An Argument Against Augmenting the Lagrangean for Nonholonomic Systems

Carlos M. Roithmayr
NASA Langley Research Center, Hampton, Virginia 23681
757 864 6778, c.m.roithmayr@larc.nasa.gov.

Dewey H. Hodges
Professor, School of Aerospace Engineering
Georgia Institute of Technology, Atlanta, Georgia 30332
404 894 8201, dhodges@gatech.edu.

Abstract

Although it is known that correct dynamical equations of motion for a nonholonomic system cannot be obtained from a Lagrangean that has been augmented with a sum of the nonholonomic constraint equations weighted with multipliers, previous publications suggest otherwise. An example has been proposed in support of augmentation and purportedly demonstrates that an accepted method fails to produce correct equations of motion whereas augmentation leads to correct equations; this paper shows that in fact the opposite is true. The correct equations, previously discounted on the basis of a flawed application of the Newton-Euler method, are verified by using Kane's method and a new approach to determining the directions of constraint forces. A correct application of the Newton-Euler method reproduces valid equations.

Introduction

Dealing with nonholonomic constraint equations within the framework of variational methods is a controversial subject. For example, Ray [1] modifies Hamilton's principle and augments the Lagrangean by adjoining a sum of nonholonomic constraint equations weighted with multipliers. Later, Ray [2] reverses himself. In the erratum he compares the correct way of dealing with constraint equations that are linear in the time derivatives of the generalized coordinates

to the incorrect approach of augmenting the Lagrangean that gives the wrong results, even when the constraint equations are linear. Saletan and Cromer [3] follow Ray and show the augmented Lagrangean gives correct equations of motion when the constraint equations are holonomic. They conclude that no such augmented Lagrangean exists in the nonholonomic case, in part because they say that there is no way to determine initial conditions needed for the integration of differential equations governing the multipliers. Rosenberg [4] presents the same demonstration as Ray's erratum and concludes that, although Hamilton's principle may be regarded as a variational principle for conservative holonomic systems, it cannot be so regarded for nonholonomic systems. In an effort to eliminate constraint violations, Rosen and Edelstein [5] make a proposal similar to that of Ray 30 years earlier; they account for nonlinear nonholonomic constraint equations in the same way that they do holonomic constraint equations. Hagedorn [6] points out that although their approach is justified in the holonomic case, it is incorrect for nonholonomic constraint equations, even when they are linear. He demonstrates this with an example and gives the well-known result for the correct way to handle linear equations, which does not come from modifying the Lagrangean. According to Hagedorn the mistake has been repeated many times over the past century and the pitfall has received attention in Refs. [7], [8], and [9]. More recently, Flannery [10] examines the problems encountered by Ray and others and, after in-depth analysis, concludes "General [nonlinear] nonholonomic constraints are completely outside the scope of even the most fundamental principle of D'Alembert. The generalization of any principle based on [D'Alembert's] to general nonholonomic constraints is without foundation." In an earlier paper, Polyakhov [11] reaches the same conclusion: "Thus, the D'Alembert-Lagrange principle as well as the Lagrange ideality condition can be derived from Newton's equations only for holonomic constraints. Therefore, the application of these results to nonholonomic systems is not justified."

In their response to his comments, Rosen and Edelstein offer a counterex-

ample purportedly showing that the approach advocated by Hagedorn leads to incorrect results. Their conclusion is based on a flawed application of the Newton-Euler method. The purpose of this paper is to demonstrate the validity of the approach taken by the latter author and bring to light mistakes made by the former authors in their development of the counterexample.

Problems With the Counterexample

The planar system featured in the counterexample is shown in Fig. 1. Two perpendicular unit vectors $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are fixed in an inertial reference frame N. A smooth rod B whose axis is parallel to unit vector $\hat{\mathbf{b}}_1$ is inclined at a constant angle α to $\hat{\mathbf{n}}_1$; B is permitted to translate along an axis parallel to $\hat{\mathbf{n}}_2$. A particle P of mass m moves along B, and the mass of B is negligible in comparison to m. It is said that no forces are exerted on P other than those necessary to prevent it from loosing contact with B.

Analysis is facilitated by working with two generalized coordinates q_1 and q_2 shown in Fig. 1, where q_1 is the displacement in a prismatic joint connecting B to N, and where q_2 is the displacement of P along the rod. Two generalized speeds [12] are introduced simply as $u_r = \dot{q}_r$ $(r = 1, 2)$. A motion constraint is to be imposed upon the velocity of P in B, expressed by the relationship

$$\epsilon\, C_\alpha u_2 - q_1 = 0 \tag{1}$$

where ϵ is a positive constant and C_α denotes $\cos\alpha$. The equivalence of this expression and the constraint equation in Ref. [6] is demonstrated presently. Now, the velocity $^N\mathbf{v}^{\overline{B}}$ in N of every point $\overline{B}$ fixed in B is given by $^N\mathbf{v}^{\overline{B}} = u_1\hat{\mathbf{n}}_2$, and the velocity $^B\mathbf{v}^P$ of P in B is given by $^B\mathbf{v}^P = u_2\hat{\mathbf{b}}_1$. Henceforth, $\overline{B}$ is taken to be the point of B that is coincident with P, and the velocity of P in N is simply $^N\mathbf{v}^P = {}^N\mathbf{v}^{\overline{B}} + {}^B\mathbf{v}^P$. The nonholonomic constraint equation (1)
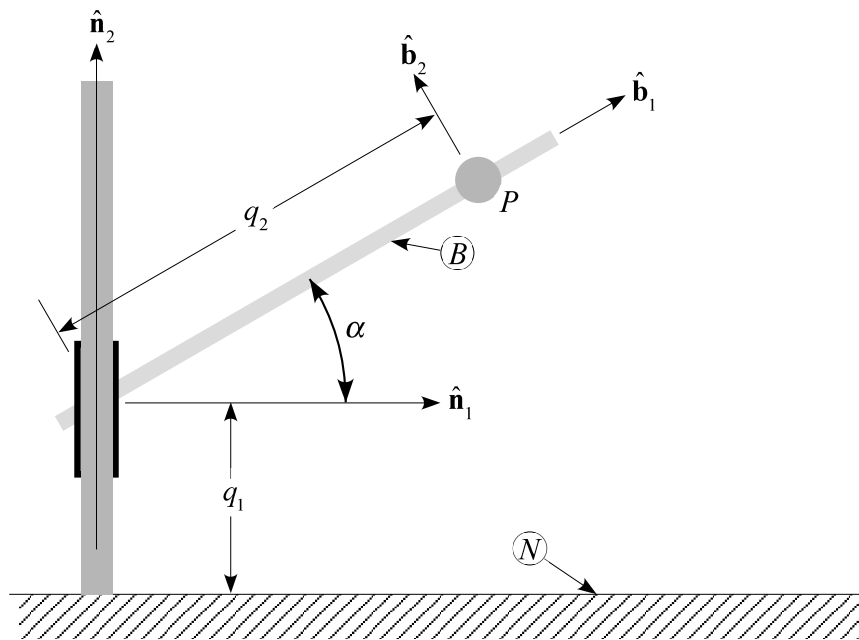
Figure 1: A Particle Moving on a Sliding Inclined Rod

can thus be written in vector form as

$$({}^{N}\mathbf{v}^{P} - {}^{N}\mathbf{v}^{\overline{B}}) \cdot \hat{\mathbf{b}}_1 - \frac{q_1}{\epsilon\, C_\alpha} = 0 \tag{2}$$

There also exists a configuration constraint that prevents P from moving in B in the direction of $\hat{\mathbf{b}}_2$; at the velocity level, the holonomic constraint equation is expressed as

$$({}^{N}\mathbf{v}^{P} - {}^{N}\mathbf{v}^{\overline{B}}) \cdot \hat{\mathbf{b}}_2 = 0 \tag{3}$$

On the basis of physical reasoning it is clear that a force normal to the rod, in other words a force of unknown magnitude and parallel to $\hat{\mathbf{b}}_2$, must be applied to P in order to satisfy Eq. (3) and keep P from losing contact with the rod. In view of the law of action and reaction, a force of equal magnitude and opposite direction is thus applied to B at $\overline{B}$. These conclusions can be reached also by a close examination of the form of Eq. (3), which is now undertaken so that the results may be used with Eq. (2) and another constraint equation to follow.

The vector form of Eq. (3) can be used to determine directions and points of application of forces necessary to impose the constraint. Upon differentiation with respect to time in N, one obtains

$$({}^{N}\mathbf{a}^{P} - {}^{N}\mathbf{a}^{\overline{B}}) \cdot \hat{\mathbf{b}}_2 = 0 \tag{4}$$

where ${}^{N}\mathbf{a}^{P}$ and ${}^{N}\mathbf{a}^{\overline{B}}$ are, respectively, the accelerations of P and $\overline{B}$ in N. (The time derivative of $\hat{\mathbf{b}}_2$ in N vanishes because B has no angular velocity in N.) Each acceleration is related by Newton's second law to the resultant of all applied contact and distance forces. Let $\mathbf{f}_P$ and $\mathbf{f}_{\overline{B}}$ denote the respective resultants of all contact and distance forces applied to P and $\overline{B}$ when their motion is not restricted by the constraint expressed in Eq. (4). In addition, let $\mathbf{C}_P$ and $\mathbf{C}_{\overline{B}}$ represent constraint forces that must be applied to P and $\overline{B}$ respectively in order to impose the constraint. By virtue of Newton's second law, Eq. (4) can be rewritten as

$$\frac{\mathbf{f}_P + \mathbf{C}_P}{m} \cdot \hat{\mathbf{b}}_2 + \frac{\mathbf{f}_{\overline{B}} + \mathbf{C}_{\overline{B}}}{m_{\overline{B}}} \cdot (-\hat{\mathbf{b}}_2) = 0 \tag{5}$$

where $m_{\bar{B}}$ is the mass of the particle at $\overline{B}$. Any component of $\mathbf{C}_P$ that is perpendicular to $\hat{\mathbf{b}}_2$ will not play a part in Eq. (5); therefore, all that is necessary is for the constraint force $\mathbf{C}_P$ to be parallel to $\hat{\mathbf{b}}_2$. Likewise, $\mathbf{C}_{\bar{B}}$ need only be parallel to $-\hat{\mathbf{b}}_2$. The first condition is expressed with the relationship

$$\mathbf{C}_2 = \mu_2 \hat{\mathbf{b}}_2 \tag{6}$$

where $\mathbf{C}_2$ is the constraint force applied to P and associated with the second of two constraint equations [Eq. (3)], and where μ_2 is an as yet unknown scalar multiplier. Furthermore, the constraint force applied to $\overline{B}$ is $\mu_2(-\hat{\mathbf{b}}_2) = -\mathbf{C}_2$. One may simply inspect Eq. (3) to make such determinations; because $^N\mathbf{v}^P$ appears in a dot product with $\hat{\mathbf{b}}_2$, it can be said that the constraint force parallel to $\hat{\mathbf{b}}_2$, $\mathbf{C}_2 = \mu_2 \hat{\mathbf{b}}_2$, must be applied to P. Likewise, because $^N\mathbf{v}^{\overline{B}}$ appears in a dot product with $-\hat{\mathbf{b}}_2$, it can be said that the constraint force $-\mathbf{C}_2 = -\mu_2 \hat{\mathbf{b}}_2$ must be applied to B at $\overline{B}$. As has been said, these results are in line with physical reasoning. In Ref. [6] the magnitude of the reaction force $\mathbf{C}_2$ normal to $\hat{\mathbf{b}}_1$ is denoted by N rather than μ_2.

Using the same reasoning, one may inspect Eq. (2) and conclude that P must be subject to a constraint force $\mathbf{C}_1$ that is parallel to $\hat{\mathbf{b}}_1$,

$$\mathbf{C}_1 = \mu_1 \hat{\mathbf{b}}_1 \tag{7}$$

whereas a force $-\mathbf{C}_1$ is applied to B at $\overline{B}$. In this case the subscript of $\mathbf{C}_1$ and μ_1 indicates they are associated with the first of two constraint equations. In practice the set of forces $\mathbf{C}_1$ and $-\mathbf{C}_1$ could be applied with a motorized gear attached to P moving on a track of gear teeth fixed in B; evidently the rod cannot be perfectly smooth as hypothesized in the problem statement, if the nonholonomic constraint equation (1) is to be satisfied.

After forming the acceleration of P in N as $^N\mathbf{a}^P = \dot{u}_1 \hat{\mathbf{n}}_2 + \dot{u}_2 \hat{\mathbf{b}}_1$, one is in a position to use Kane's method [12] and form two equations of motion for the system S composed of P and B, $F_r + F_r^\star = 0$ $(r = 1, 2)$. The holonomic

generalized active forces are given by

$$F_r = {}^N\mathbf{v}_r^P \cdot (\mathbf{C}_1 + \mathbf{C}_2) + {}^N\mathbf{v}_r^{\overline{B}} \cdot (-\mathbf{C}_1 - \mathbf{C}_2) \qquad (r = 1, 2) \tag{8}$$

and the holonomic generalized inertia forces $F_r^\star$ are constructed according to

$$F_r^\star = - {}^N\mathbf{v}_r^P \cdot m {}^N\mathbf{a}^P \qquad (r = 1, 2) \tag{9}$$

The required holonomic partial velocities are

$${}^N\mathbf{v}_1^P = \hat{\mathbf{n}}_2, \qquad {}^N\mathbf{v}_2^P = \hat{\mathbf{b}}_1, \qquad {}^N\mathbf{v}_1^{\overline{B}} = \hat{\mathbf{n}}_2, \qquad {}^N\mathbf{v}_2^{\overline{B}} = \mathbf{0} \tag{10}$$

so that the dynamical equations of motion for S in N are found to be

$$m(\dot{u}_1 + S_\alpha \dot{u}_2) = 0 \tag{11}$$

$$m(S_\alpha \dot{u}_1 + \dot{u}_2) = \mu_1 \tag{12}$$

where S_α denotes $\sin\alpha$. The nonholonomic constraint force $\mathbf{C}_1$ contributes to the holonomic generalized active forces, whereas the holonomic constraint force $\mathbf{C}_2$ does not. A third equation is needed to solve for the three unknowns $\dot{u}_1$, $\dot{u}_2$, and μ_1; it is provided by the nonholonomic constraint equation (1) expressed at the acceleration level,

$$\dot{u}_2 - \frac{u_1}{\epsilon\, C_\alpha} = 0 \tag{13}$$

An analytical solution is then available,

$$\dot{u}_1 = -\frac{\tan\alpha}{\epsilon} u_1 \tag{14}$$

$$\dot{u}_2 = \frac{u_1}{\epsilon\, C_\alpha} \tag{15}$$

$$\mu_1 = \frac{mC_\alpha}{\epsilon} u_1 \tag{16}$$

It is worth noting that Eqs. (7), (15), and (16) together contradict the statement in Ref. [6] preceding Eq. (16) therein. The acceleration of P along the path (the rod) is in general nonzero if the proposed nonholonomic constraint equation is to be satisfied; it vanishes only in the special case when the rod is stationary ($u_1 = 0$).

The differential equation (14) yields a closed form solution

$$u_1 = K_1 e^{-\rho t} \tag{17}$$

where $\rho \triangleq (\tan \alpha)/\epsilon$ as defined in Ref. [6], and the constant of integration K_1 is the value of u_1 at $t = 0$. This solution in turn facilitates integration of Eq. (15), yielding

$$u_2 = -\frac{K_1}{S_\alpha} e^{-\rho t} + K_2 \tag{18}$$

where K_2 is determined once the value of u_2 at $t = 0$ is specified. Integration of the two kinematical differential equations $\dot{q}_r = u_r$ $(r = 1, 2)$ produces solutions for the generalized coordinates.

$$q_1 = -\frac{K_1}{\rho} e^{-\rho t} + K_3, \qquad q_2 = \frac{K_1}{\rho S_\alpha} e^{-\rho t} + K_2 t + K_4 \tag{19}$$

where the constants of integration K_3 and K_4 can be evaluated on the basis of the initial conditions of q_1 and q_2. It must be noted that the initial values of u_2 and q_1 have to satisfy Eq. (1); a similar recognition appears in Ref. [6].

One is now in a position to show that Eqs. (17)–(19) verify the results attributed in Ref. [6] to Hagedorn's approach. First, relationships between the Cartesian coordinates x and y and the generalized coordinates q_1 and q_2 are established.

$$x = C_\alpha q_2, \qquad \dot{x} = C_\alpha u_2, \qquad \ddot{x} = C_\alpha \dot{u}_2 \tag{20}$$

$$y = q_1 + S_\alpha q_2, \qquad \dot{y} = u_1 + S_\alpha u_2, \qquad \ddot{y} = \dot{u}_1 + S_\alpha \dot{u}_2 \tag{21}$$

Appropriate substitution from these relationships shows that the original form of the nonholonomic constraint equation given in Ref. [6], $y - x \tan \alpha - \epsilon \dot{x} = 0$, gives way to Eq. (1). Furthermore, Eqs. (12) and (14a) in Sec. 3 of Ref. [6] are recovered from Eqs. (19) here.

$$x = C_\alpha q_2 = \frac{K_1}{\rho \tan \alpha} e^{-\rho t} + K_2 C_\alpha t + K_4 C_\alpha \triangleq D_3 e^{-\rho t} + D_4 t + D_5 \tag{22}$$

$$y = q_1 + S_\alpha q_2 = -\frac{K_1}{\rho} e^{-\rho t} + K_3 + \frac{K_1}{\rho} e^{-\rho t} + K_2 S_\alpha t + K_4 S_\alpha \triangleq D_1 t + D_2 \tag{23}$$

It is clear that because of the constraint described by Eq. (1), only three of the constants of integration K_1, K_2, K_3, and K_4 are independent, and only three of the five constants $D_1, \ldots, D_5$ used in Ref. [6] are independent.

The second order differential equations (11a) and (11b) in Sec. 3 of Ref. [6] can also be recovered from Eqs. (14) – (16) here. Dealing with the differential equation for y is straightforward,

$$\ddot{y} = \dot{u}_1 + S_\alpha \dot{u}_2 = -\frac{\tan \alpha}{\epsilon} u_1 + S_\alpha \frac{u_1}{\epsilon \, C_\alpha} = 0 \qquad (24)$$

The differential equation for x can be rewritten as

$$\ddot{x} = C_\alpha \dot{u}_2 = C_\alpha \frac{u_1}{\epsilon \, C_\alpha} = \frac{\mu_1}{m C_\alpha} \qquad (25)$$

A relationship between the multiplier μ_1 used here and the multiplier λ used in Ref. [6] is required, and can be obtained by rewriting the original nonholonomic constraint equation as

$$-\epsilon \dot{x} + y - x \tan \alpha = {}^N\mathbf{v}^P \cdot (-\epsilon \hat{\mathbf{n}}_1) + y - x \tan \alpha = 0 \qquad (26)$$

Inspection of this equation indicates that a constraint force parallel to the vector $-\epsilon \hat{\mathbf{n}}_1$ must be applied to P. Consequently,

$$\mathbf{C}'_1 = -\lambda \epsilon \hat{\mathbf{n}}_1 \qquad (27)$$

The projection of $\mathbf{C}'_1$ onto $\hat{\mathbf{b}}_1$ must be the same as that of $\mathbf{C}_1$; therefore,

$$\mathbf{C}_1 \cdot \hat{\mathbf{b}}_1 = \mathbf{C}'_1 \cdot \hat{\mathbf{b}}_1 = \mu_1 = -\lambda \epsilon C_\alpha \qquad (28)$$

Hence, Eq. (25) is rewritten

$$\ddot{x} + \frac{\lambda \epsilon}{m} = 0 \qquad (29)$$

in agreement with Eq. (11a) of Ref. [6] when m is taken as unity.

Rosen and Edelstein reject the preceding differential equations for x and y, and the closed form solutions, on the basis of their results obtained with the Newton-Euler method. With the analysis already performed here it is evident

that their application of the method is flawed, and the point in their development where the mistake was made can be identified immediately. In what follows, a correct application of the Newton-Euler approach is shown to yield the foregoing results.

In their Eqs. (2a) and (2b), Rosen and Edelstein do not account for the constraint force $\mathbf{C}'_1$ needed to ensure satisfaction of their nonholonomic constraint equation; they only consider $\mathbf{C}_2$ required to bring about the configuration constraint. Upon writing $\mathbf{C}'_1 + \mathbf{C}_2 = m\,^N\mathbf{a}\,^P$, it is seen that Eqs. (2) should be stated

$$m\ddot{x} = -NS_\alpha - \lambda\epsilon, \qquad m\ddot{y} = NC_\alpha \tag{30}$$

or

$$mC_\alpha \dot{u}_2 = -NS_\alpha + \mu_1/C_\alpha, \qquad m(\dot{u}_1 + S_\alpha \dot{u}_2) = NC_\alpha \tag{31}$$

Application of Newton's second law to P must be accompanied by its application to B. In addition to the reaction forces $-\mathbf{C}'_1$ and $-\mathbf{C}_2$ acting at $\overline{B}$, a reaction force $\mu_3\hat{\mathbf{n}}_1$ is applied at the prismatic joint, therefore we write $\mu_3\hat{\mathbf{n}}_1 - \mathbf{C}'_1 - \mathbf{C}_2 = m_B\,^N\mathbf{a}\,^{\overline{B}}$; however, the mass of B is neglected in comparison to m so $\mu_3\hat{\mathbf{n}}_1 - \mathbf{C}'_1 - \mathbf{C}_2 = \mathbf{0}$. That is,

$$\mu_3 + \lambda\epsilon + NS_\alpha = 0, \qquad -NC_\alpha = 0 \tag{32}$$

The first of these can be used if μ_3 is of interest but the relationship is not important in what remains to be done. The second of these reveals that $N = 0$ (so $\mu_2 = 0$) and, as an immediate consequence, $\ddot{y} = 0$, in agreement with what has been previously shown. Equations (31) can now be simplified, and the nonholonomic constraint equation (13) at the acceleration level is again brought to bear so that

$$mC_\alpha \dot{u}_2 - \frac{\mu_1}{C_\alpha} = 0, \qquad \dot{u}_1 + S_\alpha \dot{u}_2 = 0, \qquad \dot{u}_2 - \frac{u_1}{\epsilon\,C_\alpha} = 0 \tag{33}$$

constitute three equations in three unknowns, $\dot{u}_1$, $\dot{u}_2$, and μ_1. They lead immediately to Eqs. (14)–(16).

Conclusion

A straightforward application of Kane's method for simple nonholonomic systems, together with identification of the constraint forces needed to impose a motion constraint and a configuration constraint, are used to verify results obtained with what is called the regular variational approach, brought to the reader's attention by Hagedorn in Ref. [6]. Further, a correct application of the Newton-Euler method reproduces those results. The conclusion by Hagedorn, Ray, Flannery, and others is thus affirmed; namely, the Lagrangean cannot be augmented by the sum of nonholonomic constraint equations weighted with multipliers, regardless of whether or not such equations are linear in the time derivatives of the generalized coordinates.

References

[1] Ray, J. R., 1966, "Nonholonomic Constraints," *American Journal of Physics*, **34**(5), pp. 406–408.

[2] Ray, J. R., 1966, "Erratum: Nonholonomic Constraints," *American Journal of Physics*, **34**(12), pp. 1202–1203.

[3] Saletan, E. J., and Cromer, A. H., 1970, "A Variational Principle for Nonholonomic Systems," *American Journal of Physics*, **38**(7), pp. 892–897.

[4] Rosenberg, R. M., 1977, *Analytical Dynamics of Discrete Systems*, Plenum Press, New York, p. 220.

[5] Rosen, A., and Edelstein, E., 1997, "Investigation of a New Formulation of the Lagrange Method for Constrained Dynamic Systems," *Journal of Applied Mechanics*, **64**(1), pp. 116–122.

[6] Hagedorn, P., 1997, "Investigation of a New Formulation of the Lagrange Method for Constrained Dynamic Systems," *Journal of Applied Mechanics*, **64**(4), pp. 1024–1027.

[7] Neǐmark, J. I., and Fufaev, N. A., 1972, *Dynamics of Nonholonomic Systems*, American Mathematical Society, Providence, Rhode Island.

[8] Pars, L. A., 1979, *A Treatise on Analytical Dynamics*, Ox Bow Press, Connecticut.

[9] Whittaker, E. T., 1964, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, 4th ed., Cambridge University Press, London.

[10] Flannery, M. R., 2005, "The Enigma of Nonholonomic Constraints," *American Journal of Physics*, **73**(3), pp. 265–272.

[11] Polyakhov, N. N., 1974, "Differential Principles of Mechanics Derived from Equations of Motion for Nonholonomic Systems," *Vestnik Leningradskogo Universiteta, Seriya Matematika Mekhanika i Astronomiya*, (13), pp. 106–114, (English translation).

[12] Kane, T. R., and Levinson, D. A., 1985, *Dynamics: Theory and Applications*, McGraw-Hill, New York, Chaps. 2, 4.