The Geometry of Uncontrolled Probabilistic Motion

Edgar E. Escultura
College of Science
University of the Phiippines
Diliman, Quezon City

ABSTRACT

*This paper upgrades classical integration to deal with set-valued functions and probabilistic motion. The notions **generalized integral, derivative** and **pulsation** are introduced and applied to differential equations with plane oscillations.*

A formulation in Rn is developed and extended to Hilbert space and the Gaussian distribution is used to describe the structure of a wave packet as generalized pulsation with amplitude. It leads to a sharper form of the Heisenberg's uncertainty principle as conjugacy relationship between probability density and diffusion.

Path integration is likewise upgraded.

1. Limits of Classical Integration

Classical integration reached its apex with the introduction of the Lebesgue integral. All other integrals thereafter -- Henstock, Denjoy, Stieltjes, etc. -- were minor improvements because they differed from the Lebesgue integral, at least in their known applications, only in a set of measure zero. None of them can solve the differential equations:

(1)　(a) $\dot{y} = \sin^n 1/x$ and (b) $\dot{y} = \cos^m 1/x \sin^n 1/x$,

where m, n are positive integers, because the functions on the right side of those equations are not measurable in the neighborhood of zero; in fact, they are set-valued there. We define their set-values by:

(2) (a) $\{f(x)\} = \lim_{x \to 0^+} \sin^n 1/x$ and

(b) $\{g(x)\} = \lim_{x \to 0^+} \cos^m 1/x \sin^n 1/x$

and denote them by

(3) (a) $\sin^n 1/0$ and (b) $\cos^m 1/0 \sin^n 1/0$,

respectively. In both cases, the limit sets are point-wise limits of sequences of arcs of the appropriate trigonometric curves. In the case of (3) (a) the graph is either of the segments [-1,1] and [0,1], depending on whether n is odd or even (See Figure 1 when n is odd).

But why are these integrals unable to solve (1)?

In its most abstract form a classical integral is a function on the space of quadruples (B,X, f(x), m) where B is a measurable subset of X, f is a function on X,　and m is a measure on the measurable subsets of X. The measure is a function on sets and its composite with the integral is also a function on sets. (There is no loss of generality in taking f real-valued since if it is vector-valued we do integration on its components). This leaves f in the quadruple as the only function in the ordinary sense, i.e., single-valued. This is the fundamental limitation of the classical integral. That is why it cannot solve (1) because both functions on the right side are set-valued.

2.　The Generalized Integral, Derivative and Pulsation

Given this fundamental weakness of the classical integral we now upgrade it by taking a leap to complete what was started by Lebesgue, namely, the extension of the integral to abstract spaces and its partial upgrading into a function on sets. This means replacing the single-valued function f(x) by a set-valued function $\{f(x)\}$. Since in mathematics a set is uninteresting and its elements, inessential, apart from their structures, we take one more step: introduce some structure in the form of a

function on sets, namely, a probability distribution function p_x (.) on $\{f(x)\}$ such that at each x, $\dot{p}_x$ (.) is a probability distribution on $\{f(x)\}$.   [The subscript x in $\dot{p}_x$ (.) means that $\dot{p}_x$ (.) is the probability distribution of the set-value $\{f(x)\}$ at x in the mapping $x \longrightarrow \{f(x)\}$ defining this function and the dot in (.) is reserved for the dummy variable along $\{f(x)\}$ for integrating this function.] Then we define a **generalized integral** as a function on the space of quintuples $(m, B, X, \{f(x)\}, p_x (.))$ where B is a measurable subset of X, $\{f(x)\}$ is a set-valued function on X and p_x (.) is a probability function on $f(x)$. We require X to be Hansdorff to admit probability measure concentrated at a point.

We do the formulation here for R^n and extend it to Hilbert space in Section 5.

Let Y be a compact subset of R^n with the standard subspace topology and its Borel sets. Let I be an open bounded n-cube with finite measure p(I), and containing Y, whose edges are parallel to corresponding coordinate axes. We partition I by a finite number of linear subspaces each of which is orthogonal to a coordinate axis so that, in effect, I is partitioned into n- dimensional rectangular blocks which we shall simple call blocks. Since we will be dealing here with probability measures including measure concentrated at a single point, we do not want the blocks to overlap. And so we admit only partitions such that the boundary on the face of a block belongs to one and only one block.

Suppose a block containing a point of Y has its boundary on two subspaces orthogonal to the Yi - axis, $i = 1, 2,..., n$, and passing through the points $(0,...,Y_{i,k},...,0)$ and $(0,...Y_{i,k+1},...,0)$, where $Y_{i,k} < Y_{i,k+1}$. We include in that block its boundary contained in the first subspace and exclude from it its boundary contained in the second subspace. The exception in each row of finite sequences of blocks induced by this ordered partitioning is the last one on the right whose boundary contained in the last bounding subspace we include in that block. Note that for any such partition the system of blocks covers Y and these blocks are non-overlapping.

Let P be a partition of I. We define the norm I PI of P as the diameter of the largest block in the partition. If P_1 is another partition of I we define the product PP_1 as the partition of I by both P and P_1. We are interested in a sequence P^i of finer partitions constructed in this manner such that $|P^i| \rightarrow 0$ as $i \rightarrow \alpha$. Sup-

pose we take a block Δil containing $y \in Y$ at a particular value of i and suppose for each j, $j > i$ we take a block Δjl, $\Delta jl \subseteq \Delta il$. Suppose further that for each Δil we take a maximal Borel sub-set Δiy of Y contained in Δil (and, naturally, containing y because of maximality). Then we have a nested sequence of blocks Δil and a sequence of maximal Borel sets Δiycontained in Δil for each i. Both sequences shrink to the point y as $IP^1I \rightarrow 0$.Since $\Delta iy \subseteq \Delta il$ and $p(l) < \alpha$ then the quotient $P(\Delta iy)/P(\Delta il)$ is uniformly bounded by 0 and 1 and hence has a limit point. Since both I and Y are measurable, that limit is unique and independent of the choice of partitions. We denote that limit by $\dot{p}(y)$ and call it a density or measure distribution of Y at y. In quantum mechanics this is called the probability density at that point 9. For small Δl, $\dot{p}(y)$ P(Δl) is an approximation to p(Δy). We denote that number by dp(y). We will normalize $\dot{p}(y)$ later and call it a probability distribution.

In the case where the measure α of a block $\Delta \mathrm{l}$ is concentrated at the point y, i.e., p ({y}) $= \alpha$ and the measure of the complement of {y} in Δl is 0, then $\dot{p}(y) = 1$ and dp(y) $= \alpha$.

Conversely, suppose we have a measure distribution on Y and let P be a partition of I. We form the sum of the measures of the blocks each of which contains an element of Y and call it an upper sum S which is bounded by the measure of I. We also take the sum of the measures of the maximal Borel sets in the blocks of the partition and call it a lower sum s. It is clear that for any partition P we have the inequality o $\leq s \leq S$. Again as IPI $\rightarrow 0$, s and S tend to some number k and, also, because of the measurability of both I and Y, that limit is unique. We call that limit the measure of Y which we denote by p(Y). We express it as an integral:

$$(3) \qquad p(Y) = \int_Y 1.dp \,(.) \; = k.$$

We normalize the measure distribution p(y) by dividing it by k and call it a probability or unit measure distribution on Y. In this case p is called a probability or unit measure of Y.

Suppose we take a point $y \in Y$ in each block Δl of a partition and we attach to y the measure p (Δy), i.e., we multiply

each component of y by p(Δy) where p is a unit measure on Y. We take the sum s among all blocks in the partition each of which contains a point of Y. This sum is uniformly bounded since Y is bounded and I has finite measure. This is a rough average of the elements of Y. We let the norm of the partition approach zero and since that sum is uniformly bounded it approaches a limit which, again, is independent of the partition. We call that limit the expectation point of Y and we denote it by:

$$(4) \qquad E(Y) = \int_Y (.)dp(.).$$

We introduce an example of a **generalized integral** over an interval [a,b] in R -- a double integral defined by:

$$(5) \qquad Q_{ab} (\{f(x)\}) = \int_a^b ((.)dp_x(.))dx.$$

The inner integral on the right side of (5) maps $\{f(x)\}$ into the well-defined measurable expectation function E(x) on [a,b] so that the outer integral becomes an ordinary Lebesque integral.

(There is no loss of generality in taking $\{f(x)\}$ a plane set and E(x) real-valued since if E(x) is vector-valued we can do integration on the components.) Thus in the xy-plane the integral Q represents the area under the curve y = E(x) from x = a to x=b (See Figure 2). If $\{f(x)\}$ is single-valued and measurable then $p_x(.)$ is concentrated at the single point f(x) at each x. In this case E(x) = f(x) and the areas under the curves y = E(x), y = f(x) from x = a to x = b coincide. Then the generalized integral reduces to a Lebesgue integral. Conversely, any Lebesgue integral can be expressed as a generalized integral. We state this as a theorem.

Theorem. The Lebesgue integral is a generalized integral with suitable integrand and probability function.

Proof. Let $\int f(x)dx$ be any Lebesgue integral, i.e., f(x) is Lebesgue integrable, and let g(x) be any bounded measurable function on [a,b], where g(x) $\geq$0. Denote by A_x the cross section of the area under the curve y = y(x), i.e., the compact vertical segment joining (x,a) and (x,g(x)). Then the function $\{g(x)\}:x \to A_x$, x $\in$ [a,b] is set-valued. Consider the set-valued function $\{h(x)\}$ defined by:

$$\{h(x)\} = f(x)A_x : = \{f(x)\ \mu/\mu \in A_x\},\ x \in [a,b],$$

and let p_x be a probability function on $\{A_x\}$ such that at each x, p_x is concentrated at the point $(x,g(x))$. Then:

$$(6) \qquad Q_{ab}(\{h(x)\}) = \int_a^b (\int_{A_x} f(x)\ (.)dpx(.))dx = \int_a^b f(x)dx. \#$$

We can also set up a generalized indefinite integral by replacing the upper limit b by x in (6) and using a dummy variable s. We can use it to solve a linear first order set-valued differential equation,

$$(7) \qquad \{ \dot{y} \} = \{ f(x) \}$$

where we assume the Borel sets on each $\{f(x)\}$ as well as a measurable probability function $p_x(.)$ on the function $\{f(x)\}$. If the initial condition is given by $y(x_o) = y_o$ then the solution is given by

$$(8) \qquad y(x) = y_o + \int_{x_o}^x (\int_{\{f(x)\}} (.)dpx(.))ds$$

which reduces to the Lebesgue integral

$$(9) \qquad y(x) = y_o + \int_{x_o}^x E(s)ds.$$

For well-defined derivative $\dot{y} = f(x)$ of an absolutely continuous function $y = F(x)$, we have the relationship

$$(10) \qquad y(x) = y_o + \int_{x_o}^x \dot{y}ds$$

or

$$(11) \qquad F(x) = F(x_o) + \int_{x_o}^x f(x)dx \text{ and } F'(x) = f(x) \text{ a.e.}$$

Note that the derivative of an ordinary indefinite integral is its integrand. We extend this idea to set-valued functions by defining a **generalized derivative** as the integrand of a generalized indefinite integral which is the expectation function. But, since that is the weighted average at each point x of the set-value of the integrand with respect to a probability distribution, differentiation reconstructs a probability distribution even if it is not unique. We, therefore, consider two probability distributions that yield the same expectation point E(x) equivalent and the equivalence class determined by E(x) we define as the **generalized derivative** $\dot{p}_x(.)$ of the set- valued function $\{f(x)\}$. Thus we are justified, on two counts, in the use of the dot notation $\dot{p}_x(.)$ for the probability distribution: by looking at $\dot{p}_x(.)$ as a generalization of the ordinary derivative and as the limit of the ratio of the measures of measurable subsets embedded in the nested sequence of blocks which is analogous to an ordinary derivative. And, of course, for single-valued function, the generalized derivative reduces to an ordinary derivative.

Finally, in this section, we define an n-dimensional generalized pulsation as the limit of pulsation in the ordinary sense (rapid expansion and contraction), as the latter becomes infinitely rapid. In the plane, it is an oscillation such as $\sin^n 1/x$. In the (n + 1) - space R^n x R a generalized pulsation is a compact subset of the subspace R^n orthogonal to R. We now drop the qualifier **generalized** for pulsation. Pulsation will refer both to the pulsation-valued function $\{f(x)\}$ and its set-value at x. We describe the structure of a pulsation by a probability distribution. If $\{f(x)\}$ is a pulsation we introduce a probability distribution function $\dot{p}_x$ such that at each x, $\dot{p}_x$ is a probability distribution on the set $\{f(x)\}$. If $\{f(x)\}$ is a singleton at x we call x an ordinary point; otherwise we call it a pulsation point for $\{f(x)\}$. At an ordinary point x the probability function $\dot{p}_x$ is concentrated at the point f(x). If there is some interval [a,b] every point of which is a pulsation point of $\{f(x)\}$ we call the latter a wild pulsation.

3.   Uncontrolled Probabilistic Motion

The motion of a particle at great speed is, by definition, probabilistic since it is difficult to pinpoint its position. This is roughly expressed by the Heisenberg's uncertainty principle. For certain kinds of motion such as pulsation the calculation of

the probability distribution can be done by assuming what we shall call a pulsation probability principle which results in a sharper version of the Heisenberg's uncertainty principle.

We will apply our formulation to uncontrolled probabilistic motion, i.e., differential equation of motion with no control parameters. This means that the geometry of motion is determined by the nature of motion of that object, free from external intervention beyond the underlying electromagnetic field.

The case of differential equation with control parameters has been formulated and solved by a number of mathematicians foremost among whom was L.C. Young (11). Young's approach utilizes unit measure distribution on a compact control set as probability weights for evaluating the expectation function of the set-valued differential equation of motion. Thus a measurable probability measure-valued control function, called chattering control, corresponding to our probability distribution function here, generates the solution of the differential equation which Young calls relaxed trajectory.

An uncontrolled trajectory, according to the Filippov Lemna (11) is a controlled trajectory. Conceivably, we can reconstruct a control set and a probability distribution on it that can serve as probability weights for finding the expectation point E(x) of the set-valued differential equation of motion. This is not very promising, however. Thus we attach instead a probability distribution function that acts on the set-values of the right side of the differential equation of motion, the set-values acting as counterparts to the control set, to find E(x). That is the rationale for the formulation in Section 2.

And since uncontrolled trajectories are also controlled trajectories (11), all theorems about relaxed trajectories are valid for trajectories of uncontrolled probabilistic motion.

4.   Application to Simple Oscillation

We take on the more challenging case of 1(a): $n > 2$ and n is even. The other cases are worked out in (2). The idea is to use a wild oscillation to approximate (1) (a). We use the differential equation with wild oscillation:

(12)
$$\{\dot{y}\} = \sin^n 1/0_x,$$

where $\sin^n 1/0_x = \lim^n 1/s \to x^+ s\text{-}x$, $-x \in [-\varepsilon', \varepsilon']$,

for some $\varepsilon' > 0$, to approximate (1) (a) near $x = 0$. We will later shrink ε' to an approximate value. If $\dot{p}_0(.)$ is the probability distribution on the set $\sin^n 1/0$ then we set up the approximating differential equation to (1) (a) and its corresponding probability distribution function as follows:

$$(13) \quad \{\dot{y}\} = \begin{cases} \sin^n 1/0_x , & x \in [-\varepsilon', \varepsilon'], \\ \sin^n 1/x , & x \notin [-\varepsilon', \varepsilon'] \end{cases}$$

and the probability distribution is defined by the probability function:

$$(14) \quad p_x(.) \begin{cases} p_0(.) , & x \in [-\varepsilon', \varepsilon'], \\ \sin^n 1/x & x \notin [-\varepsilon', \varepsilon'], \end{cases}$$

which is clearly measurable. The second line of (14) means that the probability measure at each $x \notin [-\varepsilon', \varepsilon']$ is concentrated at the point $\sin^n 1/x$ and the probability distribution on the interval $[-\varepsilon', \varepsilon']$ is constant, i.e., at each $x \in [-\varepsilon', \varepsilon']$ the probability distribution on the oscillation there is equal to $\dot{p}_0(.)$.

Now we calculate the probability distribution $\dot{p}_0(.)$ using the geometry of the topologist's sine curve $y = \sin^n 1/x$. We effect, in accordance with our general formulation above, a partition by closed-open non-overlapping intervals (the blocks reduce to intervals on the segment) of the form $[s, s+dy]$, the exception being the topmost segment of any partition which will be closed with the adjunction of the upper end-point of the oscillation at $x = 0$. We effect a change of variable $w = 1/x$. This does not reproduce the entire topologist's sine curve since w is not defined at $x = 0$. But we shall use the resulting function:

$$(15) \quad y = \sin^n w$$

to approximate certain features of $\sin^n 1/0$ by a suitable subarc of (15) in some ε'- neighborhood of $x = 0$.

Consider the geometry of (14) as shown in Figure 3. At any point $W \in [0, \pi]$ whose subtending arc on the unit circle has end point P', there is some point P whose ordinate is $\sin^n w$. If $w \in [\pi, 2\pi]$, its radius vector at Q' will have a reflection P' on the upper semicircle and since n is even, $\sin^n w$ will be the ordinate of some point P on the upper semicircle. As w increases uniformly, i.e., dw/dt = constant, P oscillates back and forth from R to T along the upper semicircle and its projection y does so also along the vertical segment [0,1] at $x = 0$. At the same time, the point $(w, \sin^n w)$ traces an arc of the curve $y = \sin^n w$ with uniform frequency. Increase dw/dt until a half-arc corresponding to half a period of length $\pi/2$ lies within an ε'-neighborhood of $x = 0$. Keep dw/dt at that rate but calibrate w so that the half arc would correspond to the half period [0, $\pi/2$] and a full sweep of y along the oscillation AB (Figure 3). The point P has a projection y on both the approximating arc and the oscillation. Consider one sweep of y along the approximating arc. We ask: given a small segment [s, s+dy) (or [s, s+dy]) in some partition, what is the probability of finding y in there? If y stops momentarily in that segment, i.e., $dy/dt = 0$, we want that probability to be 1. If y does not stop there and its average velocity is large then that probability must be small. Thus there is some conjugacy relationship between that probability and the speed of the oscillating point. We, therefore, assume, as an axiom, what we call the **oscillation probability principle**: the speed dy/dt represents the probability that the oscillating point y is not in the segment [s, s+dy). If we denote that number by $\dot{q}$, since it is a derivative and the probability that the oscillating point lies in [s, s+dy] by $\dot{p}$, then, after normalizing $\dot{q}$ suitably so that its values lie between 0 and 1, we must have:

(16) $\dot{p} + \dot{q} = 1$ or $\dot{p} = 1 - \dot{q}$

We can look at $\dot{p}$ as the probability distribution along [0,1] and q as the speed distribution. They are both relative values or distribution of the value 1. Thus, division of $\dot{p}$ or $\dot{q}$ by a constant does not alter that distribution.

We have,

(17) $dy/dt = dy/dw \cdot dw/dt.$

Since dw/dt is constant, dy/dt is proportional to dy/dw. And we can replace dy/dt by dy/dw as the conjugate of $\dot{p}$ which we shall normalize.

From (17) we have:

$$dy/dw = n\sin^{n-1} w\cos w.$$

We normalize dy/dw by dividing it by its maximum a_n in the interval $[0, \pi/2]$. (A trivial calculation by differentiation yields $a_n = \sqrt{n}(n-1/n)^{n-1/2}$).

Set

(18) $\dot{q} = 1 - n/a_n \sin^{n-1} w\cos w$

and normalize $\dot{p}$ by dividing it by:

(19) $\int_0^{\pi/2}(1 - n/a_n \sin^{n-1} w\cos w)dw = a_n \pi - 2/2a_n ,$

using the fact that the probability that the oscillating point y is in the segment [0,1] is 1. Then we obtain the normalized probability distribution:

(20) $\dot{p} = 2a_n/a_n \pi - 2 (1 - n/an \sin^{n-1} w\cos w).$

We now write the approximating differential equation and the corresponding probability distribution function:

(21) $\{\dot{y}\} = \begin{cases} \sin^n 1/0_x , & x \in [-\varepsilon', \varepsilon'], \\ \sin^n 1/x , & x \notin [-\varepsilon', \varepsilon'], \end{cases}$

(22) $\dot{p}_x(.) = \begin{cases} 2a_n/a_n \pi - 2 (1 - n/a_n \sin^{n-1} w_x \cos w_x), & x \in [-\varepsilon', \varepsilon'], \\ \sin^n 1/x , & x \notin [-\varepsilon', \varepsilon']. \end{cases}$

Since

$$dp_x(w_x) = 2a_n/a_n \pi - 2(1 - n/a_n \sin^{n-1} w_x \cos w_x)dw_x = \dot{p}_x(w_x)dw_x,$$

we can apply the generalized integral Q on $\sin^n 1/0_x$ subject to some initial condition $y(x_0) = y_0$:

$$(23) \int_{x_0}^{x} 2a_n/(a_n \pi - 2) \int_{0}^{\pi/2} \sin^n w_s (1 - n/a_n \sin^{n-1} w_s \cos w_s) (dw_s) \, ds.$$

This integral can be computed by a trivial algorithm. It can be shown **(2)**, **(5)** that, as expected, the expectation function of the set-valued function $\sin^n 1/0_x$, $x \in [-\varepsilon', \varepsilon']$ is some number β which lies between 0 and 1. With the transformation effected by (23), the approximate differential equation becomes an ordinary differential equation:

$$\dot{y} = \begin{cases} \beta, & x \in [-\varepsilon', \varepsilon'], \\ \sin^n 1/x, & x \notin [-\varepsilon', \varepsilon']. \end{cases}$$

We find, one at a time, the solutions in the intervals $[0, \varepsilon']$ $[-\varepsilon', 0]$, $x \geq \varepsilon'$, $x \leq -\varepsilon'$ and join together an absolutely continuous solution. For the first interval we take $y(0) = 0$ as the initial condition to obtain:

$$y(x) = \int_0^x \beta \, x \, ds = \beta x, \qquad x \in [0, \varepsilon'].$$

With the same initial condition and using the symmetry of $\sin^n 1/x$ with respect to the y-axis we obtain another piece of the solution:

$$y(x) = -\beta x, \qquad x \in [-\varepsilon', 0].$$

For the other parts of the solution, we note that $\int_{\varepsilon'}^{x} \sin^n 1/s \, ds$ has

a least upper bound, say α, in $x \geq \varepsilon'$. We choose, as initial condition, $y(\varepsilon) = \beta\varepsilon < \alpha$ for some small ε, $0 < \varepsilon < \varepsilon'$. Hence, the solution for $x \geq \varepsilon$ is:

$$y(x) = \beta\varepsilon + \int_{\varepsilon}^{x} \sin^n 1/s \, ds.$$

Again, using symmetry with respect to the y-axis, we obtain the full approximate solution to the differential equation: (1),

$$(24) \quad y(x) = \begin{cases} \beta x, & x \in [0, \varepsilon], \\ \beta\varepsilon + \int_{\varepsilon}^{x} \sin^n 1/s \, ds, & x \geq \varepsilon, \\ \beta\varepsilon + \int_{x}^{-\varepsilon} \sin^n 1 \quad ds, & x \leq -\varepsilon, \\ -\beta x, & x \in [-\varepsilon, 0]. \end{cases}$$

Its graph is shown in Figure 4.

5. Application to Compound Oscillation

We next apply our method to the differential equation 1(b). As before, we effect a change of variable $w = 1/x$ to obtain the approximating arcs of the functions

$$(25) \quad y = \sin^n w \text{ and } y = \cos^m w.$$

The values of m and n affect the cycle of the compound oscillation as well as the shape of the graph of each of the factors of 1(b). To illustrate our method we take the case m, n both even. (The other cases are taken up in (3). The relevant graphs are shown in Figure 4, where the flatness of the graphs at their maximum points and their steepness away from those points depend on the magnitudes of m and n. It can be shown that the function $y = \cos^m w \sin^n w$ has a unique maximum at $w = \tan^{-1}\sqrt{n/m}$, has minimum at $w = 0, \pi/2$, and is strictly increasing and decreasing in the intervals $[0, \tan^{-1}\sqrt{n/m}]$ and $[\tan^{-1}\sqrt{n/m}, \pi/2]$, respectively. Thus a suboscillation which would have required a different analysis is ruled out.

Unless $m = n$, the function $y = \cos^m w \sin^n w$ has no symmetry with respect to the vertical line $w = \tan^{-1} \sqrt{n/m}$. The functions $\sin^n w$ and $\cos^m w$ have the same periodicity and, hence, a full cycle of their product is repeated at every interval of length $\pi/2$ corresponding to a full sweep of the oscillating point y. (We assume the non-trivial case $m \neq n$).

We now calculate the probability distribution function for the right side of 1(b) at $x = 0$. Note that the oscillating point y is the projection of a point t in arc C_1 and a point r in arc C_2 of the graph of $y = \cos^m w \sin^n w$. Therefore, the probability that y is in some small interval $[s, s+dy)$ (or$[s, s+dy]$) is the sum of the probabilities that t and r are in the corresponding vertical intervals dy at the arcs C_1 and C_2, respectively, of the graphs of the product function.

The calculation of these probability distributions is similar to that for the simple oscillation above. The reader is referred to [3] for detailed calculation and the solution of this problem for the other values of m and n.

By differentiation we obtain the velocity functions for $y = \cos^m w \sin^n w$ in the intervals $[0, \tan^{-1} \sqrt{n/m}]$ and $[\tan^{-1} \sqrt{n/m}, \pi/2]$ which are given by:

(26)

$$\dot{q}_1 (w) = n\cos^{m+1} w \sin^{n-1} w - m\cos^{m-1} w \sin^{n+1} w, \ w \in [0, \tan^{-1}\sqrt{n/m}\,],$$

$$\dot{q}_2 (w) = m\cos^{m-1} w \sin^{n+1} w - n\cos^{m+1} w \sin^{n-1} w, \ w \in [\tan^{-1} \sqrt{n/m}\,],$$

respectively. Let their maxima be a_{mn} and b_{mn}, respectively. Then the normalized velocity functions are:

(27)

$$_1 (w) = 1/a_{mn} (n\cos^{m+1} w \sin^{n-1} - m\cos^{m-1} w \sin^{n+1} w), \ w \in [0, \tan^{-1} \sqrt{n/m}\,],$$

$$_2 (w) = 1/b_{mn} (m\cos^{m-1} w \sin^{n+1} w - n\cos^{m+1} w \sin^{n-1} w), \ w \in [\tan^{-1}\sqrt{n/m}, \pi/2]$$

To simplify our notation, we write

(28)
$$./q(w) = \begin{cases} \dot{q}_1(w), & w \in [0, \tan^{-1}\sqrt{n/m}], \\ 0, & w \in [\tan^{-1}\sqrt{n/m}, \pi/2], \end{cases}$$

$$./q_2(w) = \begin{cases} \dot{q}_2(w), & w \in [\tan^{-1}\sqrt{n/m}, \pi/2], \\ 0, & w \in [0, \tan^{-1}\sqrt{n/m}]. \end{cases}$$

The probability distribution for the compound oscillation AB is given by:

(29) $\quad \dot{p}_o(w) = (1 - ./q_1(w) + (1 - ./q_2(w)) = 2 - ./q_1(w) - ./q_2(w)$

which we normalize by dividing by the constant:

(30) $\quad a = \displaystyle\int_0^{\tan^{-1}\sqrt{n/m}} (1 - ./q_1(w))dw \ + \ \int_{\tan^{-1}\sqrt{n/m}}^{\pi/2} (1 - ./q_2(w))dw.$

Thus, the normalized probability distribution is given by:

(31) $\quad \dot{p}(w) = 1/a \, (2 - ./q_1(w) - ./q_2(w) \,).$

For our approximate differential equation we take:

(32) $\quad \{\dot{y}\} = \begin{cases} \cos^m 1/0_x \sin^n 1/0_x, & x \in [-\varepsilon', \varepsilon'], \\ \cos^m 1/x \sin^n 1/x, & x \notin [-\varepsilon', \varepsilon'], \end{cases}$

with probability distribution

(33) $\quad \dot{p}(w_x) = \begin{cases} \dot{p}(w), & x \in [-\varepsilon', \varepsilon'], \\ \cos^m 1/x \sin^n 1/x, & x \notin [-\varepsilon', \varepsilon']. \end{cases}$

Applying the generalized integral Q with initial condition $y(0) = 0$ we obtain the solution in $[0, \varepsilon]$:

(34)
$$y(x) = \int_0^x (\int_{\lceil s} (.) \, p \, (w_s) \, dw_s) \,) \, ds,$$

where $\lceil x = \cos^m 1/0_x \sin^n 1/0_x$. From the symmetry of the wild oscillation (32) with respect to the y-axis and taking also the initial condition $y(0) = 0$ we obtain the solution of (32) in $[-\varepsilon', 0]$:

(35)
$$y(x) = \int_x^0 (\int_{\lceil s} (.) \, \dot{p} \, (w_s) \, dw_s) \,) \, ds.$$

Outside these intervals the solutions are ordinary Lebesgue integrals subject to the initial conditions:

(36) $y(\varepsilon) = \beta\varepsilon$ and $y(-\varepsilon) = \beta\varepsilon$,

Where $y = \beta$ is the constant expectation function of the wild oscillation

$$\{\dot{y}\} = \cos^m 1/0_x \sin^n 1/0_x$$

and ε satisfies $0 < \varepsilon < \varepsilon'$ so chosen to obtain an absolutely continuous global solution of (32) as we did with our first application above.

The full solution of (32), which is our approximate solution to 1(b) for m, n both even, is given by:

(37)
$$y(x) = \begin{cases} \beta\varepsilon + \int_\varepsilon^x \cos^m 1/s \sin^n 1/s \, ds, & x \in [\varepsilon, \infty] \\ \beta x, & x \in [0, \varepsilon] \\ -\beta x, & x \in [-\varepsilon, 0], \\ \beta\varepsilon + \int_x^{-\varepsilon} \cos^m 1/s \sin^n 1/s \, ds & x \in [-\infty, -\varepsilon]. \end{cases}$$

Its graph is shown in Figure 5.

6. Application to Quantum Mechanics

There is consensus among physicists that the motion of subatomic or elementary particles is probabilistic. The evi-

dences are their great velocities, wave characteristics and some indeterminacy of certain measurements such as position and momentun as expressed by the Heisenberg's uncertainty principle. Therefore, there is validity in assuming that such motion is probabilistic and in using ordinary pulsation in Hilbert space, the setting for quantum mechanics today, to describe such motion. Pulsation at great rapidity propagates waves in all directions including the direction of motion and is probabilistic at the same time. The frequency can be determined from the quantized nature of energy state which we will go into later.

We do some calculations on the generalized pulsation towards which the ordinary pulsation approaches. Physicists introduced the notion of probability density, i.e., the limit of the probability of finding the particle in motion in an element of volume in H. We take the limit as that element shrinks to a point x (t) $\in$ H. We will call that limit the probability distribution at that point and denote it by $\dot{G}(x(t))$. We introduce the notion of diffusion as well, i.e. the rate of flow of the points in that element of volume. We shrink that volume to a point so that the rate of flow approaches instantaneous velocity. We call it velocity distribution. In an analogous manner (to oscillation), we assume that that is the limit of the probability that the particle in motion is not in that element of volume as that volume shrinks to a point. We call it the **pulsation probability principle**. Let us denote that limit by $\dot{D}(x(t))$ and assume that it has been suitably normalized so that its value lies between 0 and 1. (This is possible since velocity is bounded, according to Relativity.) Then we have the conjugacy relationship

$$\dot{G}(x(t)) + \dot{D}(x(t)) = 1$$

which, along the direction of motion, is a sharper form of the Heisenberg's uncertainty principle for pulsation since velocity is simply momentum divided by mass. (The Heisenberg's uncertainty principle says that if Δx is the uncertainty in pinpointing the position of the particle and Δp the uncertainty in measuring momentum then $\Delta x. \Delta p \geq h/2\pi$, where h is the Planck's constant).

The probability distribution $\dot{G}$ describes the structure of the pulsation which we take to be a wave packet with a probability

distribution[1] obtained by modifying the Gaussian distribution given by:

$$(38) \qquad \dot{G}(x(t)) = \propto(t)e^{-\beta(t)|x(t)-\gamma(t)|^2}$$

where x(t) is a variable point in the subspace H of HxR orthogonal to R, $\alpha(t)$ and $\beta(t)$ are positive real numbers, $\gamma(t)$ is the point at which $\dot{G}(x(t))$ is maximum at time t and | | is a norm in H. There is no loss of generality in assuming H to be the space l_2 of square summable sequences since this space is isomorphic to any Hilbert space (**6**). (The choice of the appropriate norm in H is left for the physicist to decide to match empirical data; one possibility is the metric $d(x,y) = \{\sum_{i=1}^{\infty}(x_i - y_i)^2\}^{1/2}$. The function $\gamma(t)$ is determined by the underlying electromagnetic field.)

We modify (38) to describe a wave packet, by a sinusoidal amplitude to obtain the modified Gaussian distribution:

$$(39) \qquad \dot{G}(x(t)) = \propto(t)\sin^n\sigma(t)e^{-\beta(t)|x(t)-\gamma(t)|^2}$$

which describes the structure of a wave packet consisting of a wild pulsation, with an amplitude $\sin^n\sigma(t)$ where $\sigma > 0$ determines the extent of a wave packet along the direction of motion and n is a positive integer. (The terms α, σ, β and η are introduced to obtain the desired properties of the wave packet, including extent and shape, to match empirical data.)

The wave packet is the support for $\dot{G}$ which we assume to be bounded, invoking Relativity and compact by taking its closure.

An ordinary path does not make sense here and in place of it we take the trace of the expectation point of the wave packet in motion. The quantized nature of energy state requires that this path be discrete, each point of which corresponds to an in-

---

[1] This idea of Gaussian distribution I owe to the timely suggestion by Professor Dick Van Dulst of the University of Amsterdam.

tegral multiple of $\pi/2$. These points are the allowable positions of the particle in accordance with the quantized energy state requirement. Between two consecutive observable points we interpret the probability measure, i.e., the integral of $\dot{G}$, as the transition probability or the probability that the particle will jump from the first to the second point.

Finally, we indicate how path integration can be carried out in this setting.

At the subatomic level of matter and at great velocity we lose the amenities of smoothness and determinacy. Position and derivative tend to be set-valued. Even time becomes elusive.

Therefore we assume that the.derivative x is some set-valued function $g(t, \Omega_t, \dot{G}_t)$ where Ω_t is the cross-section of the wave packet and $\dot{G}_t$ is the Gaussian distribution on Ω_t at time t. We find the expectation point $E(t,x)$ of $g(t, \Omega_t, G_t)$,

where

$$(40) \quad g(t, \Omega_t, G_t) = \left\{ g(t,w) \mid w \in \Omega_t \right\}$$

and

$$(41) \quad E(t,x) = \int_{\Omega_t} g(t,(.)) \, d\dot{G}_t(.).$$

The differential equation of motion becomes:

$$(42) \quad \dot{x} = E(t,x) \quad \text{a.e.}$$

$\dot{G}(t)$ is included in the argument of (40) only as a matter of notation to indicate that $\dot{G}$ is the probability distribution on Ω_t.

We shall not go into the requirement on g to insure existence of solution subject to some initial condition; we simply assume that E is Lipschitzian in x to insure uniqueness. Let $L(t, \Omega_t, \dot{G}_t)$ be the set-valued Lagrangian subject to the same conditions as g and let its expectation point at time t be denoted by $L(t, x)$. Since the dimensionality of the problem is at our disposal, we can adjoin in (42) the differential equation:

(43) $\dot{x}_0 = L(t,x)$

to obtain the differential equation:

(44) $\dot{z} = (\dot{x}_0, \dot{x}) = (L(t, x), E(t, x))$ a.e.

Subject to some initial condition $z(t_0) = z_0$ and suitable conditions on g(2), (48) has a unique solution which is called a generalized curve. The path integral from t_1 to t_2 reduces to the difference $x_0(t_2) - x_0(t_1)$, where $x_0(t)$ is the first component of the solution of (44).

If the path integral is a minimization problem (an optimization problem can be reduced to a minimization problem), then it reduces to the minimum of $x_0(t_2) - x_0(t_1)$, which, by a simple translation of the origin of the coordinate system to $x_0(t_1)$, further reduces to the minimum of $x_0(t_2)$.

Note that in this formulation, $\dot{G}_t$ is the generalized derivative of Ω_t and Ω_t is the compact support of $\dot{G}_t$ which coincides with the wave packet. The path integral is taken along the expectation curve z(t) whose derivative is a generalized derivative given by $\dot{G}(z(t))$.

There are many interesting properties of the solution z(t) of (44) one of which being that it can be approximated by an ordinary simplicial path (11).
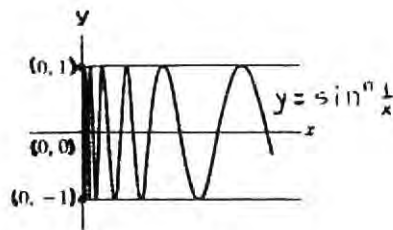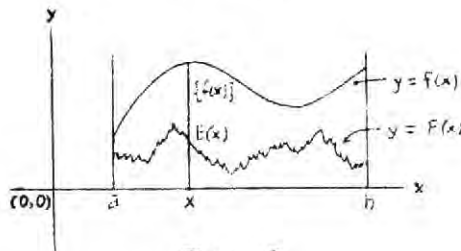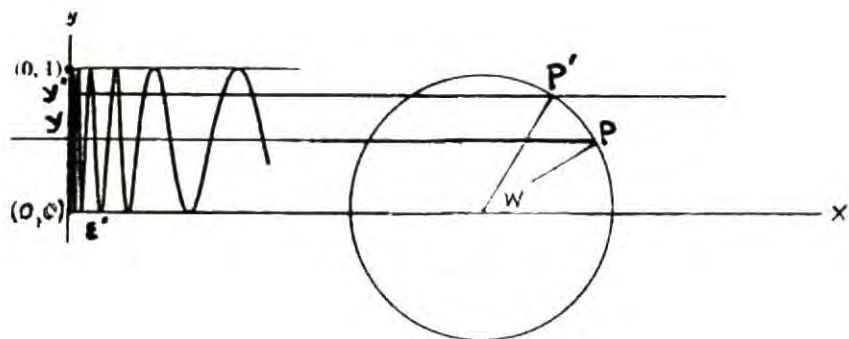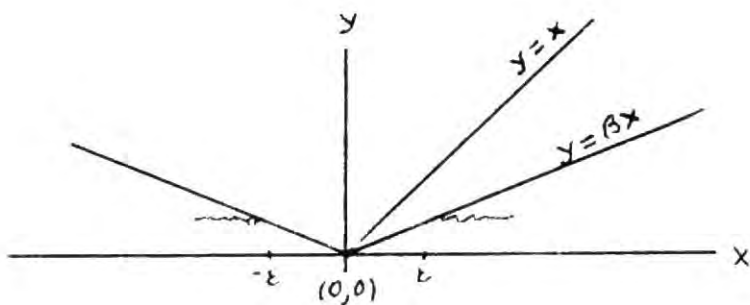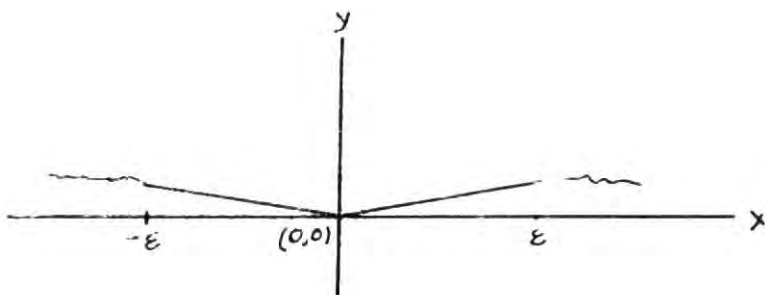

Figure 1.

Figure 2.

Figure 3.

Figure 4.

Figure 5.

REFERENCES

Caratheodory, C. 1965. Calculus of Variations and Partial Differential Equations of the First Order (English Translation, Dean, R.B. and J.J. Brandstatter). Holden Day, San Francisco.

Escultura, E.E. 1991. Introduction to Qualitative Control Theory. Kalikasan Press, Q.C.

_____. 1990. Solution of First Order Differential Equation with Compound Oscillation. Technical report to the Natural Sciences Research Institute, College of Science, U.P. Diliman, Q.C.

_____. 1991. A Rescue from Obsolescence. Technical report on Project B-74, "A Geometry of Relaxed Trajectories," under the National Research Council of the Philippines.

Escultura, E.E. and T. Suzuki. 1990. Integration of Generalized Oscillation, in Proceedings of the First Asian Mathematical Conference, Hongkong (to appear).

Lang, S. 1983. Real Analysis. Second edition, Addison-Wesley Publishing Company, Reading, MA.

Merzbacher, E. 1970. Quantum Mechanics. John Wiley & Sons, Inc., New York, N.Y.

Saks, S. 1964. Theory of the Integral. Second Edition, revised edition, Dover, New York, N.Y.

Thankappan, V.K. 1985. Quantum Mechanics. John Wiley & Sons, Inc., New York, N.Y.

Young, L.C. 1937. Generalized Curves and the Existence of an Absolute Minimum in the Calculus of Variations in *Comptes Rendus de la Societe des Sciences et des Lettres de Varsovie, Classe III*, 30, 212-234.

_____. 1969. Lectures on the Calculus of Variations and Optimal Control Theory I and II. W.B. Saunders, Philadelphia, PA.