PHYSICS EDUCATION

Connections

Freedom

Highlights:

- Marbles and Boxes Illustrate Irreversibility and Recurrence
- Motion of a Rod Pushed in a Weightless Environment in Space
- On Circular Motion of John Warren's Car
- Generation and Detection of Optical Vortex
- An Erlangen Program for Physics

Volume 34, Number 1

In this Issue

Marbles and Bottles-or Boxes illustrate Irreversibility and Recurrence

Anna Maria Aloisi[1] and Pier Franco Nali[2]

[1]IPSIA A. Meucci, Cagliari, Italy (Ret.).

annamaria.aloisi@istruzione.it

[2]Via Tempio 29, 09127 Cagliari, Italy.

ampfn@tiscali.it

Submitted on 08-02-2018

Abstract

Imaginary assemblies of intercommunicating bottles-or boxes, in which marbles circulate performing random walks, can help modeling (and hence understanding) slow and long-lasting diffusion processes and enable easy evaluations of typical times involved in their dynamics. Compared to the more traditional approaches to explaining diffusion inspired by the Ehrenfest's urn problem, these simple random walk models offer a more straightforward way to illustrate irreversibility/recurrence issues.

1   Introduction

Ideal models based on "Marbles and Boxes" are well suited to acquaint students with the basics of the diffusion and transport dynamics in gases and other media. Mostly implemented by computer algorithms, these abstract artifacts can sometimes have a material counterpart suitable for developing hands-on classroom activities. An example is the "Marble Game", [1] proposed some years ago in the U.S. in connection with the reforms of the STEM curriculum [1]. In this game N marbles are distributed between two boxes. Marbles jump between boxes (in both directions) at a constant rate, with a random extraction rule (e.g., rolling a multi-sided die) that decides which marble will next jump to the other box: if the rolled number is less than or equal to the number of marbles in *box1*, then a marble is moved from *box1* to *box2*, otherwise a marble is moved from *box2* to *box1*.

The modeling with the "Marble Game" is a kinetic Monte Carlo (kMC) simulation of the classic Ehrenfest's two-urn model system (also known as dog-fleas model) and can be physically realized if the number

---

[1]Not to be confused with the popular LEGO Marble Maze Game (Labyrinth).

N of marbles is not too large (say $N = 10$). This allows the possibility for a classroom activity with a physical representation of the game, on which the students can play moves by hand according to set rules, and gives support in a sensory form to the idea that diffusion and molecular transport phenomena are marked by the evolution of a dynamical system. Once the process has started, its evolution is examined through the concepts of probability, statistical equilibrium and randomness of the states reached by the system. By playing the game, and at a later time running computer simulations, students discover that equilibrium is the result of a dynamical process yielding to an irreversible tendency, and recognize the key role that randomness plays in the diffusion mechanism (ruled by Fick's first law) modeled by the game.

In this paper, we consider a "Marbles and Bottles-or Boxes" model explaining diffusion, reported in the 90's by Clifford Pickover, a researcher at IBM Watson Research Center, who first introduced it in form of a math puzzle connecting time to probability [2, 3]. This model - that is as a matter of fact the well-known random walk in one dimension - leads to a startling representation of time intervals in terms of solutions to simple random walk and diffusion problems, and offer an alternate view to the more classical approaches, based on the "Marble Game" or similar models, to explain macroscopic irreversibility.

The structure of this paper is organized as follows. In section 2, we outline how the "Marbles and Bottles-or Boxes" model operates and what kind of physical systems it could represent. In section 3, we introduce some very elementary concepts about random walks and stochastic processes, required in order to appreciate the statistical arguments developed in subsequent sections. In sections 4, 5, and 6, we present three sample applications of the above concepts and offer a proof, only relying on elementary probability theory and finite difference equations, of a formula to calculate the first-passage time through a generic node of a linear chain assembly. In sections 7 and 8, in order to illustrate irreversibility/recurrence issues, we introduce a variant of the basic random walk model applicable to various physical situations and briefly recapitulate the Ehrenfest's urn model along with generalizations thereof. Some remarks on the discrepancy between recurrence and irreversibility at the macroscopic level are developed in sections 9 and 10, where simple simulation results are also summarized. In section 11 a gas effusion process is presented as a microscopic analogue of the model outlined in section 2. Conclusions are drawn in section 12, along with a brief discussion of the pedagogical value of the above-mentioned models and their limitations.

2   Time in a bottle-or a box

What happens if we place a marble into a large glass bottle that had a small opening at its top and we began to shake the bottle randomly? How long would it take for the marble to leave the bottle through the tiny aperture if we were to continue shaking? The average time for the marble to get out of the bottle will obviously depend on the size of the hole (and the size of the marble). Let us say the marble popped out after 1 hour of constant bouncing around in the bottle.

What would happen now if we were to place a series of bottles together so that only a small opening connected the bottles: Bottle 1 connects to Bottle 2, Bottle 2 connects to both Bottle 1 and Bottle 3 (see figure 1), and so on. Nothing prevents us from imagining that more and more bottles could be connected using their small openings, all in the same way to make up a linear assembled chain. The last bottle of the sequence (say Bottle n) opens to the outside world. (Assume this is an ideal system: it has no friction, gravity, etc.). How long would it take for the marble to exit Bottle n? We must bear in mind that in each of the intermediate bottles the marble, in its random motion, has just as likely a chance of moving into a previous bottle as it does moving forward. Let us also assume that it takes one hour for the marble to find an opening as it did in the single-bottle experiment.[2]

It can be easily shown, virtually starting from scratch, that the average bottle number reached in a given time approaches a constant. This suggests a method for drawing diagrams, of immediate visual effectiveness for a student, connecting bottles-or boxes (or chambers or other kind of container) in long chains representing large expanses of time, in fact so large that the flow of marbles through them appears as essentially irreversible.

In addition, the imaginary assembly of bottles just described could represent a macroscopic model demonstrating the diffusion of an extremely rarefied gas in a network of high vacuum flasks connected by highly selective porous seals.

Finally, it could be used as a simulation tool suitable to give just an idea of the difference between a macroscopic and a microscopic system: how fast (and how small) should the marble be, in order to reduce the "diffusion" time to values that can actually be experienced even for this kind of relatively complex system?

3   Random walks and Poisson processes

The concept of random walk (or drunk man walking) was introduced by the English philosopher and statistician Karl Pearson (1857-1936) in a letter to Nature, dated

---

[2]All models involve some simplification. Models as those considered in this paper, in which imaginary marbles of a proper size are placed in imaginary boxes, are well suited for high-speed computer simulations and were used in Physics since the pioneering work by Berni Adler [4] on hard sphere systems.

July 27, 1905 [5], where the following problem was presented for the first time: *"A man starts from a point O and walks l yards in a straight line; he then turns through any angle whatever and walks another l yards in a second straight line. He repeats this process n times. I require the probability that after these n stretches he is at a distance between r and r + dr from his starting point, O.".*

Intensively studied in the 20th century, examples of random walk are ubiquitous, from Brownian motion to diffusion processes in chemical-physics, biology and sociology [6]; other examples, frequently quoted, concern the motion of a body through a series of adjacent passages or, occasionally, in confined geometries, with various applications to transport and separation processes [7–12]. As a special kind of stochastic process, the random walk is a mathematical model that schematizes, using probabilistic-statistical methods, the time course of a random phenomenon.

In the following, it is assumed that the search for the exit in the random motion of the marble inside a bottle is a stochastic process, more specifically a Poisson process, in which events occur in time completely at random at intermittent times, like incoming calls to a telephone. The mathematical description of natural phenomena such as radioactive disintegration, and of lots of demographic, economic and industrial production processes are based on the Poisson model [6–9].

In the Poisson model the probability

Δp for the marble to get to the opening in a time interval Δt is proportional to same Δt, namely $\Delta p = \lambda \Delta t$, where λ is a constant whose dimensions are the inverse of time and represents the probability per unit time for the marble to leave Bottle 1. It can be shown that the average time the marble would take to find the opening is λ^{-1}.[3] From now on, we will denote the above quantity by m_1 (1 hour in our example), which corresponds to the time step in the random walk. If the bottle includes more openings, the exit probability increases according to their number M (i.e., $\Delta p = M\lambda \Delta t$), while the average time for the marble to get out follows the inverse proportion, namely $(M\lambda)^{-1} = m_1/M$.

Very simple statistical reasonings allow us to analyze the motion of the marble within the bottle chain and derive an equation for the average time Bottle n is exited for the first time.[4] The above problem is widely and well-known in Statistics, and reduces to the simple random walk on $\mathbb{N}$ (the set of natural numbers) of length n. As we do not require the reader to have any prior knowledge on random walk processes, our analysis was conducted with the intent of achieving the goal by steps, and to this end it took into consideration three cases:

1. Only one free-flowing connection between the n bottles (this reduces to the

---

[3]More exactly, it can be shown that the number of openings found in time t is, on average, $\mu = \lambda t$ (See, e.g., [6,9]). For $\mu = 1$ it follows the assertion.

[4]In the stochastic jargon this is defined as first-passage time or hitting time.

trivial case of unidimensional diffusion of a free walker).

2. The addition of backward openings (case of a biased walker with equal step size in both directions but unequal probability of taking forward and backward paths).

3. Addition of backward openings, after considering that the first bottle is different from the rest (due to the reflecting barrier at the starting node).

Further generalizations could be imagined considering more complex network topologies constituted by nodes having connections through both forward and backward paths in various combinations of step sizes and direction probabilities.

4   Random walk and diffusion in a linear chain assembly

In the simplest case (1.) remember from section 2 that in each of the intermediate bottles (none at the end of the chain assembly) the marble has just as likely a chance of passing to the next bottle as it does to regressing into a previous one.

Let m_k represent the expected amount of time the marble would take to pass from Bottle k to Bottle $k+1$ (including any regressions to previous bottles). Then, for $k > 1$, there is a 50 percent chance that the marble will go directly from Bottle k to Bottle $k+1$, incurring an average time m_1, and a 50 percent chance that it would regress to Bottle

$k-1$, in which case the average time to return to Bottle k and then to move to Bottle $k+1$ would be $m_{k-1} + m_k$. This leads to the following difference[5] equation involving the averages of stochastic variables:

$$m_k = \frac{1}{2}m_1 + \frac{1}{2}(m_{k-1} + m_k), \qquad (1)$$

which simplifies to $m_k = m_{k-1} + m_1$. By induction, its solution is: $m_k = km_1$.

So, we get the simple but interesting result that the average time to pass from Bottle k to Bottle $k+1$ is k times the expected time to exit Bottle 1. The average time m to move from Bottle 1 to Bottle n would be $1 + 2 + \ldots + (n-1) = n(n-1)/2$ times the average time m_1 to find an opening (For a large number of bottles the latter equation might be approximated by $n^2/2$). We ignore here the fact that the first bottle is different from the rest of the chain as well as any complications depending upon whether or not continuous space or discrete space settings were assumed. Depending on the case, the results are slightly different when the assemblage contains just a few chambers. However, as stated, if a large number of chambers are considered, m increases according to n^2. It can be shown, in a strict form, that the first-passage probability for the discrete random walk and the continuum diffusion in transmission mode are asymptotically identical [13].

According to the stochastic jargon, the free state of the marble is an absorbing state

---

[5]Following a widely used definition, the term difference equation is treated here as synonymous with recurrence relation.

(the state in which the marble is removed from the system) and the m quantity is the average absorption (or escape) time. The chain has a reflecting node (supposed at the origin) and ends on an absorbing barrier. At the start of the experiment, the marble is placed in the reflecting node. It should be noted that if the connections between the bottles were equipped with valves preventing the marble from making regression, the average time to traverse the chain and reach the absorbing (free) state would be the viable minimum, equal to n times m_1.

5 Random walk with addition of backward connectors

Let us now examine the case (2.), for $k > 1$, with $M + 1$ possible exits from a bottle (except Bottle 1), one forward and M backward (see figure 2). In other words, the system is ruled so that one opening is free flowing, while the remaining have a one-way valve allowing the marble to only travel in a backward direction (a direction away from the opening of final egress).

Assuming as before that finding an exit

is a Poisson process, the average time to find the first of $M + 1$ exits is $m_1/(M + 1)$. There is a probability $1/(M + 1)$ that this exit will be forward, in which case no additional transition time is required. Also, there is a probability $M/(M + 1)$ that the first exit will regress to Bottle $k - 1$, in which case the additional average time to return to Bottle k and then to progress to Bottle $k + 1$ would be $m_{k-1} + m_k$. This leads to the modified difference equation:

$$m_k = \frac{1}{M+1}m_1 + \frac{M}{M+1}(m_{k-1} + m_k), \quad (2)$$

(remember that eq. (2) is valid for $k > 1$), which simplifies to $m_k = Mm_{k-1} + m_1$. Its solution (verifiable by induction) is:

$$m_k = (\textstyle\sum_{j=0}^{k-1} M^j)m_1$$
$$= \begin{cases} km_1, & \text{if } M = 1, \\ \dfrac{M^k - 1}{M - 1}m_1, & \text{if } M > 1. \end{cases} \quad (3)$$

(Note that eq. (3) is also valid for $k = 1$). The average time to exit the nth bottle (i.e. the average absorption time of the chain) is thus:

$$m = \textstyle\sum_{k=1}^{n} m_k = \begin{cases} [n(n+1)/2]m_1, & \text{if } M = 1, \\ \dfrac{M^{n+1} - M - n(M-1)}{(M-1)^2}m_1, & \text{if } M > 1. \end{cases} \quad (4)$$

The second line in the equation array (4)

might be approximated by

$$m \approx \frac{M^{n+1}}{(M-1)^2} m_1 \tag{5}$$

for large n (say $n \geq 10$).

6   Simple random walk model

We have assumed that the rate of exit from each hole is the same and hence the rate of exit from a bottle increases according to the number of holes, namely it is $M+1$ times the rate of exit from Bottle 1, which has only one hole (this makes the first bottle different from the rest). Then the marble will find an exit with a probability per unit time $\Delta p / \Delta t = (M+1)\lambda = (M+1)/m_1$, incurring an average time $m_1/(M+1)$.

If it is assumed (case 3.) that the average time to leave a bottle is the same for all bottles including the first (the simple random walk model), then the constant in the difference equation (probability of leaving Bottle 1) will increase by a factor of $M+1$. This can be obtained, for example, by enlarging the hole of Bottle 1 by the same factor. Then, relative to the first bottle, the remaining will have an exit probability $1/(M+1)$ for each opening. It follows that the marble will take directly the forward exit with a probability per unit time $\Delta p / \Delta t = \lambda/(M+1) = 1/[m_1(M+1)]$ incurring an average time $(M+1)m_1$. Substituting in equation (2) the constant m_1 with $(M+1)m_1$ the solution then changes to:

$$m_k = (2\sum_{j=0}^{k-1} M^j - 1)m_1 = \begin{cases} (2k-1)m_1, & \text{if } M = 1, \\ \dfrac{2M^k - M - 1}{M-1}m_1, & \text{if } M > 1. \end{cases} \tag{6}$$

The average time required to get past the nth bottle becomes:

$$m = \sum_{k=1}^{n} m_k = \begin{cases} n^2 m_1, & \text{if } M = 1, \\ \dfrac{2(M^{n+1} - M) - n(M^2 - 1)}{(M-1)^2}m_1, & \text{if } M > 1. \end{cases} \tag{7}$$

The second line in the equation array (7) might be approximated by

$$m \approx \frac{2M^{n+1}}{(M-1)^2} m_1 \tag{8}$$

(for large n).

The latter is the approximate equation reported by Pickover in [2, 3]. It was derived in 1991 by Shriram Biyani, Pickover's colleague at IBM, using the statistical arguments exposed above [14]. Various other approaches to first-passage problems on $\mathbb{N}$ have been devised, and some of them can be found on the Internet (see, among others, [15]). The number of steps to go from 0 to n in a simple random walk on the line of natural numbers, with probability p in the forward direction (except at the starting position, where the probability is 1) and with probability $1 - p$ in the opposite direction, is expressed as

$$\frac{n}{2p-1} + \frac{2p(1-p)}{(1-2p)^2}\left(\left(\frac{1-p}{p}\right)^n - 1\right),$$

which reduces to the second line in the equation array (7) for $p = 1/(M+1)$ and $M > 1$.

We preferred proposing here the Biyani's derivation because of its greater simplicity compared to other approaches. So, equation (8) (as its avatar (5) reported in the previous section) may be used for $n \gg 1$ and $M > 1$ - that is for large chains in the presence of one-way backward connectors in addition to the free flowing connection - and gives the average time until the nth bottle is exited for the first time (first-passage or hitting time). As stated in section 2, it turns out that the average bottle number reached in a given elapsed time approaches a constant.

7 Random walk with restarts: The Sisyphus force

Many variants of the simple random walk model could be developed as the bottles can connect to each other through thin tubes in lots of combinations, scaling up the connections from simple linear chains to complex networks. Among the most interesting, the introduction of one-way backward connectors directly to Bottle 1 (figure 3), bypassing the rest of the chain. This is a special case, susceptible of a simplified mathematical treatment, of the random walk with random restarts, in which, starting at the origin, the walker faces two choices: either moving forward to the next node, or jumping back to the starting node with a "restarting probability" r,[6] acting like a restoring force that will tend to bring the walker back to origin. This force is known as "Sisyphus force", and the walking process as "Sisyphus random walk" [16], by analogy with the Greek myth of Sisyphus. The effect considered underlies, in the context of Doppler laser cooling, the behavior of some physical mechanisms by which alkali atoms climbing from the ground level to higher excited states experience an increasing probability of being optically pumped into a minimum potential energy state from where the process restarts.

In this situation, if the marble expe-

---

[6]In the random walk with restarts (RWR) standard algorithm, the walker is allowed to move to a randomly chosen neighbor (with a certain probability $p = 1 - r$), or to jump back to the origin with probability r.

riences a regression, suddenly going back to Bottle 1 through any of the additional tubes, then the process restarts afresh from

the source node, and the marble must re-cross all intermediate nodes in order to re-turn eventually to Bottle k. This is expressed by the equation:

$$m_k = \frac{1}{M+1}m_1 + \frac{M}{M+1}(m_1 + \ldots + m_{k-1} + m_k), \qquad (9)$$

(valid for $k > 1$), which simplifies to

$$m_k = m_1 + M(m_1 + m_2 + \ldots + m_{k-1}).$$

Hence, it is obtained, by induction,

$$\begin{aligned} m_{k+1} &= m_1 + M(m_1 + \ldots + m_{k-1} + m_k) \\ &= m_1 + M(m_1 + \ldots + m_{k-1}) + Mm_k \\ &= (M+1)m_k, \end{aligned}$$

and at last $m_k = (M+1)^{k-1}m_1$, valid for $M \geq 1$. ($M = 1$ this time is not a special case). The average absorption time is:

$$\begin{aligned} m &= \sum_{k=1}^{n} m_k = \frac{(M+1)^n - 1}{M}m_1 \\ &\approx \frac{(M+1)^n}{M}m_1 \end{aligned} \qquad (10)$$

(for large n). Assuming the same average exit time for all bottles, including the first one, we have to replace in the equation (9) the constant m_1 by $(M+1)m_1$ and make the necessary simplifications[7], thus getting

$$\begin{aligned} m &= \frac{(M+1)^n - 1}{M}(M+1)m_1 \\ &\approx \frac{(M+1)^n}{M}(M+1)m_1 \end{aligned} \qquad (11)$$

(for large n).

All the former results might be further generalized for a generic p probability for the marble to directly take the forward exit (and a generic restarting probability r) by replacing $M = 1/p - 1 = r/p$ in the expressions given above.[8]

8 Ehrenfest's urn experiment and recurrence time

The Marble Game we encountered in section 1 is a kMC simulation of the Ehren-fest's urn experiment. This thought exper-iment, originally conceived in the wake of the grand debate about the apparent con-tradiction between second law of thermo-dynamics and Boltzmann's kinetic theory of gases, was authoritatively defined by Kac "one of the most instructive models in the whole of physics" [17]. In the simplest ver-sion of the experiment [18], one starts with N marbles in the left urn (or *urn0*). The mar-bles are numbered from one to N and a third urn exists, containing N cards with a natural

---

[7]The simplifications are straightforward and are left as an exercise for the more energetic appenders.

[8]$m = \frac{p^{-n}-1}{1-p}m_1 \approx \frac{p^{-n}}{1-p}m_1$. (Cfr. eq. (18) in ref. [16]).

number from one to N printed on each. The procedure considers drawing a card at random from the cards-urn, looking at the number printed on it, drawing the marble with the corresponding number from its urn, and putting it in the other urn. (The cards are returned to the card-urn after each observation). The procedure is carried out repeatedly, virtually endlessly. Then, simply relying on the counting of the possible configurations and forcing the process to be repeated on a fairly large number of moves, we expect the initial state with the whole of the N marbles in the *urn0* to be reproduced with a probability of 2^{-N}.

The literature also reports a number of interesting multi-urn extensions/ modifications of the original Ehrenfest's dog fleas model, which would be worth exploring [19–25]. In latter models, N marbles circulate in the network relying on either directed or random mechanisms that distribute the marbles until some specific condition is reached, or the first-passage to some special state is detected.

In the case of a multi-urn experiment (*M+1* urns from *urn0* to *urnM*), the probability of the special state with the whole of the N marbles in the *urn0* would be $(M + 1)^{-N}$ [19]. The corresponding so called average recurrence time, that is the time after which the system regains periodically (to arbitrary closeness) its initial state (Poincar cycle) is then calculated from the Kac's lemma [26,27], taking the inverse of this probability: $m_{[N \to N]} = (M + 1)^N m_1$, being now m_1 the

time slice spent to force a marble to change urn.[9]

9 Emergence of irreversible behavior

In the Marbles and Bottles-or Boxes model subject of this paper we have seen that, in the simplest case of a single two-way connector joining the bottles (figure 1), the number of bottles reached on average by the marble at a given time increases with the square root of the same time: $n \propto \sqrt{m}$, as expected for unidimensional random walks.[10]

The introduction of additional one-way backward connectors, as in figures 2 and 3, leads to an increase in disorder in the path of the marble. Now, as the number of backward connectors increases, the marble is more likely to undergo regression; so, on average, it would take longer to reach a given number of chain nodes. This number grows very slowly with time, in fact more slowly

---

[9]See, e.g., [27] for a proof of this. The value reported obviously depends on the peculiarity of the macroscopic state considered, namely the state of minimum probability (the N-state or zero-entropy state). Each macroscopic state has its own probability (known in the literature as stationary probability or Markov probability measure), ranging from a minimum in the N-state to a maximum in the $N/2$-state (the equilibrium state or maximum-entropy state). As the equilibrium state is reached, the probability to deviate from it is so much smaller than N is bigger. In this near-to-equilibrium regime, the probability is well approximated by a Gaussian centered on $N/2$.

[10]In the unidirectional forward motion without regression the growth would be linear: $n \propto m$.
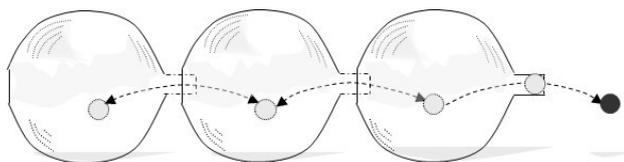
Figure 1: A chain of three bottles communicating through free flowing (two-way) openings between them. This is called a $C(3,1)$ assembly, denoting a chain of 3 bottles, with one forward opening and one backward opening ($M = 1$) each (except the first bottle that has only one opening). If the marble takes one hour to leave the first bottle, it will take about $3^2 = 9$ hours of shaking to get the marble out of the bottle chain.

than any growing function described by a power law. We found that for $M > 1$ (and very long chains) the relationship between the length of the chain assembly and the expected elapsed (absorption/escape) time (or between a given node number and the corresponding average first-passage time) exhibits a logarithmic pattern.
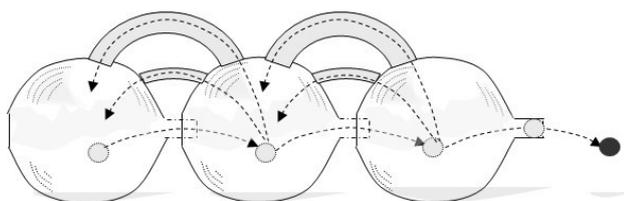
Figure 2: A $C(3,2)$ assembly: three bottles with two backward connectors ruled by one-way valves (always with the exception of Bottle 1). It will take an average of 19 hours of uninterrupted shaking for the marble to escape. (The enlarging of the hole of Bottle 1 required to make the first bottle equal to the rest is not shown).

Using either equation (8) or its avatar (5), valid for suitably long chains (say $n >$
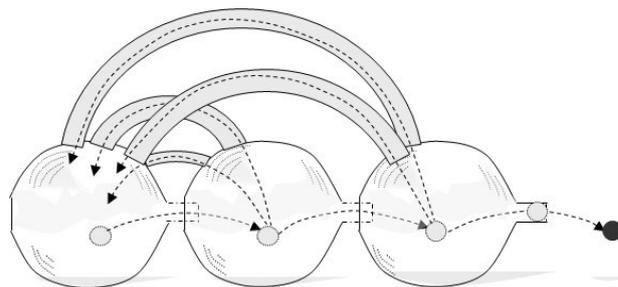
Figure 3: A variant of $C(3,2)$ with one-way tubes bringing directly the marble back into the first bottle. The marble escapes after 39 hours of enduring shake. (Additional tubes connecting Bottle 1 to itself, in order to make it equal to the rest, are not shown).

10), students can be encouraged to draw diagrams representing large time stretches, by varying the number of retrograde connections (represented by arcs) and / or the number of bottles-or boxes in the chain assembly. Some pictorial examples of such diagrams have been reported by Pickover, who also introduced – in order to facilitate the discussion of the startling characteristics of these chains – the symbol $C(n, M)$ to represent a generic chain assembly with n chambers and M backward connectors between them [3].[11] Each connector is represented by a line. Using Excel, you can easily map typical times of "improbable" processes, as the escape of a walker from such an intricate maze as that of figure 4.

It was apparently Smoluchowski (cited by Kac in [26]) who advanced the rule that a process started in a state with long recurrence time (that is - roughly speaking -

[11]Bear in mind that there is always one forward connector per node.
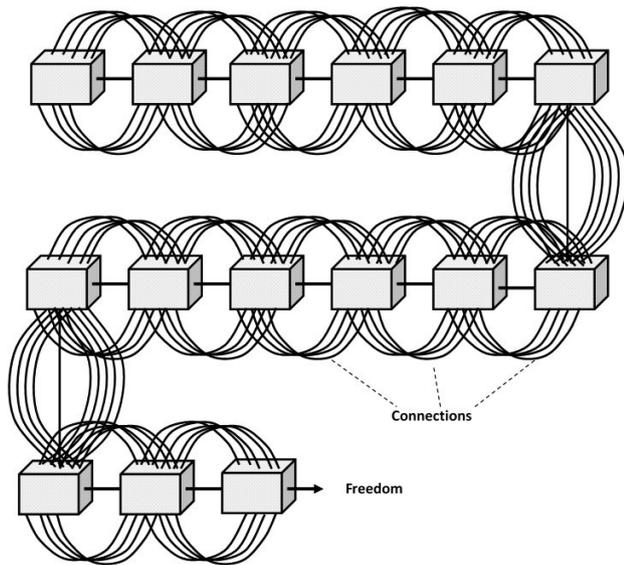
Figure 4: A $C(15, 9)$ assembly, consisting of 15 boxes with 9 backward connectors each, represents about 6.6 billion years, far beyond the age of the solar system.

the time to wait, on average, for the same state to recur) will appear as irreversible. On the other hand, a short mean recurrence time makes it meaningless to speak about irreversibility. Other different notions of (ir)reversibility and recurrence, and of their interrelationships, have characterized the many different formulations of classical thermodynamics in the last two centuries, leading to non-univocal meanings and sometimes confusing expositions of the same concepts [28]. Entering such subtle distinctions, however, is beyond the scope of this paper.

Marbles and Bottles-or Boxes models described in this paper are suitable for illustrating "essentially" irreversible processes because they are nothing more than absorbing chains of finite size, and all absorbing chains are actually not recurrent, being their exit probability (i.e. the probability that the walker eventually terminates at a particular node corresponding to an absorbing state) equal to 1 [29]. Ultimately, this non-recurring behavior has to be traced back to the eventual removal of the walker as it hits the absorbing barrier, so that the system can not be longer considered insulated. Differently, and perhaps surprisingly, infinite non-absorbing chains can exhibit recurrence, transience or ergodicity under sufficient conditions, while random walks on (finite) circular paths have the less restrictive recurrence conditions [30, 31]. The Sisyphus random walk considered in section 7 is recurrent and ergodic on an infinite chain, as intuitively expected considering that the reset mechanism will prevent the walker from being driven too far off from the origin. Every point is reached infinitely often and the mean recurrence time is given by:

$$m_{[n \to n]} = \frac{(M+1)^n}{M}(M+1)m_1 \qquad (12)$$

(Cfr., for a proof, eq. (27) and eq. (11) in ref. [16]). Note that equation (12), valid for an infinite chain, looks the same as the asymptotic expression of equation (11) for the absorption time of the finite chain. It turns out, as intuitively expected, that the average time to surpass position n and the mean recurrence time of same position both grow exponentially with the distance.

So, the emergence of irreversible behavior in systems modeled by finite - yet very large - absorbing chains is intended in a still

stronger sense than the Smoluchowski's criterion did, as the elapsed times involved are so long, compared to any ordinary experience, that the guess of the irreversibility (viz non recurrence) of the flow of marbles through these chains can be fully trusted. For example, a $C(15,9)$ assembly, that is an assemblage of fifteen chambers connected by one forward connector and nine backward connectors (see diagram of figure 4), represents a span of time of about $6.6 \cdot 10^9$ years – far beyond the age of the solar system (about $4.5 \cdot 10^9$ years) – because the marble would spend that time to escape. With the addition of a single retrograde path, we realize a $C(15,10)$ assembly, which represents about $2.8 \cdot 10^{10}$ years or twice the age of the universe (about $1.4 \cdot 10^{10}$ years).

10 Irreversibility Vs. Recurrence

To illustrate recurrence, the standard pedagogical approach contemplates urns containing numbered objects that are forced to change urn by some withdrawal mechanism. The relationship between the number N of marbles returning on average to the zero-entropy state (the whole of the marbles in *urn0*), the number *M+1* of urns (*2* in the basic version of the experiment), and the expected recurrence interval $m_{[N \to N]}$ takes the form $m_{[N \to N]}/m_1 = (M+1)^N$, being m_1 the time slice spent to force a change of urn and make an observation. Note that the left-hand side of latter relation may be interpreted as the average number of obser-

vations (or time steps) expected until a recurrence incurs. This also means that for a given number of observations, the marbles ever-returning are limited on average to $N \leq \log_{M+1}$ [number of observations].

Unfortunately, a classroom activity on this subject is workable only for trivial values of N. For example, assuming it would take 5 seconds to carry out one observation in the simplest two-urn experiment (extracting, moving, counting, etc.), the work required to observe the return of 10 marbles requires about 5,120 seconds, far beyond 1 hour. However, simple Excel simulations conducted by other authors have shown that all the main features of the Ehrenfest's model can be tested in a few seconds [32].

Computer simulations suffer, however, serious limitations as the number of marbles begins to increase, due to the exponential growth of the computer's time steps required: if we start, for example, with $N = 100$ marbles in *urn0*, then the expected return time is $2^{100}m_1$. Even if each simulation step required a billionth of a second, the entire run would take about $4 \cdot 10^{13}$ years to complete, or roughly 3,000 times the length of time that the Universe has existed thus far! Simulation runs with $N \cong 100$ (or higher) are therefore stopped at a very early stage, but even a limited number of run-steps will suffice to highlight the convergence towards equilibrium, which is attained rather quickly. For N values within this order of magnitude, simulations give evidence that the amplitude of the fluctua-

tions decreases with N, so there is no hope of incurring a return. For a workable compromise, it would need to drop to $N \cong 40$ (or lower) to observe recurrence.

Despite the above limitations, the simulations allow the following key points to be highlighted:

- the equilibrium state is the state of maximum probability;

- the attractor character of the equilibrium state on systems far from it (A far-from-equilibrium system tends to evolve towards equilibrium);

- the inverse process (spontaneous system transition from a more probable to a less probable state) is always possible, yet improbable.

Concrete classroom experiences conducted in past years by other authors proved that two-urns simulations can support the appropriation of the concepts of statistical equilibrium, irreversibility, entropy, and unidirectionality of time[12] [33].

Although with some differences, urn and random walk models have various similarities and share important characteristics. It was Kac, in 1947, who pointed out the equivalence between the Ehrenfest's two-urn problem and the discrete random walk

formulation of the Brownian motion of an elastically bound particle, when the excess over $N/2$ of marbles in *urn0* is interpreted as the displacement of the particle [26].[13]

With no doubt both models contain the same main message: the emerging discrepancy between irreversibility and recurrence observed at the macroscopic level, when the statistical behavior of the system takes relevance. To realize it, it is worth comparing characteristic "diffusion" times of the two models by running simple simulations. Using Excel we have carried out a sample mapping of elapsed times for random walkers transiting across unidimensional chains (typically non-recurrent) and recurrence times in urn-like and Sisyphus models. It turns out that all these characteristic times tend to be comparable with each other as the number N of marbles initially in *urn0* is interpreted as the distance n covered by the walker, the random walk is increasingly asymmetrical (viz M increasing) and the marbles expand in M urns before returning to *urn0*. Summary results of our simulations are plotted in figures 5 and 6 (where unit time steps are assumed).

11 A physical argument: effusion of a low-pressure gas

Returning at last to the question at the bottom of section 2, concerning the passage

---

[12]Simulations of this type were carried out without a computer directly with students: each student represents a particle in an initial state (*urn0*), the teacher takes the name of a student at random and pushes him to move… but these experiments are (obviously) limited!

[13]Kac himself attributed to Schrödinger and Kohlrausch in 1926 the original insight about the connection between the two models (see [26], p. 380).
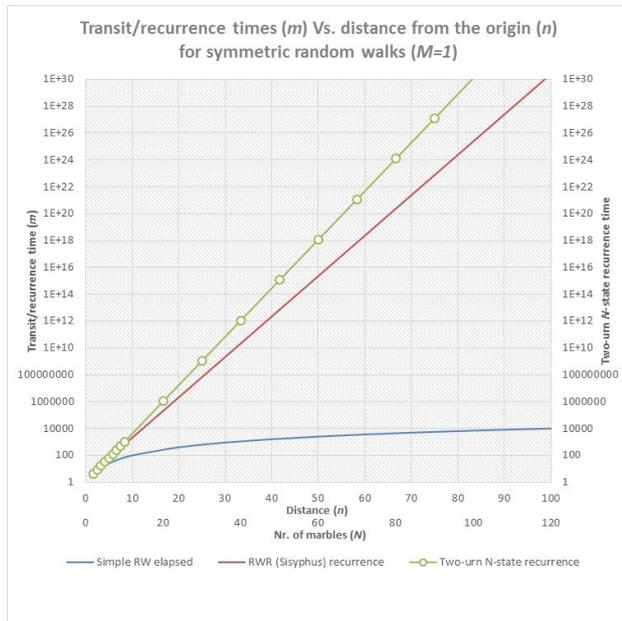
(a) Symmetric ($M = 1$) RW elapsed (blue) and RWR recurrence (red) times compared with two-urn recurrence time (dotted line, secondary scales).
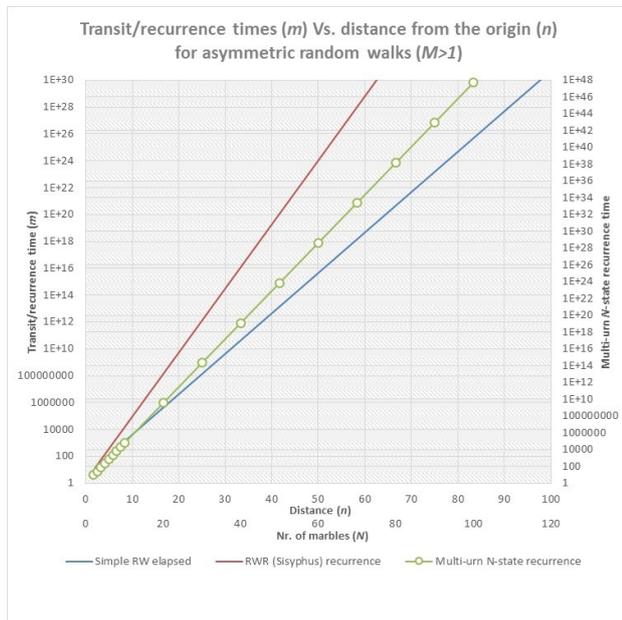
(b) The same as (a) for the asymmetric case ($M > 1$). Data are plotted for $M=2$. The marbles expand in two urns before returning to *urn0*.

Figure 5: Logarithmic plot of transit/recurrence time versus distance of the walker from origin for different systems.
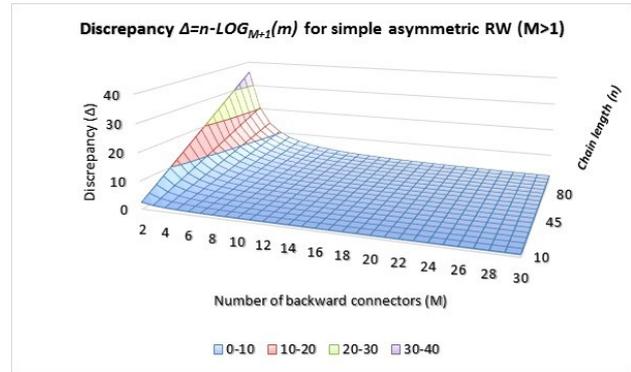
Figure 6: For the simple asymmetric random walk ($M > 1$), the discrepancy from a pure exponential growth ($\Delta = n - \log_{M+1}$ [elapsed time], vertical axis) is plotted versus the number of backward connectors (M, horizontal axis) and of the chain length (n, depth axis). The resulting surface rises steeply for high n and flattens out for high M.

from the experiment with marbles and bottles to the microscopic scale, let us consider, in place of marbles, particles of molecular size in motion at thermal velocities in sealed containers (imagine that there is a vacuum outside the container). If a pinhole large enough for particles to fit through is punctured in a wall of the container, a "gas" leakage from the pinhole is observed, with a characteristic effusion time.

Particle escape times from confined regions have been evaluated by various methods, either numerically by means of classical mechanics (see [34] for an example) or by direct random walk Monte Carlo simulations of the effusive process [35]. A rough idea of the typical time of a particle escape process can be attained through a parallel with low-pressure (say < 0.1 Torr) effusion

experiments with gases [36]. The characteristic effusion time for air in a 1 liter container (in a vacuum) at room temperature (25 °C) punctured by a 1 mm^2 hole turns out to be about 8.6 s.[14]

Even if one brings down the characteristic time m_1 to the time scale of seconds, thereby consistently alleviating the particle trapping effect in the relatively simple assemblies of figures from 1 to 3, yet for chains of increasing length and complexity (like that of figure 4) the escape time quickly reaches values experimentally inaccessible.

12   Conclusions

The models discussed in this paper can be a valuable educational tool, fostering a basic understanding of the statistical nature of the irreversible behavior of macroscopic systems. Urn-like models are traditionally considered advantageous for acquainting students with the concept of thermodynamic equilibrium and with the statistical origin of macroscopic irreversibility. On the other hand, the escape of a random walker out of a long chain of communicating compartments - similar to the escape from a maze - exemplifies in a more immediate sense

an "improbable" process, possible in principle but that would require an unreachable amount of time in order to be actually experienced, close to - and in some sense stronger than - the Smoluchowski's conception of irreversibility (irreversible process = non-recurrent initial state in any conceivable experiment).

Taken together, the two models give evidence that macroscopic irreversibility does not manifest itself at a fundamental level but as a result of statistical nature embodied in the macroscopic approach. Oppositely, the reversible character of phenomena, when examined at the microscopic level, appears to be ineluctable.[15]

Another advantage of these models is that both are very intuitive and do not require that students have any prior knowledge of the subject. In addition, the number of variables to be understood is extremely limited: only one variable - the initial number N of marbles in *urn0* (or *box1* in the Marble Game) -, or two - the number n of chain nodes and the number M of backward connectors - for the chain assembly.

Finally, both models give an occasion for rich disciplinary and interdisciplinary insights on the meaning of time, irreversibility, time arrow and other advanced topics. However, a dose of caution is needed. It should be kept in mind that if the level of subjects covered by these topics is very high,

---

[14]The characteristic time of effusion (or "relaxation time") is given by $\tau = 4V/A\bar{v}$, where: $V =$ Volume of the container, $A =$ Area of the pinhole, $\bar{v} =$ Average speed of molecules (467 m∕s for air at 25 °C), provided that the pressure outside the pinhole is essentially zero (in practice $\lesssim 10^{-5}$ Torr); also $\bar{v} = \sqrt{8RT/\pi M}$, where $M =$ Molar mass (28.97 g for air) [36].

[15]Be careful not to overemphasize reversibility at the micro level: examples were given of extremely simple, reversible, insulated systems that exhibit irreversible statistical behavior [37].

they cannot be developed in all types of secondary schools and careful consideration must be paid to the skill level and interest of students.

Acknowledgments

References

[1] P. H. Nelson, Biophys J 104, 532a (2013).

[2] C. A. Pickover, Mazes for the Mind: Computers and the Unexpected (St Martins Press, New York, 1994), p. 272.

[3] C. A. Pickover, Time: A Traveler's Guide (Oxford University Press USA, New York, 1998), p. 206.

[4] B. J. Alder and T. K. Wainwright, Sci. Am. 201, 113 (1959).

[5] K. Pearson, Nature 72, 294 (1905).

[6] L. Daboni, in Enciclopedia della scienza e della tecnica (Mondadori, Milan, 1980), v. 11, p. 695.

[7] G. H. Weiss, Am. Sci. 71, 65 (1983).

[8] G. H. Weiss, Aspects and Applications of the Random Walk (North-Holland, Amsterdam, 1994).

[9] A. DasGupta, Probability for Statistics and Machine Learning (Springer, New York, 2011), p. 437.

[10] E. J. Saltzman and M. Muthukumar, J. Chem. Phys. 131, 214903 (2009).

[11] P. S. Burada, P. Hänggi, F. Marchesoni, G. Schmid, and P. Talkner, ChemPhysChem 10, 45 (2009).

[12] Z. Xu, U. Bagci, B. Foster, A. Mansoor and D. J. Mollura, in Med Image Comput Comput Assist Interv., Nagoya, 2013, edited by K. Mori, I. Sakuma, Y. Sato, C. Barillot and N. Navab (Springer-Verlag, Berlin Heidelberg, 2013), v. 2, p. 559.

[13] S. Redner, A guide to first-passage processes (Cambridge University Press, New York, 2001), p. 76.

[14] C. A. Pickover, private communication (Sep. 8, 2007).

[15] Leg (https://math.stackexchange.com/users/228265/leg), Random walk on natural number, URL (version: 2015-05-17): https://math.stackexchange.com/q/1283488.

[16] M. Montero and J. Villarroel, Phys. Rev. E 94, 032132 (2016).

[17] M. Kac, Probability and Related Topics in Physical Sciences (Interscience, London, 1959).

[18] P. Ehrenfest and T. Ehrenfest, Phys. Z. 8, 311 (1907).

[19] Y.-M. Kao and P.-G. Luan, Phys. Rev. E 67, 031101 (2003).

[20] P.-G. Luan and Y.-M. Kao, Phys. Rev. E 69, 022102 (2004).

[21] J. Clark, M. Kiwi, F. Torres, J. Rogan and J. A. Valdivia, Phys. Rev. E 92, 012103 (2015).

[22] S. Song and Q. Yao, arXiv:1610.09745 (2016).

[23] J. Güémez, S. Velasco, and A. Calvo Hernández, Am. J. Phys. 57, 828 (1989).

[24] E. D. Banta, J. Math. Phys. 8, 387 (1967).

[25] M. J. Klein, Phys. Rev. 103, 17 (1956).

[26] M. Kac, Am. Math. Mon. 54, 369 (1947).

[27] M. Falcioni and A. Vulpiani, Meccanica Statistica Elementare (Springer-Verlag Italia, Milan, 2015).

[28] W. M. Haddad, V. Chellaboina and S. G. Nersesov, Nonlinear Anal Real 9, 250 (2008).

[29] J. G. Kemeny, J. L. Snell and A. W. Knapp, Denumerable Markov Chains, (Springer-Verlag, New York, 1976), p. 112.

[30] M. Denny, Eur. J. Phys. 21, 421 (2000).

[31] W.-K. Ching and M. S. Lee (2005), Internat. J. Math. Ed. Sci. Tech., 36, 680 (2005).

[32] D. Neri and V. Innocenti Sedili, Didattica delle Scienze 241, 24 (2006).

[33] A. Marie-Jeanne, F. Beau and M. Janvier, in Actes en ligne du colloque International ITEM, Reims, 2003, edited by J. B. Lagrange & al. (IUFM,Reims,2003), <edutice-00001351>.

[34] R. De Luca, Rev. Bras. Ens. Fis. 32, 3306 (2010).

[35] Y. Levin, M. A. Idiart and J. J. Arenzon, Physica A 354, 95 (2005).

[36] D. P. Shoemaker, C. W. Garland and J. I. Steinfeld, Experiments in Physical Chemistry. 3rd edition (McGraw-Hill, New York, 1974), p. 126.

[37] A. Hobson, Am. J. Phys. 34, 411 (1966).

Multi-facets of one dimensional δ-function potential

Debnarayan Jana[1]

[1]Department of Physics

University of Calcutta

92 A PC Road, Kolkata- 700009, India.

djphy@caluniv.ac.in

Submitted on 23-01-2018

Abstract

In this article, we would like to discuss the implications of bound state and scattering states of δ function potential of non-relativistic quantum mechanics in one dimension. The various expectation values in this potential have been illustrated to satisfy the virial theorem. The results of Kronig-Penny model for one dimensional lattice can be nicely demonstrated via this potential. Later on, some perturbation problems in this system have been critically examined to connect the completness relation. The results obtained have been compared with the exact results.

Keywords: δ function potential, bound states, scattering states, expectation values, Kronig-Penny Model

1   Introduction

Historically, this distribution rather than the function was introduced for the modelling of the density of idealized point mass or point charge. The Dirac delta-function denoted as $\delta(x)$ is strictly speaking not a function but a limit of sequence in which it is zero everywhere except at $x = 0$. At $x = 0$ it is infinite in such a way that the area under the curve is equal to unity. Here it is to be noted that during the integration the point $x = 0$ is to be included. It is clear from the above definition that no known real potential does indeed satisfy the criteria set by the delta function potential but can be used judiciously as an approximation in many real physical problems [1, 2, 3]. Besides due to the identity $\int_{-\infty}^{\infty} \delta(x)\, dx = 1$ we can conclude that $\delta(ax) = \frac{1}{|a|}\delta(x), \delta(-x) = \delta(x)$ and the dimension of $\delta(x)$ is $M^0 L^{-1} T^0$. One may also imagine the emergence of delta function in the following way. If one considers a finite potential well of strength α existing between $x = -a/2$ to $x = a/2$. Now, if we take the width a as a limiting

process going to zero in the context of the depth α tending to infinity in such a way that the product $a\alpha = V_0$ remains constant. In this limit, the potential turns out to be a $V(x) = -V_0\delta(x) = -\delta(x/V_0)$. Again, pure dimensional analysis [4, 5] confirms that the binding energy $E_B \propto -\frac{mV_0^2}{\hbar^2}$. The larger the value of the strength of the potential V_0, the lower is the energy. If we change the strength to double the present value, then the binding energy becomes four times the previous one. Therefore, the numerical value of V_0 essentially indicates how deeply the particle is bound in a stationary state.

This being a one-dimensional problem, there is no degeneracy associated with it. Moreover, the upper limit of the number of bound states $N < 1 + \frac{2m}{\hbar^2} \int_{-\infty}^{\infty} |x| \, V^-(x) \, dx$ with $V^-(x)$ being the absolute value of $V(x)$ indicates there are indeed finite number of bound states [6, 7, 8, 9, 10, 11] in contrast to infinite number bound states observed in particle in a box or harmonic oscillator potential. A well-known theorem for one-dimensional quantum mechanics based on varational principle [6, 12] indicates that no matter how small V_0 is, the potential well is able to support at least one bound state. Besides, the Hamiltonian is invariant under the transformation $x \rightarrow -x$ so that the Parity operator commutes with Hamiltonian. As a result, by the condition of the non-degeneracy in one dimension, the eigenfunctions must be simultaneous eigenfunctions of both Hamiltonian as well as Par-

ity operator. In other words, the energy eigen functions must be even or odd. The Schrödinger equation for this one dimensional potential is given by

$$-\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2} - V_0\delta(x)\Psi(x) = E\Psi(x) \qquad (1)$$

Note that except at $x = 0$ where the δ function fires, the Schrödinger equation is simple $\frac{\partial^2\Psi}{\partial x^2} = k^2\Psi$. As a result, one can choose the wave function as $\exp(\pm kx)$ for $x \rightarrow -\infty$ and $+\infty$ respectively. The first derivative of the wave function also possesses a discontinuity at the origin which can be easily seen by integrating the equation (1) over an infinitesimal interval. In this way the exact calculation demonstrates [13, 14, 15, 16] that there is only one bound state of magnitide $E_B = -\frac{mV_0^2}{2\hbar^2}$ and the ground state $\Psi_0 = \sqrt{\frac{mV_0}{\hbar^2}}\exp\left(-\frac{mV_0}{\hbar^2}|x|\right)$ like other quantum mechanical problem is nodeless. A careful look into the problem indicates that there is indeed a length scale $a_0 = 1/K_0$ ($K_0 = \frac{mV_0}{\hbar^2}$) associated with the problem. In terms of this length scale, the ground state energy as well as the wave function remarkably match with those of hydrogen atom problem ($V(r) = -\frac{V_0}{r}$, $E_n = -\frac{mV_0^2}{2\hbar^2}\frac{1}{n^2}; \Psi_{100} = \frac{1}{\sqrt{\pi a_0^3}}e^{-r/a_0}$).

For the positive energy case ($E > 0$), the relevant wave function with free particle energy $E_k = \frac{\hbar^2 k^2}{2m}$ with odd parity can be chosen[17, 18, 19] as

$$\Psi_k(x) = \frac{1}{\sqrt{\pi}}\sin(kx) \qquad (2)$$

This form of the wave function originates from the nomalization condition

$$\lim_{\epsilon \to 0} A(k) A(k') \int \sin(kx) \sin(k'x) e^{-\epsilon |x|} \, dx$$
$$= \delta(k - k') \qquad (3)$$

The pecularity associated with the above wave function is that it is dimensionless. Typically, in any one dimensional bound state problem, the wave function must have the dimensions of $M^0 L^{-1/2} T^0$. This is due to the cancellation of Dirac-delta from both sides of the above equation. This odd form of the eigen function is important for non-zero dipole matrix element between ground state and continumm state. There also exists the even (parity) state positive energy [17, 18, 19] solution given by

$$\Psi_k(x) = \frac{1}{\sqrt{\pi(k^2 + k_0^2)}} [K_0 \sin k|x| - k \cos(kx)] \qquad (4)$$

with $K_0 = \frac{mV_0}{\hbar^2}$. The importance of these states will be demonstrated later related to perturbation problems.

Two-dimensional delta function potential has served one of the toy model in different branches of theoretical physics including regularization and renormalization methods in particle physics context [20, 21, 22, 23, 24]. The paper is organized as follows. In section 2, we briefly discuss the alternative paths leading to bound state energy solutions. The famous Kronig-Penny model in solid state physics is connected with this potential in section 3. The various expectation values connected with the

stationary bound state energy solutions are demonstrated in section 4. In section 5, few important results of three different perturbating potentials are highlighted. The conclusions are noted in section 6.

2. Review of bound state and scattering state results

We would like to approach the problem in two ways - one by using Fourier transform [1, 25, 26, 27, 28] and another one through scattering matrix [29]. We can recast the Schrödinger equation (1) as

$$\frac{\partial^2 \Psi}{\partial x^2} + \lambda_0 \delta(x) \Psi(x) = B_0 \Psi(x) \qquad (5)$$

with λ_0 is a dimensionless constant given by $\frac{2mV_0}{\hbar^2}$ while $B_0 = -\frac{2mE}{\hbar^2}$ is indeed a constant having dimensions. By defining the Fourier transform of $\Psi(x)$ as

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \int e^{-ikx} \Phi(k) \, dk$$
$$\Psi(0) = \frac{1}{\sqrt{2\pi}} \int \Phi(k) \, dk \qquad (6)$$

we find the Fourier transform of the equation (5) as

$$- k^2 \Phi(k) + \lambda \Psi_0 = B_0 \Phi(k) \qquad (7)$$

As a result, the Fourier transform $\Phi(k)$ can be written as

$$\Phi(k) = \frac{\lambda \Psi_0}{k^2 + B_0} \qquad (8)$$

Integrating over k and throwing out the irrelevant term $\Psi(0)$ from both sides, we obtain

$$\frac{1}{\lambda_0} = \int \frac{dk}{k^2 + B_0} = \frac{m}{B_0 \hbar^2} \qquad (9)$$

Thus, the binding energy reduces $E = -\frac{mV_0^2}{2\hbar^2}$. The normalized ground state wave function can be obtained from the Fourier transform as

$$\Psi_0 = \sqrt{\frac{mV_0}{\hbar^2}} \exp\left(-\frac{mV_0}{\hbar^2}|x|\right) \qquad (10)$$

We can also obtain the bound state from the scattering matrix formulation as follows. For $x < 0$, the wave function can be written as

$$\Psi_I(x) = A_1 e^{ikx} + B_1 e^{-ikx} \qquad (11)$$

while for $x > 0$

$$\Psi_{II}(x) = A_2\, e^{ikx} + B_2\, e^{-ikx} \qquad (12)$$

The continuity of the wave function at the orgin $x = 0$ gives us the condition

$$A_1 + B_1 = A_2 + B_2 \qquad (13)$$

while the discontinuity of the first derivative of the wave function restricts the coefficients as

$$ik(A_2 - B_2) - ik(A_1 - B_1) = \frac{2mV_0}{\hbar^2}(A_1 + B_1) \qquad (14)$$

The above two equations can be rearranged in a matrix form given by

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = D(V_0, k) \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \qquad (15)$$

with the transfer matrix $D(V_0, k)$ is given by

$$D = \begin{pmatrix} 1 + \frac{mV_0}{i\hbar^2 k} & \frac{mV_0}{i\hbar^2 k} \\ -\frac{mV_0}{i\hbar^2 k} & 1 - \frac{mV_0}{i\hbar^2 k} \end{pmatrix} \qquad (16)$$

It is interesting to note that $\det D(V_0, k) = 1 = \det D(-V_0, k)$ which indicates the conservation of probabilities. Besides it is important to find that

$$D(V_0, k)D(-V_0, k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad (17)$$

This equation (17) indicates that the delta function spike and a delta function potential at the same point will cancel each other as if there is no potential ($D(0, k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$). From, the transfer matrix, the reflectvity coefficient R can be read as

$$R = \frac{\frac{mV_0^2}{2\hbar^2 E}}{1 + \frac{mV_0^2}{2\hbar^2 E}} \qquad (18)$$

while the transmission coefficient T is related with the parameters of the model as

$$T = 1 - R = \frac{1}{1 + \frac{mV_0^2}{2\hbar^2 E}} \qquad (19)$$

It is amazing to note that both R and T become infinitely large at the binding energy $E = -\frac{mV_0^2}{2\hbar^2}$. We also notice that the reflection and transmission coefficients are a function only of the square of V_0, so that we obtain the same result as for the potential well too. In other words, the transmission and reflection coefficients are identical [29] with a delta-function well or barrier.

3. Relation with Kronig-Penny Model

As an application of the delta function potential, we would like to investigate the sta-

tionary energy eigenstates of electrons in a one dimensional crystal. In such a lattice electrons invariably observe a periodic potential due to the presence of the atoms through the well-known Coulomb potential (shown in Fig 1). In the literature of the basic solid state course, this model is nothing but the Kronig-Penny model [29, 30]. Because of the long range nature of Coulomb potential, one can even simply take the above potential by some constant periodic potential as is done in text book of solid state; here we reduce the width of potential to zero to represent them as a series attractive delta functions of same strength, each of them being separated by the lattice constant a. Therefore, simplified potential energy can be expressed as

$$V(x) = -V_0 \sum_{n=-\infty}^{\infty} \delta(x - na) \qquad (20)$$

Note that in this one-dimensional model, interactions between electrons are neglected. However, one important feature of this model with the realistic solids is that the electrons are moving in a periodic potential given by $V(x + a) = V(x)$. For a sample of size $L = Na$, N being the number of atoms, the periodicity can be verified if the periodic delta function is defined as

$$\delta(x) = \frac{1}{2\pi} \sum_m e^{2\pi imx/L} \qquad (21)$$

In figure 1, these two potentials are compared. As a result, the translational symmetry remains intact only the form of the potential is changed so that one can get the
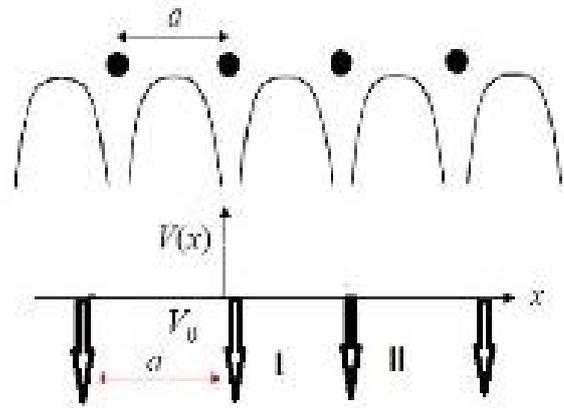
Figure 1: Schematic diagram of periodic potential (top) and representative periodic delta function potential (bottom) in a one dimensional lattice.

final solution easily. Due to this periodic property of the potential, Bloch's theorem predicts that each energy eigenstate of the Schrödinger equation must satisfy

$$\Psi(x + a) = e^{iKa}\Psi(x) \qquad (22)$$

for some value of K. In fact, it is this property which remarkably simplifies the above problem and helps to connect the eigenstates to find out the allowed energy eigen values of the problem. In region I, the free particle wave function can be written as

$$\Psi_I(x) = Ae^{ikx} + Be^{-ikx}, k^2 = \frac{2mE}{\hbar^2} \qquad (23)$$

Now, using Bloch Theorem, we can write down the wave function in region II as

$$\begin{aligned} \Psi_{II}(x) &= \Psi_I(x - a)e^{iKa} \\ &= \left[Ae^{ik(x-a)} + Be^{-ik(x-a)} \right] e^{iKa} \end{aligned}$$
$$(24)$$

Now, to find out the coefficients A and B, we use the continuity of the wave functions and discontinuity of the first derivatives of the wave functions in the two regions. The continuity requirement gives us

$$A(e^{iKa} - e^{ika}) = B(e^{-ika} - e^{iKa}) \qquad (25)$$

Similarly, the discontinuity of the derivatives of the wave functions is satisfied by another equation given by

$$A\left[ike^{iKa} - ike^{ika} - \frac{2mV_0}{\hbar^2}e^{iKa}\right]$$
$$= B\left[ike^{iKa} - ike^{-ika} + \frac{2mV_0}{\hbar^2}e^{iKa}\right]$$
$$\qquad (26)$$

Using equation (25) and (26), we eliminate the coefficients A and B to obtain the equation

$$\cos(Ka) = \cos(ka) + \frac{2mV_0a}{\hbar^2}\frac{\sin(ka)}{ka} \qquad (27)$$

Interestingly, the combination $\frac{2mV_0a}{\hbar^2} = U_0$ is a dimensionless parameter of the model that characterizes the strength of the periodic potential. It is to be noted that the LHS is bounded in the region $[-1,1]$. This eventually puts on restriction on the allowed values of k. Thus, like constant periodic potential, this periodic potential also opens gaps in the dispersion relation (as shown in figure 2). Moreover, in the limit of $V_0 \to 0$ and $V_0 \to \infty$, $E = \frac{\hbar^2 k^2}{2m}$ and $E = \frac{n^2\pi^2\hbar^2}{2ma^2}$ respectively are recovered.

However, at this stage we would like to play with the equation (27) to rediscover some important features about electrons in
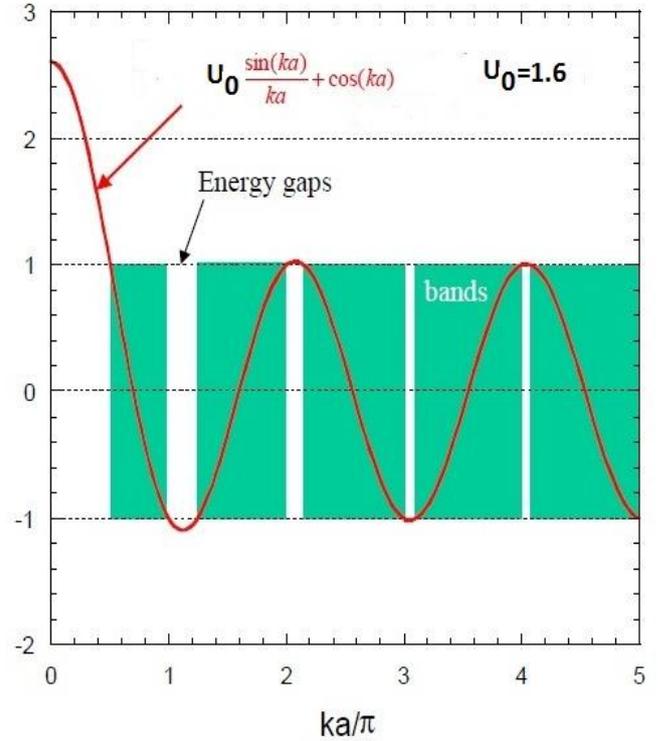
Figure 2: Graphical solution of Kronig-Penny model with emergence of energy band gaps.

a periodic system. In the limit of very large value of U_0, one can expand the RHS of equation (27) near zeroes of $\frac{\sin x}{x}$ function. The first zero of this function occurs at $x = \pi$, therefore we can choose $ka = \pi - \delta$ with $\delta \ll 1$. As a result to first order in δ, we can approximate the RHS as

$$\cos(Ka) = -1 + \frac{U_0}{\pi}\delta$$
$$k = \frac{\pi}{a}\left[1 - \frac{1}{U_0}(1 + \cos(Ka))\right]$$
$$\qquad (28)$$

Therefore, the allowed energy can be written as

$$
\begin{aligned}
E &= \frac{\pi^2\hbar^2}{2ma^2} - \frac{\pi^2\hbar^4}{m^2a^3U_0}[1+\cos(Ka)] \\
&= E_0 - 2J(1+\cos(Ka)) \quad (29)
\end{aligned}
$$

with $E_0 = \frac{\pi^2\hbar^2}{2ma^2}$ and $J = \frac{\pi^2\hbar^4}{2m^2a^3U_0}$. The equation (29) reminds us the energy solution of a typical tight-binding Hamiltonian [30]

$$
H = -t\sum_i(c_{i+1}^\dagger c_i + h.c.) + V\sum_i c_i^\dagger c_i \quad (30)
$$

in one dimension, with $c_i^\dagger (c_i)$ being the creation (annihilation) operator of electron at site i and V is a constant potential. The operators do satisfy the following anticommutations relations given as:

$$
\begin{aligned}
\{c_i, c_j^\dagger\} &= \delta_{ij} \\
\{c_i^\dagger, c_j^\dagger\} &= 0 \\
\{c_i, c_j\} &= 0 \quad (31)
\end{aligned}
$$

These relations ensure that the number operator $n_i = c_i^\dagger c_i$ can take only 0 and 1 values satisfying Pauli exclusion principle. The identification of the parameters yields that $t = J$ and $V = E_0 - 2J$. Thus, in the limit $U_0 \gg 1$, we rediscover the free particle energy $E_0 \gg J$. Besides, we can also rewrite the above dispersion relation (29) as

$$
\begin{aligned}
E &= E_0 - 2J - 2J\cos(Ka) \\
&= E_0 - 4J + 2J(1-\cos(Ka)) \\
&= E_0' + 2J(1-\cos(Ka)) \quad (32)
\end{aligned}
$$

Thus suitably redefining the zero of the energy at E_0', we can recover the famous tight binding dispersion result in one dimension

$$
E = 2J(1-\cos(Ka)) \quad (33)
$$

with bandwidth (the difference between the maximum and minimum energy) of $4J$. For Dirac electrons, transmission and conductance properties have been studied through delta function barriers [31].

4. Expectation values in δ function potential

Before we verify the expectation values associated with the δ function potential, let us discuss some generic feature in connection with stationary energy states of any potential in quantum mechanics. First of all, the expectation value of the momentum in any energy eigenstates vanishes. This can be easily visualized in the following way. We know that

$$
[H, x] = \left[\frac{p^2}{2m} + V(x), x\right] = -\frac{i\hbar}{m}p \quad (34)
$$

Now, considering the expectation value of the above expression from both sides, we identify that

$$
-\frac{i\hbar}{m}\langle\Psi_n|p|\Psi_n\rangle = (E_n - E_n)\langle\Psi_n|x|\Psi_n\rangle = 0 \quad (35)
$$

Considering the bound ground state wave function of the potential as depicted in the introduction section, we note that

$$
\begin{aligned}
&< \Psi_0|x|\Psi_0 > \\
&= \left(\frac{mV_0}{\hbar^2}\right)\int_{-\infty}^{\infty} xe^{-2mV_0|x|/\hbar}\,dx \\
&= 0 \quad (36)
\end{aligned}
$$

while

$$< \Psi_0 | x^2 | \Psi_0 >$$
$$= \left(\frac{mV_0}{\hbar^2} \right) \int_{-\infty}^{\infty} x^2 \, e^{-\frac{2mV_0|x|}{\hbar^2}} \, dx$$
$$= \frac{\hbar^4}{2m^2 V_0^2} \qquad (37)$$

In the same spirit, it is easily verified that

$$< \Psi_0 | p | \Psi_0 >= 0; \quad < \Psi_0 | p^2 | \Psi_0 >= \frac{m^2 V_0^2}{\hbar^2} \qquad (38)$$

With these expectation values we can easily verify the uncertainty relation:

$$\Delta x \Delta p = \frac{\hbar}{\sqrt{2}} \qquad (39)$$

We can use Feynman-Hellmann (FH) theorem [32] to calculate the expectaion values of the kinetic energy and potential to verify the Virial Theorem as follows. FH theorem predicts that

$$\left\langle \Psi_0 \left| \frac{\partial H}{\partial V_0} \right| \Psi_0 \right\rangle = -\frac{mV_0}{\hbar^2}$$
$$\left\langle \Psi_0 \left| \frac{\partial H}{\partial m} \right| \Psi_0 \right\rangle = -\frac{V_0^2}{2\hbar^2} \qquad (40)$$

A quick inspection reveals that

$$2 \langle T \rangle + \langle V \rangle = 0 \qquad (41)$$

This kind of relation is valid for generic Coulomb type potential. This relation also can be verified directly with the help of the bound ground state wave function explicitly.

In the same way, we note the quantum mechanical covariance

$$[H, x^2] = -\frac{i\hbar}{m} (xp + px) \qquad (42)$$

The similar approach of the expectation values in the energy eigen states yields:

$$\langle \Psi_n | (xp + px) | \Psi_n \rangle = 0 \qquad (43)$$

Let us explicitly verify the above identity for the delta function potential. A direct calculation demonstrates that

$$< \Psi_0 | xp | \Psi_0 >$$
$$= \left(\frac{mV_0}{\hbar^2} \right) \int_{-\infty}^{\infty} x e^{-mV_0|x|/\hbar^2} p e^{-mV_0|x|/\hbar^2} dx$$
$$= \frac{i\hbar}{2} \qquad (44)$$

and similarly

$$\langle \Psi_0 | px | \Psi_0 \rangle = -\frac{i\hbar}{2} \qquad (45)$$

This immediately verifies the fundamental operator relation (43). This on the other hand confirms the fundamental commutation relation in quantum mechanics $\langle \Psi_0 | xp - px | \Psi_0 \rangle = i\hbar$. All these expectaion values are important and useful to verify the various sum rules involved in non-relativistic quantum mechanics [33, 34].

5. Perturbation Results

Coupling delta function potential with some exactly soluble potential can yield exact results. Besides, the strength of the delta function can be used as an expansion parameter to yield the result of certain infinite series [35].

We would like to discuss three cases of perturbation on the delta function potential in one dimension. The first one is trivial one.

Suppose if we add a constant potential U to this problem, then the exact calculation yields $E = U - \frac{m^2 V_0^2}{2\hbar^2}$. If we treat this problem from time-independent non-degenerate perturbation theory, then the first order correction gives us $E_0 = < 0|U|0 > = U$ indicating that the total energy matches with the exact calculation. Then, it remains to show at least other higher order calculations should vanish. Indeed, it can be shown that matrix element

$$< 0|U|k >$$
$$= \sqrt{\frac{mV_0^3}{\pi\hbar^2}} \int_{-\infty}^{\infty} e^{-mV_0|x|/\hbar^2} \sin(kx)dx$$
$$= 0 \qquad (46)$$

Since, all the higher order energy corrections including second order involve this matrix elements, therefore all the higher order energy corrections are identically zero.

Let us add a perturbation of the form $H' = -g\delta(2x) = -\frac{g}{2}\delta(x)$ to the orginal delta function potential $-V_0\delta(x)$. The reason for choosing this particular Hamiltonian is two fold. Firstly, this model Hamiltonian can be solved exactly so that we can compare the success of the perturbation results with it. Secondly, we would like to illustrate the importance of positive energy solution for the second order perturbation calculations.

Interestingly, it turns out that this prob-

lem can be solved exactly as follows:

$$H = \frac{p^2}{2m} - V_0\delta(x) - \frac{g}{2}\delta(x)$$
$$= \frac{p^2}{2m} - (V_0 + g/2)\delta(x) \qquad (47)$$

Therefore, the bound state energy eigenvalue will be simply

$$E_B = -\frac{m(V_0 + g/2)^2}{2\hbar^2}$$
$$= -\frac{mV_0^2}{2\hbar^2} - \frac{mgV_0}{2\hbar^2} - \frac{mg^2}{8\hbar^2}$$
$$= E_0^{(0)} + E_0^{(1)} + E_0^{(2)} \qquad (48)$$

Now let us review the results from the perspectives of perturbation theory. The first order correction to the ground state will be

$$E_0^{(1)} = \langle \Psi_0|H'|\Psi_0 \rangle$$
$$= -\frac{g}{2}\frac{mV_0}{\hbar^2} \int_{-\infty}^{\infty} e^{-\frac{2mV_0}{\hbar^2}|x|} \delta(x)dx$$
$$= -\frac{g}{2}\frac{mV_0}{\hbar^2} \qquad (49)$$

Similarly, the second order correction to the energy becomes

$$E_0^{(2)} = \sum_{a,b} \int_{-\infty}^{\infty} \frac{|\langle \Psi_0|H'|\Psi_k^{a,b} \rangle|^2 \, dk}{E_0^{(0)} - E_k^{(0)}} \qquad (50)$$

For non-vanishing matrix elements, we have to use the positive energy even state eigen function(4). The energy difference in the second order correction to the energy can be written as

$$E_k - E_0 = \frac{\hbar^2}{2m}(k^2 + K_0^2) \qquad (51)$$

By noting that

$$-\frac{g}{2}\left\langle \Psi_0|H'|\Psi_k^+\right\rangle = \frac{\sqrt{K_0}(gk)}{2\sqrt{\pi(K_0^2+k^2)}} \quad (52)$$

we obtain the second order correction to the ground state as

$$\begin{aligned}
E_0^{(2)} &= -\frac{mK_0g^2}{2\pi\hbar^2}\int_0^\infty \frac{k^2dk}{(K_0^2+k^2)^2} \\
&= -\frac{mg^2}{8\hbar^2} \quad (53)
\end{aligned}$$

Therefore, the perturbation results upto second order match with the exact results shown in equation (48).

As a third example, we adopt another Hamiltonian designed as

$$H = \frac{p^2}{2m} - V_0\delta(x) + \frac{\lambda p}{m} \quad (54)$$

The above Hamiltonian can be recast as

$$H = \frac{(p+\lambda)^2}{2m} - V_0\delta(x) - \frac{\lambda^2}{2m} \quad (55)$$

with the binding energy $E_B = -\frac{mV_0^2}{2\hbar^2} - \frac{\lambda^2}{2m}$. It is interesting to note that the first order correction to energy $E_0^{(1)} = \langle \Psi_0|H'|\Psi_0\rangle = 0$ as can be seen from the expectation value of the momentum operator. The first order correction to the ground state wave function can be obtained as

$$\begin{aligned}
\Psi_0^{(1)} &= \sum_k \frac{\langle \Psi_0|H'|\Psi_k\rangle}{E_0^{(0)}-E_k^{(0)}}\Psi_0 \\
&= \left(\frac{\lambda}{m}\right)\left(\frac{mV_0}{\hbar^2}\right)^2\left(\frac{4im}{\hbar\sqrt{\pi}}\right) \\
&\quad \times e^{-\frac{mV_0|x|}{\hbar^2}}\int_0^\infty \frac{kdk}{(K_0^2+k^2)^2} \\
&= \left(\frac{2\lambda i}{\sqrt{\pi\hbar^2}}\right)e^{-\frac{mV_0|x|}{\hbar^2}} \quad (56)
\end{aligned}$$

Considering upto first order correction, the wave function looks like

$$\Psi(x) = e^{-\frac{mV_0|x|}{\hbar^2}}\left[\sqrt{\frac{mV_0}{\hbar^2}} + \frac{2\lambda i}{\sqrt{\pi\hbar^2}}\right] \quad (57)$$

Thus, the condition for the validity of the perturbation seems to be

$$\lambda^2 \ll \frac{m\pi V_0}{4} \quad (58)$$

To compute the second order correction of energy, we have to take into account the odd state positive energy solution (2) for the requirement of the non-vanishing of the matrix elements. The non-vanishing matrix element is given by

$$\langle 0|p|\Psi_k^-\rangle = -(2ki\hbar)\sqrt{\frac{mV_0}{\pi\hbar^2}}\frac{mV_0}{\hbar^2}\frac{1}{(K_0^2+k^2)} \quad (59)$$

Therefore, the second order correction to the energy reduces to

$$\begin{aligned}
E_0^{(2)} &= \frac{\lambda^2}{m^2}\int_{-\infty}^\infty \frac{|\langle\Psi_0|p|\Psi_k^-\rangle|^2\,dk}{E_0^{(0)}-E_k^{(0)}} \\
&= -\frac{\lambda^2}{2m} \quad (60)
\end{aligned}$$

The interested reader can try the Hamiltonian $H = \frac{p^2}{2m} - V_0\delta(x) + \Lambda p^2$ for the perturbation calculation to compare with the exact result $E_B = -\frac{mV_0^2}{2(1+2m\Lambda)\hbar^2}$. For perturbation proportional to x, the appropriate important integral would be

$$\begin{aligned}
&< 0|x|k > \\
&= \sqrt{\frac{mV_0}{\pi\hbar^2}}\int_{-\infty}^\infty e^{-\frac{mV_0|x|}{\hbar^2}}x\sin(kx)\,dx \\
&= 4\sqrt{\frac{K_0^3}{\pi}}\frac{k}{(K_0^2+k^2)^2} \quad (61)
\end{aligned}$$

Another important perturbation $H' = b|x|$ $(b \geq 0)$ when it is applied to this delta function potential, then the first order correction adds to the energy level to give

$$E = E_0^{(0)} + \frac{bm}{2V_0\hbar^2} \qquad (62)$$

In literature, there exists also an exact calculation of a particle in a box with a delta function potential by the factorization method [36, 37] and harmonic oscillator with a delta function potential[38].

6.Conclusions

To conclude, we have discussed the various features of delta function potentials and the role of bound state and scattering states in dealing with matrix elements related to perturbation theory.

Acknowledgments

The author would like to acknowledge the post-graduate students for asking several interesting questions related to this subject which prompted me to write this paper.

References

[1] M. Belloni and R.W. Robinett, *Physics Reports*, **540**, 25 (2014).

[2] Sydney Geltman, *Journal of Atomic, Molecular, and Optical Physics*, **Volume 2011**, Article ID 573179, (Hindawi Publishing Corporation).

[3] T. Busch, B. Englert, K. Rzazewski, and M. Wilkens, *Foundations of Physics*, **28**, 549 (1998).

[4] D. Jana, *Physics Education*, **25**, 35 (2008).

[5] D. Jana, *Dimensional Analysis: Modern Perspectives* (Lambert Academic Publishing (LAP), Germany, 2011)

[6] Maxim Olshanii, *Back -of-the envelope quantum mechanics* (World Scientific, Singapore, 2014).

[7] K. Chandan, N. N. Khuri, A. Martin and T. T. Wu, *J. Math. Phys.*, **44**, 406 (2003).

[8] T. Aktosun, M. Klan and C. van der Mee, *J. Math. Phys.*, **39**, 4249 (1998).

[9] B. Simon, *Ann. Phys.*, **97**, 279 (1976).

[10] M. Sassoli de Bianchi and M. Di Ventra, *J. Math. Phys.*, **36**, 1753 (1995).

[11] R. G. Newton, *J. Operat. Theor.*, **10**, 119 (1983).

[12] D. Jana, *Physics Education*, **29**, (January-March, Article No. 4), (2013).

[13] R. Shankar, *Principles of Quantum Mechanics*, (Plenum, NY, 1980) p 466.

[14] S. Gassiorowicz, *Quantum Mechanics* (2nd edition, John Wiley and Sons, NY, 1986) p 277.

[15] H. Bethe and R. Jackiw, *Intermediate Quantum Mechanics* (2nd edition, Benjamin, NY, 1978), Chap II.

[16] E. Merzbacher, *Quantum Mechanics*, (John Wiley and Sons, NY, 1961) pp 445-447.

[17] K. R. Brownstein, *Am. J. Phys.*, **43**, 173 (1974); W. C. Damert, *ibid.* **43**, 531 (1974);

[18] S. H. Patil, *Am. J. Phys.*, **68**, 712 (2000)

[19] M. Belloni and R. W. Robinett, *Am. J. Phys.*, **76**, 798 (2008) and references therin.

[20] Su-Long Nyeo, *Am. J. Phys.*, **68**, 571(2000).

[21] L. R. Mead and J. Godines, *Am. J. Phys.*, **59**, 935 (1991).

[22] P. Gosdzinsky and R. Tarrach, *Am. J. Phys.* **59**, 70 (1991).

[23] T. Padmanavan, *Resonance*, 510 (June, 2008).

[24] R. M. Cavalcanti, *Rev. Bras. Ens. Fis*, **21**, 336 (1999)

[25] I. R. Lapidus, *Am. J. Phys.*, **37**, 930 (1969).

[26] P. B. James, *Am. J. Phys.*, **38**, 1319 (1970).

[27] M. Lieber, *Am. J. Phys.*, **43**, 486 (1975)

[28] A. Jahan and M. Jafari, *Iranian Journal of Science & Technology*, Transaction A, **31**, 435 (2007).

[29] D. J. Griffiths, *Introduction to Quantum Mechanics* (Prentice Hall, Upper Saddle River, NJ, 2004)

[30] C. Kittel, *Introduction to Solid State Physics*, (5th ed.John Wiley and Sons, New York, 1976).

[31] M. Barbier, P. Vasilopoulos, and F. M. Peeters, *Phys. Rev. B*, **80**, 205415 (2009).

[32] D. Jana, *Physics Education*, **25**, 227 (2008)

[33] Bing-Lu Zhou, Jiong-Ming Zhu, Zong-Chao Yan, *Phys. Rev. A*, **73**, 014501 (2006).

[34] D. Jana, *Physics Education*, Oct-Dec 2017 (33/4/05)

[35] N. Bera, K. Bhattacharya, and J. K. Bhattacharjee, *Am. J. Phys.*, **76**, 250 (2008).

[36] Y. N. Joglekara, *Am. J. Phys.*, **77**, 734 (2009).

[37] P. Pedram and M. Bhavi, arXiv:1001.0311v2[quant-ph].

[38] J. Viana-Gomes and N. M. R. Peres, arXiv:11009.3077v1[quant-ph].

Motion of a rod pushed at one point in a weightless environment in space

Ashok K. Singal

Astronomy and Astrophysics Division, Physical Research Laboratory

Navrangpura, Ahmedabad 380 009, India.

`ashokkumar.singal@gmail.com`

Submitted on 28-12-2017

Abstract

We analyze the motion of a rod floating in a weightless environment in space when a force is applied at some point on the rod in a direction perpendicular to its length. If the force applied is at the centre of mass, then the rod gets a linear motion perpendicular to its length. However, if the same force is applied at a point other than the centre of mass, say, near one end of the rod, thereby giving rise to a torque, then there will also be a rotation of the rod about its centre of mass, in addition to the motion of the centre of mass itself. If the force applied is for a very short duration, but imparting nevertheless a finite impulse, like in a sudden (quick) hit at one end of the rod, then the centre of mass will move with a constant linear speed and superimposed on it will be a rotation of the rod with constant angular speed about the centre of mass. However, if force is applied continuously, say by strapping a tiny rocket at one end of the rod, then the rod will spin faster and faster about the centre of mass, with angular speed increasing linearly with time. As the direction of the applied force, as seen by an external (inertial) observer, will be changing continuously with the rotation of the rod, the acceleration of the centre of mass would also be not in one fixed direction. However, it turns out that the locus of the velocity vector of the centre of mass will describe a Cornu spiral, with the velocity vector reaching a final constant value with time. The mean motion of the centre of mass will be in a straight line, with superposed initial oscillations that soon die down.

1   Introduction

Consider a uniform rod of length l and mass m, freely floating in space in a weightless condition. Suppose a force f is applied at some point on the rod, in a direction perpendicular to the length of the rod. What will be the motion of the rod? The question whether any such rod in space, when pushed at say,

one end of the rod, will have only a linear motion or have only a rotation or possess both, has been argued in various forums on the web [1, 2, 3, 4]. This is a problem involving linear momentum of the centre of mass as well as moment of inertia, angular momentum and rotation about the centre of mass of the system. If the force is applied at the centre of mass C of the rod, then from the law of conservation of momentum, the rod would gain a linear motion in the direction of the applied force. However, if the same force is applied at a point other than the centre of mass, then in addition to the motion of the centre of mass as before, there could also be a rotation of the rod about C, due to a finite torque [5, 6, 7]. If the force is applied continuously, then any rotation of the rod would imply a continuous change in the direction of the applied force and a consequential change in the direction of acceleration of the centre of mass. One expects the combined motion to be quite complicated.

We could imagine the rod (as well as the observer) to be freely floating in space, say, in a weightless environment within a satellite orbiting the Earth (and thus freely falling in Earth's gravitational field) which can then be considered to be an inertial frame, provided any tidal effects over the system dimensions due to Earth's gravitation field could be ignored. We take the length of the rod to be short enough so that we may not be bothered about any light-travel time effects. All movements are also assumed to be slow enough so that no special relativistic effects come into picture. Even the sound speed within the rod, with which one part of the rod material may communicate with other parts, i.e., the speed with which any influence within the rod may travel, is taken to be fast compared to any translational or rotational speeds of the rod for whatever temporal intervals we may be concerned with. In order to provide a continuous force perpendicular to the length of the rod, we could strap a tiny rocket to the rod at a point of our choosing. We further suppose that the rocket system, providing the thrust, makes only an imperceptible, if any, change in the mass of the rod. Further, we take the force, acceleration etc., though perpendicular to the rod, but to be always in the x-y plane, so that the torque, angular momentum and angular velocity vectors will all be along the z-axis, therefore we need to consider only the magnitudes of such vectors and as we shall see, it does not give rise to any ambiguities.

2   An impulse given to one of a pair of independent masses

In a composite system, comprising two or more particles whose relative coordinates may or may not be governed by any constraints between them, any net force which is not passing through the centre of mass of the composite system, will create torque and thereby impart angular momentum to the system about its centre of mass [6].

In order to demonstrate that a linear motion of the centre of mass alone might not
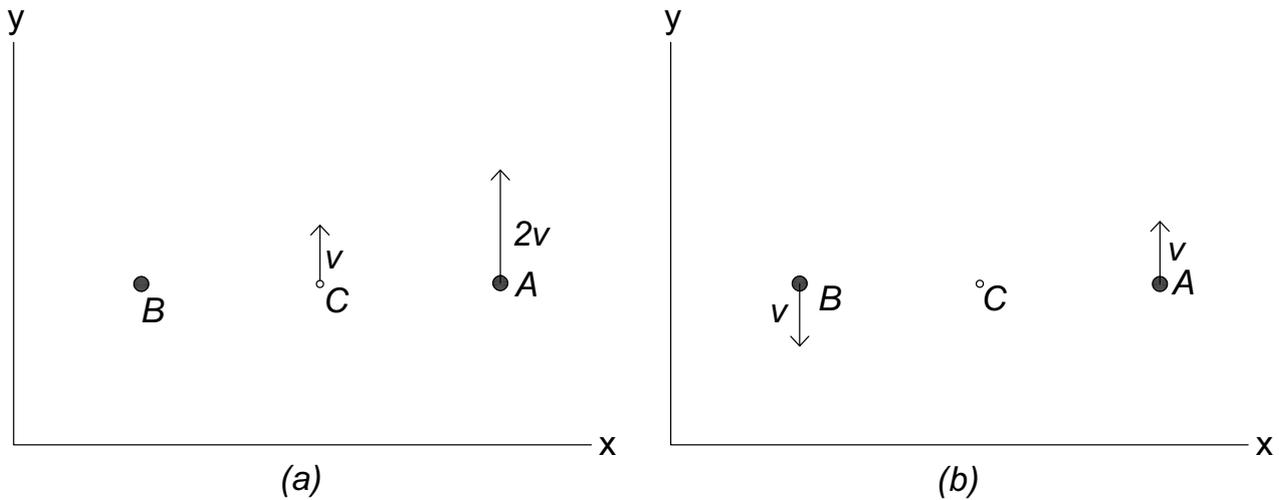
Figure 1: (**a**) Particle A, with mass $m/2$, is given an impulse p, so that it moves with a constant velocity $v_A = 2p/m$ along the y direction. The centre of mass C then moves with a linear velocity $v = p/m$ in y direction. (**b**) With respect to C, A is moving with a speed v along the y direction while B is moving with the same speed v in the opposite direction, implying that the system comprising particles A and B possesses an angular momentum $mlv/2$ about its centre of mass C.

suffice to describe the dynamics of a system even in a weightless environment in space (vacuum!) we first consider a case where the mass of the system is in the form of two independent, equal point masses. Let two particles, A and B, each of mass $m/2$, lie initially a distance l apart, parallel to the x-axis. Let us now give a push to A, say along the y-axis, i.e., in a direction perpendicular to its separation from B. For this we consider a force f along y-axis, applied for a short (infinitesimal!) duration Δt, nevertheless imparting a finite impulse $p = f\Delta t$, like in a sudden (quick) hit on particle A. As a consequence of the impulse given, particle A with mass $m/2$ gains a velocity $v_A = 2p/m$ along y direction, having a kinetic energy $K = p^2/m$, which is the total kinetic energy

of the system, since B is stationary. From the conservation of momentum, the centre of mass C of the system moves with a velocity $v = p/m$ parallel to the y-axis (Fig. 1a), with a kinetic energy of translation

$$K_1 = p^2/2m. \tag{1}$$

But this accounts for only half of the total kinetic energy of the system, i.e., $K_1 = K/2$.

Actually, the system possesses additional motion apart from the linear motion of its centre of mass. This is because while A has a velocity $2v$ along y direction, B is stationary, and C has a velocity v along y direction. Therefore with respect to C, A has a speed v along y direction while B has a speed v in the opposite direction (Fig. 1b) and these two together constitute an angu-

lar momentum

$$J = 2(m/2)(l/2)v = mlv/2 \qquad (2)$$

about C.

The kinetic energy associated with the motion about C is

$$K_2 = 2(m/2)v^2/2 = p^2/2m, \qquad (3)$$

which yields $K_1 + K_2 = K$.

Thus, when a system is given a push, the liner motion of the centre of mass alone may not describe the total dynamics of the system, which might as well, in addition, possess a motion about its centre of mass, irrespective of whether the system is on Earth or in space.

3   A push given to a rod with a uniform distribution of mass

3.1   Force applied for a short duration

A rod with a uniform distribution of mass m and length l has a moment of inertia about its centre of mass C as $I = ml^2/12$ [5, 6, 7]. Let us suppose that the rod, to begin with, is lying along the x-axis and we apply, at a distance δ from C, a force f along the y-axis, i.e., in a direction perpendicular to the length of the rod, for a short duration Δt, imparting a finite impulse, $p = f\Delta t$ to the system, like in a sudden (quick) hit at one end of the rod. As a result, the centre of mass C of the system moves with a velocity $v = p/m$ along y direction, with a kinetic energy of translation $K_1 = mv^2/2 = p^2/2m$. But in addition

there is a torque $N = f\delta$, about C, for time Δt that gives rise to an angular momentum $J = I\omega = ml^2\omega/12 = N\Delta t = p\delta$, about C. From this we can readily see that the rod would rotate around C with an angular speed $\omega = 12p\delta/ml^2$. The kinetic energy of rotation about C would be $K_2 = I\omega^2/2 = 6p^2\delta^2/ml^2$, with total kinetic energy of the system being

$$K = K_1 + K_2 = (p^2/2m)(1 + 12\delta^2/l^2). \quad (4)$$

If the force is applied at one end of the rod, then $\delta = l/2$ and $\omega = 6p/ml$. The kinetic energy of rotation will then be three times that of translation, with total kinetic energy $k = 2p^2/m$. The centre of mass C of the rod will be moving along the direction of the given impulse with a constant velocity $v = p/m$ with the rod simultaneously spinning about C with an angular frequency $\omega = 6v/l$ (Fig. 2).

Of course, if the force is applied at the centre of mass C, then $\delta = 0$ and $\omega = 0$, i.e., the rod does not rotate and from Eq. (4) the total kinetic energy of the system is $K = K1 = (p^2/2m)$.

3.2   The force applied continuously

If a finite force f is applied continuously at a distance δ from C, in a direction perpendicular to the rod, then C gets accelerated at a rate $a = f/m$, while due to the torque $N = f\delta$, there is an increasing angular momentum, $\dot{J} = N$ or $I\dot{\omega} = f\delta$, implying an angular acceleration, $\dot{\omega} = f\delta/I$, about C. The rod will rotate about C with an angular
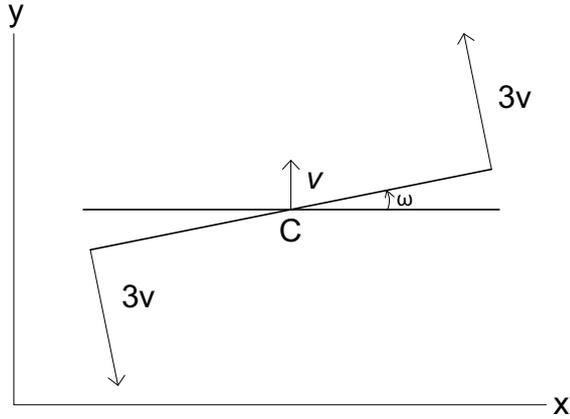
Figure 2: After a force is applied for a short duration at one end of the rod, the centre of mass C of the rod moves with a constant linear velocity v along the y direction. In addition, the rod rotates with a constant angular speed $\omega = 6v/l$ about C.

speed of rotation $\omega = tf\delta/I$ and a rotation angle, $\phi = t^2 f\delta/2I$, assuming $\phi = 0$ at $t = 0$.

With the rotation of the rod, the direction of force f, which is assumed to be always perpendicular to the rod, will change continuously. Decomposing the force vector along x and y directions [8, 9], motion of C can then be obtained from

$$ma_x = -f \sin\phi = -f \sin(t^2 f\delta/2I), \quad (5)$$
$$ma_y = f \cos\phi = f \cos(t^2 f\delta/2I). \quad (6)$$

Integrating with time we get

$$v_x = -(f/m) \int_0^t \sin(t^2 f\delta/2I) \, dt, \quad (7)$$
$$v_y = (f/m) \int_0^t \cos(t^2 f\delta/2I) \, dt, \quad (8)$$

where we have assumed the system to be at rest (**v**=0) at $t = 0$. Writing $k = \sqrt{f\delta/I\pi}$

and with a change of variable $t = u/k$, we can write

$$v_x = \frac{-f}{mk} s(u), \quad (9)$$
$$v_y = \frac{f}{mk} c(u), \quad (10)$$

where s(u) and c(u) are the famous Fresnel's integrals encountered in Fresnel diffraction in optics [10] or elsewhere [11]

$$s(u) = \int_0^u \sin(\pi u^2/2) \, du, \quad (11)$$
$$c(u) = \int_0^u \cos(\pi u^2/2) \, du. \quad (12)$$

For large t, Eqs. (9) and (10) yield a final constant value of **v**

$$v_x = \frac{-f}{2mk} = \frac{-f}{2m}\sqrt{\frac{I\pi}{f\delta}}, \quad (13)$$
$$v_y = \frac{f}{2mk} = \frac{f}{2m}\sqrt{\frac{I\pi}{f\delta}}. \quad (14)$$

We can describe the behaviour of the velocity vector in physical terms, following [9]. With the rotation of the rod, the direction of the force (which is applied always perpendicular to the rod) and hence that of the acceleration, will change continuously and go through cycles of 2π angle each. However, during each cycle, the speed of rotation will be slower at the beginning than at the end. Initially, since the force is pointing in the y direction, there is a bit more velocity gained in y direction. But with the rotation of the rod, as the direction of force turns towards the $-x$ direction, C picks up velocity in that direction. These velocity gains will be substantial in the very first cycle due to
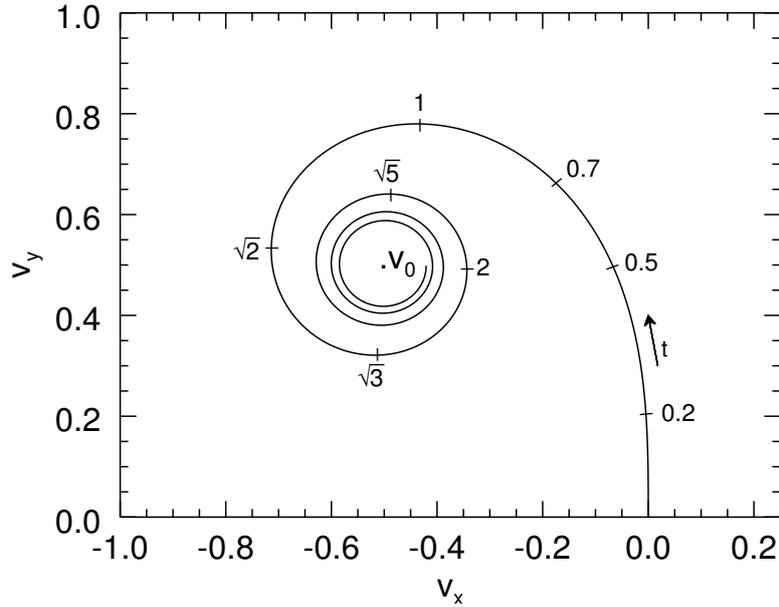
Figure 3: Velocity vector **v** of C, as a function of time, describes a Cornu spiral in the v_x–v_y plane, converging to $\mathbf{v}_0 = -0.5, 0.5$ (in units of f/mk). The arrow indicates direction of increasing time, t, with tick marks on the curve showing t (in units of $1/k$).

the low rotation speed (Fig. 3). In later cycles, as the rod rotates faster and faster, any velocity gains during each cycle will be relatively smaller and the velocity of C would soon stabilize to a constant, $\mathbf{v}_0$, at the centre of the Cornu spiral, as seen in Fig. 3.

As the centre of mass, C, of the rod will be moving with a constant final linear velocity $\mathbf{v_0}$, it means that the kinetic energy of translation will stabilize to a terminal value $K_1 = mv_0^2/2 = \pi f l^2/48\delta$. On the other hand the kinetic energy of rotation, $K_2 = I\omega^2/2 = 6f^2\delta^2 t^2/ml^2$, will be increasing indefinitely with time, with the rod spinning faster and faster about C.

Assuming the centre of mass C of the rod to be at rest (i.e., $\mathbf{v} = 0$) at the origin

$(x = 0, y = 0)$ at $t = 0$, the position of C, as a function of time, can be determined from the generic formula

$$\mathbf{x}(t) = \int_0^t \mathbf{v}\, dt = \mathbf{v}(t)\, t - \int_0^t \mathbf{a}\, t\, dt. \quad (15)$$

Again, with a change of variable $t = u/k$ in Eq. (15) and substituting **a** and **v** from Eqs. (5), (6), (9) and (10), we arrive at

$$x = \frac{-f}{mk^2}\left[u\, s(u) + \frac{1}{\pi}\cos\frac{\pi u^2}{2} - \frac{1}{\pi}\right], \quad (16)$$

$$y = \frac{f}{mk^2}\left[u\, c(u) - \frac{1}{\pi}\sin\frac{\pi u^2}{2}\right]. \quad (17)$$

From Eqs. (16) and (17), for large t, C will follow a straight line path, with superposed initial oscillations that soon die down, as
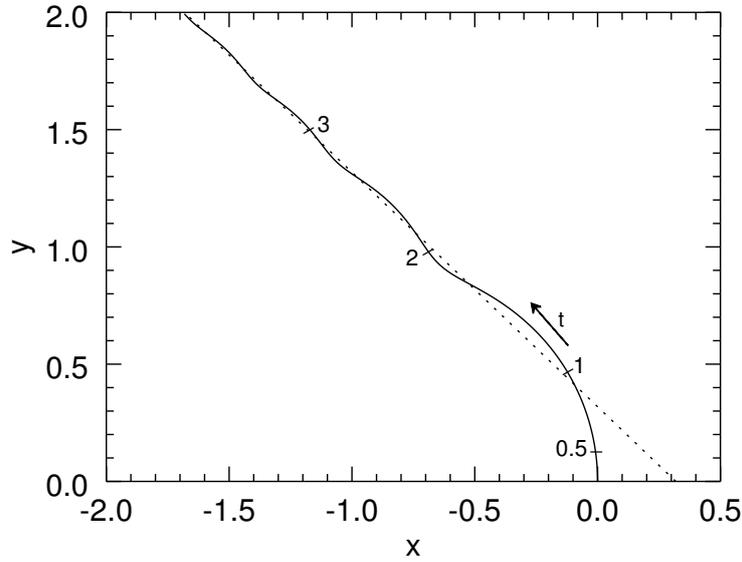
Figure 4: Movement in *x-y* plane of the centre of mass *C*, assumed initially to be at $(0,0)$. It starts along a curved path which asymptotically becomes $y = -x + 1/\pi$, shown as a dotted line. The distance scales are in units of f/mk^2. The arrow indicates direction of increasing time, *t*, with tick marks on the curve showing *t* (in units of $1/k$).

seen in Fig. 4.

$$
\begin{aligned}
x &= \frac{-ft}{2mk} + \frac{f}{mk^2\pi} \\
&= -\sqrt{\frac{\pi f l^2}{48m\delta}}\, t + \frac{l^2}{12\delta}, \quad (18) \\
y &= \frac{ft}{2mk} = \sqrt{\frac{\pi f l^2}{48m\delta}}\, t, \quad (19)
\end{aligned}
$$

and the trajectory in the *x*–*y* plane would be $y = -x + l^2/12\delta$.

According to Ferris-Prabhu [8], the average motion of *C* is along $y = -x$ (according to the conventions adopted here). It need to be pointed out that the equation (6) of Ferris-Prabhu [8], giving the trajectory of the centre of mass, is neither in agreement with the figure (3) given there, nor is it consistent with the initial condition that it starts

from origin $(x = 0, y = 0)$ at $t = 0$. This is because the last term on the right hand side in Eq. (18), i.e., $l^2/12\delta$, is missing from the equation (6) of Ferris-Prabhu [8].

Actually, the velocity vector at the beginning points predominantly in the *y* direction (Fig. 3), therefore motion of *C* will also be initially along the *y* direction. However, as the velocity vector reaches its final constant value v_0, the curved path of *C* will asymptotically coincide with a straight line, $y = -x + l^2/12\delta$, with $l^2/12\delta$ being mostly the initial gain along the *y* direction (Fig. 4).

If the force applied is at one end of the rod, i.e. $\delta = l/2$, then the rod will rotate about *C* with an increasing angular speed, $\omega = 6ft/ml$. The motion of *C* will be still

given by Eqs. (16) and (17) but with $k = \sqrt{6f/\pi ml}$, and of course, $u = kt$. For large t, the path of C will follow a straight line, $y = -x + l/6$.

However, if the force is applied at the centre of mass C, with $\delta = 0$, then from Eqs. (7), (8) $v_x = 0$, $v_y = ft/m$ and from Eqs. (16) and (17) $x = 0$, $y = ft^2/2m$. Only in such a case, the rod will then be moving linearly along the y direction, albeit at an ever increasing speed, but without any accompanying rotation.

Acknowledgements

References

[1] https://physics.stackexchange.com/questions/235482/if-i-push-or-hit-an-object-in-space-will-it-rotate-or-move-along-a-straight-line

[2] http://www.madsci.org/posts/archives/2007-03/1174873330.Ph.q.html

[3] http://www.madsci.org/posts/archives/2007-03/1174873330.Ph.r.html

[4] https://helios.gsfc.nasa.gov/qa_sp_ms.html#sptorque

[5] C. Kittle, W. D. Knight and M. A. Ruderman, Mechanics - Berkeley Phys. Course Vol. 1 (McGraw-Hill, New York, 1965) Ch. 8

[6] S. Datta, Mechanics, (Pearson, Chennai, 2013) Ch. 12

[7] H. Goldstein, Classical Mechanics (Addison-Wesley, MA, 1968) Ch. 5

[8] A. V. Ferris-Prabhu, Am. J. Phys. 38 (1970) 1356-1357

[9] https://physics.stackexchange.com/questions/9096/rocket-engine-perpendicular-to-the-end-of-stick-in-space

[10] F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th edn. (McGraw-Hill, Singapore, 1981) Ch. 18

[11] C. Hazard, Mon. Not. R. astr. Soc. 124 (1962) 343-347

Generation and Detection of Optical Vortex Using Spiral Phase Plate and Spatial Light Modulator

Gulfaroz Patel [1]

[1] Department of Physics, Electronics and Space Sciences,
Gujarat University,
Ahmedabad 380009, India

Abstract

In this article, we have experimentally demonstrated an efficient, simple and cost effective method for generation and detection of optical vortices, phase singularities in the light field. Optical vortices are light beams that carry orbital angular momentum. In order to achieve phase singularity at the center of the beam, we generate and analyze optical vortices using two methods - (i) a spiral phase plate (SPP) and (ii) a spatial light modulator (SLM). The topological charges of the generated optical vortices are verified by tilted-lens method. Also, optical vortices of different orders were generated numerically. The experimental results are in good agreement with the numerical results.

1. Introduction

In recent years, due to growing range of applications of optical vortices, there has been enormous amount of interest for studying helical wave-fronts in the optical wave field, which has an undefined phase in a region of zero amplitude [1]. Optical vortices was observed and extensively studied by Nye and Berry, around three decades ago [2]. In an optical field their structures were systematically described in a subsequent book by Nye [3]. Optical vortices have a ring shaped (donut shaped) intensity profile and at the vortex center intensity of the light drops to zero. The phase of the wave is undefined at the point of singularity (also called branch point), hence the term phase singularity. It appears as a single dark spot against a bright ring. Wave-front changes by an integral multiple of 2π at the phase singularity. At the vortex center, the wave-front increases from 0 to $2l\pi$, leading to a spiral phase structure, where l is an integer number referred as the topological charge of the optical vortex. The wave-front of the vortex beams spirals around the axis of propagation like a corkscrew as it propagates. The sign of the vortex can be positive or negative and it is defined by the phase rotation around the singularity. Both real and imaginary parts of the complex field amplitude are zero, because at the singular point phase is undefined [4].

In modern optical physics, phase singularity became a separate area of investigation. It is an important part of singular optics. The phase of any wave-front with a vortex will be discontinuous, about the point of singularity. Optical vortices are important in studying orbital angular momentum of light fields. Light beams carrying orbital angular momentum of $l\hbar$ per photon possess an azimuthal phase term $l\varphi$, which describes a helical phase pattern (phase twist) of light beams [6]. The number of twists that the beam experiences in one wavelength of propagation is described as topological charge, l. Topological charge 'l' is intertwined helical phase distribution around the point of singularity and it is described by

exp($il\varphi$).The sign of the topological charge determines the handedness (clockwise or anticlockwise) of the vorticity. If the number of twists is higher, then light will revolve faster around the axis.

Several techniques for generating optical vortices have been developed over years. A commonly used method for generating optical vortex is using spatial light modulator (SLM) or spiral phase plate (SPP). Detection of an optical vortex becomes most significant in the field of singular optics. Single and multiple topological charge vortices can be created with the help of computer generated hologram and spatial light modulator method. Optical vortices have found broad applications in many fields. It can be widely used in optical communication [7], optical tweezers [8], cooling and trapping of neutral atoms [9], singular optics [10], manipulating small microscopic particles [11], quantum entanglement of photons [12] and Bose-Einstein condensates [13]. These applications need high efficiency and good generation of optical vortex field.

Allen and co-workers [6] recognized that paraxial Laguerre–Gauss (LG) laser beams carry an orbital angular momentum with a helical wave-front.

Laguerre-Gaussian (LG_p^l) modes signify a beam with helical wave-fronts, whereas l indicates the number of intertwined helices and p as the number of radial nodes [14]. Intertwined helices execute a number of 2π cycles of phase in azimuthal direction. Orbital angular momentum rises from a beam with helical wave-front spiraling around the beam axis. In this paper we have generated optical vortices which are Laguerre-Gaussian LG_p^l modes with a zero radial index.

2. Numerical Study

2.1 Generation of an optical vortex using LG_p^l modes.

In cylindrical coordinates the solution of the paraxial wave equation can be expressed in terms of the product of a Laguerre polynomial, which has radial index p and azimuthal index l, a Gaussian envelope and a phase term. Laguerre–Gauss (LG) solution is generally known as an optical scalar vortex. In cylindrical coordinates, a Laguerre Gaussian (LG_p^l) mode describes the OAM modes. The solution of the paraxial equation in cylindrical coordinates is given by [15, 16],

$$L_p^l(r, \theta, z) = \frac{A}{\omega(z)} \left(\frac{\sqrt{2}\,r}{\omega(z)}\right)^l L_p^l \left(\frac{2r^2}{\omega(z)^2}\right) e^{-r^2/\omega(z)^2} e^{ikr^2/[2R(z)]} e^{il\varphi} e^{-i\xi(z)} \qquad (1)$$

where,

$$A \quad = \text{Amplitude,}$$
$$R(z) = \text{Radius of curvature;}$$
$$\omega(z) \ = \text{Gaussian beam size;}$$
$$L_p^l(x) \ = \text{Associated Laguerre polynomials;}$$
$$l \ = \text{Topological charge;}$$
$$p \ = \text{Number of modes in radial direction;}$$
$$\xi(z) = \text{Arctan}\frac{z}{z_R}: \text{Gouy phase shift;}$$
$$Z_R = \pi\omega_0^2/\lambda: \text{Rayleigh range;}$$

This solution has an important aspect that the amplitude of the wave depends only on the radial nodes. In this paper, numerically we have generated optical vortices with zero radial index

using LG_p^l modes. Laguerre Gaussian (LG) beam reduces to fundamental Gaussian beam when $l=p=0$. LG_p^l modes possess a vortex phase term $e^{il\emptyset}$ when the number of intertwined helices is nonzero. In cylindrical coordinates Laguerre

Gaussian (LG) beams are solutions to the wave of Gaussian beam. A wave function is generally represented as a Gaussian, propagating in one dimension. A pure Gaussian is the zeroth order of the LG beam. The two main components of the Gaussian are its amplitude and phase. Amplitude of a Gaussian beam can be expressed by,

$$A = \exp[\frac{-(x^2+y^2)}{w^2}\;] \qquad (2)$$

where, w is the width of the beam, also known as the beam's spot size.
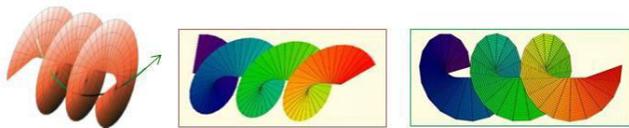
Figure 1: (a) helical wave-front with the Poynting vector indicated by a curved line, (b) helical structure for LG_0^{-1} (c) helical structure for LG_0^{+1}.

Figure 1 shows a helical phase pattern in the LG beam along with the Poynting vector, where the wave-front has a wave-vector which is spiral around the beam axis. As shown in figure 1(b & c) for LG_0^{-1} it spirals in the anticlockwise direction and for LG_0^{+1} it spirals in the clockwise direction around the singularity. A wave-front can be described as points of the beam having same phase when the wave propagates. The helical structure gives rise to OAM of the light beams. The amplitude of the wave becomes zero because of the destructive interference of phases [16].
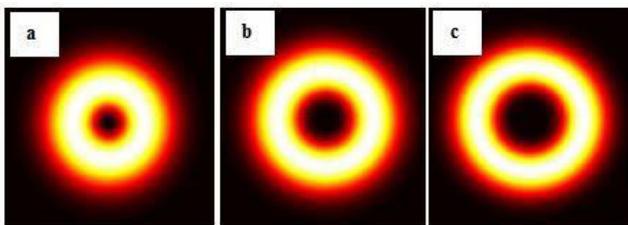
Figure 2: Numerical Intensity distributions (a)LG_0^1(b) LG_0^2 and (c) LG_0^3 mode of vortex beams.

equation. LG beam is the most fascinating families

Figure 3: Helical phase pattern (a)LG_0^1 (b) LG_0^2 and (c)LG_0^3 mode of vortex beams.

Figure 2, shows the intensity profile has a hollow center for LG_p^l mode of vortex beams. Here, the radial index is zero. The beam contains an isolated ringed "donut" shape with zero radial index and nonzero azimuthal index. A singularity in the phase gives rise to donut shape intensity distribution at the center of an optical vortex. A phase jump occurs from 0 to 2π about the singularity point for LG_0^1 mode (l=1). Size of the dark hollow center increases, when the intertwined helices (l) increase in going towards the LG_0^2 mode (l=2), and the phase changes from 0 to 4π about the singular point [15]. As shown in figure 2(c) for l=3, which is LG_0^3 mode, the diameter of the ring is sizeable compared to that for l=1, hence it confirms that the optical vortex diameter increases as the topological charge increases. For l=1 a helical surface is produced in a vortex. Similarly, for l=2, the vortex will see a double helix. Topological charge is described by the helical movement around the beam axis, where the change of phase through one revolution is $2l\pi$. Therefore, when the intertwined helices (l) increase to 2 the topological charge is 2, while (during which) the change in phase is 2π around the beam axis. Topological charge of an optical vortex is represented by the number of phase shifts. We have generated optical vortices of order 1, 2, and 3 numerically for He-Ne laser (λ=632.8nm) and Verdi V 10, a solid state laser (λ=532nm) by using equation (1) in MATLAB software. MATLAB code for generating an optical vortex of Laguerre-Gaussian (LG_p^l) modes is same for both lasers, only difference is wavelength. We can generate optical vortices of all orders for other lasers also just by changing the wavelength in

MATLAB code corresponding to computer generated holograms (CGH).

2.2 Holograms for generating an optical vortex.
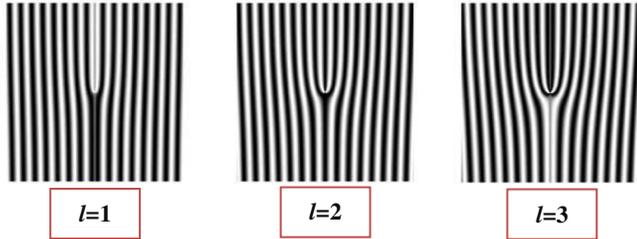
Figure 4: Fork pattern CGH with topological charge *l*=1, 2, and 3 associated with their spiral phase.

Figure (4) shows a forked hologram which we have generated numerically by using equation (3) in MATLAB software which is used in experiment for creating a reliable optical vortex. A computer generated hologram (CGH) is used in the experiments with spatial light modulator (SLM). We have taken a He-Ne laser (λ=632.8nm) for generating optical vortices of order 1, 2 and 3 by using a fork like CGH and SLM. Hologram ("fork") can also be called a computer-generated diffraction pattern, which is used in experiment to obtain a beam having a diffraction pattern in a required configuration (shape). Generally, the CGH is used for generating a monochromatic beam having particular amplitude and phase distribution. Computer generated hologram (linear diffraction grating structure) shows the superposition of plane wave and the spiral phase wave. The equation of the CGH is,

$$R = |g_1 + g_2|^2 = |\exp(ikx) + \exp(-il\varphi)|^2 \quad (3)$$

where, $g_1 = \exp(ikx)$ represents the plane wave and $g_2 = \exp(-il\varphi)$ represents the spiral phase wave. A variation in the number of fringes can be observed from the central singularity between the upper and lower side of the CGH, which confirms the topological charge of an optical vortex. The charge of the vortex can be determined by the number of lines of the fork subtracted by one.

3.Experimental Study

Optical vortex beams can be created by introducing a helical phase distribution on an incident radiation. Optical vortex with a helical phase profile can be easily generated using spiral phase plate (SPP) and spatial light modulator (SLM).

3.1 Generation of optical vortices using spiral phase plate.

We demonstrate experimentally that a spiral phase plate (SPP) can convert a plane wave-front into a helical wave-front. A change in phase will be experienced by a fundamental mode beam, when it passes through the plate. The helical surface of SPP gives a helical character to the incident beam. SPP is a transparent plate and it functions by directly imposing an azimuthal phase shift on the incident light. SPP has a spiral phase, an optical element whose thickness increases linearly along the direction of the azimuthal angle. In the center of the plate SPP thickness increases proportional to the azimuthal angle [16]. It is capable to take high power of lasers because it is made of a glass material. The propagation direction of the beam does not change as the SPP enables a direct conversion of Gaussian beam into a LG beam.

Experimental Setup.
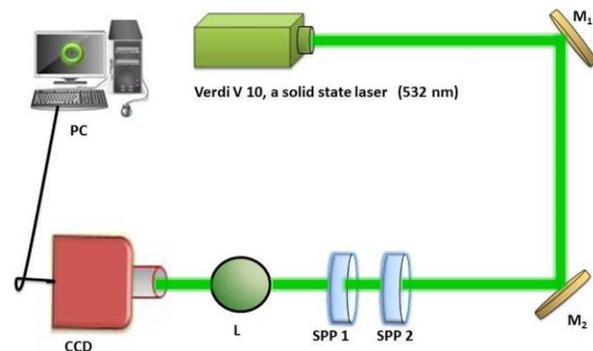
Figure 5: Experimental set up for generation and detection of an optical vortex: M_1, M_2 mirror; SPP1, spiral phase plate of order 1; SPP2, spiral phase plate of order 2; L, plano convex lens; CCD, charge coupled device; PC, computer.

Figure (5) shows the experimental setup to produce vortices with the help of spiral phase plate (SPP). In the setup, we have used two mirrors i.e. M_1 and M_2. Here, we have used a solid state laser whose wavelength is 532nm. First, we align the laser with the help of two mirrors. A highly monochromatic green laser beam falls on M_1 mirror. It reflects from M_1 mirror and falls on M_2 mirror. From M_2 mirror it reflects and falls on spiral phase plate of order 2. An order 2 vortex beam passes onto spiral phase plate of order 1. Afterwards, it passes through an imaging lens (f=30 cm) and falls on the screen. Thus, we generate an optical vortex of an order 3 and record it on a CCD camera. In the same way we have generated an optical vortex of order 2 by removing a spiral phase plate of order 1and keeping only SPP of order 2. Similarly, we have generated an optical vortex of order 1 by placing a spiral phase plate of order1 in the experimental set up. In the experiment a charge coupled device (CCD) camera was used for taking images of optical vortices of all orders. The image of an optical vortex can be seen in the computer (PC). As a consequence, a Gaussian input profile is converted into donut-shaped intensity distribution with zero intensity at the center. Furthermore, amplitude of the wave is zero around the beam axis because of the mode conversion. Thus, by using a spiral phase plate (SPP) we have generated an optical vortex of an order 1, 2, and 3(as shown in figure 6).
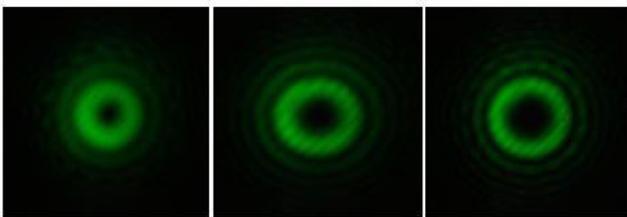
Figure 6: Experimental intensity distribution of an optical vortex with topological charge (a) l=1; (b) l=2; (c) l=3.

3.2 Generation of optical vortices using spatial light modulator

We demonstrate experimentally that a spatial light modulator (SLM) can convert a Gaussian beam into the Laguerre Gaussian beam. SLM is a pixelated device that "spatially modulates" a coherent beam of light. In the experimental study we have used a phase SLM that modulates the phase of the light beam. It is a common technique to generate an optical vortex. It is controlled by a computer. SLMs are expensive, but are quickly reconfigurable. The wave-front of an incident beam is changed by imposing an incident beam onto the SLM [16]. Here, a Gaussian beam is diffracted by the SLM to generate an optical vortex with topological charge (l) 1, 2, and 3.

Experimental Setup.
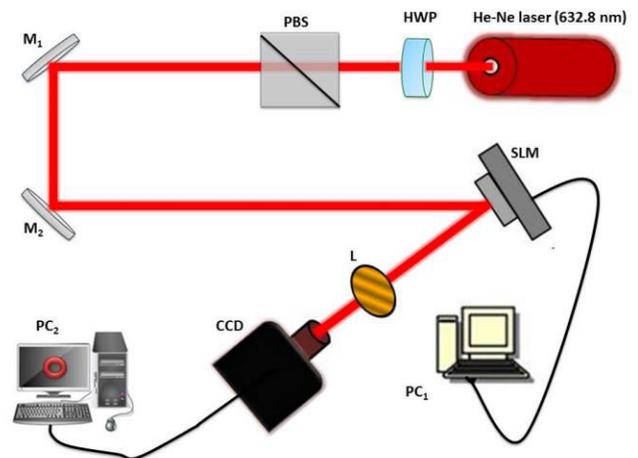
Figure 7: Experimental set up for generation and detection of an optical vortex: M_1, M_2 mirror; PBS, polarization beam splitter; HWP, half wave plate; SLM, spatial light modulator; L, plano convex lens; CCD, charge coupled device; PC_1&PC_2, computer.

Figure (7) shows the experimental setup to produce vortices. The experimental study was carried out with the help of SLM and computer generated hologram (CGH). First, we have displayed the CGH (fork pattern) onto the SLM in real time with the help of computer (PC_1). Here, we have used a He-Ne laser whose wavelength is 632.8 nm. The red laser light passes through the

half wave plate and then it goes to polarization beam splitter. After passing from the beam splitter it is reflected by the mirror M_1 and again by M_2 then falls on the SLM. The wave-front of light is modified by the SLM. Therefore, when an incident Gaussian beam is reflected from SLM, it is converted into the Laguerre Gaussian beam which passes through the aperture and falls on the plano convex lens (f=30 cm). After passing from the lens light goes to the CCD. The SLM changes the phase of the incident He-Ne laser beam by giving a helical phase to the incident light beam. If the incident He-Ne laser beam has a topological charge, its sign and order can be detected by loading a hologram (diffraction grating) on the SLM. In the experiment we have used a charge coupled device (CCD) camera which is used to take the image of optical vortices of all orders. We have generated optical vortices of order 1, 2, and 3 (as shown in figure 8). The image of optical vortex captured by the CCD can be seen in the computer (PC$_2$).

Figure 8:Experimental Intensity distribution of an optical vortex with topological charge (a) *l* =1; (b) *l* =2; (c) *l* =3.

4. Detection

Figure 9: Intensity distribution of optical vorticesof order 1, 2, and 3 after passing through a tilted plano convex lens with solid state laser at 532 nm.
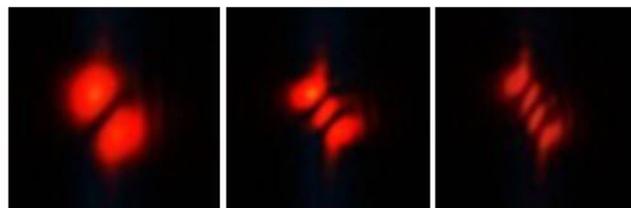
Figure 10: Intensity distribution of optical vorticesof order 1, 2, and 3 after passing through a tilted plano convex lens with He-Ne laser (632.8 nm).

The topological charge of an optical vortex beam shows a specific eigenstate of an orbital angular momentum (OAM). Every eigenstate is orthogonal to each other. Hence, every optical vortex beam with individual topological charge will be orthogonal to each other [21]. To find out the OAM state of light an accurate measurement of the topological charge is necessary. For this measurement we have used a tilted lens. After passing through a tilted lens, the vortex beam gives lobe structure that is characteristic of the topological charge carried by the beam. The obtained lobe structures by tilting a plano convex lens are shown in figures 9&10. For topological charge *l* we get *l*-1 lobes. Suppose *l*=2, then number of lobes will be 1. Using a tilted lens technique, we have measured the topological charge of the generated optical vortices of order 1, 2 and 3.

5. Conclusion

In this article, optical vortex beams have been introduced and studied. Concepts and methods of generation and detection are briefly discussed. Using a solid state laser (λ=532nm) and a He-Ne laser (λ=632.8nm) we have generated optical vortices of orders 1, 2, and 3.Using tilted-lens technique we have measured the topological charge of the generated optical vortices produced by spatial light modulator and spiral phase plate. Optical vortex beam can be used as an information carrier because it carries discrete orbital angular momentum (OAM). The success of detecting topological charges of an optical vortex beam may find an important potential application in optical communications. Scalar vortex beams has been

proposed as an information channel for an optical communication. The experimentally obtained optical vortex shows a good agreement with the numerically obtained optical vortex.

Acknowledgement

References

1. C. Jun, et al. "Generation of optical vortex using a spiral phase plate fabricated in quartz by direct laser writing and inductively coupled plasma etching." Chin. Phys. Lett. 26.1 (2009): 014202.

2. J. F. Nye, and M. V. Berry, "Dislocations in wave trains." Proc. R. Soc. A 336 (1974): 1605.

3. J. F. Nye, "Natural focusing and fine structure of light: caustics and wave dislocations." CRC Press (1999).

4. D. P. Ghai, et al. "Detection of phase singularity using a lateral shear interferometer." Opt. Laser Eng. 46.6 (2008): 419-423.

5. K. Murphy, and C. Dainty, "Comparison of optical vortex detection methods for use with a Shack-Hartmann wavefront sensor." Opt. Express 20.5 (2012): 4988-5002.

6. L. Allen, et al. "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes." Phy. Rev. A 45.11 (1992): 8185.

7. G. Gibson, et al. "Free-space information transfer using light beams carrying orbital angular momentum." Opt. Express 12.22 (2004): 5448-5456.

8. D. G. Grier, "A revolution in optical manipulation." Nature 424.6950 (2003): 810.

9. S. Kuppens, et al. "Polarization-gradient cooling in a strong doughnut-mode dipole potential." Phy. Rev. A 58.4 (1998): 3068.

10. M. S. Soskin, et al. "Topological charge and angular momentum of light beams carrying optical vortices." Phy. Rev. A 56.5 (1997): 4064.

11. H. He, et al. "Direct observation of transfer of angular momentum to absorptive particles from a laser beam with a phase singularity." Phy. Rev. Lett. 75.5 (1995): 826.

12. A.Vaziri, et al. "Entanglement of the Angular Orbital Momentum States of the Photons." Nature 412.313 (2001).

13. E. M.Wright, et al. "Toroidal optical dipole traps for atomic Bose-Einstein condensates using Laguerre-Gaussian beams." Phy. Rev. A 63.1 (2000): 013608.

14. M. Padgett, and L. Allen, "Light with a twist in its tail." Contemp. Phy. 41 (2000): 275-285. {http://people.physics.illinois.edu/Selvin/PRS/498IBR/Twist.pdf}

15. W. Cheng, "Optical vortex beams: Generation, propagation and applications". University of Dayton(2013).{https://etd.ohiolink.edu/!etd.send_file?accession=dayton1375370902&disposition=inline}

16. E. J. Galvez, "Gaussian beams." Dept. of Physics and Astronomy, Colgate Univ. (2009).{http://www.colgate.edu/portaldata/imagegallerywww/98c178dc7e5b4a04b0a1a73abf7f13d5/imagegallery/Gaussianbeams.pdf}, {http://cms.colgate.edu/portaldata/imagegallerywww/98c178dc7e5b4a04b0a1a73abf7f13d5/ImageGallery/gaussian-beams-in-the-optics-course.pdf}

17. P. Vaity, et al. "Measuring the topological charge of an optical vortex by using a tilted convex lens." Phys. Lett. A 377.15 (2013): 1154-1156.

Work Done by Internal Forces on a Multibody System with Constraints

Yukio Kobayashi

Department of Information Systems Science, Faculty of Science and Engineering, Soka University, 1-236 Tangi-machi, Hachioji-shi, Tokyo 192-8577, Japan.
`koba@t.soka.ac.jp`

Abstract

The work done by internal forces on a multibody system is determined by the constrains which limit the motion of the bodies in the system. At college level, however, some instructors and students do not recognise the distinction between external forces exerted on the bodies of a system and internal forces inherent to the system. Therefore, the work done by internal forces on a system is sometimes confused with that exerted by external forces. Here, the author illustrates by using three examples an approach to the systematic study of the relationship between the constraints imposed on the motion of a multibody system and the work done on the system by internal forces. Even if the absence of external forces causes a constant total momentum in the system, the mechanical energy is not necessarily constant given the work done by internal forces on the system. This difference is explained from temporal and spatial viewpoints of physical phenomena.

1 Introduction

In a two-year elementary college physics course for students majoring in physics, even those who have learned the concepts of work and kinetic energy do not necessarily understand the mechanism of the work done by internal forces on a multibody system with constraints [1]. Unlike the concepts of work and kinetic energy applied to single particles, some questions arise even when considering a system composed of a pair of particles:

1. Does the law of conservation of mechanical energy hold for two bodies showing different displacements with respect to the rest frame?

2. Does the law of conservation of mechanical energy hold despite internal forces such as normal reaction being non-conservative?

3. Why is the total kinetic energy of a

system not always constant even when the law of conservation of total linear momentum holds?

In this article, the author considers that the change in total kinetic energy between two bodies is restricted by a positional relationship. Hence, it is necessary to consider the constraints which limit the motion of the bodies composing a system. Unexpectedly, the change in total kinetic energy produced by internal forces has not been thoroughly described in textbooks on elementary physics (e.g. the Berkeley Physics Course [2]). Two key points must be considered in a multibody system with constraints:

1. The work done by internal forces depends on the relative displacement between bodies.

2. The relative displacement is restricted by constraints imposed on the motion of the bodies.

The conservation of mechanical energy is systematically explained for various examples on these two points throughout this article.

2. Constraints on motion of particles

The systematic study considering first the motion on a plane and then the motion on an inclined surface is essential to explain the relationship between the work done by internal forces and motion constraints. In the following three examples, internal forces are not relevant to the total kinetic energy of the systems, because the motion of the bodies in the corresponding systems is restricted by constraints.
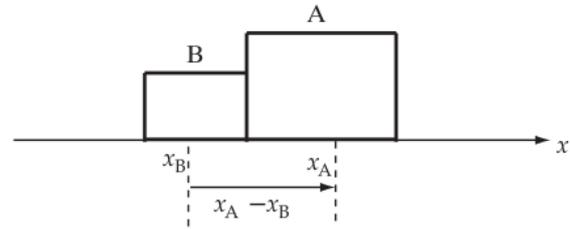
Figure 1. Two bodies in contact.

Example 1

First, as a simple example of plane motion, let us consider two bodies, namely A and B, in contact on a horizontal smooth plane. An external force is exerted on body B. The internal forces in the system composed by bodies A and B, and shown in Figure 1, are the force exerted on B by A and that exerted on A by B. These forces are equal in magnitude and opposite in direction as determined by the law of action–reaction. The two bodies move under the constraint that A and B are in contact, and thus the position difference, $x_A - x_B$, between the bodies is constant.

Work W done on A by the force exerted on A by B, and work w done on B by the force exerted on B by A are given by

$$W = F dx_A,$$
$$w = (-F) dx_B,$$

where F is the magnitude of the internal forces, and x_A and x_B are the horizontal positions of representative points in bodies A and B, respectively. Thus, in an infinitesimal time interval dt,

$$W + w = F(dx_A - dx_B)$$
$$= 0 \, \text{J}.$$

Hence, the internal forces do not contribute to the change in total kinetic energy in this system, which depends on the change in the relative position between bodies A and B, i.e. $d(x_A - x_B)$. Internal forces F and $-F$ cancel each other by the law of action–reaction. In contrast, work components W an w cancel each other by constraint $dx_A = dx_B$ expressed by in the rest frame. Thus, the work done on this system by the internal forces is determined by the constraint on motion.

Example 2

Next, let us consider a system consisting of a body and an inclined surface on a smooth plane, where friction forces can be ignored.
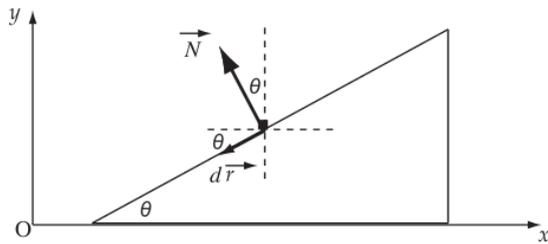
Figure 2. System composed by body and inclined surface.

The relative position of the body with respect to the inclined surface, $\vec{r}$, is given by $\vec{r}_{body} - \vec{r}_{surface}$, where $\vec{r}_{body}$ and $\vec{r}_{surface}$ are the positions of representative points in the body and the surface with respect to the rest frame, respectively. When moving, $\vec{r}$ changes by $d\vec{r}$ or equivalently (dx, dy). The constraint on the body moving along the inclined surface is expressed by $dy/dx = \tan\theta$, where θ is the angle between the inclined surface and the horizontal line. This constraint can also be expressed by $dx = |d\vec{r}|\cos\theta$ and $dy = |d\vec{r}|\sin\theta$. When the body moves along the inclined surface toward the plane, the values of dx and dy are negative.

The force exerted on the body by the surface, $\vec{N}$, and the that exerted on the surface by the body, $-\vec{N}$, compose an action–reaction pair. Work W done on the body by the force exerted on the body by the surface, and work w done on the surface by the force exerted on the surface by the body are given by

$$W = \vec{N} \cdot d\vec{r}_{body},$$
$$w = (-\vec{N}) \cdot d\vec{r}_{surface}.$$

Thus,

$$\begin{aligned} W + w &= \vec{N} \cdot (d\vec{r}_{body} - d\vec{r}_{surface}) \\ &= \vec{N} \cdot d\vec{r}. \end{aligned}$$

In terms of components, we obtain

$$\begin{aligned} \vec{N} \cdot d\vec{r} &= N_x dx + N_y dy \\ &= (-|\vec{N}|\sin\theta)|d\vec{r}|\cos\theta \\ &\quad + (|\vec{N}|\cos\theta)|d\vec{r}|\sin\theta \\ &= 0\,\text{J}, \end{aligned}$$

which, as illustrated in Figure 2, indicates that $\vec{N} \cdot d\vec{r} = 0\,\text{J}$, provided that the internal force and the change in the relative position satisfy $\vec{N} \perp d\vec{r}$. Hence, the internal force does not exert work on the system, and is thus irrelevant to the change in total kinetic energy in the system, although the kinetic energy of the system is changed by the work

done on the body by gravity, which is an external force. Similar to Example 1, the work done on the system by the internal forces is determined by the constraint on motion.

In addition, we should consider the constraint of the distance between any two given points on the surface to remain constant over time, regardless of external forces exerted on it, because the surface is modelled as a rigid body.

Example 3

Finally, for a similar problem to the motion of a body on an inclined surface, let us consider the swing of a pendulum attached to the ceiling of a train in motion, as illustrated in Figure 3. The problem is to determine whether the force exerted on a body by a rope does work on the body. To understand this problem, it is essential to note that the external and internal forces exerted on the body depend on the system constitution. For instance, if we consider the body as a system, and the rope and train as surroundings of the body, the tension of the rope is an external force exerted on the body. However, in a three-body system composed by the body, the rope, and the train, the action–reaction force pair exerted on the body by the rope and on the rope by the body correspond to internal forces.

Let us consider the change in the kinetic energy of the body. The tension exerted on the body by the rope does work in the rest frame, because the displacement of the body, $\vec{r}_Q$, is not perpendicular to the

tension, as shown in Figure 3. Furthermore, the displacement perpendicular to the tension corresponds to that of the body with respect to the point of contact between the ceil and the rope, at end P of the rope. The work done on the body by the force exerted on the body by the rope is comparable to W of both the two-body system shown in Figure 1 and the system composed by a body and an inclined surface shown in Figure 2. Hence, the kinetic energy of the body changes by as much as W.
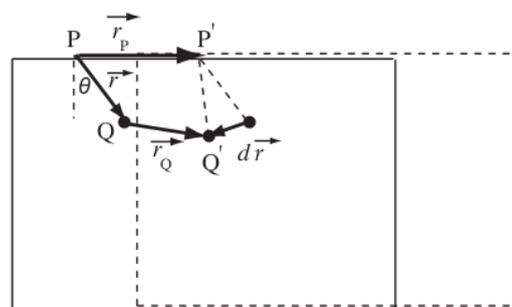
Figure 3. Pendulum attached to the ceil of train in motion. The pendulum moves from Q to Q' and one end of the rope from P to P'. The angle between the vertical line and the rope is denoted by θ.

Let us proceed to the three-body system composed by the body, the rope, and the train, and consider whether the work done by the tension changes the total kinetic energy of this system to ignore the mass of the rope. This system is an extended model of the two-body system in Example 2 that is composed by a body and an inclined surface. Thus, internal forces do no work on the considered three-body system. Both the

work done on the rope by the force exerted on the rope by the train and the work done on the train by the force exerted on the train by the rope cancel each other. Similarly, both the work done on the rope by the force exerted on the rope by the body and the work done on the body by the force exerted on the body by the rope cancel each other. Therefore, these internal forces are irrelevant to the total kinetic energy of the system. Furthermore, the work done on the system by internal forces is also determined by constraints on motion in this example.

3. Physical meaning of relative displacement from the viewpoint of frames of reference

There is another perspective for the work done by internal forces in the two-body or three-body systems. For the two-body system shown in Figure 1, we can consider that it consists of body A in the reference rest frame at body B. Likewise, we can consider the two-body system shown in Figure 2 as consisting of one body in the reference rest frame at the inclined surface. In the former system, the relative velocity of A with respect to B, $\vec{v}_A - \vec{v}_B$, can be regarded as the velocity of A measured in the reference rest frame at B. In the latter, the relative velocity of the body with respect to the inclined surface, $\vec{v}_{body} - \vec{v}_{surface}$, can be regarded as the velocity of the body measured in the reference rest frame at the inclined surface. The examples presented in this article aim to highlight the fact that work is

a physical quantity independent from reference frames. Still, even if the motion of a system is observed in the rest frame, some instructors and students confuse W with $W + w$ in Examples 1 and 2 [3]. This confusion arises from the failure to distinguish between the system and its surroundings, and it might cause problems to understand thermodynamics concepts.

Let us consider again the three-body system composed by the body, the rope, and the train shown in Figure 3. In the reference rest frame at the train, the motion of the body is circular around point P assuming that the rope imposes a constant length constraint. The tangential direction (i.e. the direction of the displacement of the body) is perpendicular to the radial direction (i.e. the direction pointed by the rope). Suppose an xy-coordinate system whose origin is P and the positive x and y axes follow the horizontal right and vertical upward directions, respectively. The position of the body is represented by $\vec{r} = (x, y)$, and thus the constraint is $|\vec{r}|^2 = x^2 + y^2 = \ell^2$, where ℓ is the length of the rope. Thus, we obtain $x dx + y dy = \vec{r} \cdot d\vec{r} = 0$, which implies $\vec{r} \perp d\vec{r}$. Tension $\vec{T}$ is parallel to the position of the body, $\vec{r}$, which is denoted by $\vec{r} \parallel \vec{T}$. Thus, the tension is also perpendicular to the change in position of the body, i.e. $\vec{T} \perp d\vec{r}$. The tension does no work on the body, which indicates that it is irrelevant to the change in kinetic energy in the body, although this energy is changed by the work done on the body by gravity, which is an ex-

ternal force.

The following discussion is important from a pedagogical point of view. The constraint of constant rope length, ℓ, restricts the path of the body to an arc of radius ℓ centred at P. In the reference rest frame at the train, the tangential component of the equation of motion of the body is

$$m\frac{1}{\ell}\frac{d}{dt}(\ell^2\dot{\theta}) = F_\theta,$$

where m and F_θ are the mass of the body and the tangential component of gravity, respectively, and θ is the angle between the vertical line and the rope. With slight rearrangements, this equation becomes

$$\frac{d(m\ell^2\dot{\theta})}{dt} = \ell F_\theta,$$

which shows that the rate of change over time of the angular momentum is equal to the torque. In addition, by the constraint on motion, the distance from the rotation axis to the body remains constant when the body moves.

4. Centre-of-mass motion and relative motion of a multibody system

The change in the total kinetic energy of a system depends on the relative displacement between its bodies. In addition, we can understand these mechanisms from the perspective of the centre-of-mass motion and relative motion of the system.

Let us consider the system composed by the body and the inclined surface shown in Figure 2. The total kinetic energy of

this system can be determined by the sum of the kinetic energies from the centre-of-mass motion and relative motion. The external forces exerted on the system are gravity $(m + M)\vec{g}$ and normal force $\vec{N}_{\text{surface}}$ in the vertical direction, where m and M are the masses of the body and surface, respectively, and $\vec{g}$ is the gravitational field. The internal forces cancel each other by the law of action–reaction. The change in kinetic energy caused by the centre-of-mass motion, K_c, is given by

$$dK_c = [(m + M)\vec{g} + \vec{N}_{\text{surface}}] \cdot d\vec{r}_c,$$

where $d\vec{r}_c$ is the change in the position of the centre of mass with respect to the rest frame. The change in the kinetic energy caused by the relative motion, K_r, is given by

$$dK_r = \left(\vec{N} - \frac{m}{m + M}\vec{N}_{\text{surface}}\right) \cdot d\vec{r}.$$

This relationship can be easily derived by

$$dK_{\text{body}} = (m\vec{g} + \vec{N}) \cdot d\vec{r}_{\text{body}},$$

$$dK_{\text{surface}} = (M\vec{g} - \vec{N} + \vec{N}_{\text{surface}}) \cdot d\vec{r}_{\text{surface}},$$

where dK_{body} and dK_{surface} are the changes of kinetic energy of the body and the surface, respectively. From the definition of centre of mass,

$$d\vec{r}_c = \frac{m}{m + M}d\vec{r}_{\text{body}} + \frac{M}{m + M}d\vec{r}_{\text{surface}},$$

and thus we obtain

$$\begin{aligned}(m\vec{g} \cdot d\vec{r}_{\text{body}} &+ M\vec{g} \cdot d\vec{r}_{\text{surface}}) \\ -(m + M)&d\vec{r}_c \\ = \quad 0\,&\text{J}\end{aligned}$$

and

$$d\vec{r}_{\text{surface}} - d\vec{r}_{\text{c}} = -\frac{m}{m+M}d\vec{r},$$

where $d\vec{r} = d\vec{r}_{\text{body}} - d\vec{r}_{\text{surface}}$. By using these equations, we obtain

$$(dK_{\text{body}} + dK_{\text{surface}}) - dK_{\text{c}}$$
$$= \left(\vec{N} - \frac{m}{m+M}\vec{N}_{\text{surface}}\right) \cdot d\vec{r}.$$

Note that

$$\vec{N}_{\text{surface}} \cdot d\vec{r}_{\text{c}} - \frac{m}{m+M}\vec{N}_{\text{surface}} \cdot d\vec{r}$$
$$= \vec{N}_{\text{surface}} \cdot d\vec{r}_{\text{surface}}$$
$$= 0\,\text{J}$$

because $\vec{N}_{\text{surface}} \perp d\vec{r}_{\text{surface}}$. Therefore, the work done on the system by the internal forces, $\vec{N} \cdot d\vec{r}$, contributes to the change in the kinetic energy from relative motion, dK_{r}, although this work is actually zero.

5. Concluding remarks

The difference between linear momentum conservation and mechanical energy conservation is an important aspect from a pedagogical point of view in elementary physics. Plenty of practice problems to be solved by applying the laws of conservation of both momentum and mechanical energy can be found in textbooks on mechanics at college level. The total mechanical energy is not always constant, although the total linear momentum is conserved, provided that there are no external forces. In most textbooks, the mechanisms for the conservation of linear momentum are explained in detail: impulses of internal forces cancel each other by

the law of action–reaction. Likewise, in the examples presented in this article, the work done by internal forces on a system is zero. However, the work done on a system by internal forces is not always zero, despite the law of action–reaction. Still, the underlying reason of this fact is not usually explained in detail. For example, in the system shown in Figure 2, the work done on the system by the internal force is not zero, provided that the internal force is not perpendicular to the surface given friction. In this case, the work done by the internal force contributes to the change in the total kinetic energy in the system.

We can consider a variation of the example shown in Figure 1 as a more advanced problem. A person moves horizontally on a plate placed over a smooth floor. Initially, she and the plate are at rest with respect to the rest frame. The force exerted on her by the plate, and that exerted on the plate by her have the same magnitude by the law of action–reaction. However, the work pair exerted on her and the plate by these internal forces does not cancel out, because the displacement of a body with a large mass is smaller than that of a light body in the rest frame. If the mass of the plate is larger than her mass, the acceleration of the plate is smaller than her acceleration. Thus, the velocity of the plate is smaller than her velocity with respect to the rest frame. Consequently, the displacement of the plate, dx_{plate}, is smaller than her displacement, dx_{person}, which implies that the change in

the relative position, $d(x_{\text{person}} - x_{\text{plate}})$, is not zero. However, impulses $(-F)dt$ and Fdt cancel each other, where F is the magnitude of the internal forces. Thus, the law of linear momentum conservation holds, because time is absolute for her and the plate.

The same physical phenomena are explained from both temporal and spatial viewpoints [4]. Moreover, it is essential to understand physical phenomena considering both viewpoints. There are different characteristics for displacement and time. In fact, displacement with respect to the rest frame is different for the bodies, whereas time does not depend on them. In addition, displacement is a vector quantity, whereas time is a scalar quantity. Thus, the sign of work depends on the displacement and force exerted on the bodies. Whether the impulses of internal forces cancel each other, and work is done by internal forces, is related to the different characteristics of the impulse and work done by force.

The laws of conservation of linear momentum, mechanical energy, and angular momentum have been found from empirical observations. In addition, the laws of motion are structured to be consistent with the laws of conservation. Hence, conservation laws can be derived from equations of motion. To understand the underlying physical meaning when solving problems by the laws of conservation of linear momentum and mechanical energy, however, we can consider the following construction based on three propositions instead of the equations of motion. These propositions can serve as starting point for elucidating mechanical phenomena [5]. From the temporal and spatial viewpoints of physical phenomena, the three propositions, which represent fundamental principles, consider: (1) the change in linear momentum of a particle caused by impulse, (2) the change in kinetic energy of a particle caused by the work done on the particle by an applied force, and (3) the change in the angular momentum of a particle caused by torque. The equations of motion are established by providing the representation of force that satisfies these three propositions simultaneously. Only phenomena restricted to temporal and spatial conditions can occur, and thus problems can be solved by equations considering the laws of linear momentum and mechanical energy.

Acknowledgments

References

[1] Y. Kobayashi J. Phys. Educ. Soc. Jpn. 63 (2015), 104–107 [in Japanese]

[2] C. Kittel, W. D. Knight, M. A. Ruderman, A. C. Helmholz, and B. J. Moyer *Berkeley Physics Course Volume 1, Mechanics*, (New York: McGraw-Hill, 1973)

[3] H. Hosoda Phys. Educ. Univ. 19 (2013), 60–63 [in Japanese]

[4] Y. Kobayashi Phys. Educ. (IAPT) 28 (2012), 6

[5] Y. Kobayashi *Mechanics Station*, (Morikita, Tokyo, 2002) [in Japanese]

An Erlangen Program for Physics: a Brief Historical Note for High School

E. Benedetto[1,2]

[1]Department of Engineering,
University of Sannio,
Piazza Roma 21, 82100--Benevento, Italy
[2]High School 'De Luca',
Via Francesco Scandone, 67,
83100 Avellino, Italy

Abstract

In this paper, I want to summarize the lecture notes prepared for my students of "De Luca" High School. One century is passed by the publication of General Relativity and this theory is considered by many physicists, mathematicians and philosophers the greatest triumph of human reason. Indeed, thanks to Einstein field equations, cosmology has been enclosed in the physical disciplines. Despite this, cosmology has still many unsolved problems. A cosmology without Einstein gravitational theory based on the Poincaré group of Special Relativity or based on an extension of relativistic group was developed in the twentieth century with two different approaches.It would be rather anachronistic to expect a similar approach to cosmology but, despite this, these theories are interesting from a historical point of view and they have become a source of new ideas in the context of modern physics. We want briefly to retrace the history from the beginning until today hoping to stimulate the curiosity of students and teachers. I hope that this manuscript can be a starting point to explore this topic.

1. Introduction

Edward Arthur Milne was an English astrophysicist and mathematician. He was born on February 14, 1896 in Hull, Yorkshire, England. He was educated at the National School at Hessleand then at Hymers College, Hull, and at Trinity College, Cambridge. From 1932 he studied the problem of the expanding Universe and in 1935, published his theory called Kinematic Relativity. This new theory still has the Poincaré group as space-time structure and it wants to be an alternative to General Relativity theory (GR) [1].This model of the Universe is based on a flat, infinite Euclidean space and, with simple kinematic considerations, Milne deduced the laws of cosmic physics. Kinematic Relativity required only two postulates and that is the costancy of speed of light and the validity of cosmological principle and that is the Universe is homogeneous and isotropic. Even though his theory met with considerable opposition, it has been widely studied by the scientific community and it is quite cited in historical papers. Another, less well known, original approach to cosmology was that of Fantappiè by using group theory. Luigi Fantappiè was born in the town of Viterbo in 1901. He was an Italian scientist, great mathematical analysis researcher and founder of analytic functionals. He studied at the Scuola Normale Superiore in Pisa and became a friend of

Enrico Fermi. Fantappiè graduated with Professor Luigi Bianchi in 1922 with a dissertation "Le formedecomponibili coordinate alleclassi di idealineicorpialgebrici". Four years after graduation, he worked at the University of Florence and, a year later, he moved to the University of Palermo. His first papers were concerning the Riemann zeta function and analytic functionals introduced by Vito Volterra. For the work on analytic functionals he obtained a number of awards and in this area he has become known in scientific world [2]. In the last years of his young life (In fact, he died July 28, 1956) he worked on his less known theory and that is the idea of realizing what he himself calls "an Erlangen program for physics", a classification of possible physical theories through their group of symmetries [3]. Fantappié started studying the group structure of Newtonian and relativistic spacetime. In classical physics we have the Galileo`s group of order 10 and it expresses, from physical point of view, the Galileo relativity principle. If, instead, we consider Special Relativity (SR), we have the so called Lorentz's proper group represented, joining spatial rotations and inertial movements, by the rotations of an Euclidean space with 6 parameters. By considering the reflections and the translations of space-time, we obtain the Poincaré's group of 10 parameters. It is well known that Poincaré's group expresses, from physical point of view, Einstein's relativity principle.In the limit case of$c \to \infty$, the geometry of relativistic space-time becomes the geometry of classical physics and Poincaré's group reduces to Galileo's group. Fantappié thought that, for mathematical physics reasons, this was the right direction and he posed the following question: can Poincaré's group be a subgroup of a more general group as well as the Galileo group is a subgroup of Poincarè group? He found a new group which had as limit Poincaré's space time transformations and the main result was his demonstration that the new group cannot be the subgroup of any continuous group of 10 parameters. That is, if we consider groups of 10 parameters and 4-dimensional

spaces, the sequence Galileo $\subset$ Poincarè $\subset$ Fantappiè can not be extended [4]. This is the reason why he called it the final group. The main feature of this "final group" is the presence of a new constant and that is the radius of space-time. In this way, with an algebraic reasoning, it is possible to find an Universe endowed with a perfect symmetry: de Sitter's Universe [5].The first scientist who used the term "de Sitter Relativity" was Freeman Dyson in a well known paper [6], in which the author assumes that physical laws may be invariant with respect to the de Sitter group, rather than the Poincaré group. He also emphasizes that this hypothesis has never been deepened and it is unknown to many physicists. For this reason Dyson says that de Sitter Relativity is one of the missed opportunities of theoretical physics. A link between Kinematic Relativity and group approach concerns the flow of time. Indeed we will see the existence of two types of time. This paper wishes to present a summary and a historical path of these scientific ideas that are comprehensible to a wider public and can lead on to more specialized studies by teachers interested in the subject. The mathematical formalism of these theories can be found in the original texts.

2.Erlangen program for physics

Mathematical thinking was deeply influenced by the speech made in 1872 by Felix Klein when he became full professor at the University of Erlangen. In fact, he illustrated his research program, now called the "Erlangen program "showing that a geometry is based on what is invariant when a transformation is applied. Fantappié, in a similar way, hypothesized an Erlangen program for physics, since the Universe is individuated by a symmetry group which let its physical laws invariant. We must only follow the path identified by classical and relativistic physics observing the passage from the group of Galileo to that of Poincarè. As we said above, Fantappié showed that the Poincaré group can be seen as a subgroup of a wider group F depending with

continuity on c and another parameter r. Furthermore, the most important demonstration is that this group cannot be furtherly extended. For this reason we have the following, uniquely determined, sequence

$$G \subset P \subset F$$

where G , P and F are respectively the Galileo, Poincaré and Fantappié (that is, de Sitter) group; when $r \to \infty$, F group becomes the P group and, when $c \to \infty$, P group becomes the G group. Fantappiè showed that such sequence is univocal. The F group represents the five-dimensional rotations of a maximally symmetric space time with a constant curvature. This space time is called de Sitter space time and it is a solution of Einsten gravitational field equations. Fantappiè, instead, obtains it without considering the gravitational field but only with an algebraic reasoning. After finding the new symmetry group of physics, it was necessary to find more general coordinate transformations than those of Lorentz-Poincaré, constructing the kinematics, dynamics, electromagnetism and thermodynamics of the new relativity. This work was done by the Fantappié's main disciple, Giuseppe Arcidiacono. His main papers were published in Italian language and, moreover, being often mathematical journals, they were not read by physicists. Arcidiacono made the following reasoning.: de Sitter space time is curved and we must consider its projection on a 4-dimensional hyperplane tangent to it at the observation point-event [7]. We must study the new relativity on this flat tangent space time and, for this reason, it is correct to call the new theory Projective Special Relativity (PSR) [8-10]. For historical accuracy we want to note that the plane representation of de Sitter space-time, called "Castelnuovo space-time", had been used by this author as far back as the early `30s [11]. The complex mathematical formalism is beyond the scope of this hystorical note and, for this reason, we do not write the Arcidiacono transformations. We only want to emphasize that in PSR there is a difference between the projected relative space time (X, T) that each observer can see, and the absolute de Sitter coordinates (x, t). The projected

space time coordinates are regulated by Arcidiacono transformations. The relations that link projected and de Sitter coordinates are the following

$$\begin{cases} T = T_b \, tgh \dfrac{t}{T_b} \\ X = r \, tg \dfrac{x}{r} \end{cases}$$

and

$$\begin{cases} t = \dfrac{T_b}{2} \log \dfrac{T_b + T}{T_b - T} \\ x = r \, arctg \dfrac{X}{r} \end{cases}$$

with $T_b \cong 13.7 \cdot 19^9 \; year$ and $r = cT_b$.

We can observe that time T, is slower than de Sitter time t , noting that we must consider intervals of many thousands years so that the two temporal scales differ only for a second. Instead when relative time comes near to the beginning of time $-T_b$, de Sitter time extends into the infinite past. Therefore the absolute Universe is infinitely old. Instead de Sitter space is smaller than the projected one. In conclusion in the Arcidiacono PSR we have an absolute Universe with finite space and infinite time and a projected Universe with finite time and infinite space. For historical accuracy we want to emphasize that the first to hypothesize existence of more time gauges were Milne and Dirac and perhaps Arcidiacono was greatly influenced by the Kinematic Relativity of Milne. Indeed Dirac speaks of this idea only in some conferences held in Trieste [12], instead Edward Arthur Milne proposed a theory, different from GR [13]. In Milne theory the Universe is supposed a flat, infinite Euclidean space and he claimed to have transformed gravitational force from the status of an empirical law to that of a mathematical result. In fact, in his theory, each physical law must be deduced from a very few principles and he believes that the expansion of the Universe could easily be explained without GR. He required only two postulates, the

constancy of the velocity of light and the cosmological principle. The Kinematic Relativity is based on the Poincarè group and there is a difference between the kinematic time t associated with atomic processes and with light frequencies and the dynamic time τ associated with gravitational phenomena. The dynamic time is connected with kinematic time by the rule

$$\tau = t_0 + t_0 \log \frac{t}{t_0}$$

A curious consequence of Arcidiacono transformations is that the relative age of the Universe is constant. In fact the temporal translation is equal in form to relativistic law of the composition of velocities. Therefore, as well as the speed of light is the same for each observer and despite being finite, cannot be exceeded, in the same way the relative age of the Universe is the same for each observer, whichever its space time position. Every observer will see the same Universe, not only from every point of space, but also in any era. He obtains the perfect cosmological principle postulated by the authors of the stationary model, who however, had to hypothesize the creation of new matter from nothing in order to verify it. Furthermore Arcidiacono showed that, differentiating temporal translations, it is possible to discover that every observer will see an expanding Universe with escape velocity proportional to the distance and he became convinced that the Hubble's law was only a geometrical effect [14]. PSR and ordinary SR coincide in the space time neighbourhood of the observer while we have totally different results on processes remote from observer and that is if x/r and $\frac{t}{r/c}$ are not negligible. Later Arcidiacono tried to find a theory that includes PSR as its empty solution, just as ordinary GR comprises SR as the "empty" limiting case. Finally, in 1964, he formulated the so called Projective General

Relativity (PGR). The Arcidiacono theory is very interesting also regarding the link with Hartle-Hawking proposal of "no-boundary" condition but the author very often used phrases that recall more metaphysics that physics. Indeed Arcidiacono stated that the hyper-spherical Universe is like a book written with seven seals (Apocalypse, 6-11), and consequently two operations are necessary to investigate its physics: 1) inverse Wick rotation and 2) Castelnuovo representation. That's the way we can completely define a relativity in de Sitter. Perhaps this way of writing contributed to the little interest shown by scientists in his papers.

3.Revival of interest in de Sitter invariant theories

The first and important developments were due to Jean-Marc Lévy-Leblond, Henri Bacry and Fock [15-16]. Before these authors, with the sole exception of Fantappié' and Arcidiacono, de Sitter space had always been conceived as an empty and therefore unphysical solution to the Einstein equations with a cosmological constant. We can say that Arcidiacono continued in his projective approach in the total scientific isolation until three Italian researchers has tried to carry out his ideas [17-21]. These authors have sought, above all, to implement Arcidiacono approach as part of the standard physics and cosmology. The fundamental dynamics equations of material point, perfect incompressible fluid and electromagnetism have been derived by Leonardo Chiatti. Furthermore, some mistakes of Arcidiacono have been corrected or eliminated [22]. Licata and Chiatti have examined the cosmological problem within PGR context and they observe that it requires additional specifications with respect to Fridman's ordinary cosmology within GR context. The two

points that distinguish the cosmological problem in modern PGR version from that in the GR version are the presence of a generalized de Sitter horizon and the ensuing fact that a part of the recessional motion of galaxies from the observer is not the result of spatial expansion but of the kinematics connected with the existence of this horizon. The recent discovery of the accelerating expansion of the universe has led to a revival of interest in de Sitter invariant theories and there is also a web page dedicated to the possible relativity in the context of de Sitter space time [23]. In the modern papers it is postulated that the empty space has de Sitter symmetry as a fundamental law of nature. This means that space time is slightly curved even in the absence of matter or energy and the acceleration of the expansion of the universe is not all due to vacuum energy, but at least partly due to the kinematics of the de Sitter group. The main difference compared to the Arcidiacono approach is that the physics is not studied on the projective Castelnuovochronotope but it is studied on the space time that Arcidiacono called absolute space time. The principal studies in this regard are made by a Chinese group [24] and a Brazilian group [25].In a recent work, three scientists apply de Sitter symmetry to solve the problem of galaxy rotation curves [26]. In recent years there are also two doctoral theses [27-28].We also can interpret de Sitter relativity as an example of Doubly Special Relativity introduced recently by Amelino-Camelia, in the context of quantum gravity[29].In conclusion we can say that the founders of modern de Sitter Relativity and Doubly Special Relativity theories were Fantappiè and his student Arcidiacono. Honestly speaking, it is evident that there are some mistakes in the original approach; no mathematical errors but from the interpretation point of view. Nevertheless

we must be equally honest acknowledging that Fantappiè and Arcidiacono were, from the hystorical point of view, the first to try to extend the Poincare group through de Sitter symmetry as well as many modern theories are doing.

Acknowlegements

References :

[1] W H McCrea *Rep. Prog. Phys.* **16** 321 (1953)

[2] N. Ciccoli, *Archive for History of Exact Sciences* 69 (3), 311-326 (2015)

[3] L. FantappièRend. *Accad. Lincei XVII*, fasc. 5 (1954)

[4] L. Fantappiè*CollectaneaMathematica XI*, fasc. 2 (1959)

[5] I. Licata *Electr. Journal Theor. Phys*. 3(10), 211-224 (2006)

[6] F.J. Dyson *Bull. Am. Math. Soc.* 78, 635--652 (1972)

[7] G. Arcidiacono *Gen. Relativ. Gravit.* 7, 885 (1976)

[8] G. Arcidiacono *Gen. Relativ. Gravit.* 8, 865 (1977)

[9] G. Arcidiacono*La TeoriadegliUniversi*, vol. I. (Di RenzoEditore, Rome 1997)

[10] G. Arcidiacono*La TeoriadegliUniversi*, vol. II. (Di RenzoEditore, Rome 2000)

[11] G. Castelnuovo*Rend. Accad. Lincei* XII, 263 (1930)

[12] P. Davies About Time: Einstein's Unfinished Revolution, (Penguin Books, Simon & Schuster 1996)

[13] E.A. Milne Relativity, Gravitation and World-Structure. (Clarendon Press, Oxford 1935)

[14] G. Iovane et al. *Chaos, Solitons & Fractals*, 22,521-528(2004)

[15] H. Bacry and J. Lévy-Leblond*Journal of Mathematical Physics* 9 (10), 1605 (1968)

[16] V.A. Fock The Theory of Space, Time and Gravitation(Pergamon Press, Oxford 1964)

[17] E. Pessa*CollectaneaMathematica* XXIV, 151-174(1973)

[18] L. Chiatti*El. Journ. Theor. Phys*. 4, 17-36 (2007)

[19] L. Chiatti*El. Journ. Theor. Phys*. 23 (7), 259-280 (2010)

[20] I. Licata and L. Chiatti*Int. Journ. Theor. Phys*.48(4),1003-1018(2009)

[21] I. Licata and L. Chiatti*Int. Jour. Th. Phys*. 49(10),2379-2402(2010)

[22] I. Licata et al De Sitter Projective Relativity (Springer Nature 2017)

[23]https://en.wikipedia.org/wiki/De_Sitter_invariant_special_relativity

[24] H.Y. Guoet al.*Mod. Phys. Lett*., A19, 1701--1710 (2004)

[25] R. Aldrovandiet al.*Class. Quantum Grav*. 24 (6), 1385-1404 (2007)

[26] A.Araujo et al. https://arxiv.org/abs/1706.06443

[27] Cristian Said Osorio Mayor" De Sitter relativity: foundations and some physical implications" Tese de doutoramento UniversidadeEstadualPaulista (2012)

[28] D. Janzen"A Solution to the Cosmological Problem of Relativity Theory" Ph. D. Thesis, University of Saskatchewan, March (2012)

[29] G. Amelino-Camelia, *Int. J. Mod. Phys. D* 11, 35 (2002)

On the Circular Motion of John Warren's Car

D. V. Sathe

Kasturba Society, Alandi Marg,

Pune, India

dvsathe@gmail.com

Abstract

According to a student of Zollman and Spears (Kansas State University) he / she did not learn anything because the teacher always answered his / her questions [1]. This comment has made profound and ever-lasting impact on my way of teaching starting from 1975 and on work in global problems in learning Newton's laws of motion at the SSC / HSC level. The most influential paper on my work was of John Warren in 1971 as I had to discuss related issues in six letters in Physics Education, UK over a span between 1984 and 2012. I regard John Warren's report of 1971 as the greatest discovery of 20th century in Newtonian mechanics. He expired on 02 December 2016 at age 93 and therefore I am describing his discovery in this article and dedicate the same to his memory with great respect.

1   Introduction

Dr. John W. Warren was born on 01 July 1923. His schooling was affected due to ill health at age 7 but this and other setbacks inculcated the habit of self-study which helped him throughout in his career as a physics educationist. He inherited the precision and rigor in teaching from his father David, who became a technical teacher after the world war. On leaving school at 16, John became a lab assistant in Illford and also completed the external B.Sc. in physics. In 1947 Prof. James Chadwisk (famous for the discovery of neutron) appointed John as demonstrator in the physics department and so John moved to Liverpool. But the health of a senior colleague caused concern that time and hence John had to take up lecturing in the university. This became a boon for him as he realized difficulties of students and the origin of those difficulties in the teaching of concepts. He also noticed that even very capable students had serious problems

which were beyond the inherent difficulties of the subject. Therefore his scrutiny of textbooks and exam papers became truly unique. This is how he became not only physics teacher but an educationist also.

In 1951, he got PhD for research in mass spectrometry and eventually moved back to London in 1953 as a lecturer in physics in the Acton Technical College. Subsequently, ATC became the Brunel university at Uxbridge. He remained there till his retirement in 1985.

In this university, he carried out students' comprehension and retention of the knowledge of circular motion in late 1970s, reported in 1971 [2].

2   Questionnaire on the uniform circular motion of a car

The task was very simple, based on the uniform circular motion of a car on a plane level ground with negligible resistance of air. Students were asked just one question: draw an arrow showing the resultant force acting on the car. He was surprised on seeing

1. Only about *10 percent* students giving the expected answer that is centripetal force.

2. About *50 percent* students giving a common and unexpected answer that is forward force of engine as the resultant force.

3. Rest of the answers were inclusive answers.

Hence John Warren was unhappy because of poor understanding of circular motion by the students who have passed the entrance exam with excellent marks. Subsequently, many other investigators repeated investigations on this line and made the same observation, thereby establishing the global feature of the poor understanding on circular motion. So let us first see what could have motivated John Warren to use the circular motion of a car.

3   Motivation for John Warren

In textbooks and exam papers, we teachers commonly use natural motions like the planetary motion or electron's motion around the hydrogen nucleus in Bohr's theory of hydrogen atom. But what about *artificial* motion? In fact, it is reasonable to say that Newton did realize the negligence of artificial motion in mechanics and therefore wanted to work on that when he was in mid-sixties, nearly 20 years after putting forth his laws of motion. Therefore I do not think that Newton forgot his own laws of motion as Anthony French said [3] - but I think that Newton made the greatest and exceptional introspection and realized some logical lacunae in his mechanics. This information could have been passed on from one generation to the next because E.E. Witmer and A.V. Bushkovitch wrote a note, title: "On the

Lack of Logic in the Literature of Physics" [4]. Even Dennis W. Sciama states in the preface of his famous book 'Physical foundation of the General Relativity' (1969) says that logical incompleteness of Newton's laws of motion leads us step by step to the General Relativity. More, over, earlier I have shown that incompleteness itself has not understood as yet and teachers still have crucial difficulties especially in teaching uniform circular motion [5]. Therefore, I think, John Warren decided to use the circular motion of car the case of artificial motion for studying students' comprehension of the subject. And the fact that nearly half of students gave one particular wrong answer proves there is a globally neglected problem in teaching circular motion. I pointed out in 1984 [6] that nearly half of these students can counter teacher's stand by questioning the need of engine in the motion of car.

4 Resemblance of students' wrong answer with angel beating wings

In the pre-Newtonian era, there was a popular idea that a planet moves forward on the orbit because of an angel keeps beating wings behind that planet, even Kepler also used this idea. This idea is seldom discussed in textbooks, except Richard Feynmann's celebrated book. So engineering students' answer is a compelling evidence that they can imagine a pre-Newtonian idea which is generally not taught. So teachers have to find out the reason for questions (i) Why students give contrasting answers, depending upon the mode of evaluation and (ii) How and Why students can imagine a pre-Newtonian idea, even though it is not taught to them, as the correct answer. Also the necessity of focusing attention and refining the present teaching is evident in following evidences gathered after John Warren's investigation.

Richard Gunstone carried out a similar investigation on circular motion in 1984 [7]. One Australian girl did not respond to the questionnaire because she found learning of this and related concepts meaningless. My own experience with graduates gained in a symposium in Wadia College, Pune, in January 1993 is an eye-opener [8]. A British girl, Josie, reacted to physics after studying for one year and then giving up like the Australian girl noted above. Therefore I guessed what could have made her to make a bitter comment on physics in 2007 [9]. Most recently, exo-planets have entered our physics and some of them have retrograde orbits reviving an old challenge in O level mechanics [10].

In view of the above discussion, I consider John Warren's investigation as the greatest discovery in physics education in the 20th century, giving an experimental support to Newton's exceptional introspection when he was in mid-sixties. Therefore, this letter dedicated in his memory, with great respect, on the occasion of his first death anniversary

on 02 December 2017.

References

[1] Zollman D.A. and Spears J.D. (1974). Am. J. Phys. 42 (12) p. 1062

[2] Warren J.W. (1971) Phys. Edu. UK, 6, p. 74

[3] French A.P. (January 1984) Am. J. Phys. 52 (1) p. 13

[4] Witmer E.E., Bushkovitch, A.V. (1937). Am. Phys. Thr. 5, (4) p. 145

[5] Sathe Dileep V. (2001) COSPAR Info. Bulletin 52, p. 53

[6] Sathe Dileep V. (1984). Phys. Edu. UK, 19 (2) p, 57

[7] Gunstone Richard (1984) Res. Sci. Edu. 14, p. 125

[8] Sathe Dileep V. (11 April 2009) http://www.science20.com/carl_wieman/scientific_approach_science_education_beliefs_guided_thinking_and_technology

[9] Sathe Dileep V. (July 2007) Phys. Edu. UK, 42 (4) p. 423

[10] Sathe Dileep V. (January 2012) Phys. Edu. UK, 47 (1) p. 132