Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2018 doi:10.20944/preprints201810.0202.v1

Article

Finding Exact Forms on a Thermodynamic Manifold

Chao Ju [1]*, Mark Stalzer [2]

[1] Minerva Schools at KGI; jj@minerva.kgi.edu
[2] California Institute of Technology; stalzer@caltech.edu
* Correspondence: jj@minerva.kgi.edu; Tel.: +44-7542-861706

Abstract: Because only two variables are needed to characterize a simple thermodynamic system in equilibrium, any such system is constrained on a 2D manifold. Of particular interest are the exact 1-forms on the cotangent space of that manifold, since the integral of exact 1-forms is path-independent, a crucial property satisfied by state variables such as internal energy dE and entropy dS. Our prior work[1] shows that given an appropriate *language* of vector calculus, a machine can re-discover the Maxwell equations and the incompressible Navier-Stokes equations from data. In this paper, We enhance this language by including differential forms and show that machines can re-discover the equation for entropy dS given data. Since entropy appears in various fields of science in different guises, a potential extension of this work is to use the machinery developed in this paper to let machines discover the expressions for entropy from data in fields other than classical thermodynamics.

Keywords: thermodynamics; entropy; artificial intelligence; differential geometry; computational physics

---

1. Introduction

We are concerned with developing a *language* that can allow computers to re-discover the expressions for entropy in the setting of classical thermodynamics. The reason we start by considering entropy from classical thermodynamics is that classical thermodynamics provides us with crucial clues of what form our language can take: the fact that entropy appears as a path-independent state variable suggests that the language be differential forms. Forms are natural objects to differentiate and integrate[3], and they fit nicely with our existing framework of vector calculus[1]. The outline of this paper is as follows. In section 2, we briefly explain classical thermodynamics as framed in differential geometry and introduce the concept of the thermodynamic manifold. In section 3, we derive the most general expression for exact 1-forms for a simple thermodynamic system. In section 4, we introduce our existing framework for automated theorem-discovery and show that the expression for entropy can be re-discovered from data. We close the paper by considering the significance of the approach used in this paper and highlight potential extensions of this work.

2. The geometry of thermodynamics

We model the system of interest using an ideal gas of a certain volume. The system is allowed to contract and expand, exchange heat with the surroundings, and do work, assuming the processes are quasi-static. We can of course represent the state of the system on a p-V diagram, but that hides much of the richness of the system. If, instead, we treat the system as a submanifold of $\mathbb{R}^3$, we will discover much structure by using the language of differential forms[2].

© 2018 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2018 doi:10.20944/preprints201810.0202.v1

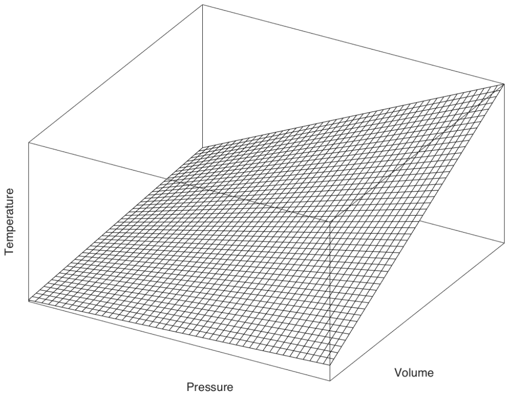

Figure 1. The mesh shows a part of the 2D thermodynamic submanifold embedded in $\mathbb{R}^3$ for an ideal gas $pV = Nk_BT$. The state of the gas is represented by a point on the submanifold.

Let that submanifold occupied by the simple thermodynamic system be denoted by M^2, the superscript indicating that the submanifold is locally $\mathbb{R}^2$. Then, the space of 1-forms at a point on M^2 has dimension

$$\dim \bigwedge^1(\mathbb{R}^2) = \binom{2}{1} = 2.$$

Therefore, on this submanifold, we can expand any 1-form in any basis consisting of the differentials of a set of 2 coordinate functions, dx and dy. For example, in Caratheodory's formulation of thermodynamics[2], the first law reads

$$dE = Q^1 - W^1, \tag{1}$$

where Q^1 is the heat 1-form and W^1 is the work 1-form. The superscript refers to the dimension of the form. Given the knowledge of entropy and work, we can expand the above equation as

$$\begin{aligned} dE &= TdS - pdV \\ &= T(S,V)dS - p(S,V)dV. \end{aligned}$$

The above example shows that indeed, dE as a 1-form can be expanded in a basis $\{dS, dV\}$ using only the variables S and V. In the following analysis, without loss of generality[1] we pick the basis of our 1-forms as $\{dp, dV\}$ with the goal of expressing every 1-form in terms of p and V. In addition to the variables p and V, we have 2 other important "constants", nR and c_v, where the first combination comes from the ideal gas law $pV = nRT$, and c_v is the heat capacity which appears in $dE = c_v dT$.

3. A general expression for exact 1-forms

By definition, the goal is to look for any 1-form f^1 such that $f^1 = dg$, where $g = g(p, V)$ is a 0-form function. Instead of enumerating all possible g and taking the differential to get f^1, we observe that on the thermodynamic manifold, all closed 1-forms are exact. This is because of the De Rham cohomology of this manifold, which has vanishing first Betti number $b_1 = 0$ (i.e. it is retractable to a point). By De Rham's theorem, all closed 1-forms on this manifold are therefore exact[3]. Since all exact 1-forms are automatically closed, to find those exact forms we can simply look for closed

---

[1] Indeed, forms are geometric objects whose properties are coordinate-independent. Exact forms in one basis stay exact in any other basis.

1-forms f^1 such that $df^1 = 0$. As we shall see, the condition $df^1 = 0$ severely constrains the form f can take, and will reduce the enumeration space enormously.

Using the $\{dp, dV\}$ basis, we can express every 1-form f^1 on the thermodyanmic submanifold as

$$f^1 = A(p,V)dp + B(p,V)dV, \tag{2}$$

where A and B are assembled from the symbols of the following set $\mathcal{S}$. Note that consistent with our previous approach[1], we exclude any transcendental functions in the language.

$$\mathcal{S} = \{p, V, nR, c_v, +, -, \times, \div\}.$$

At first sight, enumerating all possible f^1 seems a daunting task, because the enumeration space is too big. However, there are two crucial pieces of information that we can harness to significantly reduce the size of the enumeration space.

First, we demand that the units of the 2 summands in equation (2) must agree. This constraint is a physical one that must be satisfied by any equation. The first consequence of this constraint is that we can leave Nk_B and c_v out of the enumeration space for a while: both of their units contain $1/[Temperature]$, which does not appear in the unit of p or V. Therefore, they must have the same power in A and B to balance out the temperature.

The second consequence of this constraint on unit is that, if we write

$$f^1 = p^\alpha V^\beta dp + p^{\alpha'} V^{\beta'} dV, \tag{3}$$

then by dimensional analysis we will obtain 2 independent linear equations

$$\alpha' = \alpha + 1, \tag{4}$$
$$-\alpha + 3\beta - 1 = -\alpha' + 3\beta' + 3. \tag{5}$$

We have now used up the information of the first constraint.

The second constraint on f^1 is closedness: recall that the goal is to enumerate closed 1-forms only. That said, we want f^1 such that

$$df^1 = 0,$$

which, from (1), is simply

$$\frac{\partial A}{\partial V} = \frac{\partial B}{\partial p}.$$

Therefore, if in (3) we assume that $\beta \neq 0$ and $\alpha' \neq 0$, then by equating the partial derivatives we obtain another linear equation

$$-\alpha + 3\beta - 4 = -\alpha' + 3\beta' + 4, \tag{6}$$

but this leads to a contradiction with (5). Therefore, we must have

$$\beta = \alpha' = 0,$$

and this combined with (5) gives us

$$\alpha = \beta' = -1.$$

Therefore, (3) becomes

$$f^1 = \frac{1}{p}dp + \frac{1}{V}dV,$$

and if we merge the previously left-out Nk_B and c_v into the constants c_1 and c_2, we obtain the final ansatz of our closed 1-form:

$$f^1 = \frac{c_1}{p}dp + \frac{c_2}{V}dV, \tag{7}$$

where c_1 and c_2 are constants of the same dimension. We shall utilize this finding in the next section.

4. Entropy

A valid thermodynamic theorem (equation) must equate n-forms to n-forms. The first law, equation (1), is one such example that equates 1-forms to 1-forms. This section is concerned with finding a thermodynamic theorem governing entropy for a simple system in equilibrium.

In our prior work on Maxwell and Navier-Stokes, we created a program to enumerate "theorems" (instantiated by equations) from a set of symbols, and then validate a certain theorem by using the output of a virtual experiment to see whether the constants in the theorem can be found. To start with, we need to create a finite set consisting of singleton theorems

$$\mathcal{H} = \{A_1, A_2, A_3, ...\},$$

where each singleton theorem A_i is associated with a certain complexity score and is represented by a linear equation

$$c_0 + c_1 A_i = 0,$$

where c_0 and c_1 are constants to be found by the program to test the validity of the theorem. A concrete example for a singleton theorem is when $A_1 = \nabla \cdot \mathbf{B}$, where $\mathbf{B}$ is the magnetic field. Then the first singleton theorem enumerated from the set is

$$c_0 + c_1 \nabla \cdot \mathbf{B} = 0,$$

and electromagnetism tells us that this is a valid theorem for $c_0 = 0$ and $c_1 = 1$.

After we input the singleton theorem set, the program takes another input N, the total complexity score, and efficiently enumerates all candidate theorems whose complexity scores are no more than N[5]. For example, suppose each A_i in the set $\mathcal{H}$ has a complexity score of 1, then theorems of complexity score 2 are of the following form:

$$c_0 + c_1 A_i + c_2 A_j = 0, \quad \forall i \neq j.$$

The program uses a smart way to validate a theorem as soon as it is enumerated by using the output of a virtual experiment. For example, the virtual experiment we used to re-discover the Maxwell equations is the far-field behavior of an oscillating electric dipole with a certain angular frequency and dipole moment[1]. From the output of this virtual experiment the program can validate theorems such as $c_0 + c_1 \nabla \cdot \mathbf{B} = 0$. The method of validating theorems involves the use of applied linear algebra, and the details can be found in [1].

The above is a summary of the essential process of enumerating and validating theorems. In the following, we shall show that an expression for entropy can be found using this process.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2018 doi:10.20944/preprints201810.0202.v1

To start with, we hypothesize that entropy S is an observable of a certain virtual experiment[2], and that its differential dS is a 1-form. Then, using the theoretical results obtained from the previous section, we can form a tentative theorem set

$$\mathcal{T} = \{dS, \frac{1}{V}dV, \frac{1}{p}dp\}.$$

One theorem that is guaranteed to be enumerated from $\mathcal{T}$ is

$$c_0 + c_1 dS + c_2 \frac{1}{p}dp + c_3 \frac{1}{V}dV = 0. \tag{8}$$

To test whether the above equation is a valid theorem or not, we must use the output from a certain virtual experiment and solve a system of linear equations to find the constants. If the constants have a unique nontrivial solution, then we conclude that (8) is a valid theorem. In this application, we shall simply use 1 mole of monatomic gas that can contract and expand as the virtual experiment, whose output for entropy has a simple mathematical expression valid for moderate temperature[4]

$$S = c_v \ln \frac{pV}{R} + R \ln V + a, \tag{9}$$

where a is a constant whose specific value is irrelevant in this application: to set up equations, we want the difference in entropy instead of its absolute value. In general, the virtual experiment can be represented by a trajectory $x(t)$ on the p-V diagram parameterized by t:

$$x(t) = (p(t), V(t)),$$

and the output of the virtual experiment is $S(t) = S(p(t), V(t))$. To validate the theorem, we need to pull back (8) onto the t variable and evaluate the integral

$$\begin{aligned} c_1 \Delta S &= c_2' \int F_t^* \{\frac{1}{p}dp\} + c_3' \int F_t^* \{\frac{1}{V}dV\} \\ &= c_2' \int_{t_1}^{t_2} \frac{1}{p(t)} \frac{dp}{dt} dt + c_3' \int_{t_1}^{t_2} \frac{1}{V(t)} \frac{dV}{dt} dt, \end{aligned}$$

where $\Delta S = S(t_2) - S(t_1)$, F_t^* is the pull-back from the p-V plane to t, and $c_2' = -c_2$, $c_3' = -c_3$. We can then merge c_1 into the other 2 constants to obtain the following equation:

$$\Delta S = c_2' \int_{t_1}^{t_2} \frac{1}{p(t)} \frac{dp}{dt} dt + c_3' \int_{t_1}^{t_2} \frac{1}{V(t)} \frac{dV}{dt} dt. \tag{10}$$

In most applications, the output data of the virtual experiment come in discrete forms:

$$\{p(t_i), V(t_i), S(t_i)\},$$

and we need to numerically integrate (10) and set up equations to find c_2' and c_3' given a trajectory. In the following, we use a simplified trajectory to finalize this example with the goal of showing the essentials while avoiding numerical integrations.

To turn (10) into a set of linear equations, we specify 3 points $A = (p_1, V_1)$, $B = (p_2, V_1)$, $C = (p_2, V_2)$. Starting at point A, we integrate (8) isochorically to point B, and then isobarically to point C.

---

[2] The assumption of entropy as an observable might be a bit far-fetched. However, just as work (which itself is not a direct observable) can be obtained by measuring force and distance, so entropy can be obtained by calculating heat and measuring temperature. The purpose here is to show how the process of finding theorems works.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 October 2018 doi:10.20944/preprints201810.0202.v1

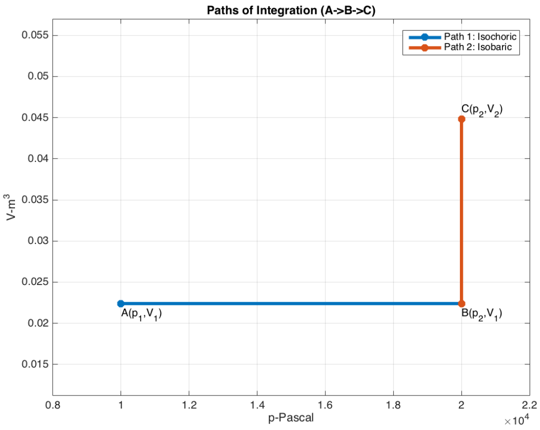

Figure 2. Path of integration from A to B to C.

The 2 equations we obtain are thus

$$S(B) - S(A) = \int_{p_1}^{p_2} \frac{c_2'}{p} dp = c_2' \ln \frac{p_2}{p_1}, \tag{11}$$

$$S(C) - S(B) = \int_{V_1}^{V_2} \frac{c_3'}{V} dV = c_3' \ln \frac{V_2}{V_1}. \tag{12}$$

Let $p_1 = 10000$ Pa, $V_1 = 22.4 \times 10^{-3}$ m^3 (this is the approximate volume of 1 mole of ideal gas at standard room temperature and pressure), and $p_2 = 2p_1$, $V_2 = 2V_1$. The virtual experiment (instantiated by (9)) gives us the following output (given $R = 8.3145$ J/(mol K) and $c_v = 3/2R$):

$$S(B) - S(A) = 8.644758 J/K,$$
$$S(C) - S(B) = 14.407931 J/K.$$

From the above output of the virtual experiments, we can then solve for c_2' and c_3' in equations (11) and (12). They are

$$c_2' = 12.47175 J/K,$$
$$c_3' = 20.78625 J/K.$$

and we conclude that (8) is a valid theorem. In fact, given the knowledge of thermodynamics, we can easily show that $c_2' = c_v$, $c_3' = c_v + R$, and the correct expression for dS for 1 mole of ideal gas is

$$dS = \frac{c_v}{p} dp + \frac{c_v + R}{V} dV + c_0,$$

where c_0 is an additive constant. In this example, we used a simplified approach to illustrate the core idea of constructing tentative theorems from a given set and the use of virtual experiment to determine the validity of a theorem. The complete process can be found in [1].

5. Conclusion

We have shown that we can greatly simplify the problem of enumerating exact 1-forms using one mathematical (closedness) and one physical (dimensional analysis) constraint. In our previous work, we dealt with re-discovering linear differential theories using the language of vector calculus. The above result shows that there is great potential to extend our previous framework to cover differential forms, which will enable us in the future to re-discover scientific theorems that can be geometrically formulated. As an example, Lott and Villani[6] have established a mathematical connection between Ricci curvature, entropy, and optimal transport. But Ricci curvature, $R_{\mu\nu}$, can be thought of as a vector-valued 1-form when the first index is raised by some metric $R^{\mu}_{\ \nu} = g^{\mu\sigma}R_{\sigma\nu}$. In addition, another vector-valued measure of curvature is $\theta^{\mu}_{\ \nu} = \frac{1}{2}R^{\mu}_{\ \nu\rho\sigma}dx^{\rho} \wedge dx^{\sigma}$. Perhaps, given a judicious choice of singleton theorem set and virtual experiment, we could find some curious functional relationship between curvature, entropy (which might also appear as a vector-valued 1-form by covariance), and other physical variables in the transport setting or gradient flow. In re-discovering old theorems, we wish to establish the robustness of this enumeration-validation framework, but the ultimate goal is to apply this framework to find new scientific laws from a wealth of data available that could shed light on scientific discovery.

References

1. Stalzer, M., Ju, C. *TheoSea: Marching Theory to Light*. arXiv:1708.04927 [cs.AI]. This work expanded with an application to Navier-Stokes equations has been submitted to *Artificial Intelligence*.
2. Frankel, T. *The Geometry of Physics: An Introduction, 3rd Edition*. Cambridge University Press, 2011.
3. Warner, F. *Foundations of Differentiable Manifolds and Lie Groups*. Springer Graduate Texts in Mathematics, 1983.
4. Fermi, E. *Thermodynamics*. Dover Publications, 2012.
5. Stalzer, M. *On the Enumeration of Sentences By Compactness*. arXiv:1706.06975 [cs.AI].
6. Lott, J., Villani, C. *Ricci Curvature for Metric-Measure Spaces via Optimal Transport*. arXiv:math/0412127 [math.DG].