Journal of the Mechanics and Physics of Solids 137 (2020) 103871

Contents lists available at ScienceDirect

Journal of the Mechanics and Physics of Solids

journal homepage: www.elsevier.com/locate/jmps

Data-driven automated discovery of variational laws hidden in physical systems

Zhilong Huang[a], Yanping Tian[b], Chunjiang Li[a], Guang Lin[c], Lingling Wu[d], Yong Wang[a,*], Hanqing Jiang[d,*]

[a] Department of Engineering Mechanics, Zhejiang University, Hangzhou, Zhejiang 310027, China
[b] College of Mechanical Engineering, Hangzhou Dianzi University, Hangzhou, Zhejiang 310018, China
[c] Department of Mathematics and School of Mechanical Engineering, Department of Statistics, Purdue University, West Lafayette IN 47907, USA
[d] School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA

A R T I C L E I N F O

Article history:
Received 2 July 2019
Revised 17 December 2019
Accepted 9 January 2020
Available online 13 January 2020

A B S T R A C T

The automated discovery of physical laws from discrete noisy data is significant for evaluating the response, stability, and reliability of dynamic systems. In contract to the existing work on the discovery of differential laws, this paper presents a data-driven method to discover the variational laws of physical systems. The effectiveness and robustness to measurement noise are demonstrated with five physical cases. Two features of variational laws, the compact form and holistic viewpoint, lead to two intrinsic advantages in the data-driven discovery of variational laws, namely, reduced data requirement and robustness to noise. The presented data-driven method can be applied to discover variational laws in real time for physical fields or more complicated social sciences, with or without prior knowledge.

1. Introduction

The extraction of knowledge from measured discrete data has played a pivotal role in scientific discovery, such as in the case of Kepler's laws of planetary motion based on Brahe's observation data. In the history of physics, scientific discoveries have established physical laws and theories, always requiring enormous costs in terms of time and effort, and, more importantly, an ingenious imagination to extract nontrivial equal or nonequal mathematical relations (i.e., the laws) from large amounts of data. Such *manual* data-driven discoveries of physical laws in history define the "giants of science". With the advancement of data science and computing power, it is becoming possible to *automatically and rapidly* discover physical laws from discrete noisy data (Wang et al., 2016, Langley et al., 1987).

There are two mainstream methods employed to discover the governing equations of physical systems in differential form, i.e., symbolic regression (Bongard and Lipson, 2007, Schmidt and Lipson, 2009, Quade et al., 2016) and sparsity-promoting optimization (Brunton et al., 2016, Wang et al., 2011, Schaeffer et al., 2017) methods. By searching a space of mathematical expressions constructed by prescribed and nearly unbounded analytical functions, algebraic operators, and constant coefficients until reaching minimized metrics of error, the symbolic regression method can extract "free-form"

* Corresponding author.
E-mail addresses: ypwang@zju.edu.cn (Y. Wang), hanqing.jiang@asu.edu (H. Jiang).

https://doi.org/10.1016/j.jmps.2020.103871
0022-5096/

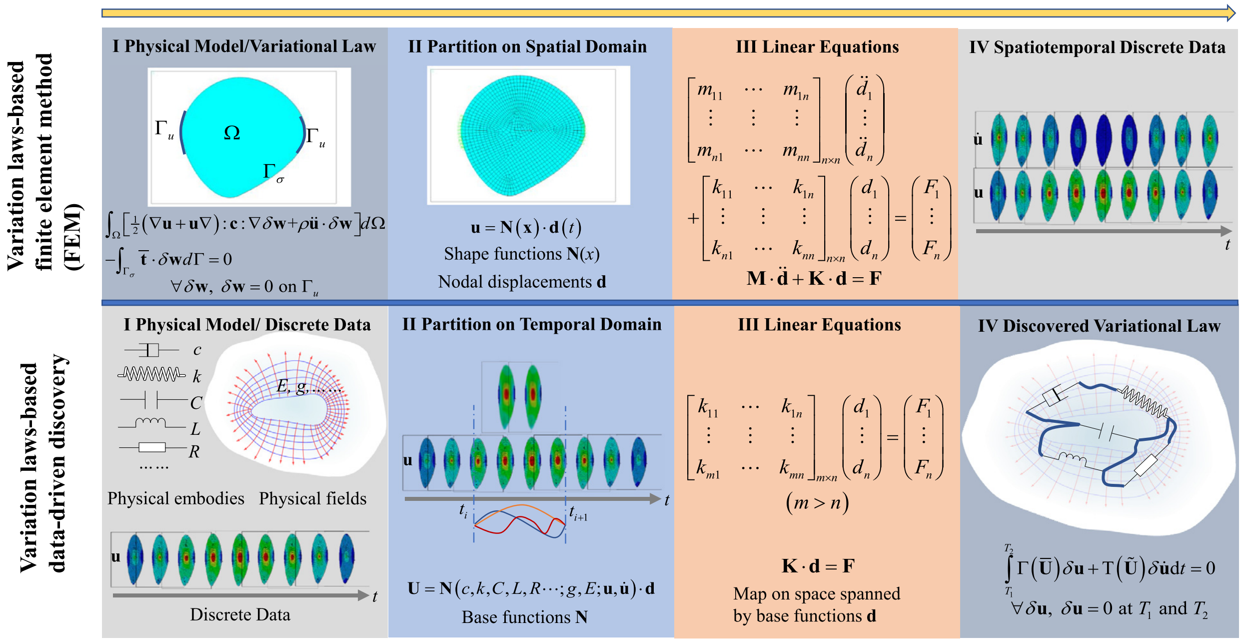

Fig. 1. Comparison of the variational laws-based finite element method and the variational laws-based data-driven discovery. The data-driven discovery (lower row) can be regarded as the reverse process of the finite element method (upper row). The finite element method starts from variational laws (I), partitions the spatial domain (II), and obtains discrete spatiotemporal data (IV) by solving linear equations (III), while the data-driven method starts from discrete data (I), partitions the temporal domain (II), and obtains the variational laws (IV) by solving linear equations (III).

physical laws, of any kind in theory, such as the Hamiltonian, Lagrangian, momentum conservation and governing equation of motion for some simple and crucial physical systems (Schmidt and Lipson, 2009); however, this process can be extremely time-consuming. The sparsity-promoting optimization method can identify first-order differential equations of nonlinear dynamical systems by expressing the first-order derivatives as linear combinations of a set of preselected basis functions and then, determining the unknown coefficients of the linear combinations to minimize the residuals between the assumed differential laws and discrete data (Brunton et al., 2016). The sparse-promoting optimization method is quite efficient and has been generalized to successfully discover partial differential equations (Rudy et al., 2017, Schaeffer, 2017, Mangan et al., 2017, Zhang and Lin, 2018), stochastic differential equations (Boninsegna et al., 2018), and model predictive control in the low-data limit (Quade et al., 2018, Brunton et al., 2016, Kaiser et al., 2017).

The differential laws of physical systems targeted by symbolic regression, sparse-promoting optimization, and their variants (Schaeffer et al., 2018, Schaeffer and McCalla, 2017) generally have an associated variational counterpart. Differential laws describe the evolutional relations between state variables at adjacent instants, while variational laws describe the concerned motions from a holistic viewpoint and are suitable for developing numerical methods to solve the differential laws, e.g., the finite element methods that have been adopted by almost all engineering subjects. Finite element methods (Belytschko et al., 2014, Fries and Belytschko, 2010) start from the variational laws of physical systems, such as mechanical, electrical, thermal or coupled systems, and obtain discrete spatiotemporal data, such as displacement, electric current, temperature or coupled quantities at discrete nodes and discrete instants by solving linear equations. As shown in Fig. 1, the reverse process of the finite element method is just a data-driven method that discovers the variational laws of physical systems directly from discrete data. This paper reports a new data-driven physical law discovery in variational form. Compared with the differential form data-driven physical law discovery, the present method has inherent significant merits, including a reduced data requirement, higher robustness to measurement noise, and an explicit embodiment of system/excitation parameters.

2. Data-Driven automated discovery of variational laws

The variational principle is considered to be the most ubiquitous principle in physics for both conservative and non-conservative systems. The fundamental assumption of the integral variational framework states the following: for a general physical system, the actual motion $\mathbf{q}$ (such as displacement or electric current, which is the output of this physical system) in any time interval $[T_1, T_2]$ obeys the variational relation (Whittaker, 1917, Lurie, 2002),

$$\int_{T_1}^{T_2} \left(\delta U_0 + U_j \delta q_j\right) dt = 0, \tag{1}$$

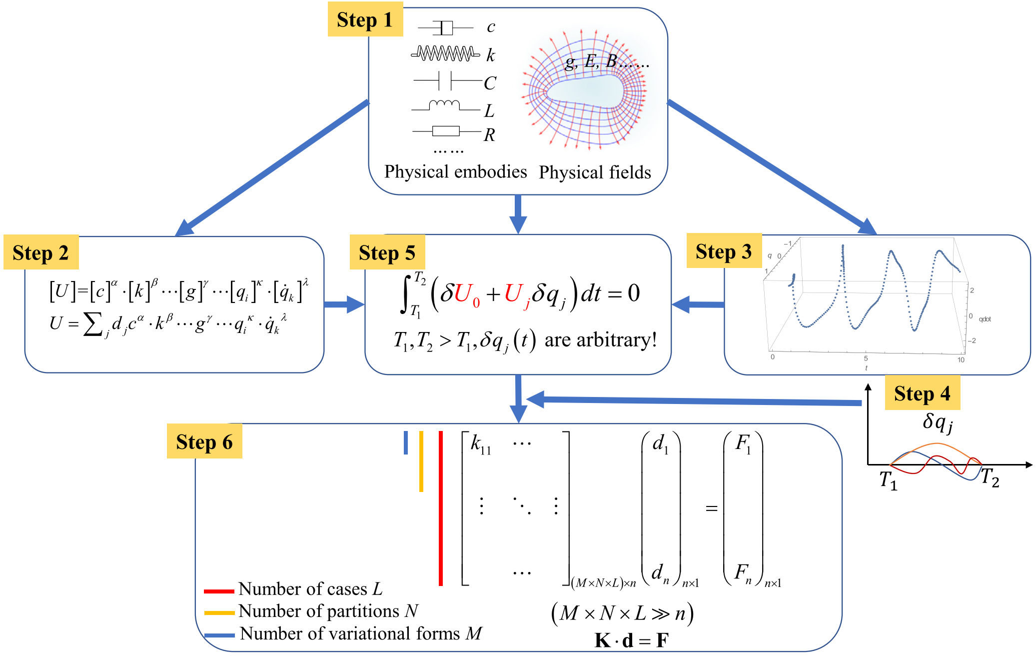

Fig. 2. Whole processes of the data-driven method used to discover the variational laws of physical systems. Step 1: Identify the system attributes **s** and the concerned motion **q**. Step 2: Construct the integrands U_0 and U_j based on dimensional analysis. Step 3: Collect the data. Step 4: Generate arbitrary variations of **q**. Step 5: Perform the integration. Step 6: Solve overdetermined linear algebraic equations to determine the coefficients of the linear combinations (the number of equations equals the product of the number of cases L, the number of partitions of each case N, and the number of variational forms M).

in which U_0 (with the unit of energy) and U_j are functions of the system attributes **s** (such as the stiffness and capacitance, as shown in Fig. 2), **q** and its time derivative $\dot{\mathbf{q}}$. $\delta q_j(t)$ represent arbitrary variations of the actual motion with the constraints $\delta q_j(T_1) = 0$ and $\delta q_j(T_2) = 0$. The Einstein summation convention has been adopted. This relation is a more general version of the principle of least action, one of the cornerstones of physics, which has been used to derive fundamental physical laws, such as the Newtonian, Lagrangian, and Hamiltonian descriptions and general relativity (Glodstein, 1980, Longair, 2003) by giants of science. The variational relation is acknowledged to be the most elegant and compact form of physical laws (Lanczos, 2015, Landau and Lifshitz, 2000), and the actual motion is completely described once the integrands (i.e., U_0 and U_j in Eq. (1)) are specified. Different from the trial-and-error method advised by Richard Feynman (Feynman et al., 2010), the data-driven method aims to automatedly determine the integrands from discrete data. Once the integrands are determined in terms of **s**, **q**, and $\dot{\mathbf{q}}$, the variational laws obeyed by physical systems are completely determined.

Fig. 2 illustrates the processes of the data-driven method implemented to discover the variational laws of physical systems. The processed consist of step 1, identifying the system attributes **s** that *may* influence the motion **q**; step 2, constructing the to-be-determined integrands U_0 and $U_j$based on dimensional analysis and linear combinations of the associated parameter clusters; step 3, collecting discrete data (i.e., $\{\mathbf{q}(t_i)\}$ and $\{\dot{\mathbf{q}}(t_i)\}$); step 4, generating arbitrary variations of **q** (i.e., $\delta q_j(t)$) based on the constraints at the initial and final instants; step 5, performing the integration given by Eq. (1); and finally step 6, solving an overdetermined linear algebraic equation to determine the coefficients of the linear combinations given by step 2 and thus formulating the integrands.

It should be noted here that prior physical knowledge is required to grasp the physical attributes, e.g., mass m, damping c and stiffness k for a mechanical system and capacitance C, inductance L, and resistance R for an electrical system, as well as the concerned physical motions, i.e., displacement for a mechanical system and electric current for an electrical system. The to-be-determined integrands $U_0(\mathbf{s};\mathbf{q},\dot{\mathbf{q}})$ and $U_j(\mathbf{s};\mathbf{q},\dot{\mathbf{q}})$ are expressed as linear combinations of the parameter clusters constituted by the physical attributes **s**, concerned physical motion **q** and its time derivate $\dot{\mathbf{q}}$. Although the linear combinations should include infinite terms, only the terms with lower-order powers are of great significance according to the parsimonious criterion (Brunton et al., 2016). Dimensional analysis is utilized to determine the power functions of the arguments (i.e., the parameter clusters).

In any physical system, a physical quantity with a unit can always be expressed by basic units (Bridgeman, 1963), such as mass [M], length [L], time [T], and electric current [A], where [·] denotes the dimension of the quantity " · ". Thus, the integrand U_0 that possesses the dimension of energy can be expressed as a linear combination of any clusters with the dimension of $[\mathrm{M}]^1[\mathrm{L}]^2[\mathrm{T}]^{-2}$ constituted by the physical attributes **s**, physical motion **q** and it time derivate $\dot{\mathbf{q}}$; i.e., $U_0(\mathbf{s},\mathbf{q},\dot{\mathbf{q}}) = A_i([\mathrm{M}]^1[\mathrm{L}]^2[\mathrm{T}]^{-2})_i$, and A_i are dimensionless coefficients. The integrands U_j can be written with the same means given that its dimension is $\frac{[\mathrm{M}]^1[\mathrm{L}]^2[\mathrm{T}]^{-2}}{[q_j]}$; i.e., $U_j(\mathbf{s},\mathbf{q},\dot{\mathbf{q}}) = B_i(\frac{[\mathrm{M}]^1[\mathrm{L}]^2[\mathrm{T}]^{-2}}{[q_j]})_i$, where B_i are dimensionless coefficients. Substituting the integrands into Eq. (1), performing the variational calculus on δU_0, and rearranging the integrands according to δq_i and $\delta \dot{q}_i$, Eq. (1) becomes

$$\int_{T_1}^{T_2} \left\{ C_i \left(\frac{[\mathrm{M}]^1[\mathrm{L}]^2[\mathrm{T}]^{-2}}{[\dot{q}_j]} \right)_i \delta \dot{q}_j + D_i \left(\frac{[\mathrm{M}]^1[\mathrm{L}]^2[\mathrm{T}]^{-2}}{[q_j]} \right)_i \delta q_j \right\} dt = 0. \quad (2)$$

Data acquisition is then performed by any means (e.g., with experiments or simulations), i.e., collecting discrete data $\{\mathbf{q}(t_i)\}$ and $\{\dot{\mathbf{q}}(t_i)\}$ in the prescribed time interval, where t_i represent discrete time instants. Before the data are applied to determine the coefficients C_i and D_i, the variations $\delta\mathbf{q}$ and $\delta\dot{\mathbf{q}}$ need to be specified. The choice of variation is arbitrary in the arbitrary subinterval of the prescribed time interval but requires vanishing variations for the initial and final instants, such as $\delta\mathbf{q}(t_j) = 0, \delta\mathbf{q}(t_{j+1}) = 0$ for the j^{th} partition $[t_j, t_{j+1}]$. For example, without loss of generality, $\delta q_i(t) = (t - t_j)^m (t - t_{j+1})^n \varepsilon$, where ε is a small parameter and m and n are positive integers. By substituting the discrete data (i.e., $\{\mathbf{q}(t_i)\}$ and $\{\dot{\mathbf{q}}(t_i)\}$) into a certain partition (e.g., $[t_j, t_{j+1}]$, a subinterval of a prescribed time interval, where $t_i \in [t_j, t_{j+1}]$) and the prescribed variations $\delta\mathbf{q}$ and $\delta\dot{\mathbf{q}} = \frac{\mathrm{d}}{\mathrm{d}t}(\delta\mathbf{q})$ into Eq. (2), one obtains a homogeneous linear algebra equation with respect to the coefficients C_i and D_i. Since both the subinterval $[t_j, t_{j+1}]$ and the variations $\delta\mathbf{q}$ can be selected arbitrarily, homogeneous linear algebraic equations with enough equations (i.e., overdetermined equations) to solve the coefficients C_i and D_i can be obtained. In solving the homogeneous linear algebraic equations, one coefficient of C_i and D_i is initially set to 1 or −1, and then one needs to solve the nonhomogeneous linear algebraic equations by various means. Once the coefficients C_i and D_i are solved, the data-driven discovery of the variational laws of physical systems is completed.

2.1. Toy example – free-falling object

Here, we illustrate the whole process of the data-driven method with a simple example of a free-falling object, a fundamental problem in Newtonian physics. Consider a particle with mass m falling under the sole influence of gravity with gravitational acceleration g, which defines the physical attributes (i.e., $s_1 = m, s_2 = g$). The coordinate of the particle $q(t)$ is the concerned output. Now, we utilize the discrete data of the output $\{q(t_i)\}$ to extract the variational law of a free-falling object. This classical problem is chosen, because the exact solution is known: i.e., the differential law $\ddot{q}(t) = g$, velocity $\dot{q}(t) = gt + \dot{q}_0$, and displacement $q(t) = \frac{1}{2}gt^2 + \dot{q}_0 t + q_0$, where q_0 is the initial displacement and $\dot{q}_0$ is the initial velocity.

After identifying the physical attributes and system output as described in step 1 in Fig. 2, the to-be-solved integrands U_0 and U_1 are expressed as linear combinations of the parameter clusters constituted by m, g, q, and $\dot{q}$ based on dimensional analysis. Considering the units of m, g, q, and $\dot{q}$, for the sake of simplicity, we select three of these four terms to construct the parameter clusters of U_0 and U_1, yielding $U_0(\mathbf{s}; q, \dot{q}) = A_1 mgq + A_2 m\dot{q}^2$ and $U_1(\mathbf{s}; q, \dot{q}) = B_1 mg$, where A_1, A_2, andB_1 are unknown constants. Substituting U_0 and U_1 into Eq. (1) and rearranging the integrands according to $\delta q, \delta\dot{q}$ yield an integral similar to Eq. (2), $\int_{T_1}^{T_2} (C_1 m\dot{q}\delta\dot{q} + D_1 mg\delta q)dt = 0$, where C_1 and D_1 are two to-be-determined coefficients. Now, step 2 in Fig. 2 is completed. It should be noted that all four terms can be used to construct the parameter clusters if three terms cannot achieve satisfactory results.

Step 3 involves obtaining the discrete data $\{q(t_i)\}$ and $\{\dot{q}(t_i)\}$, which can be achieved by experimental observations. Here, to demonstrate the process of this data-driven method, $\{q(t_i)\}$ and $\{\dot{q}(t_i)\}$ are first obtained analytically. **Figs. S1** and **S2** show the discrete data in the time interval [0, 10]s with the sampling interval 10^{-3} s (see Supporting Information for details). Step 4 applies the arbitrary variation $\delta q(t)$ to the integral. The following is adopted in this example: $\delta q(t) = (t - t_j)(t - t_{j+1})\varepsilon$, and $\delta\dot{q}(t) = (2t - t_j - t_{j+1})\varepsilon$, where t_j and t_{j+1} are the initial and final instants for subinterval $[t_j, t_{j+1}]$ and $\varepsilon = 1$. In step 5, the discrete data (i.e., $\{q(t_i)\}$ and $\{\dot{q}(t_i)\}$) and the variations $\{\delta q(t), \delta\dot{q}(t)\}$ are substituted into the integral. Numerical integration is performed to yield the homogeneous linear algebraic equation(s) with respect to undetermined coefficients $C_1$and D_1. Setting $D_1 = 1.0$, the undetermined coefficient C_1 is determined by solving the nonhomogeneous linear algebraic equation(s).

In this example, only one coefficient C_1 needs to be determined. Only one partition is used, i.e., $t_1 = 0$ and $t_2 = 10$. Substituting the discrete data and infinitesimal variation into the variational formula and performing the numerical integration of $\int_{t_1}^{t_2} (C_1 m\dot{q}\delta\dot{q} + D_1 mg\delta q)dt = 0$ yield $C_1 = 1.000$. Thus, the variational law obeyed by the motion of the free-falling particle is obtained as,

$$\int_{T_1}^{T_2} (1.000 \times m\dot{q}\delta\dot{q} + 1 \times mg\delta q)dt = 0,$$

which rebuilds the well-known variational law of a free-falling object in Newtonian physics, i.e., $\int_{T_1}^{T_2} \delta(\frac{1}{2}m\dot{q}^2 + mgq)dt = 0$. As shown in the Supporting Information (**Figs. S3** and **S4**), if the discrete data (i.e., $\{q(t_i)\}$ and $\{\dot{q}(t_i)\}$) are accurate, even for a few data points (for example, only three data points at t_i = 0, 5, 10 s, instead of 10,001 data points in the time interval [0, 10]s), this method can still obtain the same variational law.

Another important characteristic of employing the data-driven method to identify physical laws is robustness to measurement noise. Here, Gaussian noise with zero mean and a standard deviation of 10 is imposed in exact dataset (**Figs. S1** and **S2**), as shown in **Figs. S5** and **S6**. Using the same method, the variational law discovered from the noisy data is $\int_{T_1}^{T_2} [(1 + 0.66 \times 10^{-3}) \times m\dot{q}\delta\dot{q} + 1 \times mg\delta q]\mathrm{d}t = 0$. This suggests that the present method is robust to measurement noise.

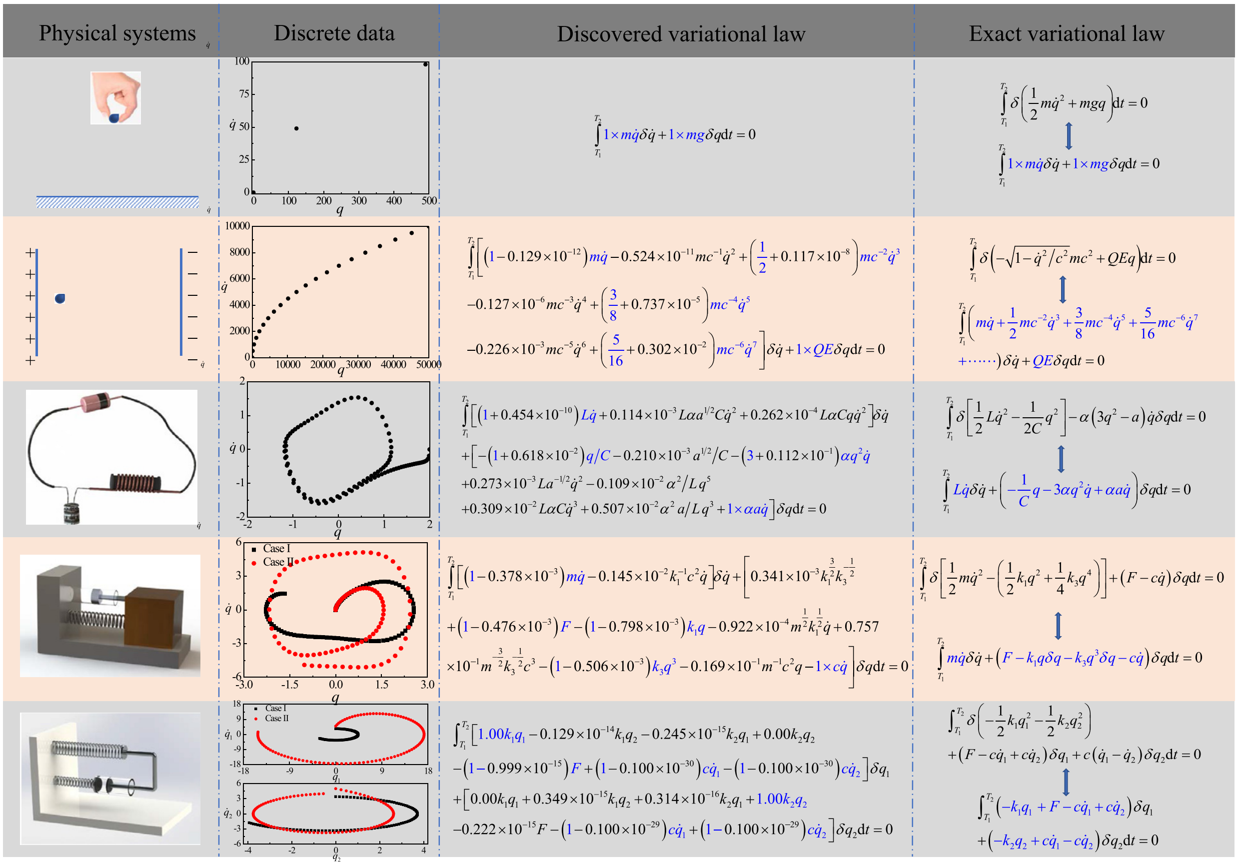

Fig. 3. Summary of the physical laws discovered from the data-driven method. First row: free-falling object. Second row: relativistic particle moving in an electric field. Third row: van der Pol oscillator. Fourth row: Duffing oscillator. Fifth row: two-degree-of-freedom dissipative system.

Fig. 3 summarizes the cases we have studied in this paper, and in the following, we further show the capability and robustness of this method. A detailed discussion can be found in the Supporting Information.

2.2. Relativistic particle moving in an electric field

Consider a small particle with charge Q and mass m rapidly moving in a homogenous electric field E along the direction of the electric field. This is a classical problem in the theory of relativity. The known differential equation is $\frac{d}{dt}(\frac{1}{\sqrt{1-\dot{q}^2/c^2}}m\dot{q}) - QE = 0$, where c is the velocity of light and $q(t)$ is the coordinate of the particle. Now, we use the data-driven method to rediscover the variational law. Prior physical knowledge is required to determine the observable output of this problem as the coordinate of the particle $q(t)$ and the physical attributes as the mass $m(=s_1)$, charge $Q(=s_2)$, intensity of the electric field $E(=s_3)$, and light velocity $c(=s_4)$. As detailed in the Supporting Information, the dimensional analysis expresses U_0 and U_1 as linear combinations, i.e., $U_0(\mathbf{s}; q, \dot{q}) = A_1QEq + A_2mc\dot{q} + A_3m\dot{q}^2 + A_4mc^{-1}\dot{q}^3 + A_5mc^{-2}\dot{q}^4 + A_6mc^{-3}\dot{q}^5 + A_7mc^{-4}\dot{q}^6 + \cdots$, and $U_1(\mathbf{s}; q, \dot{q}) = B_1QE + B_2QEc^{-1}\dot{q} + B_3QEc^{-2}\dot{q}^2 + B_4QEc^{-3}\dot{q}^3 + B_5QEc^{-4}\dot{q}^4 + \cdots$, in which A_i, B_i are constant coefficients. Substituting linear combinations into Eq. (1) and rearranging based on Eq. (2) yield the following equation:

$$\int_{T_1}^{T_2}\left[\sum_{i=1} C_i mc\left(\frac{\dot{q}}{c}\right)^{i}\delta\dot{q} + \sum_{i=1} D_i QE\left(\frac{\dot{q}}{c}\right)^{i-1}\delta q\right]\mathrm{d}t = 0 \tag{3}$$

The same variational function $\delta q(t)$ as in the previous case is used. The data are collected in the time interval [0, 10]s with the sampling interval 10^{-3} s. Here, ten identical segments are utilized in this time interval; i.e., Eq. (3) leads to ten equations. Thus, ten linear algebraic equations are used to determine the coefficients C_i and D_i.

It should be noted that the actual variational form for this problem is $\int_{T_1}^{T_2}\delta[-\sqrt{1-(\frac{\dot{q}}{c})^2}mc^2 + QEq]dt = 0$, which can be expanded based on the relative velocity of the particle (i.e., $\frac{\dot{q}}{c}$); i.e., $\int_{T_1}^{T_2}[(m\dot{q} + \frac{1}{2}mc^{-2}\dot{q}^3 + \frac{3}{8}mc^{-4}\dot{q}^5 + \frac{5}{16}mc^{-6}\dot{q}^7 + \cdots\cdots)\delta\dot{q} + QE\delta q]\mathrm{d}t = 0$. This explains the infinity terms in Eq. (3). As shown in the Supporting Information, depending on the orders of the coefficient of $\delta\dot{q}$ truncated in Eq. (3), the present method can accurately reproduce the actual variational form. For example, when truncating at $(\frac{\dot{q}}{c})^7$, the present method gives

$$\int_{T_1}^{T_2} \Big[\left(1 - 0.94 \times 10^{-14}\right) m\dot{q} - 0.296 \times 10^{-13} mc^{-1}\dot{q}^2 + \left(\frac{1}{2} - 0.513 \times 10^{-9}\right) mc^{-2}\dot{q}^3 - 0.422 \times 10^{-9} mc^{-3}\dot{q}^4$$
$$+ \left(\frac{3}{8} + 0.133 \times 10^{-5}\right) mc^{-4}\dot{q}^5 - 0.666 \times 10^{-5} mc^{-5}\dot{q}^6 + \left(\frac{5}{16} + 0.918 \times 10^{-3}\right) mc^{-6}\dot{q}^7 \Big] \delta\dot{q} + 1 \times QE\delta q \mathrm{d}t = 0,$$

which agrees very well with the expansion form of the actual variational form.

We also study the sensitivity of the quality of the discovered variational law to the amount of data. Instead of using 10^{-3} s as the sampling interval, a sampling interval of 0.5 s is used in the same [0, 10] s time interval. In other words, only 21 data points are provided, instead of the original 10,001 data points. The same ten identical segments are used, i.e., only three data points for each segment. As detailed in the Supporting Information, if the data are accurate enough, the variational laws discovered from fewer data points still agree very well with those discovered using a large amount of data.

2.3. van der Pol oscillator and Duffing oscillator

Now, we consider two nonlinear dynamic examples, namely, a van der Pol oscillator (van der Pol, 1920) consisting of a semiconductor component with threshold a and strength α for energy injection/ dissipation, an inductor L and a capacitor C, and a Duffing oscillator (Nayfeh and Mook, 1979) consisting of a block with mass m, a spring with linear stiffness k_1, a nonlinear stiffness k_3, and a damper c, excited by a time-varying external force $F(t)$, as shown in **Figs. S11** and **S24**, respectively. A van der Pol oscillator is a nonconservative oscillator with nonlinear damping and has been extensively studied in the physical and biological sciences (Fitzhugh, 1960, Cartwright et al., 1999). The known governing equation in differential form is $L\ddot{q} + \alpha(3q^2 - a)\dot{q} + q/C = 0$, where q is the current. The Duffing oscillator is an important dynamical system exhibiting chaotic and frequency hysteresis behaviors (Novak and Frehlich, 1982). The known governing equation in differential form is $m\ddot{q} + c\dot{q} + k_1q + k_3q^3 = F(t)$, where q is the displacement of the block. As detailed in the Supporting Information, the present data-driven method can discover the variational laws of the van der Pol oscillator and the Duffing oscillator with good accuracy, even with a small amount of exact data.

Moreover, the influence of measurement noise on the quality of the discovered variational laws is also studied. The noisy data are generated by adding Gaussian noise with zero mean and standard deviation σ into the exact data by numerically solving the differential equations. It is found that even in the cases with strong noise, the data-driven method discovers variational laws with acceptable accuracy. Specifically, for a standard deviation of $\sigma = 0.05$, the mean values and standard deviations of the error of the discovered variational laws are $E[\text{Error}] = 0.390 \times 10^{-2}$ and $\sigma[\text{Error}] = 0.481 \times 10^{-2}$ for the van der Pol oscillator, and $E[\text{Error}] = 0.526 \times 10^{-3}$ and $\sigma[\text{Error}] = 0.401 \times 10^{-3}$ for the Duffing oscillator, respectively (the error is defined in Eq. (8) in the Supporting Information).

Undoubtedly, noisy data from the same system with different parameter values and different initial conditions (i.e., the same system with different cases) can provide more information with respect to the system itself and may reduce the influence of measurement noise on the results. The accuracy of the discovered variational laws will be improved dramatically by utilizing data from several cases simultaneously. The values of the parameters and initial conditions for five cases are listed in **Table S3** (for the van der Pol oscillator) and **Table S5** (for the Duffing oscillator). The relations of the error of the discovered variational laws with respect to the number of cases are depicted in **Figs. S20** (for the van der Pol oscillator) and **S31** (for the Duffing oscillator). As expected, the error of the discovered variational laws monotonically decreases with an increase in the number of cases.

We also discuss the influence of the number of partitions of the time interval, the length of the time interval and the selection of variation styles on the quality of the discovered variational laws. As shown in **Figs. S21** (for the van der Pol oscillator) and **S34** (for the Duffing oscillator), the error of the discovered variational laws depends on the number of partitions of the time interval, and there exists an optimal length of the segment, approximately 1/4–1/8 of the oscillation period, to minimize the error. This length of the segment balances the information included in each segment and the number of linear equations available. Notably, more information can be included in the model by using a longer time interval. Thus, the error of the discovered variational laws decreases monotonically with an increase in the length of the time interval, as shown in **Figs. S22** (for the van der Pol oscillator) and **S35** (for the Duffing oscillator). It is also found that the error of the discovered variational laws only slightly fluctuates with changes in the variation style, which suggests an insensitivity of the data-driven method to the variation styles.

2.4. Two-Degree-of-Freedom dissipative system

To demonstrate the application and efficacy of the data-driven method to multi-degree-of-freedom dissipative systems, here we study a specific mechanical system in which a linear spring (with stiffness coefficient k_2) is in series with a dashpot (with damping coefficient c) and then in parallel with another linear spring (with stiffness coefficient k_1) and excited by a time-varying external force $F(t)$, as shown in **Fig. S37**. This mechanical system is inspired by the Maxwell representation of a standard linear solid model. The known governing equations in differential form are $-k_1q_1 + F - c(\dot{q}_1 - \dot{q}_2) = 0$ and

$-k_2q_2 + c(\dot{q}_1 - \dot{q}_2) = 0$, where q_1 describes the position of the free endpoint, and q_2 describes the position of the connection point between the dashpot and the spring.

Similar to the aforementioned examples, the variational law for this two-degree-of-freedom system can be discovered using exact or noisy data. Again, the robustness to measurement noise and the need for small amount of data have been shown for this example. Furthermore, the effects of the other parameters, such as the number of partitions and the length of the time interval have been studied. Different from the four one-degree-of-freedom examples in which only one differential equation associated with the variational law is discovered, a two-degree-of-freedom problem can lead to the discovery of two differential equations associated with the variational law, and these two differential equations can be expressed as an equivalent transformation of the exact differential equations. For more details, please refer to the Supporting Information.

2.5. One-Dimensional continuous system

The present data-driven method is not limited to discrete systems, although the aforementioned examples are only discrete systems. The method can be readily generalized to discover the variational laws of continuous systems. Here, we adopt a one-dimensional continuous system as a simple example to illustrate the generalization and implementation of the method.

For one-dimensional continuous systems, the fundamental assumption of the integral variational framework is the following: the actual motion $q(x, t)$ in any time interval $[T_1, T_2]$ obeys the variational relation:

$$\int_{T_1}^{T_2} \left\{ \delta\left[\int_{L_1}^{L_2} U_0\big(\mathbf{s}; q, \dot{q}; q', q''...\big)\mathrm{d}x\right] + \int_{L_1}^{L_2} \left[\bar{U}\big(\mathbf{s}; q, \dot{q}; q', q''...\big)\delta q\right]\mathrm{d}x \right\}\mathrm{d}t = 0 \tag{4}$$

where the integrands $U_0(\mathbf{s}; q, \dot{q}; q', q''...)$ (with the unit of energy density) and $\bar{U}(\mathbf{s}; q, \dot{q}; q', q''...)$ are functions of the system attributes $\mathbf{s}$, actual motion q, its time derivative $\dot{q}$ and various-order spatial derivatives (q', q''...). $\delta q(x, t)$ represents arbitrary variations of the actual motion satisfying the temporal constraints $\delta q(x, T_1) = 0$ and $\delta q(x, T_2) = 0$ and the spatial constraints at the boundaries $x = L_1$ and $x = L_2$.

The ultimate objective is to obtain the analytical expression of the to-be-determined integrands U_0 and $\bar{U}$ from a set of exact or noisy discrete data. First, we construct the parameter clusters with the dimensions of the integrands U_0 and $\bar{U}$, and express the integrands as linear combinations of the associated parameter clusters. By assuming arbitrary variations of $q(x, t)$ (i.e., $\delta q(x, t)$) based on the temporal and spatial constraints and performing the integration, one obtains a set of overdetermined linear algebraic equations with respect to the coefficients of the linear combinations. Once the coefficients of the linear combinations are determined, the variational law of the continuous system is then discovered.

As shown in the Supporting Information, the proposed data-driven method successfully discovers the variational law of a free-vibrating straight beam with simply supported constraints (as shown in **Fig. S50**).

3. Discussion

In this paper, we have established a data-driven method to automatedly discover variational laws hidden in physical systems from discrete data, which can be exact or noisy. As the most ubiquitous principle and the counterpart of differential laws, variational laws have two features, namely, the compact form and holistic viewpoint to the description of physical behaviors, that provide the data-driven method with great advantages. The compact form with fewer to-be-determined quantities requires a small amount of data for the discovery of physical laws; the holistic viewpoint with integral calculus leads to robustness to measurement noise. The advantages of these two features, i.e., the reduced data requirement and robustness to noise, enable real-time discovery of physical laws with the data-driven method, which is of great significance to time-sensitive prediction and control of dynamic events, such as in earthquake and extreme climate scenarios.

The present data-driven method utilizes some prior physical knowledge (i.e., the dimensions of the system attributes) to simplify the selection of the integrands. We must emphasize that it is not necessary to include dimensional analysis in the method used to discover variational laws. As shown in the Appendix, one can simply expand the integrands as linear combinations of functions (which are not limited to polynomial functions and can be sinusoidal or exponential functions) of the concerned physical motion and its time derivative, determine a series of coefficients for each function for the same system with different values of attributes, and then derive explicit relations between each coefficient and the system attributes by symbolic regression. In this way, the presented data-driven method can be extended to new fields with little prior knowledge. In fact, such new fields may not necessarily be physical fields; they can be fields in the social sciences (Dale and Bhat, 2018), since the concerned motions (e.g., the response of a specific population to an event) are often time-dependent.

In summary, the presented data-driven method in variational form possesses two intrinsic advantages, namely, a reduced data requirement and robustness to measurement noise, which enable its applicability to the discovery of governing equations in real time for physical fields or more complicated social sciences, with or without prior knowledge.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

CRediT authorship contribution statement

Zhilong Huang: Supervision, Writing - review & editing. **Yanping Tian:** Conceptualization, Methodology, Software. **Chunjiang Li:** Software. **Guang Lin:** Supervision. **Lingling Wu:** Writing - review & editing, Visualization. **Yong Wang:** Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Hanqing Jiang:** Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing.

Acknowledgments

YW acknowledges the National Natural Science Foundation of China under Grant Nos. 11872328 and 11472240. HJ acknowledges the Guang Biao Professorship of Zhejiang University. ZL acknowledges the National Natural Science Foundation of China under Grant Nos. 11532011 and 11621062. GL would like to acknowledge the support from the National Science Foundation (DMS-1555072, DMS-1736364, and DMS-1821233). YW and HJ would like to thank Prof. JQ Lu and Prof. XL Jin from Zhejiang University for helpful discussions.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jmps.2020.103871.

Appendix. Discovery of a Variational Law Without Prior Physical Knowledge

The generalized data-driven method can discover the variational law of a system without prior physical knowledge (i.e., the dimensions of the system attributes) and includes the following seven steps: step 1, determining the attributes and measures $\mathbf{s}$ and the concerned motion $\mathbf{q}$; step 2, expanding the integrands U_0 and U_j as polynomials of physical motion $\mathbf{q}$ and its time derivative $\dot{\mathbf{q}}$ with undetermined coefficients; step 3, collecting the data of the system with different values of attributes; step 4, generating arbitrary variations of $\mathbf{q}$; step 5, performing the integration; step 6, solving overdetermined linear algebraic equations to determine a series of coefficients for the system with different values of attributes; and finally step 7, establishing explicit relations between the coefficients and attributes by symbolic regression.

Here, we illustrate the whole process of the generalized data-driven method through the one-dimensional motion of a particle subjected to a constant mechanical action. Step 1: select the system attributes (i.e., $s_1 = m$ to measure the weight of the particle and $s_2 = F$ to measure the constant action) and the coordinate of the particle $q(t)$ to describe the concerned motion. Step 2: express the to-be-determined integrands U_0 and U_j by the 2nd-order polynomials of the coordinate $q(t)$ and its time derivative $\dot{q}$, i.e., $U_0 = A_1 q + A_2\dot{q} + A_3 q\dot{q} + A_4 q^2 + A_5\dot{q}^2$, $U_1 = B_1 + B_2 q + B_3\dot{q} + B_4 q\dot{q} + B_5 q^2 + B_6\dot{q}^2$. Here for the sake of simplicity, only the 2nd-order polynomials are considered, though different orders can be adopted. Substituting U_0 and U_1 into the variational relation in Eq. (1) and rearranging the integrand according to $\delta q, \delta\dot{q}$ yield

$$\int_{T_1}^{T_2} (C_1 + C_2\dot{q})\delta\dot{q} + \left(D_1 + D_2 q + D_3\dot{q} + D_4 q\dot{q} + D_5 q^2 + D_6\dot{q}^2\right)\delta q dt \tag{A1}$$

in which, C_i and D_i are the to-be-determined coefficients.

Step 3: collect the discrete data $\{q(t_i)\}$ and $\{\dot{q}(t_i)\}$ from the exact solution of the particle motion (i.e., $q(t) = \frac{1}{2}\frac{F}{m}t^2 + \dot{q}_0 t + q_0$ and $\dot{q}(t) = \frac{F}{m}t + \dot{q}_0$) in the time interval [0, 1]s with a sampling interval of 10^{-2} s, where the initial displacement is $q_0 = 0$ and the initial velocity is $\dot{q}_0 = 0$. Several sets of values of attributes are listed in **Table S10**, and the discrete data associated with these sets of values of attributes are depicted in **Figure S51**.

Step 4: set the arbitrary variation as $\delta q(t) = (t - t_j)(t - t_{j+1})\varepsilon$, and then use$\delta\dot{q}(t) = (2t - t_j - t_{j+1})\varepsilon$, where t_j and t_{j+1} are the initial and final instants for the subinterval $[t_j, t_{j+1}]$ and $\varepsilon = 1$. Step 5: separate the time interval [0, 1]s into ten equal segments. Substituting the discrete data and the variations $\{\delta q(t), \delta\dot{q}(t)\}$ into the variational relation and completing the integration yield one set of overdetermined linear algebraic equations for each set of values of attributes. Step 6: set $D_1 = -1.0$, and solve the overdetermined linear algebraic equations to determine a series of coefficients for the functions of the system with different values of attributes, as shown in **Table S11**.

Step 7: establish the explicit relations between the coefficients and the attributes by symbolic regression, as follows:

$$C_1 = D_2 = D_3 = D_4 = D_5 = D_6 = 0, D_1 = -1, C_2 = -\frac{m}{F} \tag{A2}$$

Thus, the discovered variational law is

$$\int_{T_1}^{T_2} -\frac{m}{F}\dot{q}\delta\dot{q} - 1 \times \delta q dt \tag{A3}$$

which rebuilds the well-known variational law of a particle under mechanical action, i.e., $\int_{T_1}^{T_2} \delta(\frac{1}{2}m\dot{q}^2) + Fqdt = 0$.

Reference

Belytschko, T., Liu, W.K., Moran, B., Elkhodary, K.I., 2014. Nonlinear finite elements for continua and structures. Wiley, Chichester.
Bongard, J., Lipson, H., 2007. Automated reverse engineering of nonlinear dynamical systems. Proc. Natl. Acad. Sci. USA 104, 9943–9948.
Boninsegna, L., Nuske, F., Clementi, C., 2018. Sparse learning of stochastic dynamical equations. J. Chem. Phys. 148, 241723.
Bridgeman, P.W., 1963. Dimensional analysis. Yale University Press, New Haven.
Brunton, S.L., Proctor, J.L., Kutz, J.N. , 2016a. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc. Natl Acad. Sci. USA 113, 3932–3937.
Brunton, S.L., Proctor, J.L. & Kutz, J.N. Sparse identification of nonlinear dynamics with control (SINDYc). arXiv:1605.06682v1 (2016b).
Cartwright, J., Eguiluz, V., Hernandez-Garcia, E., Piro, O., 1999. Dynamics of elastic excitable media. Internat. J. Bifur. Chaos 9, 2197–2202.
Dale, R., Bhat, H.S., 2018. Equations of mind: data science for inferring nonlinear dynamics of socio-cognitive systems. Cogn. Syst. Res. 52, 275–290.
Feynman, R.P., Leighton, R.B., Sands, M., 2010. Lectures in physics. Basic Book, New York.
Fitzhugh, R., 1960. Impulses and physiological states in theoretical models of nerve membranes. Biophys. J. 1, 445–466.
Fries, T.P., Belytschko, T., 2010. The extended/generalized finite element method: an overview of the method and its applications. Int. J. Numer. Meth. Engng. 84, 253–304.
Glodstein, H., 1980. Classical mechanics. Addison-Wesley Publishing Company Inc, Massachusetts.
Kaiser, E., Kutz, J.N. & Brunton, S.L.Sparse identification of nonlinear dynamics for model predictive control in the low-data limit. arXiv:1711.05501v1 (2017).
Lanczos, C., 2015. The variational principles of mechanics. Dover Publications Inc, New York.
Landau, L.D., Lifshitz, E.M., 2000. Mechanics. Butterworth-Heinemann, Oxford.
Langley, P., Simon, H.A., Bradshaw, G.L., Zytkow, J.M., 1987. Scientific discovery: computational explorations of the creative processes. The MIT Press, Cambridge.
Longair, M., 2003. Theoretical concepts in physics. Cambridge University Press, Cambridge.
Lurie, A.I., 2002. Analytical mechanics. Springer-Verlag, Berlin.
Mangan, N.M., Kutz, J.N., Brunton, S.L., Proctor, J.L., 2017. Model selection for dynamical systems via sparse regression and information criteria. Proc. R. Soc. A 473, 20170009.
Nayfeh, A.H., Mook, D.T., 1979. Nonlinear oscillations. John Wiley, New York.
Novak, S., Frehlich, R.G., 1982. Transition to chaos in the Duffing oscillator. Phys. Rev. A 26, 3660–3663.
Quade, M., Abel, M., Shafi, K., Niven, R.K., Noack, B.R., 2016. Prediction of dynamical systems by symbolic regression. Phys. Rev. E 94, 012214.
Quade, M., Abel, M., Kutz, J.N., Brunton, S.L., 2018. Sparse identification of nonlinear dynamics for rapid model recovery. Chaos 28, 063116.
Rudy, S.H., Brunton, S.L., Proctor, J.L., Kutz, J.N., 2017. Data-driven discovery of partial differential equations. Sci. Adv. 3, e1602614.
Schaeffer, H., McCalla, S.G., 2017. Sparse model selection via integral terms. Phys. Rev. E 96, 023302.
Schaeffer, H., Tran, G. & Ward, R. Learning dynamical systems and bifurcation via group sparsity. arXiv:1709.01558v1 (2017).
Schaeffer, H., Tran, G., Ward, R. & Zhang, L.Extracting structured dynamical systems using sparse optimization with very few samples. arXiv:1805.04158v1 (2018).
Schaeffer, H, 2017. Learning dynamical systems via data discovery and sparse optimization. Proc. R. Soc. A 473, 20160446.
Schmidt, M., Lipson, H., 2009. Distilling free-form natural laws from experimental data. Science 324, 81–85.
van der Pol, B., 1920. A theory of the amplitude of free and forced triode vibrations. Radio Rev. 1, 701–710.
Wang, W.X., Yang, R., Lai, Y.C., Kovanis, V., Grebogi, C., 2011. Predicting catastrophes in nonlinear dynamical systems by compressive sensing. Phys. Rev. Lett. 106, 154101.
Wang, W.X., Lai, Y.C., Grebogi, C., 2016. Data based identification and prediction of nonlinear and complex dynamical systems. Phys. Rep. 644, 1–76.
Whittaker, E.T., 1917. A treatise on the analytical dynamics of particles and rigid bodies with an introduction to the problem of three bodies. Cambridge University Press, London.
Zhang, S., Lin, G., 2018. Robust data-driven discovery of governing physical laws with error bar. Proc. R. Soc. A 474, 20180305.