

# Report: Some little Points on Physics and Stochastics

Kenn K. Q. Zhang

## Abstract

Two major categories of little points are presented here, one category about errors or doubts in established knowledge and the other category about unified or simplified presentation of established knowledge. Some of these points are illustrated via rigorous quantitative derivations while others via rigorous qualitative inference. Occasionally an unreliable but could-be inspiring metaphysical fashion is adopted, but it will be indicated whenever and wherever this takes place. Most materials presented in this report may misleadingly seem trivial, but actually contain finest subtlety, demand precision thinking, and could be of great value in the future.

## 1 Classical Thermodynamics: differentiation and integration

### 1.1 An example

Consider a typical equation of state as an example

$$p = \frac{RT}{v-b} + c, \quad (1)$$

where  $R$ ,  $b$ , and  $c$  are constants. If temperature  $T$  and specific volume  $v$  are regarded as independent variables, with assumption of sufficient smoothness we have total differentiation

$$dp = \frac{\partial p}{\partial T}dT + \frac{\partial p}{\partial v}dv = \frac{R}{v-b}dT + (-1)\frac{RT}{(v-b)^2}dv \quad (2)$$

### 1.2 A 150+ years old illegal mathematical operation

To my knowledge, here is a 150+ years old mathematical operation witnessed by many textbooks on thermodynamics such as [1]. Given

$$dp = \frac{R}{v-b}dT + (-1)\frac{RT}{(v-b)^2}dv \Rightarrow \int dp = \int \frac{R}{v-b}dT + \int (-1)\frac{RT}{(v-b)^2}dv, \quad (3)$$

with immediate consequence of

$$p = 2\frac{RT}{v-b} + f_1(v) + f_2(T),$$

which is wrong since it is bounded to fail to recover Eq. (1).

### 1.3 The correct operation

Given

$$\begin{aligned} dp &= \frac{R}{v-b}dT + (-1)\frac{RT}{(v-b)^2}dv \Rightarrow \frac{\partial p}{\partial v} = (-1)\frac{RT}{(v-b)^2} \\ &\Rightarrow \int \frac{\partial p}{\partial v}dv = \int (-1)\frac{RT}{(v-b)^2}dv \Rightarrow p = \frac{RT}{v-b} + f(T), \end{aligned}$$

which is substituted into

$$\frac{\partial p}{\partial T} = \frac{R}{v-b} \Rightarrow f = \text{const}$$

Therefore, the integration yields

$$p = \frac{RT}{v-b} + c,$$

which correctly goes back to Eq. (1).

### 1.4 Comments

So, which operation in Eq. (3) went wrong? Equation (2) tells that the total differentiation of the dependent variable,  $dp$ , is contributed by that due to variation of independent variable  $T$  and that due to variation of independent variable  $v$ . It is indeed tempting to integrate on both sides of equation, since we know the fundamental meaning of integration is simply summation of all little differentiations. However, as in the following integration

$$F(T, v) = \int^T f(T, v)dT,$$

the integration is always a *partial integration*, and its inverse operation is *partial derivative*. In other words, in general there is no *total integration* (under special conditions yes there is a total integration). Therefore, the following operation in Eq. (3) is not well defined in general

$$\int dp,$$

if  $p$  is a dependent variable.

Certainly, when  $p$  is an independent variable the right above integration is well defined. This is one difference between dependent variables and independent variables. To see more subtle difference between independent variables and dependent variables let us look at the meaning of partial derivative and total derivative

$$\frac{\partial p}{\partial T} \equiv \lim_{\Delta T \rightarrow 0} \frac{p(T + \Delta T, v) - p(T, v)}{\Delta T}, \quad \frac{dp}{dT} \equiv \lim_{\Delta T \rightarrow 0} \frac{p(T + \Delta T, v + \Delta v) - p(T, v)}{\Delta T},$$

from which we notice that there is no difference between *partial change* and *total change* of the independent variable, in contrast to the case for dependent variable. We may introduce  $\partial p^T$  to

represent partial change of  $p$  due to  $T$ , that is,  $\partial p^T = p(T + \Delta T, v) - p(T, v)$ . For dependent variable, the following is well defined

$$\int \partial p^T = p + \text{const}$$

## 1.5 Summary

Integration is always *partial* (it becomes *total* in reduced cases where there is only one variable).

For independent variables, *partial change/differentiation* and *total change/differentiation* are the same, and inverse operation of the partial integration in respect to the independent variable is partial derivative in respect to the same independent variable.

For dependent variables, *partial change/differentiation* and *total change/differentiation* are different, and  $\int \partial p^T$  is well defined but  $\int dp$  is not.

## 2 Classical Thermodynamics: heat death

Next let us carefully and rigorously recheck how the conclusion of heat death of the universe is drawn. Basically, an *isolated* system with *finite* volume is studied and it is discovered that the entropy of the system always increases. Then it is claimed that at some time point in the future, very very far future, the universe will end up with a heat death.

There is a fundamental mathematical error in the above line of logical inference. When a system is finite then we can find a “bag” and put the system completely into the bag, so that system has no mass and energy interactions with the outside-bag world. Then we do find an ever increasing of entropy which leads to the heat death conclusion. But how about the universe? Can we find a “bag” at the first place to put the universe in?

In mathematics, there are fundamental differences among finite, countably infinite, and infinite numbers. So we shall be aware that conclusion drawn for a finite system may be completely not applicable to an infinite system. The universe is very likely infinite though some contemporary theories insist the finiteness of the universe. In case the universe turns out to be finite indeed, then the heat death conclusion can be drawn. However, the heat death of universe assertion survived all those years when there were no theories about the finiteness of the universe, and today the heat death assertion is still presented without any reference to the possible finiteness of the universe. In other words, the bottom line is, the logical process of heat death conclusion is fundamentally flawed.

As a matter of fact, if the time is infinite, then heat death probably never will happen: since the

time is infinite, if the heat death could happen then it must have happened (then we would not see such a wonderful world).

Actually the right above logical inference is clear-cut if the time is infinite while the space is finite. However, when both are infinite then it becomes difficult to judge. But if let me take a metaphysical guess, I would like to pick that even for both infinite time and infinite space the heat death will never happen.

### 3 Classical Thermodynamics: more on differentiation

Let us see more subtlety in derivation of mathematical relations in thermodynamics. Suppose we want to calculate compressibility  $\gamma$

$$\gamma \equiv \frac{\kappa_T}{\kappa_S}, \quad \text{where } \kappa_{T,S} \equiv -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{T,S}$$

An identity to be employed is

$$\left( \frac{\partial P}{\partial V} \right)_T \left( \frac{\partial V}{\partial T} \right)_P \left( \frac{\partial T}{\partial P} \right)_V = -1$$

Also,

$$\left( \frac{\partial S}{\partial V} \right)_P \left( \frac{\partial V}{\partial T} \right)_P = \left( \frac{\partial S}{\partial T} \right)_P, \quad \text{and } \left( \frac{\partial V}{\partial S} \right)_P = \frac{1}{\left( \frac{\partial S}{\partial V} \right)_P}$$

Now let us proceed as in a typical textbook

$$\gamma \equiv \frac{\kappa_T}{\kappa_S} = \frac{\left( \frac{\partial V}{\partial P} \right)_T}{\left( \frac{\partial V}{\partial P} \right)_S} = \frac{\left( \frac{\partial P}{\partial V} \right)_S}{\left( \frac{\partial P}{\partial V} \right)_T} = \frac{-\left( \frac{\partial V}{\partial T} \right)_P \left( \frac{\partial T}{\partial P} \right)_V}{-\left( \frac{\partial V}{\partial S} \right)_P \left( \frac{\partial S}{\partial P} \right)_V} = \frac{\left( \frac{\partial S}{\partial V} \right)_P \left( \frac{\partial V}{\partial T} \right)_P}{\left( \frac{\partial S}{\partial P} \right)_V \left( \frac{\partial T}{\partial V} \right)_V} = \frac{\left( \frac{\partial S}{\partial T} \right)_P}{\left( \frac{\partial S}{\partial T} \right)_V} = \frac{c_p}{c_v} \quad (4)$$

The derivation seems solid. But since here two different processes are involved, we shall index the isothermal process as process I and the adiabatic process as process II. Then for the isothermal process we have

$$\kappa_T \equiv -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_T^I, \quad \text{where } V = V_I(P, T)$$

and for the adiabatic process we have

$$\kappa_S \equiv -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_S^{II}, \quad \text{where } V = V_{II}(P, S)$$

Note that for both processes the dependent variable  $V$  is the same, but the mapping  $V_I$  and  $V_{II}$  are different. Similarly as in Eq. (4), we can arrive at

$$\gamma \equiv \frac{\kappa_T}{\kappa_S} = \frac{\left( \frac{\partial S}{\partial V} \right)_P^{II} \left( \frac{\partial V}{\partial T} \right)_P^I}{\left( \frac{\partial S}{\partial P} \right)_V^{II} \left( \frac{\partial T}{\partial V} \right)_V^I}$$

Now, here is the question: can we proceed further as in Eq. (4)?

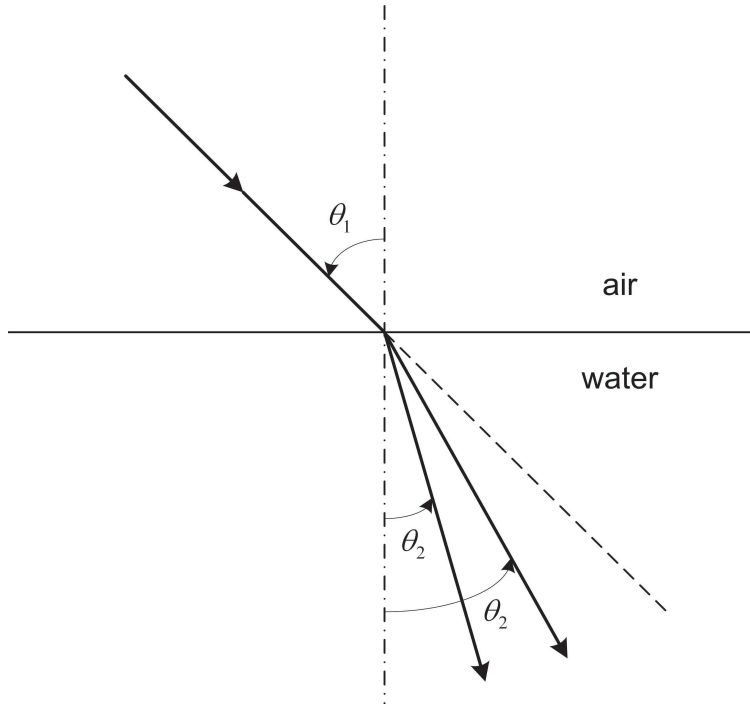


Figure 1: Refraction of a mixed (red and green) beam of light from air into water. The air is denoted as medium 1 and the water as medium 2. Some interesting questions could be: which light bends more, red or green? Why bends? Why one light bends more than the other?

## 4 Optics: light refraction

### 4.1 Light refraction

Let us consider light refraction from air into water as shown in Fig. 1. The light beam in the air is a mixture of red and green, while in the water they are separated. Quantities in the air (incident angle, index factor, speed of light, frequency of red light, frequency of green light, wave length of red light, wave length of green light, energy of red light, and energy of green light, respectively) are the reference:

$$\theta_1, \quad n_1 \equiv 1, \quad v_1, \quad f_1^{red}, \quad f_1^{green}, \quad \lambda_1^{red}, \quad \lambda_1^{green}, \quad E_1^{red}, \quad E_1^{green}$$

Quantities in the water (transmit angle of red light, transmit angle of green light, index factor of red light, index factor of green light, speed of light, frequency of red light, frequency of green light, wave length of red light, wave length of green light, energy of red light, and energy of green light, respectively) are to be compared with quantities in the air or compared between red and green lights in the water:

$$\theta_2^{red}, \quad \theta_2^{green}, \quad n_2^{red}, \quad n_2^{green}, \quad v_2, \quad f_2^{red}, \quad f_2^{green}, \quad \lambda_2^{red}, \quad \lambda_2^{green}, \quad E_2^{red}, \quad E_2^{green}$$

An analysis begins with the following mutually independent and complete mathematical relations or given conditions:

1.  $\lambda f = v$  (a purely kinematical relation)
2.  $f_1 = f_2$  (why true?)
3.  $v^{red} = v^{green}$  (a fact in the same medium)
4.  $v_2 < v_1$  (due to resistance)
5.  $n_1 \sin \theta_1 = n_2 \sin \theta_2$  (Snell's law, is this fundamental?)
6.  $E = hf$  (why?)
7.  $n_1 \equiv 1, \theta_1$  given,  $f^{red} < f^{green}$  (fact)

There are three questions in the above mathematical relations or conditions. First of all, why  $f_1 = f_2$ ? The invariance of frequency in reflected and transmitted waves from incident wave are typically taken for granted without explanation. Here an explanation is provided: the constancy of frequency of waves is due to continued supply from the light source. Another way to explain this is, if we assume  $E = hf$  (which is a fundamental assumption of quantum mechanics), now since energy of single-frequency light beam across the interface remains constant, then frequency must remain constant.

The second question is about the Snell's law  $n_1 \sin \theta_1 = n_2 \sin \theta_2$ , where  $n_1 \equiv 1$  and  $\theta_1$  is given. Suppose we want to know what  $\theta_2$  is, then Snell's law tells us what kind of equation  $\theta_2$  has to obey. However, in this equation the index factor  $n_2$  is not known either, thus Snell's law just drop the ball from one unknown quantity to another. Nevertheless, Snell's law reveals a structure of the relation between two unknown quantities,  $n_2$  (a material property) and  $\theta_2$ , so that with help of experiments we can base on Snell's law to calculate  $\theta_2$ . Therefore, Snell's law is practically useful, but not fundamental and does not address the reason behind light refraction. As an application of Snell's law, since experimentally it shows that  $n^{green} > n^{red}$  then the green light must bend more. The third question about energy quantization  $E = hf$  is postponed to section 4.2.

The above mathematical relations or conditions are complete. For instance, that  $E_1^{red} < E_2^{green}$  can be inferred from  $E = hf$  (item 6) and  $f^{red} < f^{green}$  (item 7). As a more advanced example, we can do a self comparison for the red light across the interface

$$v \searrow, \quad f \longrightarrow, \quad \lambda \searrow, \quad E \longrightarrow$$

Now, since the wavelength  $\lambda$  shortens across the interface, what will be the color? Still red? Indeed, the wavelength can be used to characterize a specific color, however, that applies only to the same medium.

For situations involved by different media, the frequency, not the wavelength, is the characteristic of the color.

## 4.2 Two metaphysical explanations

Now we want to do a much deeper, unfortunately metaphysical, analysis. We want to think:

- Why  $E = hf$ ?
- Why bends?
- Why green light bends more than red?

$E = hf$  is a fundamental assumption of quantum mechanics, which is experimentally verified but not explained in established physics. In physics, that “light takes the shortest path” is believed to be able to explain why bends and why green light bends more. However, why light wants to take the shortest path? More, how does light figure out the destination at the first place (otherwise talking about the shortest path is meaningless)? Therefore, as the rampant practice in physics, by saying “light takes the shortest path”, a hard ball is dropped from one question to another. In this report, we attempt not to do a ball dropping.

## 5 Stochastics

### 5.1 Notation

In the field of stochastics (probability, stochastic processes, and statistics) the notation is predominantly over cautious. For example, the probability mass function of poisson distribution reads

$$Pr\{X = x\} = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0, 1, 2, \dots,$$

where  $X$  is a random variable and  $x$  is a supplied independent variable. Since  $X$  shows up in the above expression only once then it should be spared, similar to the computer programming situation where a variable shows up only once (besides definition of the variable) then the compiler will issue at least a warning. Note that in audacious mathematical notation

$$y = y(x)$$

the  $y$  on the left differs from the  $y$  on the right, yet we use the same  $y$ . In contrast, the notation in stochastics is too cautious and cumbersome. Another factor in notation is the default. In a predominantly deterministic environment a random variable should be especially pointed out, in contrast, in

a predominantly random environment a random variable is the default. Let us try to simplify the poisson probability mass function to

$$Pr\{x\} = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0, 1, 2, \dots,$$

where  $x$  is a random variable, possibly taking values at non-negative integers.

In stochastics, the probability mass function has the “dimensionality” of probability for the discrete random variable, in contrast, for the continuous situation we have the probability density function whose integration in respect to the independent variable has the “dimensionality” of probability. Such a separation and inconsistency between discrete and continuous random variables may be easily resolved by introducing Dirac Delta function. For example, we may write the probability density function of poisson distribution as

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!}\delta(x - k) \quad k = 0, 1, 2, \dots,$$

where the random variable  $x \in (-\infty, \infty)$  and  $k$  is a supplied parameter. Obviously, the integration of this above density function gives cdf (cumulative distribution function), which is closely associated with probability as in continuous case. These defined notations are subject to future extensive tests, such as the derivatives.

## 5.2 Canonical definitions of mean, standard deviation, fluctuation

Suppose we have  $n$  number of *same measurements*  $\{x_j | j = 1, 2, \dots, n\}$ . The *mean* should be defined, canonically, in the arithmetic sense

$$\bar{x} \equiv \frac{1}{n} \sum_{j=1}^n x_j \quad (5)$$

There is a reason why the arithmetic mean, not, say, geometric mean, is chosen; this reason will be investigated in the future. As the number of data increases,  $\bar{x}$  approaches to the limit  $\mu$ , the expectation value. Here the mean is considered more canonical than expectation, because the mean is always doable. The canonical *standard deviation* measures the spread of data from the *mean*  $\bar{x}$

$$\sigma \equiv \sqrt{\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x})^2} \quad (6)$$

In other words, here the sample standard deviation is considered more canonical than population standard deviation, because the former is always doable. As  $\bar{x}$  approaches the limit  $\mu$ , the standard deviation approaches the limit

$$\tilde{\sigma} \equiv \sqrt{\frac{1}{n} \sum_{j=1}^n (x_j - \mu)^2}$$

$\tilde{\sigma}$  better reflects quality of experimental data, however, due to experimental errors and the finiteness of data number, the true value of the expectation is not known and this makes direct calculation of  $\tilde{\sigma}$  difficult. Fortunately, it turns out

$$\tilde{\sigma} = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2} \quad (7)$$

Again, since the random variable  $x$  (i.e., the precise value of each measurement  $x_j$ ) is not known, the remedy is to provide a *fluctuation* (also named as confidence interval or uncertainty)  $\Delta x$  along with the mean value

$$x = \bar{x} \pm \Delta x$$

The standard deviation can be chosen as the fluctuation, that is,  $x = \bar{x} \pm \sigma_x$ . The fluctuation can also be expressed in percentage format as  $\bar{x}(1 \pm \text{err}\%)$ .

### 5.3 Error propagation

For two independent random variables  $x = \bar{x} \pm \sigma_x$  and  $y = \bar{y} \pm \sigma_y$ , let us consider error propagation

$$u = f(x, y)$$

It can be shown through series expansion, under assumption of small errors, that

$$u = f(\bar{x}, \bar{y}) \pm \sqrt{f_x^2 \sigma_x^2 + f_y^2 \sigma_y^2}$$

As a special case, for the mean value  $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$  the error is

$$\bar{\sigma} = \frac{\sigma}{\sqrt{n}}$$

### 5.4 Why so few surnames in China?

China has 1.3 billion of Han populations, but only with approximately hundreds of surnames (there is 100 millions of people on this planet bearing the last name “Zhang”). Why is that? Two primary reasons are:

- According to the culture, a child will bear the surname of the father. Hence, a surname is either passed to the next generation or terminated, but not created.
- China is an ancient civilization without interruption.

That “a surname is either passed to the next generation or terminated, but not created” gives an impression that a surname always declines. This is actually not true because a surname could multiply, so that while no new surname can be created the population with existing surnames could increase which makes the extinction of a surname more difficult. Then, how many children born to a pair of parents becomes an important factor.

So, even with such a simplified model without consideration of important initial conditions such as how many surnames began with and the uniformity of the beginning time, an exact analysis is quite challenging. Hence we would like to resort to numerical simulations. Some more interesting questions we seek are, what is the long term behavior of this process? How many surnames will be left eventually, that is, the asymptotic solution?

## References

- [1] M. J. Moran and H. N. Shapiro. *Fundamentals of Engineering Thermodynamics*. Wiley, 5th edition, 2003. chapter 11.