EML 6229: Introduction to Random Dynamical Systems

Mrinal Kumar©
Jan 06, 2016

Initial Impressions and the History of Probability Theory

- The subject of probability has been around for well over two thousand years. Through the ages, it has gone through the complete cycle of scholastic evolution - from being nothing more than pure guesswork, to being at best a half-baked branch of natural science meant only for explanation of empirical data, to its present state in which it is a firmly established abstract field of mathematics.

- Even after the rigorous theoretical foundations of probability theory were laid down, it was not accepted by all in the scientific and mathematics communities as a legitimate field of mathematics. The primary controversy was one borne out of metaphysical concerns - how could the mathematical reality underlying nature be written in the language of chance? Even today, there are researchers who cringe from it, preferring to believe in the deterministic view of the universe. A popular example is Albert Einstein, who stated that "God does not play dice". Ironically, part of the work for which he won the Nobel prize (his 1905 seminal paper on Brownian motion) became the foundation for the mathematical model of ***white noise***, which is used universally for describing random perturbations in engineering systems.

- In light of all this fuss, we would first like to know if probability theory is actually a legitimate field of mathematics?! This requires us to ask the following question: *When does a field of study become a well-founded branch of mathematics?* In other words, what is the ***mathematical method?***

- We have all studied about the ***scientific method*** at some point of our educational training: it is an approach of studying nature that involves the following steps:

  1. Observation of physical phenomena.
  2. Formulation of hypothesis to explain observations.
  3. Use of hypothesis to predict future behavior of the observed phenomena.
  4. *Most important:* Continuous testing of the formulated hypothesis by repeated experimentation, performed independently by several different researchers.

  The above steps qualify without doubt astronomy as a field of science, but reduce astrology to mere speculation. So then, what is the mathematical method? How should probability theory be formulated so that it may unequivocally be regarded as a field of mathematics?

- We can say that a set of concepts becomes a well-founded area of mathematics if it has the following underlying structure, known as the ***axiomatic framework***:

1. At the bottom, it contains a set of "undefinables" which form the very basic building blocks for more sophisticated ideas to follow. These cannot really be defined, simply because there are no quantities simpler than them in terms of which a definition could be written. Examples include a point, a line, an element of a set, etc.
2. Using the undefinables, a set of *axioms* or postulates is laid down. These are "undeniable" properties. As undefinables cannot be defined, the undeniables cannot be proved. The axioms result purely from intuition and do *not* involve mathematical reasoning. It is required that the set of axioms be consistent because they will be used as the starting point to build the theory. Consistency is very important in developing axioms.
3. Using the axioms, *logic* is employed to develop the *theory.* Today, this approach is easy to follow because it uses the definition-theorem-proof style. Note that definitions come after the axioms have been laid down.

The above procedure is called the axiomatic approach of mathematics and in fact, *is* mathematics.

- When there are only axioms and nothing else (not even definitions), we are in the realm of pure abstraction. As we define objects and build theorems using them, we start giving these abstract ideas concrete form. When the number of axioms underlying a mathematical theory is small, the scope of the concrete results that follow is broad. The best example is set theory - which has the fewest number of axioms. Therefore in set theory alone, there are only a few theorems to prove; but, these theorems are widely applicable. Consequently, set theory serves as a precursor to many other branches of mathematics, including probability theory. As more axioms are added, more and more theorems can be proved. This results in additional branches of mathematics, such as topology (how close are two sets?), algebra (how can we add two sets?), geometry (topology + algebra) and of course, probability theory. As mentioned above, there are a lot more theorems to prove in these new branches of mathematics because there is a greater number of axioms, but their applicability is narrower.

- It was **Andrey Nikolaevich Kolmogorov** in the early 1930's, who established probability as a branch of mathematics by building an axiomatic framework for it. Not only did he formalize the process of computing probabilities, his work led to significant generalization of previously held "notions", making them formally applicable to a much broader set of objects. *Most significantly, before Kolmogorov's arrival probability was only applicable to countable sets.* Under his axiomatic paradigm, it became possible to extend it to uncountable sets as well. This will be of crucial importance to us in this course.

- As the title of this course suggests, we are interested in studying probability theory *in the context of dynamical systems.* The history of dynamical systems can be traced all the way back to Isaac Newton and his Newtonian mechanics. We will mostly be concerned with dynamical systems evolving in continuous time, a general mathematical model for which can be given by the following vector differential equation:

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) \tag{1}$$

where, $\mathbf{x} \in \Re^N$ is called the ***state*** of the dynamical system. The vector function $\mathbf{f}(t, \mathbf{x}) : [0, \infty) \times \Re^N \to \Re^N$ has desirable properties of smoothness. Note that Eq.1 is a set of N first order ordinary differential equations (ODE's). This form of a dynamical system is called its state-space form. For example, consider a damped linear spring:

$$\ddot{x} + c\dot{x} + kx = 0 \tag{2}$$

The above equation represents a dynamical system model for a linear spring damper assembly, where x is the difference between the current and the natural lengths of the spring. Note that Eq.2 is *not* in state-space form, because it is a second-order ODE. In order to reduce it to state-space form, we have the following developments:

$$\dot{x}_1 = x_2 \tag{3a}$$

$$\dot{x}_2 = -cx_2 - kx_1 \tag{3b}$$

In the above equations, we have defined new states $x_1 \doteq x$ and $x_2 \doteq \dot{x}$. This allows us to reduce the second order ODE of Eq.2 to two first order ODE's. We thus get the state-space form, and x_1 and x_2 are the two states of the spring-damper system. In terms of physics, note that they correspond to position and velocity respectively. The two equations can be written together in vector form as follows:

$$\begin{Bmatrix} \dot{x_1} \\ \dot{x_2} \end{Bmatrix} = \begin{Bmatrix} x_2 \\ -cx_2 - kx_1 \end{Bmatrix} \quad \left(= \begin{Bmatrix} f_1 \\ f_2 \end{Bmatrix} \right) \tag{4}$$

The above system is of the type given in Eq.1, with $\mathbf{x} = \{x_1, x_2\} \in \Re^2$.

- What is missing in the dynamical system model of Eq.1? This is a typical question asked in every introductory lecture on mechanics! An ODE is written down, following which the instructor asks with excitement - *what did I miss?* And someone gives the answer with a bored look on their face - the initial conditions..

- Indeed, initial conditions are required to complete the description. Let us therefore write the whole thing again:

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}), \quad \mathbf{x}(t = t_0) = \mathbf{x}_0 \tag{5}$$

The solution to the above equation is given by:

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \mathbf{f}(\tau, \mathbf{x}) d\tau \tag{6}$$

- The above model of a physical process is called a ***deterministic dynamical system***. It is called deterministic because complete knowledge about the states *can* be obtained for all times, i.e. $\mathbf{x}(t)\ \forall t \in [t_0, \infty)$, without any room for doubt or uncertainty. This assumes the following:

1. Eq.5 captures the complete description of the dynamical system. There are no other residual disturbing forces such as "noise" left unmodeled.
2. The precise value of the initial conditions, i.e. $\mathbf{x}(t_0)$ is known.
3. All the system's parameters are known without any uncertainty. In the discussed model of a linear spring, we have two parameters: c and k. Both of them should be known perfectly.
4. The integral in Eq.6 can be computed, either analytically or numerically.

- Therefore, given a fixed initial condition and a fixed integrator, the same trajectory i.e. $\mathbf{x}(t)$ will be obtained irrespective of how many times the experiment is repeated. *This is the essence of determinism.* See Fig. 1 for an example.

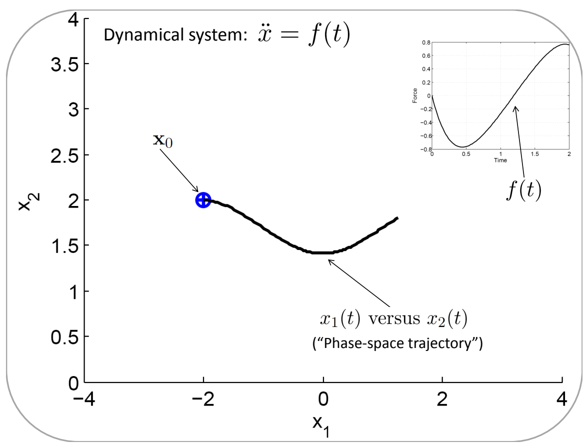


Figure 1: A two dimensional ($N = 2$) example of a deterministic dynamical system.

- Quite clearly, the deterministic model is extremely powerful! Laplace was probably the most vocal proponent of the deterministic model of the universe. He said that in principle, if at a given time we had the knowledge of position and velocity of all particles that exist, we could predict deterministically the future of the universe. As you can guess, determinism entails several philosophical implications. One pertinent question is the following: *Is such a deterministic mathematical model of the real world possible?*

- We all know that the question of determinism has caused numerous battles among scientists and mathematicians over the past several decades. At the microscopic level, the long held view of a deterministic universe came under question when Heisenberg propounded his uncertainty principle - the fact that observation changes the outcome of an experiment, thus making determinism impossible. Even though this is true, such collapse of determinism occurs only at the microscopic level. At the macroscopic level in which most of us engineers operate, the Heisenberg uncertainty principle is too weak to meaningfully influence the outcome.

- In this course, we will look at macroscopic systems and study how the deterministic view of the universe is inadequate or impractical for them.

- What are the possible sources of uncertainty in a macroscopic dynamical system? It will be useful to re-examine the list given in Pg. 4 as conditions of determinism, while also looking at Eq.5. First, let us consider only the first two items on this list. When items 1 and 2 in the list fail, we obtain the following two main sources of uncertainty:

  1. Presence of random perturbations, or, *noise* in the system, which directly influences the evolution of state dynamics. This perturbation is present *in addition* to the "deterministic part" of the model, namely $\mathbf{f}(t, \mathbf{x})$. For now, we are assuming everything about the function $\mathbf{f}(t, \mathbf{x})$ is completely known.
  2. It is also possible that the initial conditions are not known precisely. In real life, this is typically the result of imperfect measurements.

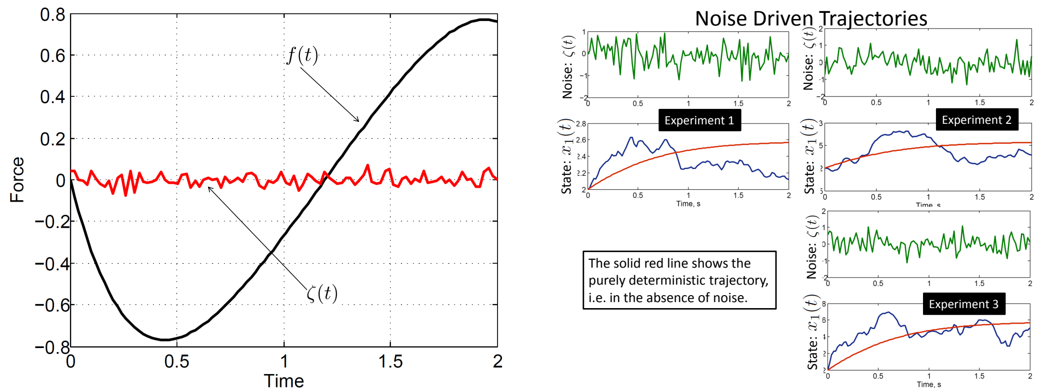
- The deterministic dynamical system model of Eq.5 can now be expanded to include the two sources of uncertainty described above. For simplicity, let us first consider the case of a scalar state-space ($N = 1$):

$$\dot{x} = f(t, x) + \zeta(t), \qquad \mathcal{W}(t_0, x) = \mathcal{W}_0(x) \tag{7}$$

  Written in the above form as an ordinary differential equation with a deterministic part and a non-deterministic "noise" part, Eq.7 is called a **Langevin equation**. In Eq.7, $\zeta(t)$ represents "noise", a random perturbation typically much smaller in magnitude than the deterministic force f. A typical comparison of relative magnitudes is shown in Fig.2(a). The symbol $\mathcal{W}$ is explained below in Pg. 6.

- We will learn in this course that it is not easy to model the random force $\zeta(t)$. It has many exotic properties, the most intriguing of which is that it is impossible to generate the same time-history of noise in two separate runs of an experiment, even under the exact same conditions and the same initial conditions. Each time-history can be referred to as a "noise sample". Obviously, this is a body blow to the concept of determinism on the macroscopic scale. Every time we run the experiment, the random forces driving the system will be slightly different, thus making it impossible to make definitive predictions about the future of the dynamical process. This situation is shown in Fig.2(b), where three integrations starting from the same initial conditions give three different time histories of the state. The noise sample for each run of the numerical experiment is also shown. For these figures, do not bother with the numbers on the y-axis. Just concentrate on the qualitative nature of the plots. The point is that different runs give different results despite using the same integrator and the same initial conditions.

- The second source of uncertainty mentioned above is the lack of precise knowledge of initial conditions. This form of uncertainty is fairly common and makes physical sense. No instrument is perfect and therefore it is never possible to know exactly what values the state has at t_0. We can at best provide a *probabilistic description*. I.E., instead of specifying the initial value of the state as $x(t_0) = x_0$, we say that the probability



(a) Typical relative magnitudes of the deterministic force and noise

(b) Noise samples are always different in different runs of the same experiment

Figure 2: The first source of uncertainty in dynamical systems: random perturbations.

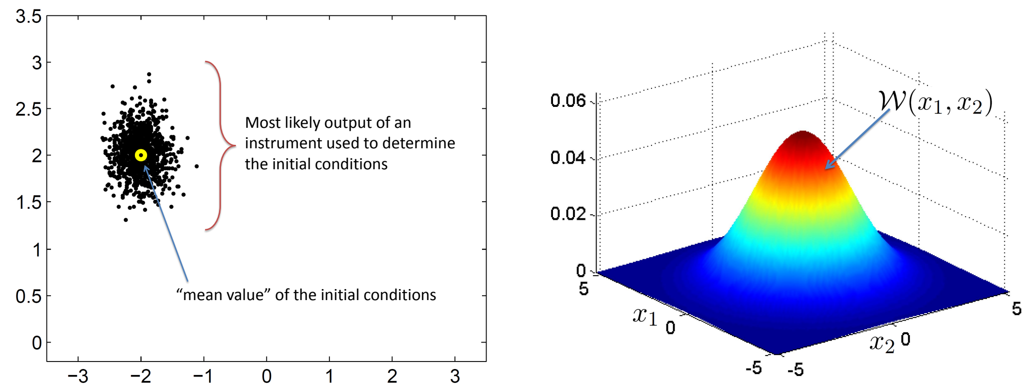
density function (pdf) of the state at t_0 is $\mathcal{W}(t_0, x) = \mathcal{W}_0(x)$. A typical output of an instrument used to measure the initial conditions is shown in Fig.3(a). The yellow circle is the "average" value of all possible initial states.

- A pdf is very much like a mass density function. You know that integrating the mass density over a volume V gives the mass of the object contained in the volume V. Similarly, integrating the pdf of the state over a region A of the state-space gives us the probability of the state assuming *some* value in the region A:

$$P(x_0 \in A) = \int_A \mathcal{W}_0(x) dx \tag{8}$$

  We will obviously elaborate more on the above equation in class. For now, you need to know that at t_0, the state could actually have *any* value in $\Re^N$, and the probability of it being in a particular region can be computed by integrating its initial probability density function.

- The pdf of the state at t_0, i.e. $\mathcal{W}_0(x)$ will depend on the characteristics of the instrument used to measure the initial conditions. We will learn about many different types of pdf's. The most common and perhaps the most important is the Gaussian pdf. A Gaussian pdf in two dimensions is shown in Fig.3(b).



(a) An estimate of the initial conditions of the state as measured by an instrument.

(b) A Gaussian pdf in two-dimensions.

Figure 3: The second source of uncertainty in dynamical systems: random initial conditions.

- Langevin's equation can be generalized to the vector case as follows:

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}) + \mathbf{g}(t, \mathbf{x})\zeta(t), \qquad \mathcal{W}(t_0, \mathbf{x}) = \mathcal{W}_0(\mathbf{x}) \tag{9}$$

  In the above equation, $\mathbf{x} \in \Re^N$ is the state and $\zeta \in \Re^M$ is a M-dimensional noise process. Clearly, $\mathbf{f}(t, \mathbf{x}) : [0, \infty) \times \Re^N \to \Re^N$ and $\mathbf{g}(t, \mathbf{x}) : [0, \infty) \times \Re^N \to \Re^{N \times M}$ is called the noise-influence matrix, containing sufficiently smooth functions. Note that both $\mathbf{f}$ and $\mathbf{g}$ can be highly nonlinear in nature.

- Besides the two sources of uncertainty described above, there is one other important source of randomness that plays havoc with the idea of determinism. This is known as ***parametric uncertainty*** and involves the parameters contained in the function $\mathbf{f}$ and perhaps even $\mathbf{g}$. This is best explained with an example. We see that in Eq.4, $f_2 = -cx_2 - kx_1$. Just like we are not sure about the initial conditions, we could similarly not be sure about the value of the spring constant k. Perhaps it can also only be described probabilistically, just like the initial conditions. This complicates matters even further.

- There exist special techniques for handling parametric uncertainty, the most popular of which is called ***polynomial chaos*** (PC). We will look at PC briefly in this course.

- In light of the above types of uncertainty, we need to ask the question - what knowledge is actually possible? Clearly, plotting trajectories as we did for the ideal deterministic system would make no sense - there are uncountably many of them! We would need to consider an infinite number of noise samples for an infinite number of possible initial conditions, possibly for an infinite number of values of the system parameters! So what do we do?

- The answer must be given in a way that makes sense to the engineer. Actually there are several answers, but according to me, there is only one that makes engineering sense. Mathematicians have tried to give absolute descriptions of the state even in such hopelessly uncertain settings. There has been some success, but the problem lies in the fact that such descriptions are not very useful, especially to the practitioner.

- When the problem set-up is so inherently probabilistic, it makes perfect sense to give estimates about the system that are also probabilistic. This can mean only one thing - instead of trying to solve for the time varying state of the system, i.e. $\mathbf{x}(t)$; solve for its time varying probabilistic description. This will obviously need a lot of explaining. But first, it will require an understanding of what we *mean* by probabilistic descriptions. A glimpse was given when we talked about the probability density function, or the pdf. We looked at the probability density of the state at time t_0, which we called $\mathcal{W}(t_0, x)$. Then, *the complete probabilistic information about the state of the random dynamical system, influenced by one or more of the described forms of uncertainty can be obtained by computing the probability density function of the state at all times, i.e.* $\mathcal{W}(t, x)$. This will always be the ultimate objective.

- In summary, we are attempting to bring together two separate fields in this course - dynamical systems and probability theory. Before we proceed to the technical details, it is important to learn a little about the history of probability theory:

  - **BC** era: Both in Greece and Rome, games of chance were popular. However, there was no scientific or mathematical development of the subject. It is speculated that the main reason was the number system used by the Greeks, which was not amenable to algebraic calculations.
  - **16$^{\text{th}}$** century: Italian mathematician Girolamo Cardano published the first book containing correct methods for calculating the probabilities in games of chance involving dice and cards. Probability was mainly considered an exercise in counting.
  - **17$^{\text{th}}$** century: Work by Fermat and Pascal stirs further interest in the field of probability, which was still being studied as a counting of frequencies.
  - **18$^{\text{th}}$** century: Jacob Bernoulli introduced the first law of large numbers after studying repeated coin tossing experiments. This was an important step towards linking empirical physical reality with conceptual probability. Contributions were made by several heavyweights of mathematics like Daniel Bernoulli, Leibnitz, Bayes, Lagrange and De Moivre. De Moivre introduced the normal distribution and proved the first form of the central limit theorem.
  - **19$^{\text{th}}$** century: Laplace published an influential book, firmly establishing the importance of probability theory as a quantitative field. He also provided a more general version of the central limit theorem. Legendre and Gauss applied probability to problems in astrodynamics through the method of least squares. Poisson developed the Poisson distribution and published an important book with numerous original contributions. Chebyshev and his students, Markov and Lyapunov studied limit theorems, leading to very important results.

- Important fact about development of probability theory thus far: In the development of the theory of probability up to this point, probability was primarily viewed as a natural science, whose primary goal was to explain physical phenomena in the context of repeatable experiments. The keyword here is "repeatable", because so far probability was considered an exercise in counting and then determining the chance of occurrence as the limit of relative frequencies. This *counting* approach to probability is today known as the ***frequentist approach*** to probability. Note that this approach relies heavily on empirical support. Also note that with this approach, it is possible only to deal with finite sized, or at most countable sets of data. Kolmogorov changed all that in the 20^{th} century.
- $\mathbf{20^{th}}$ century: Andrey Kolmogorov is the key figure among several giants who made contributions to this field. It was Kolmogorov however, who established probability theory as a pure branch of mathematics, thus banishing its status as a natural science. He did this by discarding the empirical frequentist approach described above and formulating the ***axiomatic framework*** of probability. Probability could now be studied as a subject *in itself*, purely on the basis of logical correctness arising out of the axiomatic structure, and without any dependance on physical phenomena. This was the turning point for probability theory and today it has become one of the most theoretically sound and unfortunately, sometimes highly esoteric fields of mathematics.

- As mentioned above, this course is about the application of probability theory to dynamical systems. It is therefore important to study that branch of the history as well:
  - The first known application of probability theory to dynamical systems was in the frequentist era (i.e. before Kolmogorov's seminal work), by well known physicists Maxwell and Boltzmann in the 1860's. They were trying to prove that heat in a medium is nothing but the random motion of the constituent gas molecules. With reference to Eq.9, they modeled the system without random perturbation (i.e. $\mathbf{g}(t, \mathbf{x}) = 0$) and only initial state uncertainty. Their work culminated in the Maxwell-Boltzmann distribution, which is the steady state distribution of the gas molecules in an undisturbed medium. Details are available in the 1896 book by Boltzmann. Their work was pioneering, but riddled with annoying paradoxes and inconsistencies. Nonetheless, they were able to account for several properties of gases.
  - Around this time, Rayleigh (1880, 1894), who among other things was also a physicist, studied (unknowingly) the problem of ***random walk*** in two dimensional space. He arrived at a partial differential equation describing the evolution of the displacement of the object performing the random walk. This was later identified as the first form of the **Fokker-Planck equation (FPE)**. FPE is a very important equation in the subject of probabilistic mechanics, having some resemblance (philosophically speaking) to Newton's second law of motion for deterministic mechanics. We will learn about FPE towards the end of this course.
  - French mathematician Bachelier was studying the problem of *gambler's ruin*, which is another manifestation of the random walk problem. He obtained a more

general form of the Fokker-Planck equation. Note that FPE did not get its name at this point because neither Fokker nor Planck have yet entered the picture. It is only in retrospect that we know the equations obtained by Rayleigh, Bachelier and others were nothing but the FPE.

– The stage was set for Albert Einstein, who in 1905 brought together the works of Maxwell and Boltzmann and the random walk approach to develop the theory of Brownian motion. He considered the following simplest possible form of a randomly perturbed system:

$$\dot{x} = \zeta \tag{10}$$

where, x is the displacement of the fluid particle performing Brownian motion and ζ is a random impulse acting on the particle due to collisions with its neighbors. Assuming the displacements to be small, he proceeded to obtain a partial differential equation for the probability density function of x, the displacement of the fluid particle. This was the simplest form of the Fokker-Planck equation. Note that he was considering only a single dimensional Brownian motion (Eq.10 is a scalar ODE).

– In the meantime, there was another significant contribution by Paul Langevin in 1908. He was the first to actually provide Eq.7. In other words, it was his idea to write the dynamics of a randomly perturbed system as an ordinary differential equation, in which the forces are split into a deterministic part and a random disturbance part appearing as an additional forcing function. This served as an amazing clarifying tool for understanding the internal workings of a random dynamical system. Its power lies in its simplicity. We will later see that Langevin's description of a random dynamical system as a system of ODE's encounters serious problems of mathematical rigor. Nonetheless, these problems do not change the fact that it still is a great visualization device for understanding the physics of the problem.

– In 1914, Adriaan Fokker - part time musician and part time physicist applied probability theory to study a first order ODE system with noise and obtained a partial differential equation. He had considered the general case in which the noise intensity was dependent on the state of the system. In 1915, Max Planck generalized Fokker's work to vector systems ($N > 1$ in Eq.9) and applied it to problems in quantum mechanics. **The partial differential equations that Fokker and Planck obtained with the state probability density function as the unknown was named the Fokker-Planck equation.** We will study this equation in detail. But first, we need to build our background in probability theory and understand what a probability density function is!

– The theory of FPE was greatly enhanced and made more abstract by Kolmogorov and his axiomatic framework of probability. He also talked about the uniqueness of the solutions of FPE. To honor Kolmogorov's seminal contributions to this area, the Fokker-Planck equation is also sometimes called the Fokker-Planck-Kolmogorov equation (FPKE), and also Kolmogorov's first equation.

- Since Kolmogorov, the focus on random dynamical systems has been channeled in two main directions: (1) consolidation of the theory of random processes and (2) application of the theory to engineering problems. Notable researchers for the former are Andronov, Uhlenbeck, Ornstein, Chandrashekhar, Smoluchowski, Stratonovich and more recently, Sean Meyn and Kushner.
- A large number of researchers started applying probability theory to study nonlinear random vibration problems in various fields of engineering in the 1960's. Those interested should look at papers by Caughey, Ariaratnam, Dienes, Crandall and Lyon for extremely readable research articles (from the 60's).
- Another chunk of research effort was directed towards the use of probability theory to solve problems in control of nonlinear systems perturbed by random forces. To learn about the early work in this direction, look for papers by Stratonovich, Bellman, Chuang, Kazda and Barrett from the 1960's.
- Today, FPE lies at the heart of numerous problems in engineering - for example, structural vibration for mechanical and civil engineering applications; uncertainty propagation in nonlinear dynamics, including astrodynamics; nonlinear filtering theory; stochastic control theory; chemical process equilibria; particle physics, etc.
- Unfortunately, we have found that FPE is a formidable problem to tackle. Analytical results have not been found except for the simplest of cases. In the 1970's, with the advent of moderately powerful computers, numerical attempts were initiated but quickly ran into multiple roadblocks. We will discuss these roadblocks later. For now, it suffices to mention that the biggest obstruction is something known as the curse of dimensionality. In the 70's, researchers used discretization techniques like finite differences (Killeen, Futch, Whitney etc.). In 1985, Langley used finite elements (FEM) for the first time, which then became the norm for about 20 years. And more recently, I have used what is known as the meshless finite elements technique to correct the shortcomings of FEM for FPE.
- In conclusion, it is important to mention that FPE is the most accurate probabilistic description of a continuous random dynamical system modeled by the Langevin equation in Eq.9. Ever since it was discovered that its analytical solution is extremely difficult to obtain, researchers started looking for alternate methods of analysis. Some of the popular approximate methods are **statistical linearization, Gaussian closure, higher order moment closures and Monte Carlo methods**. Actually, Monte Carlo method is not really an approximate technique - it converges to the true answer (the one that FPE would provide), but only in the limit as the computation effort tends to infinity. The rate of convergence has also been a controversial issue so we will look at Monte Carlo as an approximate technique. We will study at least some of these alternate methods in this course.

References:

1. Arch W. Naylor & George R. Sell, *Linear Operator Theory in Engineering and Science*, Springer Series in Applied Mathematical Sciences # 40, Springer, New York, 1982. ISBN: 0-387-95001-X

2. Dimitri Bertsekas & John N. Tsitsiklis, *Introduction to Probability*, Athena Scientific; 2nd edition (July 15, 2008); ISBN-10: 1-88-652923-X

3. H. Risken, *The Fokker Planck Equation: Methods of Solution and Applications*, Springer Series in Synergetics

4. A. T. Fuller, *Analysis of Nonlinear Stochastic Systems by Means of the Fokker-Planck Equation*, International Journal of Control , Vol. 9, No. 6, 1969, pp. 603655