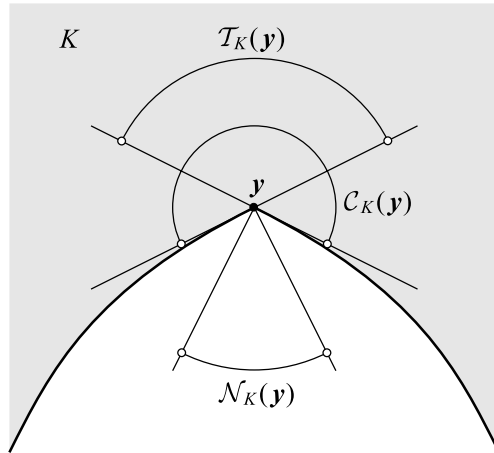
however, requires a number of cones. We first introduce these cones in the Euclidean $\mathbb{R}^n$. Let K be a closed non-empty subset of the $\mathbb{R}^n$. We define the distance from y to K as

$$d_K(y) = \inf_{y^* \in K} \|y - y^*\|, \tag{1}$$

with $\|y\| = \sqrt{\sum_i y_i^2}$. We assume that K is *derivable* (Aubin and Frankowska 1990; Rockafellar and Wets 1998) in the sense that $\lim_{t \downarrow 0} d_K(y + t\xi)/t$ exists for all $y \in K$ and all $\xi \in \mathbb{R}^n$. The set

$$\mathcal{C}_K(y) = \left\{ \xi \,\middle|\, \lim_{t \downarrow 0} \frac{d_K(y + t\xi)}{t} = 0 \right\}, \tag{2}$$

is the adjacent cone of K at y and has the following characterisation in terms of sequences

$$\xi \in \mathcal{C}_K(y) \iff \forall t_k \downarrow 0, \exists \xi_k \to \xi \quad \text{such that} \quad y + t_k \xi_k \in K \ \forall k. \tag{3}$$

If the set K is derivable, then the adjacent cone agrees with the so-called contingent cone (Aubin and Frankowska 1990). In the following, we will always assume that the set K is derivable and simply speak of the contingent cone $\mathcal{C}_K(y)$. Loosely speaking, the contingent cone is the cone which is obtained if one zooms in on the point y after a shift of the set K such that y is at the origin (see Fig. 1). The contingent cone is closed but not necessarily convex. The tangent cone $\mathcal{T}_K(y)$ defined by

$$\mathcal{T}_K(y) = \left\{ \xi \,\middle|\, \lim_{t \downarrow 0, y^* \to_K y} \frac{d_K(y^* + t\xi)}{t} = 0 \right\}, \tag{4}$$

where $\to_K$ denotes the convergence in K, is a closed convex sub-cone of the contingent cone. We define the normal cone $\mathcal{N}_K(y)$ to be the polar cone of the tangent cone $\mathcal{T}_K(y)$

$$\mathcal{N}_K(y) = \left\{ z \mid \xi^{\mathrm{T}} z \leq 0 \ \forall \xi \in \mathcal{T}_K(y) \right\}. \tag{5}$$

If the tangent cone $\mathcal{T}_K(\boldsymbol{y})$ is a sub-space of $\mathbb{R}^n$, then the normal cone $\mathcal{N}_K(\boldsymbol{y})$ is the annihilator and, therefore, a sub-space of its dual $(\mathbb{R}^n)^*$. The normal cone $\mathcal{N}_K(\boldsymbol{y})$ and tangent cone $\mathcal{T}_K(\boldsymbol{y})$ are closed convex cones. A set K is called tangentially regular, if the contingent cone $\mathcal{C}_K(\boldsymbol{y})$ agrees with the tangent cone $\mathcal{T}_K(\boldsymbol{y})$. Tangential regularity excludes reentrant corners. If K is a convex set, then it is also tangentially regular.

If $\mathcal{M}$ is a differentiable manifold embedded in the $\mathbb{R}^n$, then $\mathcal{T}_{\mathcal{M}}(\boldsymbol{y})$, which equals $\mathcal{C}_{\mathcal{M}}(\boldsymbol{y})$, is the tangent space to $\mathcal{M}$ at $\boldsymbol{y}$. Let $\mathcal{K}$ be a closed non-empty subset of the differentiable manifold $\mathcal{M}$. The tangent cone $\mathcal{T}_{\mathcal{K}}(\boldsymbol{y})$ is a subset of the tangent space $\mathcal{T}_{\mathcal{M}}(\boldsymbol{y})$, whereas $\mathcal{N}_{\mathcal{K}}(\boldsymbol{y})$ is a subset of the cotangent space $(\mathcal{T}_{\mathcal{M}}(\boldsymbol{y}))^*$.

The previously defined cones can be generalised to infinite dimensional vector spaces (see, for instance, Aubin and Frankowska 1990). Let $(X, \|\cdot\|)$ be a Banach space and X^* its dual space with the duality pairing $\langle\cdot,\cdot\rangle$. Let $\mathcal{K}$ be a closed, derivable subset of the X and $d_{\mathcal{K}}(y) = \inf_{y^*\in\mathcal{K}}\|y - y^*\|$ be the distance form y to $\mathcal{K}$. We define the contingent cone, tangent cone, and normal cone as (see Aubin and Frankowska 1990)

$$\mathcal{C}_{\mathcal{K}}(y) = \left\{\xi \in X \,\Big|\, \lim_{t\downarrow 0} \frac{d_{\mathcal{K}}(y + t\xi)}{t} = 0\right\}, \tag{6}$$

$$\mathcal{T}_{\mathcal{K}}(y) = \left\{\xi \in X \,\Big|\, \lim_{t\downarrow 0, y^*\to_{\mathcal{K}} y} \frac{d_{\mathcal{K}}(y^* + t\xi)}{t} = 0\right\}, \tag{7}$$

$$\mathcal{N}_{\mathcal{K}}(y) = \left\{z \in X^* \mid \langle \xi, z\rangle \leq 0 \,\forall \xi \in \mathcal{T}_{\mathcal{K}}(y)\right\}. \tag{8}$$

The cones on infinite dimensional vector spaces can be related to the cones on finite dimensional space. Let the space $X = L^\infty(I, \mathbb{R}^n)$ of real-valued n-dimensional essentially bounded functions be defined on the compact interval $I \subset \mathbb{R}$ with the duality pairing

$$\langle \boldsymbol{g}, \boldsymbol{f}\rangle = \int_I \big(\boldsymbol{g}(x)\big)^{\mathrm{T}} \boldsymbol{f}(x)\, \mathrm{d}x$$

$$= \sum_{i=1}^n \int_I g_i(x) f_i(x)\, \mathrm{d}x, \quad \boldsymbol{f} \in L^\infty\big(I, \mathbb{R}^n\big), \boldsymbol{g} \in L^1\big(I, \mathbb{R}^n\big), \tag{9}$$

and norm $\|\boldsymbol{f}\|_\infty = \operatorname{ess\,sup}_{x\in I}\|\boldsymbol{f}(x)\|$. Let $\mathcal{K}$ be a closed subset of $X = L^\infty(I, \mathbb{R}^n)$ of the form

$$\mathcal{K} = \big\{\boldsymbol{y} \in X \mid \boldsymbol{y}(x) \in K \,\forall x \in I\big\}, \tag{10}$$

where K is a closed subset of $\mathbb{R}^n$. It is straightforward to prove the equivalence between the contingent and tangent cones in $L^\infty(I, \mathbb{R}^n)$ and $\mathbb{R}^n$, i.e.

$$\boldsymbol{\xi} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y}) \iff \boldsymbol{\xi}(x) \in \mathcal{C}_K\big(\boldsymbol{y}(x)\big) \quad \text{for almost all } x \in I, \tag{11}$$

$$\boldsymbol{\xi} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{y}) \iff \boldsymbol{\xi}(x) \in \mathcal{T}_K\big(\boldsymbol{y}(x)\big) \quad \text{for almost all } x \in I, \tag{12}$$

and, using the dual pairing (9), we also have

$$z \in \mathcal{N}_{\mathcal{K}}(\boldsymbol{y}) \iff z(x) \in \mathcal{N}_K\big(\boldsymbol{y}(x)\big) \quad \text{for almost all } x \in I. \tag{13}$$

3 Variational Inequalities

The aim of this section is to define weak and strong local extrema as variational inequalities in infinite dimensional space. First, a motivation for variational inequalities is given by showing that the sub-stationarity condition of a constrained variational minimisation problem leads to a variational inequality. Furthermore, it is explained that the sub-stationarity condition in terms of a Gâteaux derivative only considers the class of comparison functions which are in a neighbourhood to the extremal with respect to a weak norm. This leads to the concept of extremals with respect to a norm.

Let $I \subset \mathbb{R}$ be a real compact interval. Consider functions $\boldsymbol{y} : I \to \mathbb{R}^n$ with the n components $y_i(x) \in \mathbb{R}$, $i = 1, \ldots, n$. Let $\mathcal{Y}(I, \mathbb{R}^n)$ be the class of functions $\boldsymbol{y} : I \to \mathbb{R}^n$ which are absolutely continuous, and of which the derivative $\boldsymbol{y}'$ is of specially bounded variation (no singular part) and, therefore, exists almost everywhere in I,

$$\mathcal{Y}(I, \mathbb{R}^n) = \{\boldsymbol{y} \in \mathrm{AC}(I, \mathbb{R}^n), \boldsymbol{y}' \in \mathrm{SBV}(I, \mathbb{R}^n)\}, \tag{14}$$

and let $\mathcal{Y}_0(I, \mathbb{R}^n)$ be the class of functions in $\mathcal{Y}(I, \mathbb{R}^n)$ which have compact support on I. Functions in $\mathcal{Y}$ may be piecewise C^1, such that the derivative is discontinuous on a Lebesgue negligible set. Functions in $\mathcal{Y}$ may also have accumulation points for which the derivative is of bounded variation. Note that if $\boldsymbol{y} \in \mathrm{AC}(I, \mathbb{R}^n)$, where I is compact, then $\boldsymbol{y}$ is bounded on I and it therefore holds that $\mathcal{Y}(I, \mathbb{R}^n) \subset L^\infty(I, \mathbb{R}^n)$. Furthermore, let the class $\mathcal{D}(I, \mathbb{R}^n) \subset \mathcal{Y}(I, \mathbb{R}^n)$ fulfill the boundary conditions $\boldsymbol{y}(x) = \boldsymbol{y}_0(x)$ on the boundary of I, i.e.

$$\mathcal{D}(I, \mathbb{R}^n) = \{\boldsymbol{y} \in \mathcal{Y}(I, \mathbb{R}^n), \boldsymbol{y}(x) = \boldsymbol{y}_0(x) \, \forall x \in \mathrm{bdry}\, I\}. \tag{15}$$

Let $\hat{\boldsymbol{y}}(\varepsilon)$ be a family of functions parameterised by ε for which it holds that $\hat{\boldsymbol{y}}(\varepsilon_0) = \boldsymbol{y}$, i.e. $\hat{\boldsymbol{y}}(\varepsilon_0, x) = \boldsymbol{y}(x) \, \forall x \in I$. If $\hat{\boldsymbol{y}}(\varepsilon, x)$ is locally differentiable with respect to ε, then the difference

$$\hat{\boldsymbol{y}}(\varepsilon, x) - \hat{\boldsymbol{y}}(\varepsilon_0, x) = \hat{\boldsymbol{y}}_\varepsilon(\varepsilon_0, x)(\varepsilon - \varepsilon_0) + \mathcal{O}\big((\varepsilon - \varepsilon_0)^2\big), \tag{16}$$

with $\hat{\boldsymbol{y}}_\varepsilon(\varepsilon, x) = \partial \hat{\boldsymbol{y}}(\varepsilon, x)/\partial \varepsilon$, gives rise to the variation $\delta \boldsymbol{y}(x)$ of $\boldsymbol{y}(x)$

$$\delta \boldsymbol{y}(x) = \hat{\boldsymbol{y}}_\varepsilon(\varepsilon_0, x)\, \delta\varepsilon \tag{17}$$

with $\delta\varepsilon = \varepsilon - \varepsilon_0$. The choice of ε_0 is immaterial and we will take $\varepsilon_0 = 0$. Moreover, it suffices to consider $\varepsilon \geq 0$.

Let $\mathcal{K}$ be a closed subset of $\mathcal{D}(I, \mathbb{R})$, which is not necessarily convex. We seek a function $\boldsymbol{y}(x)$ which is the solution of the constrained minimisation problem

$$\min_{\boldsymbol{y} \in \mathcal{K}} J(\boldsymbol{y}), \tag{18}$$

where $J : \mathcal{D}(I, \mathbb{R}^n) \to \mathbb{R}$ is the functional

$$J(\boldsymbol{y}) = \int_I f(x, \boldsymbol{y}, \boldsymbol{y}')\, \mathrm{d}x. \tag{19}$$

We assume that the functional J is Gâteaux differentiable. We consider all functions $\boldsymbol{v} \in \mathcal{Y}_0(I, \mathbb{R}^n)$ for which for each sequence $t_k \downarrow 0$ there exists a sequence $\boldsymbol{v}_k \to \boldsymbol{v}$ with $\boldsymbol{y} + t_k \boldsymbol{v}_k \in \mathcal{K}$. In other words, using the characterisation (3) of the contingent cone (6), it holds that $\boldsymbol{v} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y})$. Because $\boldsymbol{y}$ minimises J in $\mathcal{K}$, it holds for this sequence that $J(\boldsymbol{y}) \leq J(\boldsymbol{y} + t_k \boldsymbol{v}_k)$. Hence, the function $\bar{h}(t_k) = J(\boldsymbol{y} + t_k \boldsymbol{v}_k)$ attains its minimum when $t_k = 0$. Consequently, it holds that (see Kinderlehrer and Stampacchia 1980)

$$\bar{h}'(0) \geq 0 \tag{20}$$

with

$$\bar{h}'(0) = \lim_{t_k \downarrow 0, \boldsymbol{v}_k \to \boldsymbol{v}} \frac{J(\boldsymbol{y} + t_k \boldsymbol{v}_k) - J(\boldsymbol{y})}{t_k} = \lim_{\varepsilon \downarrow 0} \frac{J(\boldsymbol{y} + \varepsilon \boldsymbol{v}) - J(\boldsymbol{y})}{\varepsilon}, \tag{21}$$

in which we recognise the (one-sided) Gâteaux derivative

$$\mathrm{d}J(\boldsymbol{y}; \boldsymbol{v}) := \lim_{\varepsilon \downarrow 0} \frac{J(\boldsymbol{y} + \varepsilon \boldsymbol{v}) - J(\boldsymbol{y})}{\varepsilon}. \tag{22}$$

In this paper, we will not distinguish between the one-sided Gâteaux derivative and the Gâteaux derivative because ε is defined to be non-negative. Using the Gâteaux derivative (22), we can reformulate the sub-stationarity condition as

$$\mathrm{d}J(\boldsymbol{y}; \boldsymbol{v}) \geq 0 \quad \forall \boldsymbol{v} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y}) \cap \mathcal{Y}_0(I, \mathbb{R}^n), \tag{23}$$

or as

$$\int_I \left(f_{\boldsymbol{y}}(x, \boldsymbol{y}, \boldsymbol{y}') \boldsymbol{v} + f_{\boldsymbol{y}'}(x, \boldsymbol{y}, \boldsymbol{y}') \boldsymbol{v}' \right) \mathrm{d}x \geq 0 \quad \forall \boldsymbol{v} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y}) \cap \mathcal{Y}_0(I, \mathbb{R}^n), \tag{24}$$

which is a variational inequality.

The sub-stationarity condition (23) is a *necessary condition* for a solution of the constrained minimisation problem (18). The sub-stationarity condition (23) is (in general) not a sufficient condition for a minimum of (18) because of two reasons:

1. The sub-stationarity condition (23) only considers the first variation which is in general not sufficient to determine a minimum.
2. The Gâteaux derivative in (23) only considers a certain class of comparison functions $\hat{\boldsymbol{y}}(\varepsilon, x)$, namely those of the form $\hat{\boldsymbol{y}}(\varepsilon, x) = \boldsymbol{y}(x) + \varepsilon \boldsymbol{v}(x)$, and nothing is said about the stationarity of the functional with respect to other comparison functions which do not fall within this class.

The second reason preludes the concept of a local minimum with respect to a norm (see Troutman 1996).

Definition 1 (Local minimum with respect to a norm) In a normed vector space $(\mathcal{Y}, \| \cdot \|_\alpha)$, a point $\boldsymbol{y} \in \mathcal{K} \subseteq \mathcal{Y}$ is said to provide J with a local minimum on $\mathcal{K}$ in the norm $\| \cdot \|_\alpha$ if $\boldsymbol{y}$ is a minimiser for J on $\mathcal{B}_r^\alpha(\boldsymbol{y})$ for some $r > 0$, i.e.

$$\exists r > 0: \ J(\boldsymbol{y}^*) \geq J(\boldsymbol{y}) \quad \forall \boldsymbol{y}^* \in \mathcal{B}_r^\alpha(\boldsymbol{y}),$$

where $\mathcal{B}_r^\alpha(\mathbf{y}) = \{\mathbf{y}^* \in \mathcal{K} \mid \|\mathbf{y}^* - \mathbf{y}\|_\alpha < r\}$ is an open neighbourhood of $\mathbf{y}$ with respect to the norm $\|\cdot\|_\alpha$.

We introduce on $\mathcal{Y}(I, \mathbb{R})$ two norms:

$$\begin{aligned} \text{weak norm} \quad & \|\mathbf{y}\|_1 = \sum_{i=1}^{n} \max_{x \in I} |y_i(x)| + \operatorname*{ess\,sup}_{x \in I} |y_i'(x)|, \\ \text{strong norm} \quad & \|\mathbf{y}\|_0 = \sum_{i=1}^{n} \max_{x \in I} |y_i(x)|. \end{aligned} \quad (25)$$

The weak norm uses the essential supremum because $y_i'(x)$ is not defined for all $x \in I$.

We speak of a weak/strong local minimum if the local minimum is with respect to the weak/strong norm. The designations of the weak and the strong norm are sometimes reversed in literature, but there is a general agreement about the terminology for weak/strong local minima. It holds that $\|\mathbf{y}^* - \mathbf{y}\|_0 \leq \|\mathbf{y}^* - \mathbf{y}\|_1$ and a point which is in the weak neighbourhood $\mathcal{B}_r^1(\mathbf{y})$ is therefore also in the strong neighbourhood $\mathcal{B}_r^0(\mathbf{y})$, i.e. $\mathcal{B}_r^1(\mathbf{y}) \subseteq \mathcal{B}_r^0(\mathbf{y})$. A strong local minimum minimises J with respect to all neighbouring points in the strong neighbourhood $\mathcal{B}_r^0(\mathbf{y})$. A weak local minimum minimises J with respect to all neighbouring points in the weak neighbourhood $\mathcal{B}_r^1(\mathbf{y})$. A strong local minimum is necessarily also a weak local minimum, because $\mathcal{B}_r^1(\mathbf{y}) \subseteq \mathcal{B}_r^0(\mathbf{y})$, but the converse is not necessarily true.

We will now give two theorems which give necessary conditions for a weak and a strong local minimum.

Theorem 1 (Weak local minimum) *A weak local minimum of* $\min_{\mathbf{y} \in \mathcal{K}} J(\mathbf{y})$ *satisfies*

$$\mathrm{d}J(\mathbf{y}; \mathbf{v}) \geq 0 \quad \forall \mathbf{v} \in \mathcal{C}_\mathcal{K}(\mathbf{y}) \cap \mathcal{Y}_0(I, \mathbb{R}^n). \quad (26)$$

Proof We will prove the theorem with a reductio ad absurdum. If the theorem does not hold, then $\mathbf{y}$ is a weak local minimiser of $\min_{\mathbf{y} \in \mathcal{K}} J(\mathbf{y})$ for which there exists a $\mathbf{v} \in \mathcal{C}_\mathcal{K}(\mathbf{y})$ such that $\mathrm{d}J(\mathbf{y}; \mathbf{v}) < 0$. The Gâteaux derivative considers the family of comparison functions $\hat{\mathbf{y}}(\varepsilon, x) = \mathbf{y}(x) + \varepsilon \mathbf{v}(x)$. This family of comparison functions converges in the weak norm $\|\cdot\|_1$, i.e.

$$\lim_{\varepsilon \downarrow 0} \|\hat{\mathbf{y}}(\varepsilon) - \mathbf{y}\|_1 = \sum_{i=1}^{n} \lim_{\varepsilon \downarrow 0} \max_{x \in I} |\varepsilon v_i(x)| + \lim_{\varepsilon \downarrow 0} \operatorname*{ess\,sup}_{x \in I} |\varepsilon v_i'(x)| = 0.$$

Hence, there exists a function $\mathbf{y}^*$, in an arbitrarily small weak neighbourhood $\mathcal{B}_r^1(\mathbf{y})$ of $\mathbf{y}$, such that $J(\mathbf{y}^*) < J(\mathbf{y})$. This contradicts the assumption that $\mathbf{y}$ is a weak local minimiser. $\qquad \square$

We will elaborate a bit more on weak local minima and their relationships with the Gâteaux derivative. For a weak local minimum, it holds that there exists a $r > 0$

Fig. 2 The functions $y_i(x)$, $v_i(x)$ and $\hat{y}_i(\varepsilon, x) = y_i(x) + \varepsilon v_i(x)$
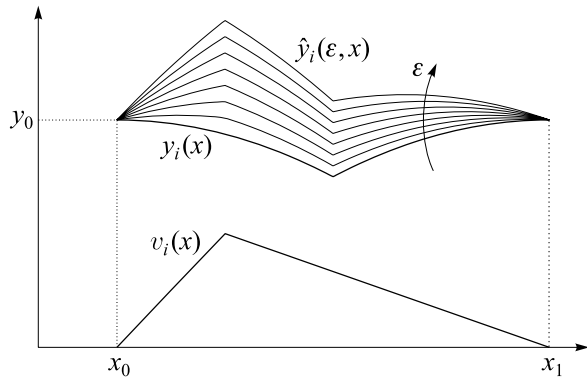
such that $J(\boldsymbol{y}^*) \geq J(\boldsymbol{y})$ for all $\boldsymbol{y}^* \in \mathcal{B}_r^1(\boldsymbol{y})$; see Definition 1. We can address each $\boldsymbol{y}^*$ in $\mathcal{B}_r^1(\boldsymbol{y})$ by considering families of comparison functions $\hat{\boldsymbol{y}}(\varepsilon)$ which satisfy

$$\lim_{\varepsilon \downarrow 0} \|\hat{\boldsymbol{y}}(\varepsilon) - \boldsymbol{y}\|_1 = 0, \tag{27}$$

with $\hat{\boldsymbol{y}}(0) = \boldsymbol{y}$ and $\hat{\boldsymbol{y}}(\varepsilon^*) = \boldsymbol{y}^*$ for some $\varepsilon^* \geq 0$.

The weak norm $\|\cdot\|_1$ measures the point-wise difference of $\hat{y}_i(\varepsilon, x) - y_i(x)$ as well as the point-wise difference of $\hat{y}_{i,x}(\varepsilon, x) - y_i'(x)$ (whenever they exist). Consequently, the vanishing of the limit (27) implies that

$$\lim_{\varepsilon \downarrow 0} \operatorname{ess\,sup}_{x \in I} |\hat{y}_{i,x}(\varepsilon, x) - y_i'(x)| = 0. \tag{28}$$

In other words, the maximal jump height of $\hat{y}_{i,x}(\varepsilon, x) - y_i'(x)$ on I has to vanish when ε is approaching zero. Proposition 1 (see the Appendix) implies that all discontinuity points of y_i' are discontinuity points of $\hat{y}_{i,x}(\varepsilon)$ for sufficiently small ε. Moreover, the jump height of those discontinuity points of $\hat{y}_{i,x}(\varepsilon, x)$ which are not discontinuity points of $y_i'(x)$ must vanish for $\varepsilon \downarrow 0$.
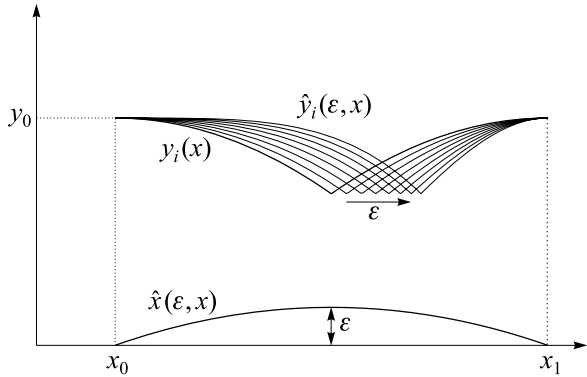
The family of comparison functions $\hat{y}_i(\varepsilon)$ is therefore of the form $\hat{y}_i(\varepsilon) = y_i + \varepsilon v_i + \mathcal{O}(\varepsilon^2)$, where $v_i \in \mathcal{Y}_0(I, \mathbb{R})$. This is illustrated in Fig. 2. The kink in the function $\hat{y}_i(\varepsilon, x)$ which is not a kink of $y_i(x)$ is due to a kink in $v_i(x)$ and, therefore, vanishes for $\varepsilon \downarrow 0$. If $\boldsymbol{y}$ is a weak local minimiser of J on $\mathcal{K}$, then it holds that $J(\hat{\boldsymbol{y}}(\varepsilon)) \geq J(\boldsymbol{y})$ for all $\hat{\boldsymbol{y}}(\varepsilon) \in \mathcal{B}_r^1(\boldsymbol{y})$ with sufficiently small $\varepsilon > 0$. The weak norm implies that we are only looking at comparison functions of the form $\hat{\boldsymbol{y}}(\varepsilon) = \boldsymbol{y} + \varepsilon \boldsymbol{v} + \mathcal{O}(\varepsilon^2) \in \mathcal{K}$. If $\boldsymbol{y} + \varepsilon \boldsymbol{v} + \mathcal{O}(\varepsilon^2) \in \mathcal{K}$ and ε is approaching 0, then $\boldsymbol{v}$ must lie in the contingent cone $\mathcal{C}_{\mathcal{K}}(\boldsymbol{y})$. Consequently, for a weak local minimum, it holds that

$$\lim_{\varepsilon \downarrow 0} \frac{J(\boldsymbol{y} + \varepsilon \boldsymbol{v}) - J(\boldsymbol{y})}{\varepsilon} \geq 0 \quad \forall \boldsymbol{v} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y}) \cap \mathcal{Y}_0(I, \mathbb{R}^n),$$

in which we recognise the Gâteaux derivative $\mathrm{d}J(\boldsymbol{y}; \boldsymbol{v})$.

Consider, for instance, the family of comparison functions $\hat{\boldsymbol{y}}(\varepsilon)$ which are of the form $\hat{\boldsymbol{y}}(\varepsilon, x) = \boldsymbol{y}(x - \hat{x}(\varepsilon, x)) + \varepsilon \boldsymbol{v}(x - \hat{x}(\varepsilon, x))$ involving a value shift $\varepsilon \boldsymbol{v}$ as well

Fig. 3 The functions $y_i(x)$, $\hat{x}(\varepsilon, x)$ and $\hat{y}_i(\varepsilon, x) = y_i(x - \hat{x}(\varepsilon, x))$

as a shift $\hat{x}(\varepsilon)$ of the argument with $\lim_{\varepsilon \downarrow 0} \hat{x}(\varepsilon, x) = 0$. If this family of comparison functions converges to y in the weak norm $\| \cdot \|_1$, then it has to fulfill

$$\lim_{\varepsilon \downarrow 0} \operatorname{ess\,sup}_{x \in I} \left| y_i'\left(x - \hat{x}(\varepsilon, x)\right) + \varepsilon v_i'\left(x - \hat{x}(\varepsilon, x)\right) - y_i'(x) \right|$$

$$= \lim_{\varepsilon \downarrow 0} \operatorname{ess\,sup}_{x \in I} \left| y_i'\left(x - \hat{x}(\varepsilon, x)\right) - y_i'(x) \right| = 0,$$

which implies that $\hat{x}(\varepsilon, x_s) = 0$ if y_i' is discontinuous at x_s. We see from this example that a family of comparison functions which involves a shift of the argument at discontinuity points can not converge in the weak norm. This is illustrated in Fig. 3. The family of comparison function $\hat{y}_i(\varepsilon, x) = y_i(x - \hat{x}(\varepsilon, x))$ converges in the strong norm to $y_i(x)$ but not in the weak norm, because the kinks do not appear for the same values of x.

Subsequently, we would like to consider a strong local minimum of the constrained minimisation problem (18). We consider therefore a family of functions $\hat{y}(\varepsilon) \in \mathcal{K}$, with $\hat{y}(0) = y$, which converge in the *strong* norm $\| \cdot \|_0$. Using

$$h(\varepsilon) = J\left(\hat{y}(\varepsilon)\right) = \int_I f\left(x, \hat{y}(\varepsilon, x), \hat{y}_x(\varepsilon, x)\right) dx, \tag{29}$$

we define the variation δJ of the functional J as

$$\delta J = h'(0)\delta\varepsilon = \delta\varepsilon \lim_{\varepsilon \downarrow 0} \frac{J(\hat{y}(\varepsilon)) - J(y)}{\varepsilon} \tag{30}$$

for comparison functions $\hat{y}(\varepsilon)$ such that $\|\hat{y}(\varepsilon) - y\|_0 \to 0$ for $\varepsilon \downarrow 0$.

Theorem 2 (Strong local minimum) *A strong local minimum of* $\min_{y \in \mathcal{K}} J(y)$ *satisfies*

$$\delta J \geq 0 \quad \forall \delta y \in \mathcal{C}_{\mathcal{K}}(y). \tag{31}$$

Proof If $\boldsymbol{y}$ is a strong local minimiser of J on $\mathcal{K}$, then the function $h(\varepsilon)$ has a strong local minimum for $\varepsilon = 0$ which gives the sub-stationarity condition

$$0 \leq h'(0) = \lim_{\varepsilon \downarrow 0} \frac{J(\hat{\boldsymbol{y}}(\varepsilon)) - J(\boldsymbol{y})}{\varepsilon} \tag{32}$$

in which we recognise the variation δJ. The comparison function $\hat{\boldsymbol{y}}(\varepsilon)$ is in $\mathcal{K}$ for infinitely small values of ε if $\delta \boldsymbol{y} = \hat{\boldsymbol{y}}_\varepsilon|_{\varepsilon=0} \, \delta \varepsilon$ is in the contingent cone $\mathcal{C}_{\mathcal{K}}(\boldsymbol{y})$. $\qquad\square$

The properties (26) and (31) of weak and strong local minima can be viewed as sub-stationarity conditions. These properties can be used as *definition* for weak and strong local extrema of the functional J on $\mathcal{K}$.

Definition 2 (Weak local extremal) A weak local extremal $\boldsymbol{y}$ of J on $\mathcal{K}$ is defined as a function fulfilling

$$\mathrm{d}J(\boldsymbol{y}; \boldsymbol{v}) \geq 0 \quad \forall \boldsymbol{v} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y}) \cap \mathcal{Y}_0\big(I, \mathbb{R}^n\big).$$

Definition 3 (Strong local extremal) A strong local extremal $\boldsymbol{y}$ of J on $\mathcal{K}$ is defined as a function fulfilling

$$\delta J \geq 0 \quad \forall \delta \boldsymbol{y} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y}).$$

Note that $\delta \boldsymbol{y} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{y})$ fulfills the boundary condition $\delta \boldsymbol{y}(x) = \boldsymbol{0}$ for $x \in \mathrm{bdry}\, I$, because $\mathcal{K} \subset \mathcal{D}(I, \mathbb{R})$. Weak and strong local extrema will play an important role in the following sections as the sub-stationarity conditions (26) and (31) constitute variational inequalities.

If $\mathcal{K}$ is of the form (10)

$$\mathcal{K} = \big\{ \boldsymbol{y} \in \mathcal{D}\big(I, \mathbb{R}^n\big) \mid \boldsymbol{y}(x) \in K \,\, \forall x \in I \big\}, \tag{33}$$

where K is a closed subset of $\mathbb{R}^n$, then we can write a variational inequality such as (23) by using the notation

$$\mathrm{d}J(\boldsymbol{y}; \boldsymbol{v}) \geq 0 \quad \forall \boldsymbol{v}(x) \in \mathcal{C}_K\big(\boldsymbol{y}(x)\big) \text{ almost everywhere on } I, \boldsymbol{v} \in \mathcal{Y}_0\big(I, \mathbb{R}^n\big), \tag{34}$$

or in the more sloppy form

$$\mathrm{d}J(\boldsymbol{y}; \boldsymbol{v}) \geq 0 \quad \forall \boldsymbol{v} \in \mathcal{C}_K(\boldsymbol{y}) \text{ a.e. on } I, \boldsymbol{v} \in \mathcal{Y}_0\big(I, \mathbb{R}^n\big), \tag{35}$$

in which the argument x has been suppressed. We will use the latter notation in the following sections.

4 The Principle of d'Alembert–Lagrange

In this section, we introduce the concept of a prefect constraint through the principle of d'Alembert–Lagrange. We first discuss the principle of d'Alembert–Lagrange for a

bilateral constraint. Subsequently, we discuss a unilateral constraint and define a perfect unilateral constraint through the principle of d'Alembert–Lagrange in inequality form (Goeleven et al. 1997; Panagiotopoulos and Glocker 1998).

Consider a point-mass m with position $\boldsymbol{r}(t) \in \mathbb{R}^3$ which is confined by a bilateral geometric scleronomic constraint $g(\boldsymbol{r})$ with $g \in C^1(\mathbb{R}^3, \mathbb{R})$. In other words, the point-mass is constrained to the two-dimensional manifold $\mathcal{M} = \{\boldsymbol{r} \in \mathbb{R}^3 \mid g(\boldsymbol{r}) = 0\}$ embedded in the $\mathbb{R}^3$; see Fig. 4a. The constraint is enforced by the constraint force $\boldsymbol{R}$. Furthermore, the mass is subjected to an external force $\boldsymbol{F}$. Newton's second law for the point-mass therefore reads as

$$m\ddot{\boldsymbol{r}} = \boldsymbol{R} + \boldsymbol{F}. \tag{36}$$

The principle of virtual work states that if the virtual work (Glocker 2001) of the system vanishes for all virtual displacements $\delta \boldsymbol{r}$, i.e.

$$\delta W = (m\ddot{\boldsymbol{r}} - \boldsymbol{F} - \boldsymbol{R})^{\mathrm{T}} \delta \boldsymbol{r} = 0 \quad \forall \delta \boldsymbol{r}, \tag{37}$$

then the system is in dynamic equilibrium (the Newton–Euler equations are fulfilled). The virtual displacements $\delta \boldsymbol{r}$ in (37) are *arbitrary* and can therefore be admissible as well as inadmissible with respect to the constraint $\boldsymbol{r} \in \mathcal{M}$, see Glocker (2001). Virtual displacements $\delta \boldsymbol{r}$ are admissible with respect to the constraint if they belong to the tangent space $\mathcal{T}_{\mathcal{M}}(\boldsymbol{r}) = \{\boldsymbol{z} \in \mathbb{R}^3 \mid \frac{\partial g}{\partial \boldsymbol{r}} \boldsymbol{z} = 0\}$. If the virtual work of the constraint force vanishes for all virtual displacements which are admissible with respect to the constraint, i.e.

$$\boldsymbol{R}^{\mathrm{T}} \delta \boldsymbol{r} = 0 \quad \forall \delta \boldsymbol{r} \in \mathcal{T}_{\mathcal{M}}(\boldsymbol{r}), \tag{38}$$

then we speak of a *perfect* bilateral constraint. A perfect constraint force of a bilateral constraint is therefore 'normal' to the constraint manifold in the sense that

$$-\boldsymbol{R} \in \mathcal{N}_{\mathcal{M}}(\boldsymbol{r}), \tag{39}$$

where $\mathcal{N}_{\mathcal{M}}(\boldsymbol{r})$ is the set of all vectors which annihilate the elements of the tangent space $\mathcal{T}_{\mathcal{M}}(\boldsymbol{r})$. Instead of taking (38) as definition of a perfect constraint, we could also have chosen to take the normality condition (39) of the constraint force as definition. The supposition of perfect bilateral constraints in classical mechanics and, therefore, the normality of the constraint force to the constraint manifold, excludes phenomena such as friction. Combining (38) and (37) gives the variational equality

$$(m\ddot{\boldsymbol{r}} - \boldsymbol{F})^{\mathrm{T}} \delta \boldsymbol{r} = 0 \quad \forall \delta \boldsymbol{r} \in \mathcal{T}_{\mathcal{M}}(\boldsymbol{r}), \tag{40}$$

which is usually referred to as the classical principle of d'Alembert–Lagrange.

We now impose in addition unilateral geometric scleronomic constraints $g_i(\boldsymbol{r}) \geq 0$, $i = 1 \ldots m$. The position of the point-mass is now constrained to the subset $\mathcal{K} = \{\boldsymbol{r} \in \mathcal{M} \mid g_i(\boldsymbol{r}) \geq 0, i = 1 \ldots m\}$, which is a sub-manifold with boundary; see Fig. 4b. Virtual displacements $\delta \boldsymbol{r}$ are admissible with respect to the bilateral and unilateral constraints if $\delta \boldsymbol{r} \in \mathcal{C}_{\mathcal{K}}(\boldsymbol{r})$, where $\mathcal{C}_{\mathcal{K}}(\boldsymbol{r})$ is the contingent cone (6) on the subset $\mathcal{K}$. In the following, we will confine ourselves to the case that $\mathcal{K}$ is a tangentially regular set
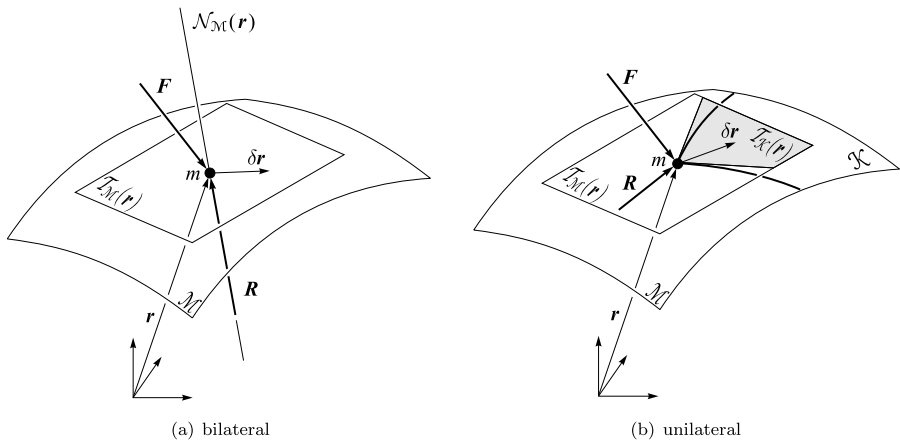
Fig. 4 Constrained point-mass

for which the contingent cone agrees with the tangent cone (7). We define a unilateral constraint as *perfect* if the constraint force satisfies the normality condition (39)

$$-\boldsymbol{R} \in \mathcal{N}_{\mathcal{K}}(\boldsymbol{r}), \tag{41}$$

where $\mathcal{N}_{\mathcal{K}}(\boldsymbol{r})$ is the normal cone (8) to the subset $\mathcal{K}$. The tangent cone $\mathcal{T}_{\mathcal{K}}(\boldsymbol{r})$ and normal cone $\mathcal{N}_{\mathcal{K}}(\boldsymbol{r})$ are polar in the sense that for all $\boldsymbol{R}$ and $\delta\boldsymbol{r}$ satisfying $-\boldsymbol{R} \in \mathcal{N}_{\mathcal{K}}(\boldsymbol{r})$ and $\delta\boldsymbol{r} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{r})$ it holds that $-\boldsymbol{R}^{\mathrm{T}}\delta\boldsymbol{r} \leq 0$. Hence, for perfect unilateral constraints, it holds that the virtual work of the constraint force is non-negative for kinematically admissible virtual displacements, i.e.

$$\boldsymbol{R}^{\mathrm{T}}\delta\boldsymbol{r} \geq 0 \quad \forall \delta\boldsymbol{r} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{r}), \tag{42}$$

which is sometimes referred to as Fourier's inequality (Lanczos 1962). Combining (42) and (37) gives the variational inequality

$$(m\ddot{\boldsymbol{r}} - \boldsymbol{F})^{\mathrm{T}}\delta\boldsymbol{r} \geq 0 \quad \forall \delta\boldsymbol{r} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{r}), \tag{43}$$

which we will refer to as the principle of d'Alembert–Lagrange in inequality form.

5 The Principle of Hamilton for Non-impulsive Motion

The principle of Hamilton is derived in this section for a system with perfect bilateral as well as unilateral constraints and is shown to be a variational inequality. In this section, the motion is assumed to be non-impulsive. As a consequence, impulses are absent and velocities are absolutely continuous functions in time. The principle of Hamilton for perfect unilateral constraints with impulsive motion will be derived in Sects. 6 and 7.

Consider a system S consisting of material points, e.g. an elastic continuum body or a multi-body system consisting of rigid bodies. For convenience, we will make

the tacit assumption that each material point has a different position vector X in the reference configuration at a reference time t_0. We can therefore address each material point in S by its position vector $X \in \mathbb{R}^3$ in the reference configuration. Each material point X has a position x in the current configuration at time t. We can define a vector field $\boldsymbol{\xi}$ which carries material points X to places $x = \boldsymbol{\xi}(X, t)$ in their current configuration. The vector field $\boldsymbol{\xi}$ is called the motion of the system and induces the velocity field $\dot{\boldsymbol{\xi}}$ and acceleration field $\ddot{\boldsymbol{\xi}}$ defined by

$$\dot{\boldsymbol{\xi}}(X, t) = \frac{\partial \boldsymbol{\xi}(X, t)}{\partial t}, \qquad \ddot{\boldsymbol{\xi}}(X, t) = \frac{\partial^2 \boldsymbol{\xi}(X, t)}{\partial t^2}, \tag{44}$$

whenever they exist. The system S is subjected to bilateral geometric scleronomic constraints, which constrain the motion $\boldsymbol{\xi}$ to an n-dimensional differentiable sub-manifold $\mathcal{M}$ of the configuration space, which is called the configuration manifold. Furthermore, the system is subjected to unilateral geometric scleronomic constraints, which constrain the motion $\boldsymbol{\xi}$ to a closed subset $\mathcal{K}$ of configuration manifold $\mathcal{M}$. The set of admissible motions $\mathcal{K}$ is therefore a sub-manifold with boundary. Let $\mathrm{d}m$ be the mass element at $\boldsymbol{\xi}(X, t)$. The mass-element $\mathrm{d}m$ is subjected to external and internal forces $\mathrm{d}F$, which are not due to the constraints (e.g. elastic forces and gravitational forces). We will assume that these forces $\mathrm{d}F$ are potential forces, which will allow us to formulate a principle of Hamilton. Furthermore, the mass-element $\mathrm{d}m$ is subjected to constraint forces $\mathrm{d}R$ which impose the bilateral and unilateral geometric scleronomic constraints.

We define the virtual work of the system S as

$$\delta W = \int_S \delta \boldsymbol{\xi}^{\mathrm{T}} (\ddot{\boldsymbol{\xi}} \, \mathrm{d}m - \mathrm{d}F - \mathrm{d}R), \tag{45}$$

which exists almost everywhere on the time-interval I. The principle of virtual work states that if the virtual work vanishes for all virtual variations $\delta \boldsymbol{\xi}$, i.e.

$$\delta W = 0 \quad \forall \delta \boldsymbol{\xi} \text{ a.e. on } I, \tag{46}$$

then the system is in dynamic equilibrium. The constraint forces $\mathrm{d}R$ are assumed to be perfect and it therefore holds that

$$\mathrm{d}R^{\mathrm{T}} \delta \boldsymbol{\xi} \geq 0 \quad \forall \delta \boldsymbol{\xi} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{\xi}) \text{ a.e. on } I. \tag{47}$$

A virtual displacement $\delta \boldsymbol{\xi}$ which is in the tangent cone $\mathcal{T}_{\mathcal{K}}(\boldsymbol{\xi})$ is kinematically admissible with respect to the constraint $\boldsymbol{\xi} \in \mathcal{K}$. The principle of d'Alembert–Lagrange in inequality form (43) for the system S therefore reads as

$$\int_S \delta \boldsymbol{\xi}^{\mathrm{T}} (\ddot{\boldsymbol{\xi}} \, \mathrm{d}m - \mathrm{d}F) \geq 0 \quad \forall \delta \boldsymbol{\xi} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{\xi}) \text{ a.e. on } I. \tag{48}$$

We now can define a chart with coordinates $q \in \mathbb{R}^n$ by choosing a one-to-one mapping $\boldsymbol{\varphi} : U \subset \mathbb{R}^n \to \boldsymbol{\varphi}(U) \subset \mathcal{M}$ which maps the coordinates $q \in U \subset \mathbb{R}^n$ in some open set U to a subset of the configuration manifold $\mathcal{M}$. The motion $\boldsymbol{\xi}(\cdot, t)$ is

therefore for each $t \in I$ uniquely determined (at least within some open set) by the generalised coordinates $\boldsymbol{q}$ through the mapping $\boldsymbol{\varphi}$, i.e.

$$\boldsymbol{\xi}(\cdot, t) = \boldsymbol{\varphi}\big(\boldsymbol{q}(t)\big). \tag{49}$$

The coordinates $\boldsymbol{q}$ form a set of generalised coordinates which are minimal with respect to the bilateral constraints. The unilateral constraints, which enforce the constraint $\boldsymbol{\xi} \in \mathcal{K}$, induce a set $K = \boldsymbol{\varphi}^{-1}(\mathcal{K})$ of admissible generalised coordinates which we assume to be tangentially regular. The tangent cone $\mathcal{T}_K(\boldsymbol{q})$ therefore agrees with the contingent cone and is a subset of the chart tangent space. Let $\frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}}$ denote the mapping from the chart tangent space to the tangent space $\mathcal{T}_{\mathcal{M}}(\boldsymbol{\xi})$ of the configuration manifold. An arbitrary virtual displacement $\delta \boldsymbol{q}$ on the chart tangent space causes a variation of the motion $\delta \boldsymbol{\xi} = \frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}} \delta \boldsymbol{q} \in \mathcal{T}_{\mathcal{M}}(\boldsymbol{\xi})$, which is a element of the tangent space $\mathcal{T}_{\mathcal{M}}(\boldsymbol{\xi})$ and which is therefore kinematically admissible with respect to the bilateral constraints. A virtual displacement $\delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q})$ induces a variation of the motion $\delta \boldsymbol{\xi} = \frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}} \delta \boldsymbol{q} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{\xi})$ which is in addition admissible with respect to the unilateral constraints.

The virtual work of the inertia forces in (48) can be rewritten as

$$\int_S \delta \boldsymbol{\xi}^{\mathrm{T}} \ddot{\boldsymbol{\xi}} \, \mathrm{d}m = \frac{\mathrm{d}}{\mathrm{d}t} \left(\int_S \delta \boldsymbol{\xi}^{\mathrm{T}} \dot{\boldsymbol{\xi}} \, \mathrm{d}m \right) - \int_S \delta \dot{\boldsymbol{\xi}}^{\mathrm{T}} \dot{\boldsymbol{\xi}} \, \mathrm{d}m, \tag{50}$$

or, by using the generalised coordinates $\boldsymbol{q}$, as

$$\int_S \delta \boldsymbol{\xi}^{\mathrm{T}} \ddot{\boldsymbol{\xi}} \, \mathrm{d}m = \frac{\mathrm{d}}{\mathrm{d}t} \left(\delta \boldsymbol{q}^{\mathrm{T}} \int_S \left(\frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}} \right)^{\mathrm{T}} \dot{\boldsymbol{\xi}} \, \mathrm{d}m \right) - \int_S \delta \dot{\boldsymbol{\xi}}^{\mathrm{T}} \dot{\boldsymbol{\xi}} \, \mathrm{d}m$$

$$= \frac{\mathrm{d}}{\mathrm{d}t} \big(\delta \boldsymbol{q}^{\mathrm{T}} \boldsymbol{p} \big) - \delta T, \tag{51}$$

where

$$\boldsymbol{p} = \int_S \left(\frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}} \right)^{\mathrm{T}} \dot{\boldsymbol{\xi}} \, \mathrm{d}m \tag{52}$$

is the generalised momentum and δT is the variation of the kinetic energy

$$T = \int_S \frac{1}{2} \dot{\boldsymbol{\xi}}^{\mathrm{T}} \mathrm{d}m \, \dot{\boldsymbol{\xi}} = \frac{1}{2} \dot{\boldsymbol{q}}^{\mathrm{T}} \boldsymbol{M}(\boldsymbol{q}) \dot{\boldsymbol{q}}, \quad \boldsymbol{M}(\boldsymbol{q}) = \int_S \left(\frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}} \right)^{\mathrm{T}} \mathrm{d}m \left(\frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}} \right). \tag{53}$$

Moreover, we introduce

$$\boldsymbol{f} = \int_S \left(\frac{\partial \boldsymbol{\varphi}}{\partial \boldsymbol{q}} \right)^{\mathrm{T}} \mathrm{d}\boldsymbol{F} \tag{54}$$

as the generalised force. The forces $\mathrm{d}\boldsymbol{F}$ are assumed to be potential forces and the generalised force $\boldsymbol{f}$ is therefore also a potential force $-\boldsymbol{f} = (\partial V / \partial \boldsymbol{q})^{\mathrm{T}}$ where $V(\boldsymbol{q})$ is the potential energy and $\delta V = -\delta \boldsymbol{q}^{\mathrm{T}} \boldsymbol{f}$. Substitution of (51) and (54) in the principle of d'Alembert–Lagrange (48) gives

$$\frac{\mathrm{d}}{\mathrm{d}t} \big(\delta \boldsymbol{q}^{\mathrm{T}} \boldsymbol{p} \big) - \delta T - \delta \boldsymbol{q}^{\mathrm{T}} \boldsymbol{f} \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}) \text{ a.e. on } I, \tag{55}$$

which is the inequality form of the well-known Lagrange central equation (Bremer 1988; Hamel 1912; Papastavridis 2002). The Lagrange central equation (55) holds at each time-instant t for which the generalised velocities $\dot{\boldsymbol{q}}(t)$ and accelerations $\ddot{\boldsymbol{q}}(t)$ exist. If the motion on I is assumed to be non-impulsive, then the generalised positions $\boldsymbol{q}(t)$ and generalised velocities $\dot{\boldsymbol{q}}(t)$ are absolutely continuous on I and the generalised accelerations $\ddot{\boldsymbol{q}}(t)$ exist for almost all $t \in I$. Hence, we can integrate the central equation (55) over a non-impulsive time-interval $I = [t_0, t_f]$ which gives

$$\left[\delta \boldsymbol{q}^{\mathrm{T}} \boldsymbol{p}\right]_{t_0}^{t_f} - \int_I \left(\delta T + \delta \boldsymbol{q}^{\mathrm{T}} \boldsymbol{f}\right) \mathrm{d}t \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}). \tag{56}$$

By defining the Lagrange function $L = T - V$ we arrive at

$$\left[\delta \boldsymbol{q}^{\mathrm{T}} \boldsymbol{p}\right]_{t_0}^{t_f} - \int_I \delta L \, \mathrm{d}t \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}). \tag{57}$$

If the boundary conditions are fixed, then the variation $\delta \boldsymbol{q}(t)$ vanishes at $t = t_0$ and $t = t_f$ and we are allowed to interchange the order of integration and variation such that

$$-\delta \int_I L \, \mathrm{d}t \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}), \qquad \boldsymbol{q}(t_0) = \boldsymbol{q}_0, \qquad \boldsymbol{q}(t_f) = \boldsymbol{q}_f, \tag{58}$$

which is the principle of Hamilton in inequality form for a non-impulsive time-interval $I = [t_0, t_f]$. From the principle of Hamilton, we can derive the Euler–Lagrange equations in inequality form by evaluating the variation in (58) as

$$-\int_I \left(\frac{\partial L}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) \delta \boldsymbol{q} \, \mathrm{d}t \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}), \qquad \boldsymbol{q}(t_0) = \boldsymbol{q}_0, \qquad \boldsymbol{q}(t_f) = \boldsymbol{q}_f. \tag{59}$$

Hence, for almost all $t \in I$, the variational inequality

$$-\left(\frac{\partial L}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right) \delta \boldsymbol{q} \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}) \tag{60}$$

should hold, which can be cast into the form

$$\left(\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} - \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}})\right)^{\mathrm{T}} \delta \boldsymbol{q} \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}) \quad \text{a.e. on } I, \tag{61}$$

where the mass matrix $\boldsymbol{M}(\boldsymbol{q})$ is defined by (53) and the vector

$$\boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) = -\left(\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{M}(\boldsymbol{q})\right) \dot{\boldsymbol{q}} + \left(\frac{\partial T}{\partial \boldsymbol{q}} - \frac{\partial V}{\partial \boldsymbol{q}}\right)^{\mathrm{T}} \tag{62}$$

contains all smooth forces. The variational inequality (61) leads to the differential inclusion

$$\boldsymbol{M}(\boldsymbol{q})\ddot{\boldsymbol{q}} - \boldsymbol{h}(\boldsymbol{q}, \dot{\boldsymbol{q}}) - \boldsymbol{f}_{\mathrm{R}} = \boldsymbol{0}, \quad -\boldsymbol{f}_{\mathrm{R}} \in \mathcal{N}_K(\boldsymbol{q}) \text{ a.e. on } I, \tag{63}$$

where $\boldsymbol{f}_{\mathrm{R}}$ is the generalised constraint force, which fulfills the principle of d'Alembert–Lagrange in inequality form

$$\boldsymbol{f}_{\mathrm{R}}^{\mathrm{T}} \delta \boldsymbol{q} \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_K(\boldsymbol{q}). \tag{64}$$

6 The Strong Principle of Hamilton

Up to now, the principle of Hamilton in inequality form (58) has been proven to hold for non-impulsive motion. Impacts and discontinuities in the velocity have therefore not been considered. Now the question rises: Under which conditions does the principle of Hamilton in inequality form (58) also hold for a time-interval with impulsive motion? We will derive necessary conditions for a strong local extremal of the action integral by making use of the general variation of a functional. This will lead to conditions which are similar to the classical Weierstrass–Erdmann corner conditions. The main idea (see Panagiotopoulos and Glocker 1998, 2000) is to split the time-interval I in two non-impulsive sub-intervals I_1 and I_2 and to set up the variational conditions at the impact time t_c which joins the two sub-intervals I_1 and I_2. This requires the variation of the action integral over time-intervals with variable begin or end time, i.e. a general variation of the action integral. This section is organised as follows. First, the strong variation of the position function $q(t)$ is introduced. Then the general variation of the action integral is derived. Subsequently, these results are used to treat the variation of the action integral over I by splitting the interval in two non-impulsive sub-intervals which leads to generalised Weierstrass–Erdmann corner conditions. Finally, we arrive at a strong principle of Hamilton in inequality form for impulsive motion.

Let $q(t) \in \mathcal{D}(I, \mathbb{R}^n)$ describe the motion of the system (on the chart) on the compact time-interval $I = [t_0, t_f]$. The velocity $u(t)$ is therefore of bounded variation, which means that we can define the left limit $u^-(t)$ and right limit $u^+(t)$ for each time t and $u(t) = \dot{q}(t)$ for almost all t. We consider the family of comparison functions $\hat{q}(\varepsilon, t) = q(t - \hat{t}(\varepsilon, t)) + \varepsilon w(t - \hat{t}(\varepsilon, t))$. The function $\hat{t}(\varepsilon, t)$, being continuous and differentiable with the property $\hat{t}(0, t) = 0$, induces a virtual time-shift $\delta t(t) = \hat{t}_\varepsilon(0, t) \delta\varepsilon$. The function $w \in \mathcal{Y}_0(I, \mathbb{R}^n)$ induces a value shift. The family of comparison functions $\hat{q}(\varepsilon, t)$ converges to $q(t)$ in the strong norm $\| \cdot \|_0$, because $\lim_{\varepsilon \downarrow 0} \hat{q}(\varepsilon, t) = \hat{q}(0, t) = q(t)$. The variation

$$\delta q(t) = \left. \frac{\partial \hat{q}(\varepsilon, t)}{\partial \varepsilon} \right|_{\varepsilon=0} \delta\varepsilon = -u(t)\,\delta t(t) + w(t)\,\delta\varepsilon \qquad (65)$$

is discontinuous and not defined for those time-instants for which $u(t)$ is discontinuous and not defined. Hence, the variation $\delta q(t)$ is of bounded variation and admits for each $t \in I$ a left and a right limit

$$\delta q^\pm(t) = -u^\pm(t)\,\delta t(t) + w(t)\,\delta\varepsilon. \qquad (66)$$

The variation $\delta q(t)$ should be understood as an infinitesimal difference between the comparison function $\hat{q}(\varepsilon, t)$ and the function $q(t)$ for the same value of t. Likewise, we can introduce the variation $\overline{\delta q}(t) \approx \hat{q}(\varepsilon, t + \delta t(t)) - q(t)$, or more explicitly using (65)

$$\overline{\delta q}(t) = q_t(t)\,\delta t(t) + \hat{q}_\varepsilon(0, t)\,\delta\varepsilon = u(t)\,\delta t(t) + \delta q(t) = w(t)\,\delta\varepsilon, \qquad (67)$$

for which also the time is varied. Clearly, the variation $\overline{\delta q}(t)$ is continuous, because $w(t)$ is continuous.

Fig. 5 The general variation of a function $q(t)$
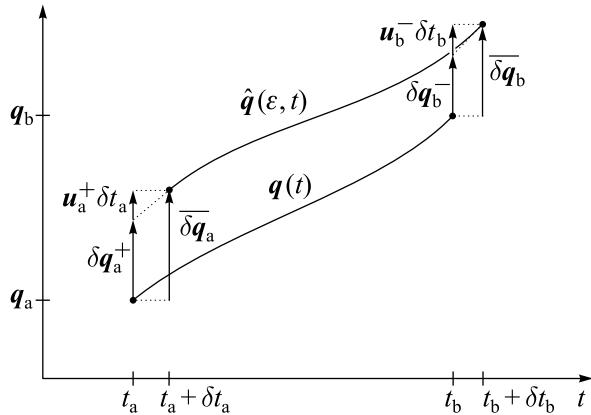
We first consider a time-interval $[t_a, t_b] \subset I$ on which the motion $q(t)$ is differentiable in its interior, i.e. $u(t)$ is continuous on the open interval (t_a, t_b). The general variation of the action integral

$$s(q) = \int_{t_a}^{t_b} L \, dt \tag{68}$$

involves a variation of the begin point $q(t_a) = q_a$ and end point $q(t_b) = q_b$ as well as a variable begin time t_a and end time t_b; see Fig. 5. We introduce the function

$$h(\varepsilon) = \int_{t_a(\varepsilon)}^{t_b(\varepsilon)} L\big(\hat{q}(\varepsilon, t), \hat{u}(\varepsilon, t)\big) \, dt, \tag{69}$$

where $\hat{q}(\varepsilon, t) = q(t - \hat{t}(\varepsilon, t)) + \varepsilon w(t - \hat{t}(\varepsilon, t))$. The general variation δs of (68) is defined as the ε-derivative $\delta s = h'(0) \, \delta\varepsilon$, i.e.

$$\delta s = \int_{t_a}^{t_b} \delta L \, dt + [L \, \delta t]_{t \downarrow t_a}^{t \uparrow t_b}$$

$$= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial u} \delta u \right) dt + [L \, \delta t]_{t \downarrow t_a}^{t \uparrow t_b}$$

$$= \int_{t_a}^{t_b} \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \frac{\partial L}{\partial u} \right) \delta q \, dt + \left[\frac{\partial L}{\partial u} \delta q + L \, \delta t \right]_{t \downarrow t_a}^{t \uparrow t_b}. \tag{70}$$

The boundary terms in (70) are due to the variation of the begin time t_a and end time t_b (see Fig. 5), and are expressed with left and right limits because the velocity u and, therefore, L may not exist for $t = t_a$ or $t = t_b$. Note that $\delta q(t)$ is continuous in the open interval (t_a, t_b). Using the generalised momentum p (which equals (52)) and the Hamiltonian function H, i.e.

$$p(q, u) = \left(\frac{\partial L}{\partial u} \right)^{\mathrm{T}}, \qquad H(q, u) = \frac{\partial L}{\partial u} u - L = T + V, \tag{71}$$

Fig. 6 The general variation of a function $q(t)$ with kink
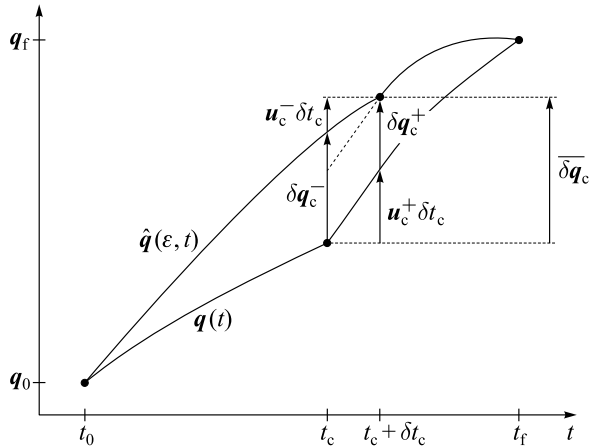
we can express the boundary terms in (70) as

$$\lim_{t \downarrow t_a}\left(\frac{\partial L}{\partial u}\delta q + L\,\delta t\right) = p\big(q_a, u_a^+\big)^{\mathrm{T}}\big(\delta q_a^+ + u_a^+\,\delta t_a\big) - H\big(q_a, u_a^+\big)\,\delta t_a, \quad (72)$$

$$\lim_{t \uparrow t_b}\left(\frac{\partial L}{\partial u}\delta q + L\,\delta t\right) = p\big(q_b, u_b^-\big)^{\mathrm{T}}\big(\delta q_b^- + u_b^-\,\delta t_b\big) - H\big(q_b, u_b^-\big)\,\delta t_b, \quad (73)$$

in which the abbreviation $\delta q^+(t_a) = \delta q_a^+$, $u^+(t_a) = u_a^+$, etc. has been used. We recognise in (72) and (73) the variation $\overline{\delta q}(t)$ of the begin and endpoint

$$
\begin{aligned}
\overline{\delta q}_a &= \delta q_a^+ + u_a^+\,\delta t_a, \\
\overline{\delta q}_b &= \delta q_b^- + u_b^-\,\delta t_b,
\end{aligned}
\quad (74)
$$

see Fig. 5. The total variation (70) simplifies to

$$\delta s = \int_{t_a}^{t_b}\left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial u}\right)\delta q\,\mathrm{d}t - \big(p_a^+\big)^{\mathrm{T}}\overline{\delta q}_a + (p_b^-)^{\mathrm{T}}\overline{\delta q}_b + H_a^+\,\delta t_a - H_b^-\,\delta t_b, \quad (75)$$

where the notation $p_a^+ = p(q_a, u_a^+)$, $H_a^+ = H(q_a, u_a^+)$, etc. has been used.

In the following, we will consider solution curves $q(t)$ which may have a kink at some point in time $t_c \in I$ (see Fig. 6). The action integral, which is a Lebesgue integral over an interval I, can be decomposed into two differentiable parts

$$s(q) = \int_I L\,\mathrm{d}t = \int_{I_1} L\,\mathrm{d}t + \int_{I_2} L\,\mathrm{d}t = s_1(q) + s_2(q), \quad (76)$$

where $I_1 = [t_0, t_c]$ and $I_2 = [t_c, t_f]$. The end points $t = t_0$ and $t = t_f$ are fixed and we require that the two differentiable parts of the function $q(t)$ join continuously at $t = t_c$, but otherwise the point $t = t_c$ can move freely. The comparison function $\hat{q}(\varepsilon, t)$ is fixed at the end points $t = t_0$ and $t = t_f$ and consists of two differentiable parts which join continuously at $t = t_c + \delta t_c$. The function $\overline{\delta q}(t)$ is therefore continuous

at $t = t_c$, whereas the variation $\delta q(t)$ is discontinuous as $t = t_c$. The variation of the impact position $\overline{\delta q}_c = \overline{\delta q}(t_c)$ can be assessed from the left and from the right which gives the equality

$$\overline{\delta q}_c = \delta q_c^- + u_c^- \, \delta t_c = \delta q_c^+ + u_c^+ \, \delta t_c, \qquad (77)$$

with the notation $\delta q^\pm(t_c) = \delta q_c^\pm$; see Fig. 6. It holds that $-u_c^- \in \mathcal{T}_K(q(t_c))$ and $u_c^+ \in \mathcal{T}_K(q(t_c))$. Therefore, if $\delta t_c < 0$ and $\delta q_c^- \in \mathcal{T}_K(q(t_c))$, we have $\overline{\delta q}_c = \delta q_c^- + u_c^- \, \delta t_c \in \mathcal{T}_K(q(t_c))$, whereas if $\delta t_c > 0$ and $\delta q_c^+ \in \mathcal{T}_K(q(t_c))$ we have $\overline{\delta q}_c = \delta q_c^+ + u_c^+ \, \delta t_c \in \mathcal{T}_K(q(t_c))$. Hence, if $\delta q_c^\pm \in \mathcal{T}_K(q(t_c))$, then it also holds that $\overline{\delta q}_c \in \mathcal{T}_K(q(t_c))$.

We introduce the functions

$$h_1(\varepsilon) = \int_{t_0}^{t_c(\varepsilon)} L\big(\hat{q}(\varepsilon, t), \hat{u}(\varepsilon, t)\big) \, \mathrm{d}t, \qquad h_2(\varepsilon) = \int_{t_c(\varepsilon)}^{t_f} L\big(\hat{q}(\varepsilon, t), \hat{u}(\varepsilon, t)\big) \, \mathrm{d}t \quad (78)$$

such that $h(\varepsilon) = h_1(\varepsilon) + h_2(\varepsilon)$ and take the variation of the action integral s by calculating the ε-derivatives $h_1'(\varepsilon)$ and $h_2'(\varepsilon)$ separately. In other words, the variation δs_1 has a variable end-point whereas δs_2 has a variable starting-point. Evaluation of $h_1'(\varepsilon)$ and $h_2'(\varepsilon)$ gives

$$\delta s_1 = h_1'(0)\delta\varepsilon = \int_{I_1} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial u} \right) \delta q \, \mathrm{d}t + (p_c^-)^{\mathrm{T}}\overline{\delta q}_c - H_c^- \, \delta t_c, \qquad (79)$$

$$\delta s_2 = h_2'(0)\delta\varepsilon = \int_{I_2} \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial u} \right) \delta q \, \mathrm{d}t - \big(p_c^+\big)^{\mathrm{T}}\overline{\delta q}_c + H_c^+ \, \delta t_c. \qquad (80)$$

Addition of $h_1'(0)$ and $h_2'(0)$ yields $h'(0)$. The variation of s can be written as

$$\delta s = \int_I \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial u} \right) \delta q \, \mathrm{d}t - \big(p_c^+ - p_c^-\big)^{\mathrm{T}}\overline{\delta q}_c + \big(H_c^+ - H_c^-\big) \delta t_c, \qquad (81)$$

where $p_c^\pm$ and $H_c^\pm$ are the pre- and post-impact values of the canonical variables.

We now return to the question of finding conditions for which the principle of Hamilton in inequality form (58) also holds for a time-interval I which contains an impact event. If (58) holds, in other words if $-\delta s \geq 0$ for all $\delta q \in \mathcal{T}_K(q)$, then (81) yields

$$-\int_I \left(\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial u} \right) \delta q \, \mathrm{d}t + \big(p_c^+ - p_c^-\big)^{\mathrm{T}}\overline{\delta q}_c - \big(H_c^+ - H_c^-\big)\delta t_c \geq 0 \qquad (82)$$

for all $\delta q(t) \in \mathcal{T}_K(q(t))$, where $t \in I \backslash \{t_c\}$, for all $\overline{\delta q}_c \in \mathcal{T}_K(q(t_c))$ and for all δt_c. The first term, $-\int_I (\frac{\partial L}{\partial q} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial u}) \, \delta q \, \mathrm{d}t$, involves the non-impulsive motion and is clearly non-negative for admissible virtual displacements; see (60). This analysis suggests to make the following two suppositions about the nature of the impact:

1. The supposition of a perfect constraint remains valid during the impact. This implies that the change in momentum must be normal to the constraint in the sense that

$$-\big(p_c^+ - p_c^-\big) \in \mathcal{N}_K\big(q(t_c)\big), \qquad (83)$$

which is equivalent to

$$\overline{\delta \boldsymbol{q}}_{\mathrm{c}}^{\mathrm{T}}\big(\boldsymbol{p}_{\mathrm{c}}^{+} - \boldsymbol{p}_{\mathrm{c}}^{-}\big) \geq 0 \quad \forall \overline{\delta \boldsymbol{q}}_{\mathrm{c}} \in \mathcal{T}_{K}\big(\boldsymbol{q}(t_{\mathrm{c}})\big). \tag{84}$$

2. The impact is assumed to preserve the total energy in the system, i.e.

$$H_{\mathrm{c}}^{+} - H_{\mathrm{c}}^{-} = 0. \tag{85}$$

These assumptions are for the unconstrained case ($K = \mathbb{R}^{n}$, $\mathcal{N}_{K}(\boldsymbol{q}(t_{\mathrm{c}})) = \boldsymbol{0}$) exactly the so-called first and the second Weierstrass–Erdmann corner conditions. Given a functional $J(y) = \int_{I} f(x, y, y')\, \mathrm{d}x$, the first Weierstrass–Erdmann condition reads as

$$\lim_{x \downarrow x_{\mathrm{c}}} f_{y'} = \lim_{x \uparrow x_{\mathrm{c}}} f_{y'}, \quad \text{or} \quad \boldsymbol{p}_{\mathrm{c}}^{+} = \boldsymbol{p}_{\mathrm{c}}^{-}, \tag{86}$$

and the second Weierstrass–Erdmann condition reads as

$$\lim_{x \downarrow x_{\mathrm{c}}} f - y' f_{y'} = \lim_{x \uparrow x_{\mathrm{c}}} f - y' f_{y'}, \quad \text{or} \quad H_{\mathrm{c}}^{+} = H_{\mathrm{c}}^{-}. \tag{87}$$

The suppositions (83) and (85) lead to the following theorem (see Panagiotopoulos and Glocker 1998).

Theorem 3 (The strong principle of Hamilton in inequality form) *Consider a conservative Lagrangian mechanical system with perfect unilateral constraints and a non-dissipative impact law. A function $\boldsymbol{q}(t) \in K$ is a motion of the system if and only if it is a strong local extremal of the action integral, i.e.*

$$-\delta \int_{I} L \, \mathrm{d}t \geq 0 \quad \forall \delta \boldsymbol{q} \in \mathcal{T}_{K}(\boldsymbol{q}) \text{ a.e. on } I, \qquad \boldsymbol{q}(t_{0}) = \boldsymbol{q}_{0}, \qquad \boldsymbol{q}(t_{\mathrm{f}}) = \boldsymbol{q}_{\mathrm{f}}, \tag{88}$$

which is the strong principle of Hamilton in inequality form for impulsive motion.

Proof The supposition of a perfect unilateral constraint implies the normality condition $-(\boldsymbol{p}^{+} - \boldsymbol{p}^{-}) \in \mathcal{N}_{K}(\boldsymbol{q})$. The supposition of a non-dissipative impact implies the energy condition $H^{+} = H^{-}$. Hence, the conditions (83) and (85) are satisfied. If $\boldsymbol{q}(t)$ is a strong local extremal of the action integral then, by definition, $-\delta s \geq 0$ for all strong variations $\delta \boldsymbol{q} \in \mathcal{T}_{K}$ (88). Using (81), we obtain the variational problem (82). The variations $\delta \boldsymbol{q}(t)$, $\overline{\delta \boldsymbol{q}}_{\mathrm{c}}$ and δt_{c} are independent. Hence they yield two variational inequalities and one variational equality:

$$-\int_{I}\left(\frac{\partial L}{\partial \boldsymbol{q}} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\boldsymbol{q}}}\right)\delta \boldsymbol{q} \, \mathrm{d}t \geq 0 \quad \forall \delta \boldsymbol{q}(t) \in \mathcal{T}_{K}\big(\boldsymbol{q}(t)\big), \text{ almost everywhere on } I, \tag{89}$$

$$\big(\boldsymbol{p}_{\mathrm{c}}^{+} - \boldsymbol{p}_{\mathrm{c}}^{-}\big)^{\mathrm{T}}\overline{\delta \boldsymbol{q}}_{\mathrm{c}} \geq 0 \quad \forall \overline{\delta \boldsymbol{q}}_{\mathrm{c}} \in \mathcal{T}_{K}\big(\boldsymbol{q}(t_{\mathrm{c}})\big), \tag{90}$$

$$\big(H_{\mathrm{c}}^{+} - H_{\mathrm{c}}^{-}\big)\delta t_{\mathrm{c}} = 0 \quad \forall \delta t_{\mathrm{c}}. \tag{91}$$

From (89) we see that the Euler–Lagrange inequality (60) holds for almost all $t \in I$. Equation (90) requires that the unilateral constraints are perfect (in the sense of impulsive motion) and (91) that the impacts are non-dissipative. Consequently, if $\boldsymbol{q}$ is a

strong extremal of the action integral, then $q(t)$ is a motion of the system. The proof can easily be followed in the opposite direction.   □

7 A Weak Principle of Hamilton

In this section we derive a weak principle of Hamilton in inequality form for impulsive motion by directly incorporating the impulsive dynamics in the principle of virtual work. Concepts of measure and integration theory appear to be very useful in this respect.

As before, let $\boldsymbol{\xi}(X, t)$ describe the motion of a mass element dm with position X in the reference configuration. The configuration at time t is denoted by $\boldsymbol{\xi}(\cdot, t)$ but we will simply write $\boldsymbol{\xi}(t)$ and denote the velocity field by $\boldsymbol{v}(t)$. We assume that the motion $\boldsymbol{\xi}(t)$ is absolutely continuous in time and that the velocity field $\boldsymbol{v}(t)$ is a function (field) of locally bounded variation without singular terms (which is sometimes called 'special bounded variation'). This implies the following:

1. At each time-instant t we can define a left and right velocity field

$$
\begin{aligned}
\boldsymbol{v}^+(t) &= \lim_{\tau \downarrow 0} \frac{\boldsymbol{\xi}(t + \tau) - \boldsymbol{\xi}(t)}{\tau}, \\
\boldsymbol{v}^-(t) &= \lim_{\tau \uparrow 0} \frac{\boldsymbol{\xi}(t + \tau) - \boldsymbol{\xi}(t)}{\tau}.
\end{aligned}
\tag{92}
$$

2. The differential measure d$\boldsymbol{\xi}$ of the configuration $\boldsymbol{\xi}(t)$ contains only a density $\boldsymbol{v}$ with respect to the Lebesgue measure dt

$$
d\boldsymbol{\xi} = \boldsymbol{v}\, dt.
\tag{93}
$$

   For almost all t, we can define a velocity field $\boldsymbol{v}(t) = \dot{\boldsymbol{\xi}}(t)$.

3. The differential measure d$\boldsymbol{v}$ of the velocity field $\boldsymbol{v}(t)$ contains a density with respect to the Lebesgue measure dt and with respect to the atomic measure dη, i.e.

$$
d\boldsymbol{v} = \dot{\boldsymbol{v}}\, dt + (\boldsymbol{v}^+ - \boldsymbol{v}^-)\, d\eta.
\tag{94}
$$

Let as before d$\boldsymbol{R}(X, t)$ be the constraint force, associated with bilateral and unilateral geometric scleronomic constraints, and d$\boldsymbol{F}(X, t)$ be the external force on a mass element dm with position $\boldsymbol{\xi}(X, t)$. The non-impulsive dynamics of the system is therefore described by the equation of motion

$$
dm\, \ddot{\boldsymbol{\xi}} - d\boldsymbol{F} - d\boldsymbol{R} = \boldsymbol{0}.
\tag{95}
$$

The impulsive dynamics is described by the impact equation

$$
dm\, (\boldsymbol{v}^+ - \boldsymbol{v}^-) - d\boldsymbol{P} = \boldsymbol{0},
\tag{96}
$$

where d$\boldsymbol{P}$ is the impulsive force field of the constraints. The introduction of the differential measure d$\boldsymbol{v}$ allows us to combine the equation of motion (95) and the impact

equation (96) in a single equation

$$\mathrm{d}m\,\mathrm{d}\boldsymbol{v} - \mathrm{d}\boldsymbol{F}\,\mathrm{d}t - \mathrm{d}\boldsymbol{R}\,\mathrm{d}t - \mathrm{d}\boldsymbol{P}\,\mathrm{d}\eta = \boldsymbol{0}, \tag{97}$$

which is an equality of measures and which should be understood in the sense of integration. The measure $\mathrm{d}\boldsymbol{v}$, which has a density with respect to the atomic measure, is by definition a mapping on the space of continuous functions, i.e. $\int_I \boldsymbol{f}^{\mathrm{T}}\mathrm{d}\boldsymbol{v}$ only makes sense if $\boldsymbol{f} \in C^0(I, \mathbb{R}^3)$. The equality of measures leads to a principle of virtual work in differential measures

$$\int_S \delta\boldsymbol{\xi}^{\mathrm{T}}(\mathrm{d}m\,\mathrm{d}\boldsymbol{v} - \mathrm{d}\boldsymbol{F}\,\mathrm{d}t - \mathrm{d}\boldsymbol{R}\,\mathrm{d}t - \mathrm{d}\boldsymbol{P}\,\mathrm{d}\eta) = \boldsymbol{0} \tag{98}$$

for all $\delta\boldsymbol{\xi} \in \mathcal{Y}_0(I, \mathbb{R}^3)$. The class $\mathcal{Y}(I, \mathbb{R}^3)$, see (14), comprises all functions on the domain I which are absolutely continuous and piecewise C^1 as well as absolutely continuous functions with an accumulation point. We put $\hat{\boldsymbol{\xi}}(\varepsilon, t) = \boldsymbol{\xi}(t) + \varepsilon\boldsymbol{\omega}(t)$ with $\boldsymbol{\omega} \in \mathcal{Y}_0(I, \mathbb{R}^3)$. For this family of comparison functions, it holds that $\delta\boldsymbol{\xi} = \hat{\boldsymbol{\xi}}_\varepsilon'(0)\,\delta\varepsilon = \boldsymbol{\omega}\,\delta\varepsilon$ is continuous in time whereas $\delta\boldsymbol{v} = \dot{\boldsymbol{\omega}}\,\delta\varepsilon$ is not. Moreover, note that $\hat{\boldsymbol{\xi}}(\varepsilon)$ converges for $\varepsilon \downarrow 0$ to $\boldsymbol{\xi}$ in the weak norm $\|\cdot\|_1$; see (25). We now make the assumption that both the constraint force $\mathrm{d}\boldsymbol{R}$ and the constraint impulse $\mathrm{d}\boldsymbol{P}$ are perfect constraint forces/impulses, i.e.

$$-\mathrm{d}\boldsymbol{R} \in \mathcal{N}_{\mathcal{K}}(\boldsymbol{\xi}), \qquad -\mathrm{d}\boldsymbol{P} \in \mathcal{N}_{\mathcal{K}}(\boldsymbol{\xi}). \tag{99}$$

The virtual work of the constraint force $\mathrm{d}\boldsymbol{R}$ and the constraint impulse $\mathrm{d}\boldsymbol{P}$ is therefore non-negative for admissible virtual displacements

$$\delta\boldsymbol{\xi}^{\mathrm{T}}\mathrm{d}\boldsymbol{R} \geq 0, \qquad \delta\boldsymbol{\xi}^{\mathrm{T}}\mathrm{d}\boldsymbol{P} \geq 0 \quad \forall\delta\boldsymbol{\xi} \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{\xi}). \tag{100}$$

This leads us to the principle of d'Alembert–Lagrange in inequality form for differential measures

$$\int_S \delta\boldsymbol{\xi}^{\mathrm{T}}(\mathrm{d}m\,\mathrm{d}\boldsymbol{v} - \mathrm{d}\boldsymbol{F}\,\mathrm{d}t) \geq 0 \quad \forall\delta\boldsymbol{\xi} = \boldsymbol{\omega}\,\delta\varepsilon \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{\xi}) \tag{101}$$

in which we explicitly write that the virtual displacements $\delta\boldsymbol{\xi}$ are of the form $\boldsymbol{\omega}\,\delta\varepsilon$. A mass element has a constant mass $\mathrm{d}m$. Taking the differential measure-in-time of the term $\delta\boldsymbol{\xi}^{\mathrm{T}}\mathrm{d}m\,\boldsymbol{v}$ and applying the chain rule gives

$$\mathrm{d}(\delta\boldsymbol{\xi}^{\mathrm{T}}\mathrm{d}m\,\boldsymbol{v}) = \delta\boldsymbol{v}^{\mathrm{T}}\mathrm{d}m\,\boldsymbol{v}\,\mathrm{d}t + \delta\boldsymbol{\xi}^{\mathrm{T}}\mathrm{d}m\,\mathrm{d}\boldsymbol{v}, \tag{102}$$

in which we used the equality $\mathrm{d}(\delta\boldsymbol{\xi}) = \mathrm{d}\boldsymbol{\omega}\,\delta\varepsilon = \dot{\boldsymbol{\omega}}\,\mathrm{d}t\,\delta\varepsilon = \delta\boldsymbol{v}\,\mathrm{d}t$. We therefore arrive at the variational inequality

$$\mathrm{d}\left(\int_S \delta\boldsymbol{\xi}^{\mathrm{T}}\mathrm{d}m\,\boldsymbol{v}\right) - \int_S \delta\boldsymbol{v}^{\mathrm{T}}\mathrm{d}m\,\boldsymbol{v}\,\mathrm{d}t - \int_S \delta\boldsymbol{\xi}^{\mathrm{T}}\mathrm{d}\boldsymbol{F}\,\mathrm{d}t \geq 0 \tag{103}$$

$\forall\delta\boldsymbol{\xi} = \boldsymbol{\omega}\,\delta\varepsilon \in \mathcal{T}_{\mathcal{K}}(\boldsymbol{\xi})$. We recognise the second term $\delta T = \int_S \delta\boldsymbol{v}^{\mathrm{T}}\mathrm{d}m\,\boldsymbol{v}$ as being the variation of the kinetic energy $T = \int_S \frac{1}{2}\boldsymbol{v}^{\mathrm{T}}\mathrm{d}m\,\boldsymbol{v}$. In the same way as before, we choose

a chart $\boldsymbol{\varphi}$ with generalised coordinates $\boldsymbol{q}(t)$ which form a minimal set of coordinates with respect to the bilateral constraints and with which we can uniquely describe the configuration $\boldsymbol{\xi}(t) = \boldsymbol{\varphi}(\boldsymbol{q}(t))$. Moreover, we introduce generalised velocities $\boldsymbol{u}(t)$, which are assumed to be of locally bounded variation, and which are such that $d\boldsymbol{q} = \boldsymbol{u}\,dt$. The family of comparison functions $\hat{\boldsymbol{\xi}}(\varepsilon, t) = \boldsymbol{\xi}(t) + \varepsilon\boldsymbol{\omega}(t)$ is generated by the family of comparison functions $\hat{\boldsymbol{q}}(\varepsilon, t) = \boldsymbol{q}(t) + \varepsilon\boldsymbol{w}(t)$. Hence, the variation $\delta\boldsymbol{q} = \boldsymbol{w}\,\delta\varepsilon$ induces the variation $\delta\boldsymbol{\xi} = \partial\boldsymbol{\varphi}/\partial\boldsymbol{q}\,\delta\boldsymbol{q}$ with $\boldsymbol{\omega} = \partial\boldsymbol{\varphi}/\partial\boldsymbol{q}\,\boldsymbol{w}$. The kinetic energy T is a function of $\boldsymbol{v}(\boldsymbol{q}, \boldsymbol{u})$ and we can therefore write T as a function $T(\boldsymbol{q}, \boldsymbol{u})$. Hence, it holds that $\delta T = T_{\boldsymbol{q}}\,\delta\boldsymbol{q} + T_{\boldsymbol{u}}\,\delta\boldsymbol{u}$. The variations $\delta\boldsymbol{q}(t) = \boldsymbol{w}(t)\,\delta\varepsilon$ is a continuous function in time whereas $\delta\boldsymbol{u}(t) = \dot{\boldsymbol{w}}(t)\,\delta\varepsilon$ is discontinuous. Using the generalised momentum

$$\boldsymbol{p}(\boldsymbol{q}, \boldsymbol{u}) = \int_S \left(\frac{\partial\boldsymbol{\varphi}}{\partial\boldsymbol{q}}\right)^{\mathrm{T}} \boldsymbol{v}\,dm \tag{104}$$

and the generalised force (54) we transform the principle of d'Alembert–Lagrange into

$$d\left(\delta\boldsymbol{q}^{\mathrm{T}}\boldsymbol{p}\right) - \delta T\,dt - \delta\boldsymbol{q}^{\mathrm{T}}\boldsymbol{f}\,dt \geq 0 \quad \forall\delta\boldsymbol{q} = \boldsymbol{w}\,\delta\varepsilon \in \mathcal{T}_K(\boldsymbol{q}), \tag{105}$$

which is the Lagrange central equation in differential measures. As before, we integrate over a time-interval $I = [t_0, t_{\mathrm{f}}]$ and consider the generalised force $\boldsymbol{f} = -V'(\boldsymbol{q})$ to be a potential force, which yields

$$\left[\delta\boldsymbol{q}^{\mathrm{T}}\boldsymbol{p}\right]_{t\uparrow t_0}^{t\downarrow t_{\mathrm{f}}} - \int_I \delta L\,dt \geq 0 \quad \forall\delta\boldsymbol{q} = \boldsymbol{w}\,\delta\varepsilon \in \mathcal{T}_K(\boldsymbol{q}), \tag{106}$$

where $L = T - V$. Finally, taking fixed boundary conditions at t_0 and t_{f}, i.e. $\delta\boldsymbol{q}(t_0) = \delta\boldsymbol{q}(t_{\mathrm{f}}) = 0$, we obtain

$$-\delta\int_I L\,dt \geq 0 \quad \forall\delta\boldsymbol{q} = \boldsymbol{w}\,\delta\varepsilon \in \mathcal{T}_K(\boldsymbol{q}). \tag{107}$$

The comparison functions $\hat{\boldsymbol{q}}(\varepsilon, t) = \boldsymbol{q}(t) + \varepsilon\boldsymbol{w}(t)$ converge in the weak norm $\|\cdot\|_1$ and lead to weak variations $\delta\boldsymbol{q}$. The variation δs of the action $s(\boldsymbol{q}) = \int_I L\,dt$ in (107) reduces therefore to the Gâteaux derivative $ds(\boldsymbol{q}; \delta\boldsymbol{q})$; see (22). Consequently, we arrive at a weak form of the principle of Hamilton in inequality form, i.e. the principle of Hamilton as weak local extremal of the action $s(\boldsymbol{q})$ with the weak norm $\|\cdot\|_1$.

Theorem 4 (The weak principle of Hamilton in inequality form) *Consider a conservative Lagrangian mechanical system with perfect unilateral constraints. A function $\boldsymbol{q}(t) \in K$ is a motion of the system if and only if it is a weak local extremal of the action integral $s(\boldsymbol{q}) = \int_I L\,dt$, i.e.*

$$-ds(\boldsymbol{q}; \delta\boldsymbol{q}) \geq 0 \quad \forall\delta\boldsymbol{q} = \boldsymbol{w}\,\delta\varepsilon \in \mathcal{T}_K(\boldsymbol{q}), \qquad \boldsymbol{q}(t_0) = \boldsymbol{q}_0, \qquad \boldsymbol{q}(t_{\mathrm{f}}) = \boldsymbol{q}_{\mathrm{f}}, \tag{108}$$

with $\boldsymbol{w} \in \mathcal{Y}_0(I, \mathbb{R}^n)$, which is the weak principle of Hamilton in inequality form for impulsive motion.

Proof If $\boldsymbol{q}(t)$ is a motion of the system, or equivalently $\boldsymbol{\xi}(t) = \boldsymbol{\varphi}(\boldsymbol{q}(t))$, then it has to fulfill the equations of motion (95) for almost all t and the impact equation (96) at the impact time-instants. This leads to the principle of virtual work in differential measures (98). The supposition of perfect unilateral constraint forces and impulses leads to the principle of Hamilton (107) with weak variations $\delta\boldsymbol{q}$. The proof can easily be followed in the opposite direction.   □

There are two interesting things to remark at this point. First of all, by making use of differential measures, we are able to treat the impulsive and non-impulsive dynamics simultaneously. This means that we do not need to split the time-interval in differentiable parts to treat an impact as has been done in Sect. 6. Secondly, as we want to use the principle of virtual work for differential measures (98), we are forced to consider comparison functions $\hat{\boldsymbol{\xi}}(\varepsilon, t) = \boldsymbol{\xi}(t) + \varepsilon\boldsymbol{\omega}(t)$ of which the variation $\delta\boldsymbol{\xi} = \boldsymbol{\omega}\,\delta\varepsilon$ is time-continuous. A comparison function of this form can have kinks (and, therefore, impacts), but this class of comparison functions does not include a family of curves which only varies the impact time. Accordingly, for this weak form of the principle of Hamilton, we only have to satisfy the 'generalised' first Weierstrass–Erdmann corner condition (83). Hence, we are able to prove the validity of Theorem 4 by only assuming that the unilateral geometric constraint is perfect which is expressed by (99). The supposition of energy conservation during the impact, i.e. the second Weierstrass–Erdmann corner condition $H_{\mathrm{c}}^{+} = H_{\mathrm{c}}^{-}$, is therefore not a necessary condition for the validity of Theorem 4.

8 Example: Numerical Discretisation

The differences between the weak and strong principle of Hamilton can be illustrated by looking at their discretized forms, i.e. the action integral is discretized, similar to what is done in the Ritz method. As we will show, both principles lead to a numerical scheme for impulsive dynamics, but an impact law does (generally) not follow from the discretized principles. As we only want to convey the basic idea, the simplest example system is chosen: that of a one-dimensional free particle.

Consider a point-mass m with coordinate $q(t)$ which is subjected to the unilateral constraint $q(t) \geq 0$. In the absence of gravitation or other external forces, the action integral reads as

$$s(q) = \int_{t_0}^{t_{\mathrm{f}}} T(u)\,\mathrm{d}t, \quad T(u) = \frac{1}{2}mu^2, \ u = \dot{q} \text{ a.e.} \tag{109}$$

with given positions $q(t_0) = q_0$ and $q(t_{\mathrm{f}}) = q_{\mathrm{f}}$ at begin and end time. The evolution $q(t)$ is approximated with approximants $q_i \approx q(t_i)$ on a discretisation $t_0 < t_1 < t_2 < \cdots < t_N < t_{N+1} = t_{\mathrm{f}}$ of the time interval $[t_0, t_{\mathrm{f}}]$. We can either choose this time-grid a priori, i.e. we choose a fixed grid, or the choice of t_i for $i = 1, \ldots, N$ has still to be determined. Correspondingly, we define approximants $u_i \approx u(t_i)$ and $T_i \approx T(u(t_i))$ for the velocity and kinetic energy at the discretisation point t_i:

$$u_i = \frac{q_{i+1} - q_i}{t_{i+1} - t_i}, \qquad T_i = \frac{1}{2}mu_i^2 = \frac{1}{2}m\left(\frac{q_{i+1} - q_i}{t_{i+1} - t_i}\right)^2. \tag{110}$$

Using the Euler quadrature $\int_0^l f(x)\,\mathrm{d}x \approx \sum_{i=0}^{N} f(x_i)(x_{i+1} - x_i)$, the action integral $s(q)$ can be approximated by the sum

$$s^N(q_1, q_2, \ldots, q_N) = \sum_{i=0}^{N} T_i(t_{i+1} - t_i). \tag{111}$$

We will now consider a discretisation of the weak principle of Hamilton and the strong principle of Hamilton for this example system.

8.1 Weak Principle of Hamilton

If a fixed grid of time discretisation points t_i is chosen, then we are only able to consider variations δq_i of the approximants q_i, which can be looked upon as approximants of the variation at time t_i, i.e. $\delta q_i \approx \delta q(t_i)$, for fixed values of t_i. As time is not varied, these (approximate) variations are *weak* and we are able to discretise the weak principle of Hamilton in inequality form (108). The Gâteaux derivative of the action integral appearing in (108) can be approximated by

$$\mathrm{d}s(q, \delta q) \approx \mathrm{d}s^N(q_1, \ldots, q_N; \delta q_1, \ldots, \delta q_N) = \sum_{i=1}^{N} \frac{\partial s^N}{\partial q_i} \delta q_i, \tag{112}$$

with

$$\frac{\partial s^N}{\partial q_i} = \frac{\partial T_{i-1}}{\partial q_i}(t_i - t_{i-1}) + \frac{\partial T_i}{\partial q_i}(t_{i+1} - t_i) = -m(u_i - u_{i-1}). \tag{113}$$

The discretized weak principle of Hamilton therefore leads to the difference scheme in the form of a finite dimensional variational inequality

$$m(u_i - u_{i-1})\delta q_i \geq 0 \quad \forall \delta q_i \in \mathcal{T}_{\mathbb{R}_0^+}(q_i), \tag{114}$$

which can be cast in the inclusion

$$m(u_i - u_{i-1}) - \lambda_i = 0, \quad -\lambda_i \in \mathcal{N}_{\mathbb{R}_0^+}(q_i), \tag{115}$$

where λ_i is the approximant for the percussion of the constraint during the time interval $(t_{i-1}, t_i]$. Note that the inclusion in (115) merely requires the inequality complementarity between the discrete percussion λ_i and the position q_i and does not include an impact law. The difference scheme (115) together with the boundary conditions q_0 and $q_{N+1} = q_\mathrm{f}$ has therefore N solutions, each with a single impact at some time discretisation point t_j ($j = 1, \ldots, N$),

$$q_i = \begin{cases} \frac{t_j - t_i}{t_j - t_0} q_0, & i \leq j, \\ \frac{t_i - t_j}{t_\mathrm{f} - t_j} q_\mathrm{f}, & i > j, \end{cases} \tag{116}$$

and one solution at constant velocity without impact $q_i = q_0 + \frac{t_i - t_0}{t_\mathrm{f} - t_0}(q_\mathrm{f} - q_0)$; see Fig. 7 in which all solutions are drawn. Each of these solutions are approximants of weak local extremals of the action integral.

Fig. 7 Weak and strong
solutions (*in bold*) of the
discretized principle of
Hamilton
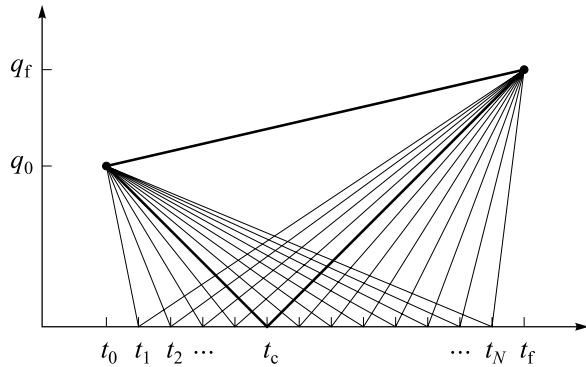
8.2 Strong Principle of Hamilton

If we consider a variable grid of time discretisation points t_i, then variations in time
δt_i induce the variations $\overline{\delta q}(t_i) \approx u_i \delta t_i + \delta q_i$ which lead to strong variations. The
strong principle of Hamilton (88) requires the sub-stationarity $-\delta s \geq 0$ of the action
integral $s(q)$ with respect all strong variations $\delta q \in \mathcal{T}_{\mathbb{R}_0^+}(q)$. The discretized form of
the strong principle of Hamilton therefore reads as

$$-\delta s^N = \sum_{i=1}^{N} \frac{\partial s^N}{\partial q_i} \delta q_i + \frac{\partial s^N}{\partial t_i} \delta t_i \geq 0 \quad \forall \delta q_i \in \mathcal{T}_{\mathbb{R}_0^+}(q_i), \ \forall \delta t_i, \tag{117}$$

with independent variations δq_i and δt_i. Hence, the partial derivatives

$$\frac{\partial s^N}{\partial t_i} = T_i - T_{i-1} \tag{118}$$

have to vanish for $i = 1, \ldots, N$, which implies the conservation of energy during
each discrete time-interval. The discretized strong principle of Hamilton in equality
form therefore leads to the difference scheme

$$m(u_i - u_{i-1}) - \lambda_i = 0, \quad -\lambda_i \in \mathcal{N}_{\mathbb{R}_0^+}(q_i),$$
$$\frac{1}{2}mu_i^2 = \frac{1}{2}mu_{i-1}^2. \tag{119}$$

The numerical scheme (119) with boundary conditions q_0 and $q_{N+1} = q_f$ has an im-
pacting solution (116) with $t_j = t_c := (q_0 t_f + q_f t_0)/(q_f + q_0)$, whereas t_i is free to
choose for $i \neq j$, as well as the non-impacting solution $q_i = q_0 + \frac{t_i - t_0}{t_f - t_0}(q_f - q_0)$
with t_i free to choose (see Fig. 7). Note that the discretized strong principle of
Hamilton (119) requires the perfectness of the constraint and the conservation of
energy, which implies in the one-dimensional case the completely elastic impact law
$u_j = -u_{j-1}$.

From the above example, it becomes clear that the weak principle of Hamilton can
be used to derive numerical integration schemes for impulsive mechanical systems on

a fixed grid. The impact law has still to be incorporated in the numerical scheme, for instance by relaxing the unilateral constraint $-\lambda_i \in \mathcal{N}_{\mathbb{R}_0^+}(q_i)$ on position level to the unilateral constraint

$$-\lambda_i \in \mathcal{N}_{\mathcal{T}_{\mathbb{R}_0^+}(q_i)}\left(\frac{1}{1+\varepsilon}u_i + \frac{\varepsilon}{1+\varepsilon}u_{i-1}\right), \qquad (120)$$

on velocity level, which is a Newton-type of impact law with a coefficient of restitution ε (Moreau 1988). More elegantly, one might try to incorporate an impact law in the weak form of Hamilton's principle by requiring the sub-stationarity of the (discretized) action integral with respect to an appropriate larger set of comparison functions, but this is still a topic of future research.

9 Conclusions

The aim of this paper is to clarify the existing forms of the principle of Hamilton as variational inequality, which have already been derived in Panagiotopoulos and Glocker (1998, 2000), by putting them within the context of weak and strong extrema.

In this paper, we have derived two different forms of Hamilton's principle as variational inequality. The strong form of Hamilton's principle has been derived in Sect. 6 using the general variation of the functional which leads to 'generalised' Weierstrass–Erdmann corner conditions. The first 'generalised' Weierstrass–Erdmann condition demands that the contact impulses are from the normal cone, i.e. the supposition of perfect unilateral constraints. The second Weierstrass–Erdmann condition requires that the collisions are completely elastic, i.e. there is no energy loss during the impact. This form of the principle of Hamilton takes all neighbouring functions into consideration for the sub-stationarity of the solution and is therefore a strong form of Hamilton's principle as variational inequality. A weak form of Hamilton's principle has been derived in Sect. 7 starting from the principle of virtual work in differential measures. The latter can only be posed for a special class of virtual displacements, namely those which are related to Gâteaux derivatives. This derivation therefore leads in a natural way to a weak form of Hamilton's principle. The weak form of Hamilton's principle only requires that the contact impulses are from the normal cone, which is a generalisation of the first Weierstrass–Erdmann condition. No restriction is posed in the weak form on the energy dissipation of the impact law.

The various forms of the principle of Hamilton as variational inequality which exist in literature can now be put within the context of weak and strong extrema. We conclude that Theorem 3 of Panagiotopoulos and Glocker (1998) is the weak form of Hamilton's principle, while Theorem 4 of Panagiotopoulos and Glocker (1998) (or Proposition 4 of Panagiotopoulos and Glocker 2000) is the strong form of Hamilton's principle. This insight clarifies why the various principles have different conditions on the impact law. The forms of the principle of Hamilton used by Tornambè (1999) and Percivale (1985, 1991) assume strong variations and the authors of these works therefore state that the principles are valid for completely elastic impact.

In Sect. 8, it has very briefly been shown on an example system that a direct discretisation of the action integral leads to a difference scheme, which can form the

basis of a numerical integration scheme. Future research will focus on the natural derivation of dedicated numerical schemes for the simulation of mechanical systems with frictionless unilateral contact by using a direct discretisation of the variational principles derived in this paper.

What is the practical/theoretical relevance of the results and insight gained in this paper? An obvious application has already been mentioned: the development of numerical schemes through a Ritz-type of method on the principle of Hamilton in inequality form. However, in the opinion of the authors, the relevance of the paper is more fundamental. One way to think about dynamics is in terms of variational principles. History proved that this way of thinking has been very rewarding. Variational principles form the foundation of classical analytical mechanics and have been essential for the development of modern physics, e.g. quantum mechanics. Furthermore, variational principles give the link to optimisation theory and put dynamics in an appropriate mathematical framework. Classical variational principles, however, are strictly valid for perfect bilateral constraints. For this reason, if we endeavour to develop a proper theoretical foundation for non-smooth dynamics, it is a promising step to go back to these principles and reformulate them in terms of variational inequalities.

Appendix: Propositions

Proposition 1 *Let $f : I \to \mathbb{R}$ be a piece-wise continuous function and $\{g(\varepsilon)\}_{\varepsilon \geq 0}$ be a family of piece-wise continuous functions $g(\varepsilon) : I \to \mathbb{R}$ with $g(0) = f$ almost everywhere. Denote the discontinuity points of f with the set $S_f \subset I$ and the discontinuity points of $g(\varepsilon)$ with the set $S_{g(\varepsilon)} \subset I$. If it holds that*

$$\operatorname{ess\,sup}_{x \in I} \left| f(x) - g(\varepsilon, x) \right| \xrightarrow{\varepsilon \downarrow 0} 0, \tag{121}$$

then there exists a number $\varepsilon^* > 0$ such that

$$\forall \varepsilon < \varepsilon^* \quad S_f \subseteq S_{g(\varepsilon)} \subset I. \tag{122}$$

Proof We will prove the proposition with a reductio ad absurdum. Assume therefore that (121) holds and that there does *not* exist a number $\varepsilon^* > 0$ such that (122) holds. In other words, for all $\varepsilon^* > 0$, there exists $\varepsilon < \varepsilon^*$ such that x_s is a discontinuity point of f which is not a discontinuity point of $g(\varepsilon)$, i.e. $\forall \varepsilon^* > 0 \, \exists \varepsilon < \varepsilon^*$: $x_s \in S_f$ and $x_s \notin S_{g(\varepsilon)}$. Hence, the function $g(\varepsilon)$ is continuous on the open interval $(x_s - \delta, x_s + \delta)$, where $\delta(\varepsilon)$ is a function of ε. We will show that this assumption leads to a contradiction. According to (121), there exists $\varepsilon > 0$ for each essential upper bound $\alpha > 0$ such that

$$\begin{aligned}
\left| f(x_s + \delta) - g(\varepsilon, x_s + \delta) \right| &< \alpha, \\
\left| f(x_s - \delta) - g(\varepsilon, x_s - \delta) \right| &< \alpha,
\end{aligned} \tag{123}$$

with sufficiently small $\delta(\varepsilon)$. Because x_s is a discontinuity point of f it holds that

$$\left| f(x_s + \delta) - f(x_s - \delta) \right| \xrightarrow{\delta \to 0} s > 0, \tag{124}$$

where s is the jump height of the discontinuity of f at x_s.

Take ε^* to be an arbitrary strictly positive value. Let ε_m and $\delta_m = \delta(\varepsilon_m)$ be such that $g(\varepsilon_m)$ is continuous on the interval $(x_s - \delta_m, x_s + \delta_m)$ and such that it holds that

$$\begin{aligned} \left| f(x_s + \delta_m) - g(\varepsilon_m, x_s + \delta_m) \right| &< \varepsilon^*, \\ \left| f(x_s - \delta_m) - g(\varepsilon_m, x_s - \delta_m) \right| &< \varepsilon^*. \end{aligned} \tag{125}$$

Using a telescoping sum and the triangle inequality, we obtain the inequality

$$\begin{aligned} \left| f(x_s + \delta) - f(x_s - \delta) \right| &\leq \left| f(x_s + \delta) - g(\varepsilon, x_s + \delta) \right| \\ &\quad + \left| g(\varepsilon, x_s + \delta) - g(\varepsilon, x_s - \delta) \right| \\ &\quad + \left| g(\varepsilon, x_s - \delta) - f(x_s - \delta) \right| \\ &< 2\varepsilon^* + \left| g(\varepsilon, x_s + \delta) - g(\varepsilon, x_s - \delta) \right| \end{aligned} \tag{126}$$

for all $\delta < \delta_m$. Moreover, the local continuity of $g(\varepsilon)$ requires that the last term vanishes for $\delta \to 0$, whereas the left-hand side converges to the jump height s. Hence, we arrive at the inequality

$$s \leq 2\varepsilon^*. \tag{127}$$

Because $\varepsilon^* > 0$ is arbitrary, it follows that the jump height s of f is zero, which contradicts the assumption. This proves the proposition. $\qquad\square$

References

Aubin, J.-P., Frankowska, H.: Set-valued Analysis. Systems and Control: Foundations and Applications, vol. 2. Birkhäuser, Boston (1990)

Bremer, H.: Dynamik und Regelung mechanischer Systeme. Teubner, Stuttgart (1988)

Brogliato, B.: Nonsmooth Mechanics. Models, Dynamics and Control, 2 edn. Communications and Control Engineering. Springer, London (1999)

Cesar, M.O.: Necessary conditions and sufficient conditions of weak minimum for solutions with corner points. Bol. Soc. Brasil. Mat. **15**(1–2), 109–135 (1984)

Cesari, L.: Optimization—Theory and Applications. Applications of Mathematics, vol. 17. Springer, New York (1983)

Cirak, F., West, M.: Decomposition contact response (DCR) for explicit finite element dynamics. Int. J. Numer. Methods Eng. **64**, 1078–1110 (2005)

Fetecau, R.C., Marsden, J.E., Ortiz, M., West, M.: Nonsmooth Lagrangian mechanics and variational collision integrators. SIAM J. Appl. Dyn. Syst. **2**(3), 381–416 (2003)

Glocker, Ch.: Set-Valued Force Laws, Dynamics of Non-Smooth Systems. Lecture Notes in Applied Mechanics, vol. 1. Springer, Berlin (2001)

Goeleven, D., Panagiotopoulos, P.D., Lebau, C., Plotnikova, G.: Inequality forms of d'Alembert's principle in mechanics of systems with holonomic unilateral constraints. Z. Angew. Math. Mech. **77**(7), 483–501 (1997)

Goeleven, D., Miettinen, M., Panagiotopoulos, P.D.: Dynamic hemivariational inequalities and their applications. J. Optim. Theory Appl. **103**(3), 567–601 (1999)

Goeleven, D., Motreanu, D., Dumont, Y., Rochdi, M.: Variational and Hemivariational Inequalities: Theory, Methods and Applications, Volume I: Unilateral Analysis and Unilateral Mechanics. Nonconvex Optimization and its Applications, vol. 69. Kluwer Academic, Dordrecht (2003a)

Goeleven, D., Motreanu, D., Dumont, Y., Rochdi, M.: Variational and Hemivariational Inequalities: Theory, Methods and Applications, Volume II: Unilateral Problems. Nonconvex Optimization and its Applications, vol. 70. Kluwer Academic, Dordrecht (2003b)

Hamel, G.: Elementare Mechanik. B.G. Teubner, Stuttgart (1912)

Hartman, P., Stampacchia, G.: On some nonlinear elliptic differential equations. Acta Math. **115**(1), 153–188 (1966)

Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and their Applications. Pure and Applied Mathematics. Academic Press, New York (1980)

Kozlov, V.V., Treshchëv, D.V.: Billiards, a Genetic Introduction to the Dynamics of Systems with Impacts. Translations of Mathematical Monographs, vol. 89. American Mathematical Society, Providence (1991)

Kravchuk, A.S., Neittaanmäki, P.J.: Variational and Quasi-variational Inequalities in Mechanics. Solid Mechanics and its Applications, vol. 147. Springer, Dordrecht (2007)

Lanczos, C.: The Variational Principles of Mechanics. University of Toronto Press, Toronto (1962)

Leine, R.I., Nijmeijer, H.: Dynamics and Bifurcations of Non-Smooth Mechanical Systems. Lecture Notes in Applied and Computational Mechanics, vol. 18. Springer, Berlin (2004)

Leine, R.I., van de Wouw, N.: Stability and Convergence of Mechanical Systems with Unilateral Constraints. Lecture Notes in Applied and Computational Mechanics, vol. 36. Springer, Berlin (2008)

May, H.O.: Generalized variational principles and nondifferentiable potentials in analytical mechanics. J. Math. Phys. **25**(3), 491–493 (1984a)

May, H.O.: Hamilton's principle as substationarity principle. Acta Mech. **52**(3–4), 177–187 (1984b)

Moreau, J.J.: La notion de superpotentiel et les liaisons unilatérales en élastostatique. C. R. Acad. Sci., Sér. A **167** (1968), 954–957

Moreau, J.J.: Unilateral contact and dry friction in finite freedom dynamics. In: Moreau, J.J., Panagiotopoulos, P.D. (eds.) Non-Smooth Mechanics and Applications. CISM Courses and Lectures, vol. 302, pp. 1–82. Springer, Wien (1988)

Panagiotopoulos, P.D.: Non-convex superpotentials in the sense of F.H. Clarke and applications. Mech. Res. Commun. **8**(6), 335–340 (1981)

Panagiotopoulos, P.D.: Hemivariational Inequalities: Applications in Mechanics and Engineering. Springer, Berlin (1993)

Panagiotopoulos, P.D., Glocker, Ch.: Analytical mechanics: Addendum I. Inequality constraints with elastic impacts. The convex case. Z. Angew. Math. Mech. **78**(4), 219–229 (1998)

Panagiotopoulos, P.D., Glocker, Ch.: Inequality constraints with elastic impacts in deformable bodies. The convex case. Arch. Appl. Mech. **70**, 349–365 (2000)

Papastavridis, J.G.: Analytical Mechanics: A Comprehensive Treatise on the Dynamics of Constrained Systems: For Engineers, Physicists and Mathematicians. Oxford University Press, New York (2002)

Percivale, D.: Uniqueness in the elastic bounce problem I. J. Differ. Equ. **56**, 206–215 (1985)

Percivale, D.: Uniqueness in the elastic bounce problem II. J. Differ. Equ. **90**, 304–315 (1991)

Rockafellar, R.T., Wets, R.-B.: Variational Analysis. Springer, Berlin (1998)

Salomon, R.: Practical use of Hamilton's principle. J. Fluid Mech. **132**, 431–444 (1983)

Tornambè, A.: Modeling and control of impact in mechanical systems: Theory and experimental results. IEEE Trans. Autom. Control **44**(2), 294–309 (1999)

Troutman, J.L.: Variational Calculus and Optimal Control: Optimization with Elementary Convexity. Springer, New York (1996)