Making Sense of the Legendre Transform

R. K. P. Zia, Edward F. Redish, and Susan R. McKay
(Dated: 9/2/2007)

The Legendre Transform (LT) is a common feature of many upper division and graduate physics classes. However, discussions of it tend to be ad hoc, poorly motivated, and confusing. As a result, the LT equations become something to be memorized without understanding. In this paper we describe a more satisfying way of looking at LT relations both mathematically and physically. Mathematically this results in highly symmetric equations that clarify the structure of the transform both algebraically and geometrically. Physically, we motivate the transform as an issue of choosing independent variables that are easily controlled and give examples drawn from classical mechanics and thermodynamics. In thermodynamics, we demonstrate how the LT arising naturally from statistical mechanics and show how use of dimensionless thermodynamic potentials lead to more natural and symmetric relations.

I. INTRODUCTION

The Legendre Transform (LT) is a commonly used mathematical tool in upper division and graduate physics courses, especially in Classical Mechanics (CM) [1], Statistical Mechanics (SM) and thermodynamics. [2]

Most physics majors are exposed to the LT first in CM, where it provides the connection between the Lagrangian $\mathcal{L}(\dot{q})$ and the Hamiltonian $\mathcal{H}(p)$, and then in SM where it creates relations between the internal energy, E and the various thermodynamic potentials, e.g., S, G, H and A.

Despite this, the LT often appears as arbitrary, unmotivated, and ad hoc. It is rarely studied in math courses taken by physics students and it is almost never described either in textbooks or lectures as a valuable general mathematical tool in the way the Fourier transform (FT) is.

In this paper we present a pedagogical introduction to the LT, specifying it as a mathematical process and motivating it in terms related to physical conditions. We discuss some of the symmetries and structures of the transform and consider some of the reasons that might be associated with why it appears difficult in physics. We present a series of increasingly complex examples beginning with CM and going through cases in SM.

This paper is not intended as a guide to pedagogical reform as we have not carried out any research on student difficulties with LT. (There are some relevant studies of student difficulties with the mathematics associated with partial derivatives and differentials in thermodynamics.[3]) It could serve as a starting point for future research and course development on the subject.

II. LT AS AN ALTERNATIVE WAY TO DISPLAY INFORMATION

A. Why does the LT seem obscure?

The way that the LT is presented in physics classes often confounds students rather than enlightens them. There are a number of reasons for this.

- The LT method tends to be poorly motivated.
- The general LT method is not often presented and its mathematical structure and symmetries are rarely displayed.
- The historically motivated choice of units hides the natural mathematical structure of the LT.[4]

A deeper reason why the LT is seen as difficult in physics is that there is a fundamental difference in the way functions are typically treated in mathematics and in many branches of physics including in lower division college physics classes.[5]

In mathematics, a function is seen as a specific rule for transforming a variable (or set of variables) in a domain into a particular result in a range. In most applications that are seen in mathematics courses taken by physics majors, both the domain and the range are simply a set of numbers.

In physics, a function is typically thought of as a relation between physical rather than mathematical quantities. Thus, when we are thinking about physical functions we tend *not* to pay particular attention to the particular functional form the mathematical function uses to encode physical information. For example, if we are describing a position as a function of time, we might write it as $x(t)$. We do not bother to change the symbol x if we decide to give t in milliseconds instead of in seconds. If we write the temperature as a function of position as $T(\vec{r})$, we do not change our symbol if we switch to a different coordinate system or measuring scale. In contrast, the LT is explicitly about how information is coded in the functional form.

In addition, the first introduction a student tends to get to the LT is the transformation in CM from the Lagrangian to the Hamiltonian. This involves the switch from the velocity to the momentum variable in the non-relativistc kinetic energy. In the context of non-relativistic particle motion with velocity independent potentials, the transform involves the kinetic energy, a quadratic function in the velocity – the most trivial function to which the LT can be applied. The result looks

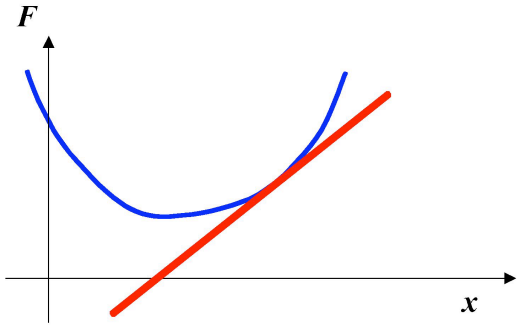

FIG. 1: The graph of a convex function.

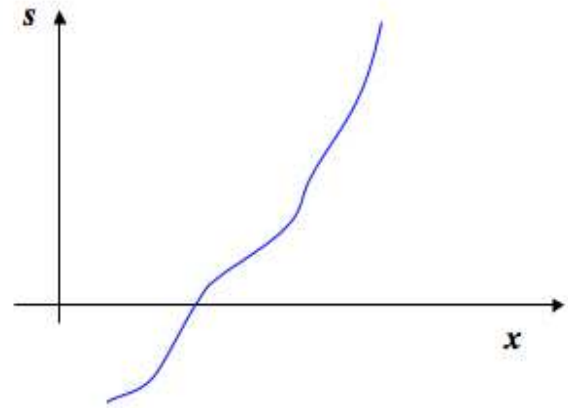

FIG. 2: The graph of the slope of a convex function.

simply like a shift in units (from v to mv as an independent variable) so that it seems pointless. Since the position variable q plays no role in the transform and typically it appears only in V, the result is often regarded as a mysterious change of sign in V ($\mathcal{L} = T - V$ vs. $\mathcal{H} = T + V$).

Typically, the next encounter a student has with the LT is in the context of thermodynamics and SM – both traditionally known as very confusing subjects. Since a general understanding of the LT has not been developed in CM, its presence in thermodynamics and SM only adds to the confusion.

In the rest of this section, we motivate the LT as a general mathematical transformation and provide a transparent description of the method that displays its general properties and symmetries.

B. The point of the LT is to express the information contained in a function in a more convenient way.

Generally, a function expresses a relation between two parameters: an independent variable or control parameter (x) and a dependent value or function (F). This information is encoded in the functional form of $F(x)$. (For clarity, we begin with a single variable, x, and consider multivariate functions later.)

In some circumstances, it is useful to encode the information contained in $F(x)$ in a different way. Two common examples are the Fourier transform and the Laplace transform. These express the function F as sums of (complex or real) exponentials, and display the information in F in terms of the amount of each component contained in the function rather than in terms of the value of the function.

Given a function $F(x)$, the Legendre Transform provides a more convenient way of encoding the information in the function when two conditions are met.

1. The function is strictly convex (second derivative never changes sign or is zero) and is smooth (has "enough" continuous derivatives)
2. It is easier to measure, control, or think about the derivative of F with respect to x than it is to measure or think about x itself.

Because of condition 1, the derivative of $F(x)$ with respect to x can serve as a stand in for x; that is, there is a one-to-one mapping between x and dF/dx. (We discuss how to relax condition 1 later.) The LT shows how to create a function that contains the same information as $F(x)$ but as a function of dF/dx.

C. The mathematics of the LT

We begin our detailed discussion of the LT by considering a single, smooth convex function of a single variable.

There are many equivalent ways to characterize convex functions. The most convenient ones for us are

- The second derivative of our function, $\partial_x^2 F(x)$ is always positive or always negative.

Note that we write ∂_x here to mean d/dx. This is meant to simplify typesetting and to visually enhance certain symmetry relations. It is not meant to stand for a *partial* derivative (although it is useful when the multivariable case is considered). A second characterization of convexity is

- The slope function

$$s(x) \equiv \partial_x F(x) \tag{1}$$

is a strictly monotonic function of x.

A graphical way to see how the value of x and the slope of a convex function can stand in for each other can be seen by considering the example in Fig 1. Here,

the curve drawn to represent F is convex. As we move along the curve to the right (as x increases), the slope of the tangent to the curve continually increases. In other words, if we were to graph the slope as a function of x, it would be a smoothly increasing curve, as shown in Fig 2. If the second derivative $\partial_x^2 F(x)$ exists (everywhere within the range of x in which F is defined; part of the condition for a "smooth" F), there is a unique value of the slope for each value of x, and vice versa. The appropriate mathematical language is that there is a *1-1* relationship between s and x:

$$x \Leftrightarrow s \ ,$$

that the function $s(x)$ is single-valued, and that it can be *inverted* to give a single-valued function

$$x(s) \ . \tag{2}$$

In this way, we can also start with s as the independent variable, use the above to get a unique value of x, and then insert that into $F(x)$ to access F as a function of s. The standard notation for such a function is $F(x(s))$.

If we insist on a new encoding of the information in F (in terms of s instead of x), this would appear to be the most "natural" way, and the student would not be subjected to the confusion that arises with the LT. Instead, the LT of $F(x)$ is defined as a different function of s, namely,[6]

$$G(s) = s \cdot x(s) - F(x(s)) \, . \tag{3}$$

Typically, this formula is presented with little motivation or explanation, and leaves the students to ponder: Why? Why the extra $s \cdot x$? Why the minus sign? Frequently, the instructor may invoke another magical relation to answer such queries. Only with this peculiar defintion can we have the property that "the slope of $G(s)$ is just x":

$$x(s) = \frac{dG}{ds} \equiv \partial_s G(s) \, . \tag{4}$$

Of course, this result also requires a careful computation.

Before providing ways to appreciate this definition of the LT, as well as how never to forget "which sign goes where" we present the graphic route to the transform. Consider a plot of F vs. x in Fig. 3. Choose a value of x, which is represented by the length of the horizontal line (black on-line) labeled by x. Go up to the value on the function curve, $F(x)$. This value corresponds to the length of the vertical line labeled by F (blue on-line). Next, draw the tangent (red on-line) to the curve at that point. The slope here is labeled s, as emphasized by the call out bubble. Extend this tangent until it hits the ordinate (the "F axis"). In this example, the intercept is negative and can be labeled as $-G$, with a positive G. This value corresponds to the length of thick vertical line (green on-line) labeled by G. This length is reproduced (thin, green on-line) just *below* the line labeled F.

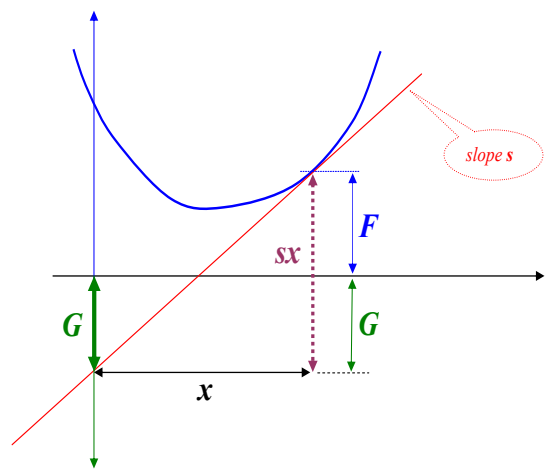

FIG. 3: The graph of the slope of a convex function.

Meanwhile, due to the slope of the tangent being s, we draw a dotted line, the length of which is sx (purple on-line). From this picture, it is clear that $sx = F + G$. In this light, the peculiar definition of the LT, Eqn. (3), appears appealingly "natural". Starting from this *symmetric* form, we can easily derive (and appreciate) the bewildering array of formulae associated with the LT.

To emphasize, the LT G carries the *same* information (as F) on a given system, only that the information is *encoded differently* - in s rather than x. We next explore some of the consequences of the LT.

D. Some mathematical properties of the LT

The geometrical construction and the resulting relations allow us to display a number of useful and elegant relations that shed light on the workings of the LT. In particular, we consider

- The inverse LT
- Extreme values
- Symmetries and derivative relations.

1. *The inverse LT*

Ordinarily, the inverse of a transformation is distinct from the transform itself. The LT distinguishes itself in that it is its own inverse. In this sense, it resembles (geometric) duality transformations. If we perform the LT a second time, we recover the original (convex smooth) function. In other words, suppose we start with the function, $G(s)$, and ask what is its LT. To be able to do this, $G(s)$ would have to satisfy our conditions: Is it convex and is it smooth? In the section on derivatives below, we see that the answer is yes.

Let us start with

$$y(s) = \frac{dG}{ds} \tag{5}$$

and invert the monotonic function $y(s)$ to $s(y)$. Next, we construct

$$H(y) = y \cdot s(y) - G(s(y)) \ , \tag{6}$$

which can be rewritten as

$$G = s \cdot y - H \ .$$

Comparing this and Eqn. (5) to Eqns. (3,4), we see that we can identify $\{H, y\}$ with $\{F, x\}$. Thus, the LT of G, which we have called H, is just F. In other words, the LT of an LT is our original function and the LT is its own inverse. This "duality" of the LT is best summarized by the symmetric form of the LT,

$$G(s) + F(x) = s \cdot x, \tag{7}$$

which displays the symmetry between $\{G, s\}$ and $\{F, x\}$ explicitly. We should emphasize that this equation should be read carefully. In particular, despite its appearance, there is only one independent variable: either s or x. These two are related to each other, through either $x(s) = \partial_s G(s)$ or $s(x) \equiv \partial_x F(x)$. Thus, a careful writing of this equation would read either $G(s) + F(x(s)) = s \cdot x(s)$ or $G(s(x)) + F(x) = s(x) \cdot x$. Starting from the first of these, say, and exploiting the chain rule, we immediately arrive at the "magical" Eqn. (4).

2. *Extreme values*

Suppose the function $F(x)$ is concave upward (such as the example in Fig. 3). It then cannot have a maximum value but it may have a minimum. Assuming it does, then the minimum is unique. Let us denote this point by

$$F_{\min} = F(x_{min}) \ .$$

Of course, the slope of the tangent vanishes here, *i.e.*, $s(x_{min}) = 0$. If we put this point into our expression defining the LT (eq. 3), we get the minimum value of F, is

$$F_{\min} = -G(0) \, . \tag{8}$$

Similarly, from the fact that F is the LT of G, we can conclude that the minimum value of G, is

$$G_{\min} = -F(0) \, . \tag{9}$$

We can use the symmetric relation eq. 7 to see what happens for general extrema. Suppose F takes on its extremal value at x_{ext}, which corresponds to a horizontal tangent, $s = 0$. We see that

$$G(0) + F(x_{ext}) = 0.$$

Similarly, G is be at its extremum at s_{ext}, where $x(s_{ext}) = 0$ due to (4), so that

$$G(s_{ext}) + F(0) = 0.$$

To appreciate the geometric meaning of this equation, we only need to inspect Fig. 3 and see that the y-intercept of the tangent to the curve $F(x)$ never reaches beyond $F(0)$.

3. *Symmetry and derivative relationships*

Since F and G are LTs of each other, we expect that a lot of highly symmetrical relations should occur. We have already seen two:

$$G(s) + F(x) = s \cdot x$$

and

$$\partial_s G = x \quad \text{and} \quad \partial_x F = s \tag{10}$$

From these, we can obtain an infinite set of relationships between G and F, by regarding the latter as an equation with s or x as the (independent) variable and taking derivatives. This leads to some very elegant and interesting relationships.

In this notation, since each function only depends on one variable, the differentials can be easily identified, from e.g., $dG = (\partial_s G)\, ds$:

$$dG = x ds \quad \text{and} \quad dF = s dx.$$

Next, we may differentiate expressions (10) again, with respect to s or x, as appropriate. The result is

$$\partial_s^2 G = \frac{dx}{ds} \quad \text{and} \quad \partial_x^2 F = \frac{ds}{dx}$$

But, $dx/ds = (ds/dx)^{-1}$, so we may write

$$\left(\partial_s^2 G\right)\left(\partial_x^2 F\right) = 1 \ . \tag{11}$$

This equation clearly illustrates the importance of (strict) convexity so that neither factor ever vanishes.

An interesting result is that the local curvatures of the LT's are inverses of each other - in a manner reminiscent of the uncertainty relation $\Delta x \Delta k \approx 1$. For simplicity, suppose F is dimensionless but x is not,[8] so that s has the dimension of $1/x$. With this convention, it is easy to check the units of Eqns. (7,10,11).

Differentiating (11) again, we can write a symmetric relationship for the third derivative:

$$\frac{\partial_s^3 G}{\left[\partial_s^2 G\right]^{3/2}} + \frac{\partial_x^3 F}{\left[\partial_x^2 F\right]^{3/2}} = 0 \ . \tag{12}$$

Notice that $\left[\partial_s^3\right] = \left[\partial_s^2\right]^{3/2}$ so that each term is dimensionless.

It is possible to arrive at an infinite set of such relations for higher derivatives, by differentiating further. Such an exercise also shows that, if F is "smooth" (with say, a well defined n^{th} derivative), then so is G. Finally, we should remark that the relationships for higher derivatives do not have forms as simple as Eqns. (7,10,11,12), but get more and more complex.

III. SIMPLE EXAMPLES OF THE LT IN SINGLE-PARTICLE MECHANICS

It is useful to provide some physical examples to illustrate these relations.

A. The quadratic function: Non-relativistic kinetic energy

The simplest example is a quadratic function $F(x) = \frac{1}{2}\alpha x^2$. For this function, we easily find that $s = \alpha x$ and $x = s/\alpha$, leading to $G(s) = \frac{1}{2\alpha}s^2$. We see immediately that the curvatures in F and G are inverses of each other as required by eq. 11. They are simply α and $\frac{1}{\alpha}$. All derivative relations beyond this level are trivial: $0 = 0$.

Indeed, the reader will immediately recognize that this example is the case in CM of a single non-relativistic particle with mass m moving in an external potential $V(q)$. There,

$$x \to \dot{q}, \quad F \to \mathcal{L}, \quad \alpha \to m, \quad s \to p, \quad G \to \mathcal{H} \tag{13}$$

Note that the potential is just a "spectator" in the LT, since the variable q is not involved. As a result, it flips its sign when it is put into F ($\mathcal{L}$) and one solves for G ($\mathcal{H}$) in the symmetric LT relation: $F + G = sx$.

B. Relativistic kinetic energy

A more interesting case is when we have relativistic kinetic energy. Here, we go the other way and start with momentum and generate a velocity as the slope of the function. The relativistic kinetic energy as a function of momentum is

$$\mathcal{H}(p) = \sqrt{p^2 + m^2} \tag{14}$$

(with $c \equiv 1$). This is a convex function and its slope at a point p is

$$\beta = \frac{d\mathcal{H}}{dp} = \frac{p}{\sqrt{p^2 + m^2}} \tag{15}$$

giving the familiar result

$$p = m\beta \Big/ \sqrt{1 - \beta^2} \ . \tag{16}$$

This leads to the Lagrangian[9]

$$\mathcal{L}(\beta) = p\beta - \mathcal{H}(p(\beta)) = -m\sqrt{1 - \beta^2} \tag{17}$$

This example can also be written in terms of the function $F(x) = cosh(\lambda x)$. The demonstration is left to the reader. (Hint: Consult Taylor and Wheeler [10].)

C. Mechanical equilibrium with an externally applied force

Let us turn to a less familiar example, one that is so trivial that it does not appear in typical textbooks. Yet it sets the stage for examining the role of the LT in equilibrium SM. Consider a particle in a (one-dimensional) convex potential well, $U(x)$, which has a unique minimum at $x_{\min}$. A good example would be the particle attached to a wall by a physical spring, with x being the distance from the point where the coils of the spring are fully compressed. The potential would be effectively infinite at $x = 0$, drop down to a minimum at its natural extension, and then rise for larger x. (We restrict our attention to positive values of x less than the breaking point of the spring.) Another example of U is the potential which binds two atoms into a molecule (though such U's are rarely convex for all separations).

Needless to say, the particle is stationary ("in equilibrium") only if it is at $x_{\min}$ for all time. If it is subjected to an additional external applied force, f, then it will reach a new stationary point, x_0, which is the solution to the equation

$$\left.\frac{dU}{dx}\right|_{x_0} = f \ . \tag{18}$$

To emphasize the dependence of this point on f, we write $x_0(f)$. We can ask the inverse question: If we want the particle to settle at $x_1 \neq x_{\min}$, what force do we need to apply? The answer is $f(x_1)$, a force that depends on which x we choose. A little thought leads us to the explicit functional form: $f(x_1) = dU/dx|_{x_1}$. Of course, there is nothing special about the subscripts here and we may just as well write

$$f(x) = \frac{dU}{dx} \tag{19}$$

and, for the inverse dependence, $x(f)$ instead of $x_0(f)$.

While Eqn. (19) gives $f(x)$ explicitly, we may ask if there is a counterpart which provides the inverse, $x(f)$, explicitly. If so, we can simply "plug f into" the expression and arrive at the new equilibrium position. The answer is the LT of U, namely,

$$V(f) = f \cdot x - U(x(f)) \ .$$

We leave it to the reader to show that

$$x(f) = \frac{dV}{df} \tag{20}$$

is the companion to Eqn. (19).

All the details can be simply worked out in the simple example of the mass on a spring, $U(x) = kx^2/2$. This is the analog of the non-relativist kinetic energy LT discussed in section A. The reader may easily demonstrate that the LT equation $U+V = fx$ becomes $(f-kx)^2 = 0$, yielding the relation between f and the new equilibrium point x.

Note that the information about our system (e.g., wall-spring-particle complex) is fully contained in either U or V. the only difference is in the coding. While U is the usual potential energy associated with putting the particle at x, V is a kind of potential, associated with the control f. In ordinary CM, such an approach seems unnecessarily cumbersome for describing the simple problems we posed. Thus, it is rightfully ignored in a course on CM. We include the example here only as a stepping stone to the LT in SM and thermodynamics. There, multiple potentials are essential ingredients and appear in courses regularly. The LT relations proliferate and in increasingly mystifying forms. The next section is devoted to displaying them in more transparent forms, as well as to showing the intimate connection between the LT and the Fourier-Laplace transform.

IV. THE LT IN STATISTICAL THERMODYNAMICS

In statistical thermodynamics, the LT appears frequently, when different variables are "traded" for their LT conjugates.[11] Often, one of the variables is easy to think about while the other is easy to control in real physical situations.

The difficulty with making sense of the LT in thermodynamics arises from two causes.

- For historic reasons, LT variables are not always chosen as conjugate pairs.
- Many variables appearing in equilibrium thermodynamics are not independent, e.g., P, V, T, N, E,etc. Some are constrained by equations of state, e.g., $PV = Nk_BT$.

As an example of the first point, the conjugate to the total energy (E) of a system is the *inverse* temperature ($\beta = 1/k_BT$). Yet, our daily experience with temperature (T) is so pervasive that T tends to be used in most of the formulae. Thus, the familiar equation

$$A = E - TS \tag{21}$$

which relates the Helmholtz free energy (A) to the entropy (S), obscures the symmetry between β and E , as well as the dimensionless nature of the LT. By contrast, if we define *dimensionless* quantities

$$\mathcal{S} \equiv S/k_B \quad \text{and} \quad \mathcal{A} \equiv \beta A \tag{22}$$

the duality (*symmetry*) between them can be beautifully expressed as

$$\mathcal{A}(\beta) + \mathcal{S}(E) = \beta \cdot E \tag{23}$$

To elaborate the second point, we typically encounter a bewildering array of thermodynamic functions (entropy, Gibbs and Helmholtz free energies, enthalpy, etc.), a slew of variables (energy, temperature, volume, pressure, etc.), as well as a jumble of thermodynamic relations (with multiple partial derivatives). In general, because of the multiple constrained independent variables, none of these examples is as simple as those above, adding to the difficulty of both teaching and learning this material.

Before turning to a discussion of the generation of the standard potentials, we show how the LT enters thermodynamics through the door of statistical mechanics, with the aid of the Laplace transform and the thermodynamic limit.

A. The foundations of statistical thermodynamics

Equilibrium statistical mechanics is founded on a bold hypothesis.[11]

> *For an isolated system, every allowed microstate (specified by the assignment of explicit microscopic variables for each particle leading to a given total energy) is equally probably. The high probability of finding a particular equilibrium macrostate(specified by the assignment of specific macroscopic variables) is due to a predominance of the number of microstates corresponding to that macrostate.*

This way of describing a closed, isolated system is known as the microcanonical ensemble.

The classic example is a gas of N identical, free,[12] non-relativistic structureless particles, confined in a D-dimensional box of volume L^D. For that system, a microstate is specified by the $2DN$ variables corresponding to the positions and momenta of each particle: $\{\vec{x}_i, \vec{p}_i\}$, with $i = 1, ..., N$. Since the total energy (E) is a constant for an isolated system, the fundamental hypothesis can be represented as

$$P(\{\vec{x}_i, \vec{p}_i\}) \propto \delta(E - \mathcal{H}(\{\vec{x}_i, \vec{p}_i\})) \tag{24}$$

where $P(\{\vec{x}_i, \vec{p}_i\})$ is the probability of finding the configuration of positions and momenta $\{\vec{x}_i, \vec{p}_i\}$ and $\mathcal{H}$ is the Hamiltonian. In this case, $\mathcal{H}$ is explicitly given by

$$\mathcal{H} = \sum_i h(\vec{x}_i, \vec{p}_i) = \sum_i \left[\frac{\vec{p}_i^2}{2m} + U(\vec{x}_i)\right] \tag{25}$$

where m is the mass of each particle and U is the confining potential, which is zero for each component of

$x \in [0, L]$ and infinite otherwise. (The interactions between the particles can be ignored, by definition of "the ideal gas.")

The normalization factor for P is therefore

$$\Omega(E) = \int_{x,p} \delta\left(E - \mathcal{H}\left(\{\vec{x}_i, \vec{p}_i\}\right)\right) \tag{26}$$

where the integral is taken over *all* $\{\vec{x}_i, \vec{p}_i\}$ from $-\infty$ to ∞. (The infinite values of U restrict the actual position integrations to the volume of the box.) We have also suppressed the other variables that Ω depends on for now: L and m.

From here, the standard approach evaluates this integral as follows. The position integral can be done explicitly, since the only dependence of the Hamiltonian on position simply confines the position integrals to the allowed volume. This yields a factor of L^{ND}. The momentum integrals are done by computing the "surface area" of a sphere in DN dimensions.

The entropy is introduced through the definition $S \equiv k_B \ln \Omega$. If exploit our "dimensionless entropy," $\mathcal{S}$, we simply write

$$\mathcal{S}(E) \equiv \ln \Omega(E). \tag{27}$$

To proceed, we have two choices: the route that emphasizes the mathematics or the physics.

1. *The route of mathematics*

To take this route, our task is straightforward: to evaluate integrals with a constraint like Eqn. (26). Needless to say, such integrals are often not easy to perform. However, exploiting the Laplace transform often renders the integrand factorizable. For example, the DN integrations in Eqn. (26) becomes just products of a single integral. Specifically, we consider the Laplace transform of $\Omega(E)$,

$$Z(\beta) \equiv \int \Omega(E)\, e^{-\beta E} dE. \tag{28}$$

Putting in (26) for $\Omega(E)$, the delta function permits us to do the E integral giving

$$Z(\beta) = \int_{x,p} e^{-\beta \mathcal{H}}. \tag{29}$$

Since $\mathcal{H}$ is a sum over the individual components, the integrand factorizes and we have the result:

$$\int_{x,p} e^{-\beta \mathcal{H}} = \int_{x,p} \prod_i e^{-\beta h(\vec{x}_i, \vec{p}_i)} \tag{30}$$

$$= \left[\int_{x,p} d\vec{x} d\vec{p} e^{-\beta h(\vec{x}, \vec{p})}\right]^N. \tag{31}$$

The remaining task is to perform an inverse Laplace transform, i.e.,

$$\Omega(E) = \int_{\mathcal{C}} Z(\beta)\, e^{\beta E} d\beta$$

where $\mathcal{C}$ is a contour in the complex β plane (running parallel to and to the right of the imaginery axis). Defining

$$\mathcal{A}(\beta) \equiv -\ln Z(\beta), \tag{32}$$

this integral can be cast as

$$e^{\mathcal{S}(E)} = \int_{\mathcal{C}} e^{-\mathcal{A}(\beta) + \beta E} d\beta.$$

To continue further, it is necessary to inject some physics. In this case, we expect to be dealing with many particles, i.e., large N. From (31), we have $\mathcal{A} \propto N$, leading us to expect that the range of E we would be interested in is also $O(N)$ (both E and $\mathcal{A}$ are "extensive"). The standard tool to deal with integrands with large exponentials is the saddle point (or steepest decent) method. Thus, we seek the saddle point in β, defined by setting the first derivative of $\beta E - \mathcal{A}(\beta)$ to zero:

$$\partial_\beta \left[\beta E - \mathcal{A}\right]\big|_{\beta_0} = 0\ . \tag{33}$$

In other words, we have

$$\left.\frac{\partial \mathcal{A}}{\partial \beta}\right|_{\beta_0} = E\ . \tag{34}$$

Let us emphasize that β_0 should be regarded as a function of E here.

Now, in this approach, the integral in equation (28) is well approximated by *evaluting the integrand* at the saddle point, so that

$$\Omega(E) \cong \exp\left[\beta_0 E - \mathcal{A}(\beta_0)\right] \tag{35}$$

or using Eqn. (27)

$$\mathcal{S}(E) + \mathcal{A}(\beta_0) = \beta_0 E \tag{36}$$

with the understanding that β_0 and E are related through (34). There is nothing significant about the subscript on β and this equation is identical to (23) above. In other words, $\mathcal{S}$ and $\mathcal{A}$ are LT's of each other. Thus, we see that (for situations involving a large parameter, N in this case) the Laplace (28) and Legendre (36) transforms, Eqs. (28, 36) respectively, are intimately related to each other as a result of the thermodynamic limit.

2. *The route of physics: interpretation of the equilibrium condition*

The great discovery of thermal physics was the realization that matter, by virtue of its temperature, contained

huge stores of energy.[13] Unfortunately, that energy is chaotic and non-directional, in contrast to the energies we are used to in classical mechanics. The critical question for extracting usable forms of these hidden energies is: Under what conditions does the internal energy of matter move from one object to another and under what conditions can it be changed to work? Part of the answer lies in understanding which way the energy will flow if we bring two very different systems into "thermal contact," i.e., allowing only energy to be transfered between the them. Does it go from the one with more energy to the one with less? No, not always: Surely, the ocean contains more energy than a teaspoon of hot water; yet the latter *loses* energy when brought into contact with the ocean! Of course, every child knows the phenomenon: Energy flows from the hot system to the cold system, *until* both have the same *temperature*. Can this phenomenon be understood in the framework of the bold hypothesis above?

Now, it is not our purpose here to recreate an entire statistical thermodynamics text in this brief article. But considering this question leads us back to the LT.

When two systems (not necessarily of the same size or energy) are in contact and the combined system isolated, we obviously have

$$E_{tot} \equiv E_1 + E_2$$

as *the* control parameter, set to whatever is the "initial condition" (i.e., when the two are first brought into contact). Meanwhile, the individual E_j's are *not* fixed, and we should ask the physics question: Starting at some initial values, how do they wind up at the final "equilibrium partition" $\{E_1^*, E_2^*\}$? The answer lies with $\mathcal{S}_{tot}\left(E_{tot}|E_1, E_2\right)$, the entropy of the combined system, *subjected* to the partition of energies $\{E_1, E_2\}$, since it carries the information of how probable a particular partition will be. In general, computing this would be no trivial task. However, if we focus on systems with extensive entropies, then we may write a good approximation:

$$\mathcal{S}_{tot} = \mathcal{S}_1 + \mathcal{S}_2$$

as well as

$$\mathcal{S}_1 = \mathcal{S}_1\left(E_1\right); \quad \mathcal{S}_2 = \mathcal{S}_2\left(E_2\right).$$

These are not trivial statements: We are injecting the physics that, under the conditions specified, the entropies of each system do *not* depend on the energy of the other. (Or at the least, we assume that whatever the dependence is on the interaction between the two systems, it is so small that it can be neglected.)

Under these assumptions, we can ask: For what partition will $\mathcal{S}_{tot}$ be maximum, or equivalently, which partition is the most probable? Writing $E_2 = E_{tot} - E_1$ and recalling that E_{tot} is fixed, this task is easy. The maximum occurs at E_1^*, where

$$\left.\frac{d\mathcal{S}_{tot}}{dE_1}\right|_{E_1^*} = 0$$

or

$$\left.\frac{d\mathcal{S}_1}{dE_1}\right|_{E_1^*} = \left.\frac{d\mathcal{S}_2}{dE_2}\right|_{E_2^*} \tag{37}$$

since $dE_1 = -dE_2$. This result is significant, when we recognize that each side does not depend on the parameters of the other system. Thus, if we associate a quantity with $d\mathcal{S}/dE$, which we define by

$$\beta\left(E\right) \equiv \frac{d\mathcal{S}}{dE},$$

then Eqn (37) becomes

$$\beta_1\left(E_1^*\right) = \beta_2\left(E_2^*\right) \ .$$

To put it in words: The most probable partition is when the β of one system equals the β of the other. How β is related to our common experience of *temperature* still remains to be clarified; but we have found a condition that can *predict* the long-time behavior of two systems when they are brought into thermal contact (i.e., allowed to exchange energy).

How do we exploit this new variable β? For any given system, we can write $\mathcal{S}\left(E\left(\beta\right)\right)$, of course; but is that useful? The answer is intimately connected to the canonical ensemble, the (Helmholtz) free energy, and the LT of $\mathcal{S}$. There is no need for us to reproduce here the standard derivation of this ensemble and the Boltzmann factor ($e^{-\beta\mathcal{H}}$) which controls the probability of finding the system in a microstate associated with Hamiltonian $\mathcal{H}$. Let us quote the result: The normalization factor here, $Z\left(\beta\right)$, is the counterpart of $\Omega\left(E\right)$ for the mircocanonical ensemble and the associated potential is just the LT of $\mathcal{S}$: $\mathcal{A}\left(\beta\right) = \beta E - \mathcal{S}\left(E\left(\beta\right)\right)$. These two thermodynamic potentials ($\mathcal{S}\left(E\right)$ and $\mathcal{A}\left(\beta\right)$) are entirely analogous to $U\left(x\right)$ and $V\left(f\right)$ in the case for a particle in a classical potential, discussed in section III.C.

B. How does the LT enter into thermodynamics?

For the convenience of the reader, let us summarize the key relations using dimensionless potentials:

$$\Omega\left(E\right) = e^{\mathcal{S}(E)}; \quad Z\left(\beta\right) = e^{-\mathcal{A}(\beta)} \tag{38}$$

$$\frac{d\mathcal{S}}{dE} = \beta; \quad \frac{d\mathcal{A}}{d\beta} = E \tag{39}$$

$$\mathcal{S}\left(E\right) + \mathcal{A}\left(\beta\right) = \beta E \ . \tag{40}$$

We can now see exactly where the LT comes in and why it is useful. The function that determines the physical condition is the entropy, $\mathcal{S}$, as a function of the energy, E. The energy is not easy to control, but the derivative

of the entropy with respect to the energy (β – essentially the temperature) is more naturally and directly controlled. We therefore perform a LT of $\mathcal{S}$, to get the same information coded in the Helmholtz free energy, $\mathcal{A}$, which is a function of β.

We should also emphasize that, in this context, the dependent variable in a thermodynamic potential should be regarded as a control parameter – or as a constraint variable.

The "slope" associated with each parameter carries significant physical information, namely, the *response* of the system to this *control*. The LT simply exchanges the role of such a pair of variables. In the case considered here, it is more familiar to think of temperature (or more precisely, β) as the control (i.e., a thermostat) and the internal energy of the system as a response. Thus, the free energy $\mathcal{A}(\beta)$ is the more intuitive potential, with $E = \partial_\beta \mathcal{A}$ being the response. In the transformed version, E is the control with the entropy $\mathcal{S}(E)$ as the potential, a setup designed for an *isolated* system. Finally, let us point out that there are many other examples of response/control pairs to which the same kind of transformation may be applied, such as particle number and chemical potential, polarizability and electric field, or magnetization and magnetic field.

Of course, in statistical physics, thermodynamic potentials depend on many variables other than just the total energy E. Each of them should be regarded as a control parameter or a constraint. Each elicits a distinct response. As we construct LT's for each of these control/response variable pairs, we generate a new potential. As a result, is a plethora of thermodynamic functions!

To help our students, we should emphasize to that all of the thermodynamic potentials carry the *same* information, but *encoded in different ways*, each of which is for our convenience as investigators of the properties of system. The next section is devoted to LT of functions with many variables.

V. LT WITH MANY VARIABLES

LTs can be applied to one or more of the variables of a multivariable function as well as to single variable functions. This is particularly important in thermodynamics where there are many linked variables. We begin this section by discussing briefly the mathematical structure of the multivariable LT and then apply it to the case of thermodynamics and SM.

A. General considerations

Suppose our function is a multivariate one:

$$F(x_1, ..., x_M). \tag{41}$$

Then, there will be M "slopes"

$$s_m = \frac{\partial F}{\partial x_m} \equiv \partial_m F \tag{42}$$

and $M(M+1)/2$ second derivatives

$$\partial_m \partial_l F \tag{43}$$

which can be regarded as a *symmetric* matrix. The convexity restriction amounts to demands that all of the eigenvalues are *positive* (or negative). A standard mathematical corollary is that the relationship between $\{x_m\}$ and $\{s_m\}$ is 1-1, so that we can "replace" any x_m by the corresponding s_m via a LT. We should remind the reader that, in the context of thermodynamics, convexity *is* the condition for stability in equilibrium systems. Put succinctly: No "eigen" response can be negative.[14]

Since we can transform any number of (up to M) the x's, we will have a total of 2^M functions! For example, if we restrict ourselves of (E, V) - the standard variables for the microcanonical ensemble of the ideal gas - there are *four* thermodynamic functions: entropy, enthalpy, Gibbs and Helmholtz free energies.

One way to picture the relationship between so many functions is to put them at the corners of an M-dimensional hypercube. Each axis in this space is associated, of course, with a particular variable pair (x_m, s_m). Going from one corner to an adjacent corner along a particular edge correspond to carrying out the LT for that associated pair. For the $M = 2$ example above, the hypercube reduces to a square, which is related, but not identical, to the square that appears in some standard texts.[11][15] Thanks to the commutativity of partial derivatives, to go from any corner to any other is a path independent process, so that the function associated with each vertex is unique. Thus, e.g., exchanging (x_ℓ, x_m) for (s_ℓ, s_m), the LT relations would be the simple generalization of (7)

$$\begin{aligned} s_\ell x_\ell + s_m x_m &= G(x_1, ...s_\ell, ...s_m, ...x_M) \\ &\quad + F(x_1, ...x_\ell, ...x_m, ...x_M) \end{aligned} \tag{44}$$

with [16]

$$\partial_\ell G = x_\ell; \quad \partial_m G = x_m \tag{45}$$

$$\partial_\ell F = s_\ell; \quad \partial_m F = s_m\ . \tag{46}$$

To be clear, we should have given this G some special notation, to denote that its variables are all $\{x\}$ except for the two that are s's. One possibility is $G^{\ell,m}$, but for the sake of simplicity, we do not pursue this issue further. However, one special LT is worth a note - the one in which *all* variables are s's. Denoting this function by H, we see that it lies at the corner of the hypercube diametrically opposed to F. For these, the LT relation reads

$$H(\vec{s}) + F(\vec{x}) = \vec{s} \cdot \vec{x} \tag{47}$$

where the vector notation for the variables should be obvious.

Generalizations for higher derivatives proceed in a similar way. For example, (11) becomes

$$\sum_m \left(\partial_\ell \partial_m H\right)\left(\partial_m \partial_n F\right) = \delta_{\ell n}$$

where δ is the unit matrix. Clearly, convexity of F guarantees that the inverse of $\partial_m \partial_n F$ exists.

B. Example in the context of thermodynamics

Let us provide an example by applying these considerations to thermodynamics of a gas. Beginning with the microcanonical partition function $\Omega\left(E,V\right)$. Thus, we consider the "mapping"

$$F\left(x_1,x_2\right) \to \mathcal{S}\left(E,V\right) \equiv \ln \Omega, \tag{48}$$

with

$$x_1 \to E, \quad x_2 \to V, \quad s_1 \to \beta, \quad s_2 \to \tilde{p}$$

Here, we choose the variable to be

$$\tilde{p} \equiv \beta P.$$

This choice comes from the traditional definition of pressure: $P = T\left(\partial S/\partial V\right)$ with E held fixed.

Carrying out the LT with respect to x_1 leads to the Helmholtz free energy. Our symmetric and dimensionless version of $A = E - TS$ is (eqn. 23 above)

$$\mathcal{A}\left(\beta,V\right) + \mathcal{S}\left(E,V\right) = \beta E.$$

Taking the LT with respect to x_2 as well, we arrive at the Gibbs free energy: $G\left(T,P\right)$ (not be confused with the G's, the generic LT of the F's). Our version is

$$\mathcal{G}\left(\beta,\tilde{p}\right) + \mathcal{S}\left(\beta,V\right) = \beta E + \tilde{p}V \tag{49}$$

where

$$\mathcal{G} \equiv \beta G$$

Rewriting (49), we see its more common manifestation:

$$G = E - TS + PV.$$

From our perspective, it is the placing of S and the use of T (instead of β) that leads to the seemingly mysterious signs on the right.

To end this subsection, we should comment on the enthalpy, which commonly appears alongside the other three potentials. For various reasons (history and/or convenience), S instead of E is chosen as the independent variable. As a result, instead of β, the natural conjugate variable is T $(= \partial E/\partial S)$. Regarding S as a control variable with which to access E is clearly conceptually difficult. However, it is common to think of transfering *heat* out-of or into a system, so that TdS appears on the scene as the means of "control." Taking the LT of $E\left(S\right)$ in the standard fashion, we would arrive at $TS - E$, which is the Helmholtz free energy, *except* for a sign. The disadvantage is clear, but there are advantages to this approach. In particular, by starting with $E\left(S,V\right)$, we naturally arrive at the ordinary pressure, $-P$, as the conjugate to V (instead of $\tilde{p}$). Note the extra minus sign here! Taking the LT with respect to V from $E\left(S,V\right)$, we arrive at $\left(-P\right)V - E$. This is the enthalpy: $H = E + PV$, but again with an extra minus sign.

Needless to say, we may continue on to add other variables, the standard example being N and the grand canonical ensemble.

VI. SUMMARY AND CONCLUDING REMARKS

In this article we have presented an analysis of the LT from both mathematical and physical points of view. Mathematically, we analyze how the LT arises from the goal of conveniently expressing the information contained in a convex function in terms of its derivative as an independent variable and relate the familiar algebraic expressions to less familiar geometry. We point out an elegant (and easy to remember) symmetry relation that allows the direct reconstruction of all of the critical results and, as a bonus, makes it obvious that the LT is its own inverse. Physically, we have motivated the LT and as facilitating a way of *thinking about* a complex problem in terms of a more naturally controllable independent variable. We present examples from classical and statistical mechanics. Our dimensionless form for the LT motivates and helps untangle the snarl of thermodynamic potentials and the confusing relations among them.

There remain two issues that we promised to discuss. For simplicity, we focused on the LT of smooth convex functions. Also due to space limitions, we did not touch upon many generalizations of the LT. Let us conclude by acknowleding just two important issues: the LT of non-convex functions and LT in domains with non-trivial topology, such as the angle on a circle.

When a function is non-convex the LT becomes multivalued or develops discontinuous first derivatives. The topic is intimately related to first order phase transitions and the Maxwell construction.

A second generalization concerns variables the domains of which have non-trivial topology, the simplest being functions defined on a circle or the surface of a sphere. The angles are the most natural variables here, but we must be mindful of the periodic nature of $\phi \in (0, 2\pi]$ and the co-ordinate singularities at the poles: $\theta = 0, \pi$. A concrete and rich example is the shape of crystals in equilibrium with its liquid (e.g., ^{4}He crystals, in coexistence with the superfluid [18]) or vapour (e.g., gold crystals [19]). Neither crystal shapes are spherical and can be described by a non-trivial function $R\left(\theta,\phi\right)$, which spec-

ifies the distance from the center of mass to a point on the crystal surface labeled by (θ, ϕ). The tangent plane at that point can be labeled by the direction of its normal, labeled by $\left(\tilde{\theta}, \tilde{\phi}\right)$. The relationship between these and the derivatives $\partial_\theta R$ and $\partial_\phi R$ clearly exists, but is far from simple. From that, a (generalized) LT of R can be constructed: $\sigma\left(\tilde{\theta}, \tilde{\phi}\right)$. It turns out that σ is also a significant physical quantity: it is the free energy per unit area (surface tension) associated with a planar interface, with normal $\left(\tilde{\theta}, \tilde{\phi}\right)$, between the crytaline and the isotropic phases of the material. An additional bonus here is that, unlike typical thermodynamic potentials like entropy and free energies, the "potential" $R(\theta, \phi)$ is not just an abstract concept; it is manifestly visualizable, being displayed explicitly as a shape in 3-D! The interested reader can find details of this intriguing connection in refs. [17].

In conclusion, we note that our approach to LT appears to us to be simpler, better motivated, and more easily connected to physical motivation than the traditional. Our sense from teaching the topic is that students often find it confusing and unnatural. It would be of considerable interest to carry out research into student understanding of the LT in physics and to explore whether our approach could improve the situation.

VII. ACKNOWLEDGEMENTS

This work is supported in part by the US National Science Foundation through grants DMR-0705152 and DUE-0524987.

---

[1] C.-C. Cheng, "Maxwell's Equations in Dynamics," *Am. J. Phys.*, **66**, 622 (1966); A.L. Fetter and J.D. Walecka, *Theoretical Mechanics of Particles and Continua* (McGraw-Hill, Inc., 1980).

[2] J. W. Cannon, "Connecting Thermodynamics to Students' Calculus," *Am. J. Phys.,* **72**, 753-757 (2004).

[3] J. R. Thompson, B. R. Bucy, and D. B. Mountcastle, "Assessing Student Understanding of Partial Derivatives in Thermodynamics," in *2005 Physics Education Research Conf. Proceedings, AIP Conf. Proc.*, **818**, 77-80 (2006); B.R. Bucy, J.R. Thompson, and D.B. Mountcastle, Student (Mis)application of Partial Differentiation to Material Properties, in *2006 Phys. Educ. Res. Conf. Proc., L. McCullough, L. Hsu, P. Heron, Eds., AIP Conf. Proc.* **883**, 157-160 (2007)

[4] The use of an inappropriate structure for historical reasons is called a "QWERTY". This comes from the arrangement of the keys on the first row of a typewrite. When the typewriter was in its infancy, typists worked too fast resulting in jamming of the mechanism. The keys were rearranged into their present format in order to slow down the typist. Although current devices do not suffer from the jamming mechanisms of early typewriters, the keyboard is retained "because so many people have learned to use it."

[5] M. Artigue, J. Menigaux, and L. Viennot, "Some aspects of students conceptions and difficulties about differentials," *Eur. J. Phys.* **11**, 262-267 (1990); E. F. Redish, "Problem Solving and the Use of Math in Physics Courses," to be published in Proceedings of the Conference, *World View on Physics Education in 2005: Focusing on Change,* Delhi, India, August 21-26, 2005. [http://www.physics.umd.edu/perg/papers/redish/IndiaMath.pdf].

[6] There's nothing special about the dot $(\cdot)$ at this stage; it's simple multiplication. It is useful, however, when we consider multivariate problems.

[7] In this example, s, x, F, G are all positive. Thus, the "G axis" actually points downwards, opposite to the "F axis."

[8] This comment can be extended to physical quantities, e.g. the Hamiltonian, with dimensions. It is just more tedious to keep track of the units like $[s] = [F] / [x]$.

[9] See, e.g., eqn (12.7) in J. D. Jackson, *Classical Electrodynamics,* 3rd ed. (John Wiley and Sons,Inc., 1999); or eqn. (7.136) in H. Goldstein, *Classical Mechanics,* 2nd ed. (Addison-Wesley, 1980).

[10] E. Taylor and J. A. Wheeler, *Spacetime Physics,* 1st ed. (Freeman, 1966).

[11] K. Huang, *Statistical Mechanics,* 2nd ed. (John Wiley and Sons, Inc., 1987); H.S. Robertson *Statistical Thermophysics* (Prentice Hall, 1997).

[12] In fact, the particles must be assumed to interact weakly at short range in order to permit a sharing of energy among the particles so that they can come to an equilibrium. As long as the interactions are short range and energy conserving, the details of their interaction do not affect the result.

[13] Similarly, a century later, matter was discovered to contain huge stores of energy by virtue of its mass.

[14] By contrast, for system in *non*-equilibrium stationary states, "negative responses" can be easily achieved. See, e.g., R. K. P. Zia, E. L. Praestgaard, and O. G. Mouritsen, "Getting more from pushing less: Negative specific heat and conductivity in nonequilibrium steady states," *Am. J. Phy.*, **70**, 384-392 (2002).

[15] This sort of construction is attributed to Born and Konig. See, e.g., W.W. Bowley, "Legendre transforms, Maxwell's relations, and the Born diagram in fluid dynamics," *Am. J. Phys.* **37**, 1066-1067 (1969).

[16] Note that these partial derivatives are taken with the understanding that all other variables are held fixed. Unfortunately, it is common (and reasonable) to consider derivatives with F or G held fixed. In this article, we avoid discussing such complications.

[17] For reviews, see, e.g., M. Wortis, "Equilibrium crystal shapes and interfacial phase transitions", in *Chemistry and Physics of Solid Surfaces*, ed. R. Vanselow (Springer, Berlin, 1988), Vol. VII, pp. 367-406; and R. K. P. Zia, "Anisotropic Surface Tension and Equilibrium Crystal Shapes", in *Progress in Statistical Mechanics*, ed. C.K. Hu (World Scientific, Singapore, 1988), pp. 303-357. See

also R. K. P. Zia and J. E. Avron, "Total Surface Energy and Equilibrium Shapes: Exact Results for the d=2 Ising Crystal", *Phys. Rev.* **B25**, 2042-2045 (1982). The connection between anisotropic surface energy and the minimizing shape was first established over a century ago: G. Wulff, "Zur Frage der Heschwindigkeit des Wachstums un. der Auflosung der Krystallflachen", *Z. Krystal. Mineral.*, **34**, 449-530 (1901).

[18] J. Landau, S. G. Lipson, L. M. Määttänen, L. S. Balfour, and D. O. Edwards, "Interface between Superfluid and Solid ^{4}He", *Phys. Rev. Lett.*, **45**, 31-35 (1980).

[19] J. C Heyraud and J. J. Métois, "Equilibrium shape of gold crystallites on a graphite cleavage surface: Surface energies and interfacial energy", *Acta Metal.* **28**, 1789-1797 (1980).