

**2007**

**S.R. Srinivasa Varadhan**



**ABEL**  
PRISEN

# Autobiography

**S.R. Srinivasa Varadhan**

According to my school records I was born on Jan 2, 1940, in the city of Madras, in the state of Madras in India, which was then a British colony. The city is now called Chennai and the state has become Tamil Nadu in the Republic of India.

My father was born in the last year of the nineteenth century, in 1899 and he married my mother in 1917, when he was eighteen and she was ten. I am an only child and my parents had been married for nearly twenty five years when I was born. Both my parents were the eldest siblings in rather large families and I have always received special attention from all my uncles, aunts, cousins, grandmothers and other assorted relatives. I was born so late that I did not really get to know either of my grandfathers.

As a child I grew up in several small towns not far from Madras. My father was a high school teacher and later principal in the county school system and was periodically transferred from one town to another within the county or district as it is called in India. That was both good and bad for me. I was treated with consideration by all my teachers. But I could not do any mischief at school without my father finding out right away.

Growing up in these small towns was fun. There was plenty of time after school to play with friends on the riverbed that was mostly dry, or play indoors on rainy days. There was very little homework, and in fact there was only minimal learning at school. I did well relative to my class, but was not challenged in any sense. It was only in the last year of the high school, that my mathematics teacher took a special interest in a small group of us and would ask us to come to his house on weekends to do some mathematics problems for fun. More than anything else he taught me that solving mathematical problems or puzzles can be fun.

I had some vague ideas of becoming a doctor as a child. But once, with a group of fellow students from the high school, I went to a medical exhibition at the local

---

S.R.S. Varadhan (✉)

Courant Institute of Mathematical Sciences, New York University, New York, NY 10012,  
USA

e-mail: [varadhan@cims.nyu.edu](mailto:varadhan@cims.nyu.edu)



(a) As a seven year old in traditional Brahmin attire



(b) As a nine year old in a group with teachers in school. I am there because my father was the principal. I am sitting at his feet. He is the one with the head dress (turban)

medical college where the medical students demonstrated their surgical skills on cadavers. That cured me of the desire to enter the medical profession.

In 1954, I moved out to live with my uncle in a suburb of Madras to attend a local college where we were required to study for two years before entering the University. This was a difficult transition. From a rather parochial school system where everything was taught in Tamil, the vernacular, I had to grasp new ideas that were being taught in English. Some of the professors were from UK and I could not quite understand their accent. I studied English and Tamil as my two languages. These were the harder subjects at college since we were expected to write critical essays, where as in school we were only required to remember facts and memorize poetry.

In addition to languages I studied Mathematics, Physics and Chemistry. I did well in all. I liked Physics the most, and I began slowly to see the connections between them.



(c) High school graduation picture. I am standing on the top row. My father as principal is there too, sitting in the front row wearing a turban

It was time to enter the University and the seats were limited and the competition was tough. I came from a Brahmin community, viewed by the government as privileged, and there was reverse discrimination. The goal was to get into a Honors program that led directly to a Master's degree in three years and saved you a year.

It was at this time that I heard of Statistics for the first time. I was told that there was only one college in the state that offered an Honours program in Statistics and they admitted only fourteen students each year. My father knew somebody who knew somebody who could help and luckily I was admitted. I had a choice between Physics and Statistics and I opted for Statistics.

Basically it was a three year program devoted to pure mathematics, probability and statistics. The Presidency College where the program was offered, was situated on the beach. You looked out the window and it was a sandy beach that was as nice as Ipanema. Meanwhile, my father had retired and moved out to the same town as my uncle and I stayed with my parents and commuted for a year. I wanted my parents to let me stay in the dorm for the last two years and they did, after much persuasion on my part.

I really enjoyed those two years. Long walks on the beach with a dozen or so close friends from my class. Lots of movies, arguments, discussions and a generally carefree student life. We are still a close group and see each other from time to time. Although I am the only academic, the others have gone on to do well in government, business and other areas. The studies themselves were relatively easy. I was learning a lot of new concepts and they all felt natural and I could do exceptionally well without much effort. The trouble with the system was that the expectations were not high and it was too rigid. We had only set courses and no electives.



(d) Picture with Kolmogorov during his visit to India (1962). I am third from the left behind Kolmogorov. With me are K.R. Parthasarathy (graduate student), B.P. Adhikari (Professor), me, J. Sethuraman (graduate student), C.R. Rao (Head of the Institute and my advisor) and P.K. Pathak (graduate student)

I graduated after three years having established a new record for grades obtained in the final examination where you are tested in one week on everything that you have learned in three years.

The usual career path at this point is to write a competitive civil service examination that is used to recruit high level government officials. My parents expected me to sit for this. If you succeed, you are set for life. I wanted to go for research. When my father saw that I was firm, he supported me and I went off to Calcutta to study at the Indian Statistical Institute. I had no idea what I was supposed to do. When I arrived in August, 1959, they gave me a desk and expected me to write a thesis in three years. No graduate courses were offered. There were seminars that were optional. I took one on point set topology from Varadarajan and one on measure theory from Bahadur. I did not know what to do and so learned to play bridge and I played a lot. But slowly over the year I met up with a few other graduate students who were there already for a year or two, and we organized our own seminars and programs of study. We lectured to each other, formulated our own problems and tried to solve them. It was a wonderful learning experience and I started to do "Research".

The atmosphere at ISI was very stimulating. We had a steady stream of distinguished visitors, during the pleasant winter months. Sir Ronald Fisher came every

year. I had just missed Norbert Wiener, who had come the year before I went there. We had tea, twice a day, when we all met at the tea room and talked informally. J.B.S. Haldane was a regular at tea, constantly puffing on his cigar. There was always some excitement in the air and some thing or other was always happening. During my second and third year I learned a lot of functional analysis and more or less finished my dissertation. I had begun to learn Markov Processes and wanted to start working on it. During my third year Kolmogorov visited us for two months and I gave a seminar on my thesis and he was in the audience asking questions. He spoke no English, none of us knew any Russian, and so we talked through an interpreter who knew French. A group of graduate students accompanied him on a two week tour of parts of India. He was a member of my thesis committee and brought a copy of my thesis with him, to Moscow, promising to send a report from there. The report came after six months, but only after Parthasarathy, a colleague who went to Moscow on an exchange program, provided a steady daily reminder.

At this point a year had gone by after my dissertation was finished, I formally got my Ph.D., and it was time to go abroad, for a postdoctoral study. In those post Sputnik days that meant USA. Varadarajan, who had just returned to India, after three years in the US, suggested that I go to NYU, to Courant Institute. He wrote a letter on my behalf that went unacknowledged for nearly three months. In the end I was offered a postdoctoral position and I arrived in New York in the Fall of 1963.

There was an international organization that greeted foreign students upon arrival, and they offered to find a place for me to stay for my first night. They met me at the old Idlewild airport and arranged for me to stay in a hotel on Times Square. When I showed up at Courant the next day and told them where I was staying, they were horrified and found me a place to stay on Tenth and University Place. Later I found myself a studio in the West Village.

Before leaving India I became engaged to be married to Vasundara and the wedding was to take place the following June, in Madras. The year went by very fast. I met a small group of Indian academics and over the years we have remained close.

Work went well at Courant. I saw a lot of Mony Donsker, who was the principal probabilist at Courant. There were lots of visitors, seminars and other activities that kept me busy during the week and I socialized with my Indian group during weekends.

Donsker and I, discussed a lot. He had many students working on an assortment of problems. This provided me with a broad perspective and I was beginning to view things from a slightly different angle that made a lot of sense to me. Donsker was interested in specific problems that he wanted to be solved and I was always interested in developing techniques that would enable one to solve a class of problems. We complemented each other and worked closely for over twenty-five years. I went to India to get married in the summer and we found a nice apartment in Washington Square Village. My wife, who was only sixteen, started undergraduate studies at Washington Square College. When my post doctoral fellowship was renewed for two more years it proved ideal for both of us. Then, I joined the faculty in the Fall of 1966, as an assistant professor.

Donsker had a student named Schilder who did a very fine thesis on Laplace asymptotics on Wiener space. I thought a lot about it, and felt that one could develop it



(e) This picture is at the airport on my way to Courant as a postdoc (Fall 64)



(f) This picture is at the airport, returning to USA with Vasu immediately after our wedding in the summer of 1964

considerably. This has been a major thrust of my research over the years and goes by the name of Large Deviation Theory. It is the technique of estimating precisely, in a logarithmic scale, how small the probabilities of certain rare events are. I have found this to be a problem that crops up in many different contexts, and have returned to it again and again during my career.

We used to have a joint seminar with Rockefeller Institute, where Mark Kac was a professor. We would go up to 68th and York often. Once on the way back, in a taxi, someone commented on a result by Cieselski, a Polish mathematician and I immediately saw connections with large deviations and diffusions. The diffusion processes are Markov processes with continuous paths, closely related to second order elliptic partial differential operators. I was able to work out this connection and it proved to be a nice result.

There was a graduate student at Rockefeller University, Dan Stroock, who came by to talk to me about this and we became close friends. He came to Courant in 1966 and stayed for six years. We worked very closely during this period and I believe we changed the way Markov processes were viewed. We introduced systematically methods based on martingales that have become more or less standard today. We jointly wrote a book that appeared in 1979 and has been received well.

In the fall of 1973, we (the family at this time included two sons Gopal born in 69 and we lost him on 9/11/01, and Ashok born 1972) had just returned after a sabbatical from Sweden and India. Donsker as usual was interested in a specific problem about a probabilistic explanation of the variational formula for the ground state energy of the Schrödinger operator. I remember going to Durham to give a talk

at Duke, and sitting in their library before my talk, I saw the connection. This led to a whole series of results on large deviation theory that Donsker and I worked on, till his untimely death in 1991.

I had become an Associate professor in 1968, won an Alfred P. Sloan Fellowship in 1969, and became a Professor in 1972. This was also a tough time at NYU, involving a financial crisis and the sale of the Bronx campus. We moved to Stanford for a year long visit in 1976–77, enjoyed the outdoor life and the open skies. But professionally I could not wait to return to New York. I also started working occasionally with George Papanicolaou, who had lots of applied problems that often required new techniques. I was invited to give a talk at the International Congress of Mathematicians, in Helsinki in 1978 which was a kind of recognition.

I served for four years from 1980–84 as Director of Courant and survived the experience. I was surprised to find that the central administration consisted for the most part of talented individuals doing their best at a difficult task. I began to appreciate for the first time that we at Courant were part of a larger University. My research slowed down some due to the demands of administrative work, but I was able to continue with it. I had started to attract a steady stream of graduate students by this time and enjoyed working with them. I became recognized as an expert on Large Deviations, lectured often on it and wrote a set of lecture notes on the topic.

This was also the time when Wall Street discovered probability. The Black and Scholes model was suddenly very popular. I started some consulting for a small company that managed pension funds and worked closely with Harry Markowitz, who was also a consultant. Harry went on to win the Nobel Prize in Economics. Richard Brignoli, who was the CEO of the company was a maverick and that made the experience all the more enjoyable. It lasted several years, until the company went under. Not because it performed poorly, but the partners had a fallout, fought in court and the lawyers ended up getting everything. Our son Gopal got his first experience working for the company in the summer and later part time during his senior year at college. He chose to stay with the financial industry and his younger brother Ashok has followed him.

I took a sabbatical in 84–85, after serving for four years as Director of the Courant Institute supported partly on a Guggenheim fellowship. Around this time I went to a conference in Marseilles and we usually walked down to the sea and back up the hill, to the campus at Luminy, after lunch. During one of the walks, George explained to me a problem of establishing what he called bulk diffusion under rescaling for interacting diffusions. The problem intrigued me and I thought it would be a simple problem. I spent my sabbatical year thinking about it and found it rather difficult. I found out later that there were a whole class of such problems generally referred to as problems of hydrodynamic scaling. I worked on the problem a lot, but made no progress. Although I had some ideas, they were not sufficient to solve the problem. We had a seminar at which the speaker Josef Fritz from Budapest, proved a result of similar type for a different model. I thought about the new model and found to my surprise, that my ideas worked well for this model and could provide a better result than Fritz. George, Guo who was a student of George, and I worked feverishly on this model and completed our work. This led me to a whole

set of problems requiring new methods and I worked closely with a very talented younger colleague, H.T. Yau. He has since developed the subject and has taken it in a variety of different directions.

The period from 1984–94 was personally very hard for me. My parents grew old and infirm. Since I had no siblings, I had the responsibility for their welfare. This meant frequent and prolonged trips back and forth between New York and Madras. This was hard on my wife, who had stayed in school part time, while raising the two boys, and earned a Ph.D. degree (her fourth degree at NYU) in 1985. My father died in 1990 and my mother in 1994. My wife lost her mother in 1991. The generation above us started thinning out slowly.

There was a bit of a problem at the office in 1992. George, who was to have been the director, left to go to Stanford and there was no time to find a replacement. I was drafted to serve for two years during which period a search committee was appointed and found an excellent choice of director in Dave McLaughlin.

This was also a time when I was receiving honors and recognition and that was gratifying. I was elected a Fellow of American Academy of Arts and Sciences and an Associate Fellow of the Third World Academy of Sciences in 1988, a Fellow of the Institute of Mathematical Statistics in 1991, a member of National Academy of Sciences in 1995, and a Fellow of The Royal Society in 1998. I received the George D. Birkhoff Prize from AMS and SIAM in 1994 and the Leroy P. Steele prize of the AMS that I shared with Dan Stroock in 1996. NYU gave me the Margaret and Herman Sokol award in 1995. I was invited to give a plenary address at the International Congress of Mathematicians in Zurich, in 1994. I was appointed as Frank J. Gould Professor of Science at NYU. In 2001, I was elected to serve as the President of Institute of Mathematical Sciences for 2002–03. I received honorary degrees from University of Paris in 2003 and Indian Statistical Institute in 2004.

Tragedy hit on 9/11/01 when we lost our older son Gopal who was working at the World Trade Center that day. I was in Paris, visiting IHP for a month and was told of the crash by my wife who called me from New York. It took four days to get back to New York. We miss him. Our grandson Gavin, born to our younger son Ashok and his wife Maggie helps take some of the edge off our grief.

Professionally I consider myself to have been fortunate, in that my career spanned a period, when science and mathematics were generously funded by the public. I had a stimulating ambience at Courant that has traditionally provided a very supportive environment for their younger faculty. I have enjoyed working closely with my colleagues as well as everyone of my nearly thirty doctoral students. Finally, I hope to have several more years of productive academic life.

# A Personal Perspective on Raghu Varadhan's Role in the Development of Stochastic Analysis

Terry Lyons

## 1 A Great Day for the Coin Flippers

I know Raghu Varadhan professionally but not personally—that is to say we have attended some of the same conferences and Oberwolfach meetings, and even the odd meal while waiting for trains home. Still, it is obvious to me, and I am sure to anyone else who comes close, that he is a person of great humanity who generates warmth and humour whenever he is in the room. A few months after the award of Fields Medals to Werner, Okounkov and Tao in Madrid, Varadhan and I were both in a group of mathematicians talking about the event. I remember clearly Varadhan's concise summary of the business as “A great day for the coin flippers”. It certainly was: all three used probability in their ground-breaking work and, for the first two, Stochastic Analysis has been a decisive part of their mathematical toolbox. We were all excited that stochastic ideas were having such a substantial effect across areas as far apart as conformal field theory, geometry and number theory. We were also delighted that these achievements were recognized. To me, Varadhan's remark seemed to capture his modesty and humour rather well. Surely it was another excellent day for the coin flippers when Varadhan was awarded the Abel Prize.

## 2 Stochastic Analysis

Stochastic Analysis is an area of mathematics that, in a little over 60 years, has grown from almost nothing to a significant field. It has importance for its intrinsic

---

T. Lyons (✉)  
Mathematical Institute and Oxford-Man Institute, University of Oxford, Oxford  
OX1 3LB, UK  
e-mail: [tlyons@maths.ox.ac.uk](mailto:tlyons@maths.ox.ac.uk)

T. Lyons  
Wales Institute of Mathematical and Computational Sciences, Swansea, SA2 8PP, UK

H. Holden, R. Piene (eds.), *The Abel Prize*,  
DOI [10.1007/978-3-642-01373-7\\_13](https://doi.org/10.1007/978-3-642-01373-7_13), © Springer-Verlag Berlin Heidelberg 2010

interest and for its contributions to other foundational “pure” areas of mathematics, as well as for its contributions to the applications of mathematics. These applications seem to be on an enormous and expanding scale, spanning engineering and climate modelling. It is totally clear to those that work in Stochastic Analysis that Varadhan has shaped the subject in hugely significant ways, both in his personal contributions—in the sense of direction he has given—and in the guidance he has provided to his many outstanding colleagues and students.

I am no historian, and my remarks assigning historical credit should be treated with caution. Given this caveat, I would like to try to place Varadhan’s contributions in context by mentioning a few of the landmark contributions that have shaped the Stochastic Analysis scenery. At least for me, the first wave in the development of Stochastic Analysis is associated with a Russian, a Frenchman, a Japanese man, and an American. The Russian is Kolmogorov, a tower of 20th century mathematics who gave us a rigorous framework for the mathematical study of probability. The Frenchman is Lévy, who made a detailed mathematical study of Brownian motion and Lévy processes. The Japanese man is Itô, who extended differential calculus to the Brownian case and so gave us stochastic calculus. The American is Doob, who gave us the ubiquitous martingale with its optional stopping theorem and up-crossing lemma, and the tools required to provide mathematical confirmation that, in a fair world, there are no free lunches. It is impossible to convey more than a few hints as to the full significance of each of these mathematicians’ contributions.

Kolmogorov’s 66 page monograph, which in 1933 set out a rigorous framework of probability, and his strong law of large numbers, form essential building blocks allowing the rigorous mathematical study of infinite-dimensional probabilistic objects, such as Brownian motion (as studied by Lévy), to flourish. Modern mobile phone technology depends on the strong law of large numbers to separate the transmissions from different phones and allow robust transmission in the context of noise and interference.

Newtonian calculus and differential equations are the classical tools for expressing interactions between evolving systems. Itô’s stochastic calculus (1942) extends the remit of differential equations to systems driven by random processes such as Brownian motion. Itô’s theory gave a direct connection between probability theory and a wide class of second order parabolic partial differential equations. It also provided the framework for many fundamental applications, for example, the continuous time Kalman–Bucy filter which revolutionized the field of estimation and was a major breakthrough in guidance technology (the lunar landing of Apollo is a well documented example). Practical application in non-linear settings requires the construction of numerical approximations to the solutions to certain non-linear PDEs (partial differential equations). Stochastic filters are used in almost all modern military and commercial control systems.

Doob took ideas that were well known in function theory about the behaviour of harmonic functions on the boundary of the disc (Fatou’s theorem, Littlewood’s theorem) and showed how they had probabilistic parallels. He introduced martingales, established optional stopping and up-crossing lemmas, quantified the oscillatory behaviour, and established convergence properties. In effect, he explained how to generalise the notion of a parabolic PDE to the context of functions defined on spaces of

paths. It was a remarkable achievement, demonstrating the enormous power a well chosen abstraction can have.

There is, and will remain, a substantial demand for financial intermediaries, who are able to supply trustworthy and economically priced products, that allow individuals and businesses to insure against financial risks in areas of their business where they have limited expertise and control. Financial intermediaries, who supply these products, must hedge their liabilities, and in general, this hedge will be dynamic and change from day to day, so as to ensure that, at the time an insurance policy matures, the intermediary has the appropriate resources to pay any claim against it.

The correction term (Itô's formula) to Newtonian Calculus to account for the volatility of market prices is at the heart of this hedging process. The provision of these basic insurance activities has become standard over the last 15 years, increasing transparency and forcing margins down. The need for insurance against interest rate and currency fluctuations on a huge scale will not disappear as a result of the current financial storm although one might expect providers to increase their margins somewhat.

Of course, these four mathematicians were not alone (Chung, Dynkin, Khinchine, McKean, Malliavin, and Meyer were some of the others). Nor had the full significance of their contributions been appreciated when Varadhan began to make his own critical interventions.

None the less, it was clear (to those in the area) that Stochastic Analysis had become a set of tools and techniques that could give insight into really quite high-dimensional systems. If one models the evolution of a population, one might use a parabolic PDE to describe the mean local density or intensity of the population as it evolved. Probability can inject additional insight into such deterministic systems. Stochastic analysts have developed tools that model the behaviour of the underlying stochastic population as well as its mean behaviour. This extra step requires more effort and genuinely probabilistic tools; but the value of this distinction, emphasizing the importance of sample paths, provides a clear advantage. In settings, such as control or modelling financial markets, where one has to understand how to interact with and respond to the actual—if uncertain—evolution of the world, the distinction is decisive. Average behaviour is interesting but is often far from the whole story.

Probabilists are not unhappy when their systems are so large and internally homogeneous that on large scales they behave as if they were deterministic. They sometimes call this the fluid or hydrodynamic limit, even though fluids frequently retain very random behaviour on normal as well as microscopic scales!

Varadhan, his collaborators, and his students have made quite fundamental contributions linking the theory of probability to the theory of PDEs, in capturing the information about systems that PDEs miss (large deviations), and in proving that interesting microscopic random systems do indeed have hydrodynamic limiting behaviour. These contributions place him at the core of these developments. I am sure that in his modesty, Varadhan would be the first to note that he is not alone in making fundamental contributions to Stochastic Analysis (for example the work of Friedlin and Ventcell was around the same time and tackles similar issues to some of Varadhan's work), but I am sure that everyone in the field was delighted when Varadhan was awarded the Abel Prize.

### 3 Varadhan

Varadhan's contributions are plentiful, original, beautiful, and surprising. They have had strategic significance in many different directions, so that no-one except perhaps Varadhan could be authoritative on them all, and I doubt if Varadhan could have guessed the range of ways they would be used. I will not (and cannot) attempt a comprehensive survey. I agreed to write this article because I thought it was a great opportunity to express my (inevitably limited) understanding for some of the ways these results have had an impact on the development of Stochastic Analysis, and to explain, through some of Varadhan's striking work, why I think Stochastic Analysis remains an exciting cornerstone of modern mathematics.

Amongst Varadhan's contributions to the development of probability and analysis, I should, at the very least, draw attention to the creation and development of the theory of Large Deviations with Donsker, upon which so much of our understanding of stochastic systems depends. I must also mention the development, alone and with Stroock, of the martingale method for characterising diffusions. This work, with its roots firmly in the probability theory developed by Doob et al., showed the power of the new methods. In a deep piece of work, Stroock and Varadhan produced the first truly satisfactory treatment of elliptic second order parabolic PDEs (in non-divergence form) with continuous coefficients; through this work they demonstrated the power of the new technologies. I should also draw attention to the beautiful work on short time behaviour of diffusions, and mention Varadhan's major foundational contributions to the theory of interacting particle systems and to the development of a theory of hydrodynamic limits for these systems. Varadhan's seminal work with Lions and Papanicolou on the homogenization of the Hamilton–Jacobi equation also has to be on any list, although, unfortunately, space and time mean that I cannot include everything I might wish.

### 4 Independent or Uncorrelated

It might seem like a contradiction, but rare events happen. They can happen frequently, and the nature of the rare events that happen can have an impact in significant ways on the systems around us. I drive a car. I only do this because I believe it is really unlikely that I will have a serious accident—I see it as a rare event. On the other hand, there are many millions of drivers, and looked at across the whole population it is very likely that there will be several car accidents in a single day, and that, sadly, some will be very serious. Car accidents are rare, but they will happen because there are lots of cars.

Now, suppose that there are two classes of accidents:  $A$  and  $B$ . Then asserting that  $A$  and  $B$  are rare events is essentially saying that  $\mathbb{P}(A)$  and  $\mathbb{P}(B)$  are both small. However, in this context it is worth appreciating that, in general, rare events are not likely to be equally rare. For example, it may be that  $\mathbb{P}(A) = 1/100$  and  $\mathbb{P}(B) = 1/1000$ . In this case it is trivial to see that, conditional on one of these two events occurring, then it is about 10 times more likely to be an event of type  $A$

than of type  $B$  which occurs. It does not take much imagination to realize that in large populations which interact in nonlinear ways, it becomes very important to understand the rare events and their relative probability. Rare events can play a key role in determining the function of the entire system.

The theory of Large Deviations provides a systematic mathematical framework for describing and estimating the probabilities of certain rare or exceptional events and predicting their consequences.

Large Deviation theory is a crucial counterbalance to the much better known Central Limit Theorem. The latter remarkable result considers the accumulated effect of repeated independent events. It is a remarkable result that says that the sample mean of the sum of independent and identically distributed variables  $X_i$  with finite variance and mean zero has a distribution that is very close to Gaussian; if  $A$  is open, then:

$$\mathbb{P}\left(\frac{\sum_{i=1}^N X_i}{\sqrt{N}} \in A\right) = \int_A \frac{e^{-\frac{x\sigma^{-1}x}{2}}}{|\sigma|^{1/2}(2\pi)^{d/2}} dx + o(1)$$

where  $\sigma$  is the covariance of  $X$ . The strength and the implications of the Central Limit Theorem cannot be overestimated. It is a hugely valuable theorem, and it can be generalised to random paths.

The limiting distribution is simple and easily described and is entirely determined by the correlations of the underlying incremental event  $X$ . However, the central limit theorem only refers to the likely events. It says that, for the vast majority of the time, the sample mean behaves as if it was a multivariate normally distributed random variable; the covariance  $\sigma$  of  $X$  is all you have to know to understand the distribution of these typical fluctuations.

However, the Central Limit Theorem has nothing to say about rare events, those that happen on the scale of probabilities  $e^{-cN}$  or for the probabilities of fluctuations of the sample mean that are  $o(1)$ , say. These rare events often have a rich behaviour, and at the same time they often play a decisive role in the behaviour of complex systems.

**4.1. A Simple Example.** Let  $X_i \in \mathbb{Z}^d$  be a sequence of independent, identically distributed, random outcomes with mean zero and finite support. Suppose the symmetric quadratic form (known as the covariance):

$$C(u, v) := \mathbb{E}((u \cdot X)(v \cdot X))$$

is finite and non-degenerate, with inverse  $g$ . Consider the random walk on  $\mathbb{Z}^d$  whose transition is  $X$ . That is to say, let  $S_i$  represented the accumulation of these events

$$S_n = \sum_{j=0}^n X_j.$$

It is a basic problem to understand the macroscopic behaviour of the process  $S_n$ , and from a certain point of view Donsker’s Functional Central Limit Theorem provides

a very complete answer (see M.D. Donsker, “Justification and extension of Doob’s heuristic approach to the Kolmogorov-Smirnov theorems”, *Annals of Mathematical Statistics*, 23:277–281, 1952). The process  $B_t^{(n)} := g^{1/2}n^{-1/2}S_{\lfloor nt \rfloor}$  is a very close approximation to Brownian motion; more precisely, its law as a measure on paths converges in the weak topology to Wiener measure.

It is standard practise in many areas of modelling where one accepts that there is a natural independence over disjoint time horizons to measure short term correlations, and then to model the process over longer time horizons using a renormalized Brownian motion. Donsker’s theorem provides a limited justification for this. But there are dangers and, in the context of rare events, it is simply unsound.

For example, a credit agency might wish to assess the probability that a product could fail. They might believe it was safe because, although each component asset was risky, they believed that the probability of a large number of the components failing at one time was very small. This would be the case, if the components behaved independently. However, it would be inordinately difficult to create an empirical test for independence of a large number of these components. However, they may well be able to carry out accurate empirical tests which justify assuming that the values of the components are uncorrelated. It might also be reasonable to assume that the incremental behaviour of these values is independent from one time step to another. Is it reasonable for the agency to use a  $d$ -dimensional Brownian motion to model the evolution of these assets? In general, there is a scale on which the answer is “yes”, but there is a larger scale on which this is a profoundly dangerous thing to do. The large deviations for the actual process might be quite different to the large deviations of a Brownian motion. In general, assuming that a number of events are independent leads to unrealistically small probabilities of rare events. There might well be strong empirical evidence that there is a 1 in  $10^4$  probability that two of the assets lose their value, and the assumption of independence would set the probability of three losing their value at about 1 part in  $10^6$ . We can see how reliable this approach is by considering an example. Suppose that  $X$  takes one of four values

$$\{1, 1, 1\}, \{1, -1, -1\}, \{-1, 1, -1\}, \{-1, -1, 1\}$$

and each is equally likely. Then it is easy to check that  $X$  has the same covariance as the increment of a 3-dimensional Brownian motion over a unit time step. Let  $X_j$  be independent with distribution equal to  $X$  and set

$$S_n = \sum_{j=0}^n X_j.$$

The central limit theorem suggests that  $S_n$  should behave like a Brownian motion and the coordinates should be independent. We can see that this is not true over large scales. Name the coordinate components of  $S_n$

$$S_n = (a_n, b_n, c_n).$$

What happens if  $a_n$  and  $b_n$  are at least  $n - m$ ? It must be the case that

$$|\{j \mid X_j = \{1, 1, 1\}\}| > n - 2m$$

and so  $c_n > n - 2m$ , and for  $m < n/2$  one sees that exceptional behaviour in the first two coordinates forces exceptional behaviour for the third coordinate, contradicting the intuition gained from the fact that the variables are uncorrelated, and the Central Limit Theorem.

If  $\hat{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  is the sample mean of the  $X_i$  and  $E$  is a Borel set with  $0 \notin \bar{E}$  then standard Large Deviations theory gives asymptotic exponential decay rates for the probability that  $\hat{X}_n$  lies in  $E$ :

$$\begin{aligned} \liminf \frac{1}{n} \log \mathbb{P}(\hat{X}_n \in E) &\geq \inf_{e \in E^o} I(z), \\ I(z) &= \sup_{\lambda \in \mathbb{R}^d} (\lambda x - \log \phi(\lambda)), \\ \phi(\lambda) &= \mathbb{E}(e^{\lambda X_1}) \end{aligned}$$

where  $I$  is the Legendre transform of the log of the Laplace transform

$$\phi((x, y, z)) := \frac{1}{4} (e^{2x} + e^{2y} + e^{2z} + e^{(x+y+z)}) (\cosh(x + y + z) - \sinh(x + y + z))$$

of the distribution of  $X$ .

We can compare this with the equivalent upper bound for the case where the increment comes from Brownian motion

$$\begin{aligned} \phi(\lambda) &= \mathbb{E}(e^{\lambda B_1}) \\ &= \exp\left(\frac{1}{2} |\lambda|^2\right) \end{aligned}$$

the function  $\phi$  is quite different; the difference between the two rate functions  $I(z) = \sup_{\lambda \in \mathbb{R}^d} (\lambda x - \log \phi(\lambda))$  would show the differences and thus this dichotomy immediately.

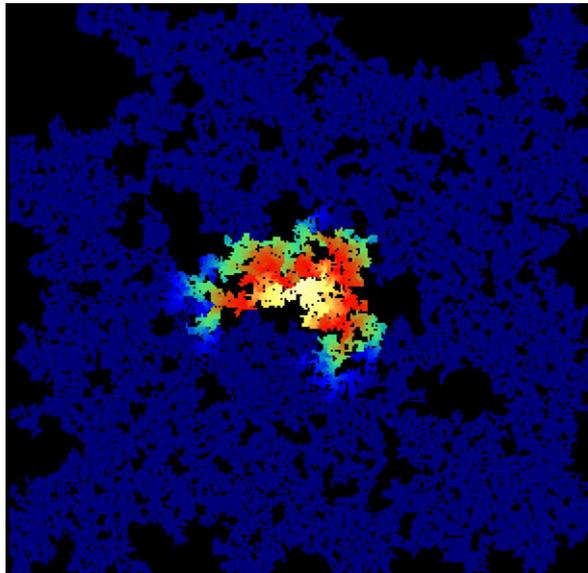
Given examples like the above (where the coordinates of the increments are even pair-wise independent, as well as uncorrelated) it is a surprise to this author that, in a wide range of contexts, one observes people quoting incredibly small probabilities for the occurrence of multiple simultaneous coincidences where there cannot be empirical evidence for these probabilities (because they are too small) and where there is no reasonable justification for the joint independence of the events. We can understand that netting makes sense for a few assets in basket of bonds but systematic collapse cannot be ruled out on the basis of empirical evidence. Similar issues arise in criminal law where any claim that the probability of an alternative explanation has *extremely* small probability, typically obtained by multiplying probabilities, is more likely to be a confirmation of the stochastic illiteracy of the person making the claim than a statement of fact.

Large deviations are not predicted by the co-variance of a system.

## 5 Diffusion Equations and PDE

When pollution spreads through soil, when heat moves away from a semiconductor, or when fluid moves through a pipe, one understands that the observed phenomena represent the averaged behaviour of a huge cloud of microscopic diffusive particles. See, e.g., Fig. 1. But at least in the first two, and for systems with low Reynolds number in the third, the randomness has been so smoothed out on normal scales (by the Central Limit Theorem) that the behaviour has become steady and predictable. Second order elliptic and parabolic PDEs are basic mathematical tools used to model bulk continuum behaviour; these tools are incredibly effective in predicting the behaviour of these systems when they can be applied. However, there are settings where the diffusion is highly inhomogeneous, such as where there are layers of insulation or impervious obstacles distorting this diffusive behaviour even into a lower dimensional flow. In the case of pollution flowing through a porous media, it is easy to imagine that the diffusion is supported on a set of fractional-dimension. Today we see a division between the Stochastic Analysis methods which are well adapted to describing bulk diffusivity on the widest classes of domain, such as fractal sets, and the PDEs methods which depend on the locally smooth nature of the domain. Progress in understanding this more general setting is steady and many questions which have remained open for considerable periods, for example about uniqueness of diffusions on regular fractals, are now being solved as deeper understandings of things like Harnack inequalities and Dirichlet Forms emerge.<sup>1</sup>

I would not want to suggest that PDE methods are in any sense redundant; they are very effective when they can be used and powerful indicators of how to proceed



**Fig. 1** Heat diffusing in a random fractal environment—as simulated by a programme by Martin Barlow

<sup>1</sup>For a survey see [1, 2].

in more general settings where they cannot be appropriate. For example, it is almost a triviality that, up to scaling, there is only one translation and rotation invariant second order differential operator, and one translation and rotation invariant Markov diffusion on  $\mathbb{R}^n$ . This observation leads one to ask for uniqueness of diffusions respecting the symmetries of a regular fractal. The Bass–Barlow proof of uniqueness for the diffusion on the Sierpinski carpet is very hard!

The flow of understanding is not one way and Varadhan, both alone and with collaborators (particularly with Stroock) demonstrated very clearly in his early work that Stochastic Analysis is a powerful tool for tackling some of the harder and more fundamental questions associated to PDEs. If PDEs model the approximately deterministic macroscopic behaviour of large diffusive populations, then there will be cases where the diffusivity of individual elements in the population remains unchanged by the evolution of neighbouring elements, but in other cases one would expect significant interaction. In this latter case the coefficients in the PDEs modelling the macroscopic diffusion will depend on the solution for their values; a priori they are unknown and the PDEs will be non-linear; and when the environment where the diffusion takes place is rapidly fluctuating on small scales the coefficients will oscillate wildly. Even within the elliptic setting, where the diffusion sees a full ball of neighbours instantly (and is not forced into, for example, cracks or fractals), the need for a priori estimates in the study of non-linear PDEs and the need to understand the large scale behaviour of diffusion in highly fluctuating media each separately justify the extension of the mathematical theory of linear parabolic PDEs to the case where the coefficients of these equations are not smooth. The precise mathematical interpretation of a PDE with less than smooth coefficients remains subtle, even in the elliptic case.

### 5.1. Elliptic PDEs in Non-divergence Form with Continuous Elliptic

**Coefficients—Uniqueness of Solution.** A central difficulty is to provide an interpretation of “solution” that is embracing enough to capture the actual diffusive possibilities, while at the same time have an equation that is precise enough that the initial data provided does indeed provide for the unique evolution of the system, and hidden variability does not remain. The approaches split according to whether or not the equations are in divergence form. A key question for the non-divergence form has always been to understand the linear backward operator

$$L := \frac{1}{2} \sum_{i,j \leq n} a_{ij}(t, x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i \leq n} b_i(x) \frac{\partial}{\partial x_i} \quad (5.1)$$

where the coefficients are just bounded. Success in such an a priori approach would allow meaning in a rich and straightforward way for the widest classes of fully non-linear equations.

A fundamental problem has been to establish the existence and uniqueness of fundamental solutions and the existence of an associated semigroup  $P_t$ . Under smoothness assumptions much was known. Nirenberg [10] had established a maximum principle (uniqueness) if the solution was  $C^2$ , while Dynkin and others had

established uniqueness and existence results that went up to the Hölder continuous coefficient case. A very important contribution came from Tanaka (1964) and independently from Krylov (1966) who established compactness results and showed the existence of (possibly non-unique) semigroups associated to the operator (5.1) in the continuous coefficient case. However, neither Krylov nor Tanaka proved uniqueness, and uniqueness is every bit as much an issue as existence; without it the standard PDEs associated to (5.1) would not be well defined.

In a beautifully written pair of papers, Stroock and Varadhan completely solved this problem in the case where the coefficient  $a_{ij}$  was uniformly elliptic, bounded and uniformly continuous, and  $b_i$  was bounded and measurable. The papers were remarkable in so many different ways, giving us new and flexible probabilistic machinery that is still used routinely today, and at the same time demonstrating its worth by solving this basic problem in the analysis of PDEs.

Varadhan [19] had been exploring exponential martingales. Then, in the remarkable papers with Stroock, [12, 13], they assigned a Markov diffusion process to the elliptic PDE (5.1). They proved that in the case where  $a$  is continuous and  $b$  is bounded and measurable, then there is at most (and so exactly) one Markov process  $X$  so that for each  $T > 0$  and each fixed  $\theta$  in  $\mathbb{R}^d$

$$e^{\int_0^t (\theta \cdot dX_s - \frac{1}{2} \theta \cdot a \theta dt - \theta \cdot b dt)}$$

is a martingale. Everything seemed to slot into place. We were given a new technology of weak solutions to stochastic differential equations, and a completely novel way, based on using exponential martingales, to prove that such solutions are unique. At the same time the results of Tanaka and Krylov were extended to show the existence of the weak solution. Stroock and Varadhan moved away from a study of solutions to a PDE (which were obscure) to looking at the action of the operator (5.1) on an exponential function (which was a local calculation and straightforward) and used stochastic integral equations to characterise the Markov process. The proofs are deep and wide ranging; in their later book on the subject they exploit techniques from topics such as singular integrals to get the estimates they need.

One can easily widen the methodology to give good local notions of  $L$ -supersolutions and  $L$ -subsolutions ([15], Remarks just after Thm 3.1), existence of solutions, and a maximum principle ([15] Thm 3.3) so solutions have the appropriate uniqueness. Viscosity-style solutions are not required for a good theory in this setting.

It seemed for many years that this intrinsic difficulty prevented these uniqueness results being extended to the uniformly elliptic bounded measurable case; but in 1997, in [9], Nadirashvili showed that the Stroock–Varadhan result is the best possible and gave the first counterexample to uniqueness in the bounded measurable case.

We emphasise that, although this is a result settling a key result in PDE theory, it, and the methodology developed around it, are important to probability. We should also not be surprised that probability seems the best route, and perhaps the only

route, to studying these PDEs in depth. A Feller diffusion processes  $X_t$  is robustly associated to its transition semigroup defined on  $C_0$  by

$$P_t(f)(x) := \mathbb{E}(f(X_t) | X_0 = x).$$

This semigroup characterises the diffusion, and its infinitesimal generator defines a closed operator  $L$  on  $C_0$ . Without stochastic differential equations, any study of the process associated to it had to go in the reverse direction. Close  $L$ , use the fundamental solutions to the closed PDE to identify the semigroup, and use the semigroup to get the process. This is really quite problematic because PDEs locally characterise the behaviour of functions  $f$ . The beauty of the martingale characterisation of  $X$  is that, in contrast to the semigroup description, it is also a local characterisation of  $X$ . With serious work and the skill of masters, this method would lead to the proof of appropriate uniqueness and existence statements for solutions to these PDEs and many other systems.

At the time of Stroock and Varadhan's work, I do not think it was at all obvious that the correct approach to proving tough questions about these PDEs was through probability. It was a tough journey they made. Now, of course, we realise that Markov diffusions are intrinsically more robust than their PDE counterparts and can be treated in a rigorous mathematical way across a wide context, including some (e.g. fractal environments and path spaces) where classical PDEs have little or no meaning.

**5.2. First-Order PDEs and Large Deviations.** It is a tautology to say that many of the most interesting features of our world and in our own personal experience evolve with time. Some evolve in a clear and predictable way. Others evolve in a less predictable and more random way. Some settle quickly to a large scale equilibrium, where the location and nature of that equilibrium is determined on a microscale by a steady incidence of rare events or large deviations. As a discipline, Stochastic Analysis is concerned with providing mathematical tools that are flexible enough to describe and study such systems, and at the same time concrete enough that they give real insight and lead one to address the real difficulties inherent in modelling them.

I would like to spend a little time discussing a beautiful paper [16] from 1966 where Varadhan is already using probabilistic methods to deeply understand a problem in non-linear first-order PDEs. The paper is fascinating for the way that, drawing on previous work of Donsker and Schilder, it forms a clear precursor result for the Large Deviations theory that would come later, and for the ease with which it brings together analytic tools that are familiar in finite-dimensional contexts (saddle point and steepest descent methods) and uses them in an infinite-dimensional setting, side-by-side with these emerging tools of Large Deviations, to create a decisive result.

Donsker had shown that the solution  $u_\varepsilon$  to the initial value problem

$$u_t = uu_x + \frac{1}{2}\varepsilon^2 u_{xx} + p(t, x)$$

converges, as  $\varepsilon \rightarrow 0$ , to the solution to

$$u_t = uu_x + p(t, x).$$

A key step was the Hopf transformation

$$V(x) := \frac{1}{C} \exp \int^x \frac{-u(y)}{2\kappa} dy$$

which results in  $V$  satisfying a third order PDE that factorises to the extent that it is obvious that if  $V$  satisfies the linear equation

$$V_t = \frac{1}{2}\varepsilon^2 V_{xx} + \frac{1}{\varepsilon^2} p(t, x)V$$

then  $u$  satisfies the original equation. Donsker understood that (at least for each fixed  $\varepsilon > 0$ ) this linear equation has an easy probabilistic interpretation as an exponential path integral. Meanwhile, his student Schilder had, in his remarkable thesis, developed Laplace's method into a rigorous asymptotic expansion for certain families of exponential path integrals (M. Schilder, *Some asymptotic formulas for Wiener integrals*, *Trans. Amer. Math. Soc.* 125 (1966), 63–85).

Varadhan took these results as his backcloth and put them together in a beautiful, clean and more abstract way, with the now famous formulation where one establishes a lower bound on open sets and upper bound on closed sets.

Using this more concise overview and further demonstrating his technical power by proving the necessary bounds for this particular problem, Varadhan established a much more comprehensive result than his predecessors.

Varadhan's result in that paper was to show that if  $f$  is convex then the solution to the initial value problem

$$u_t = [f(u)]_x + p(t, x)$$

can be easily constructed or approximated by successive solutions to linear problems.

The paper is remarkable and influential. Laplace's method has been used in the study of tail behaviour for iid sums since Cramer, and now it was a nontrivial tool that could be effectively used in infinite-dimensional settings—and even applied to give insight into the behaviour and origin of a class of nonlinear PDEs.

The notes from Courant [19] set out the importance of using exponential martingales to study the behaviour of general diffusion processes. This is an idea that is obviously at the heart of the work on the martingale characterization of diffusion processes and its applications to the theory of PDEs with continuous coefficients—one of his major works. Meanwhile [20] is an early example of solving a PDE with homogenization of the coefficients.

**5.3. PDEs and Probability in the 60s.** At the time of this early work described in this section of my article, I do not think it was at all obvious that the correct approach

to answering outstanding questions about these PDEs was through probability; there were other pioneers for sure: Itô and McKean for example. But I am sure it was a tough journey they all made and, as I have mentioned, the results and methodologies stood the test of time.

Forty years later, as we mentioned above, we now understand that Markov diffusions and these methodologies are intrinsically more robust and general than their PDE counterparts and can be treated and exploited in a rigorous mathematical way across a wide context.

The distinction between the large deviations of the two Markov processes with the same co-variance introduced in Sect. 4 vanishes in the diffusion/PDE limit, although it might be critically important to financial regulators.

Viewing macroscopic physical quantities as local integrals of functionals defined on spaces of paths is a deep and powerful concept.

Diffusions can exist in fractal like environments where there are no local charts and classical PDEs have no meaning. They are often given meaning through a theory of Dirichlet forms and weak formulations. Pollutants evolving through soil might be modelled by such methods. At the same time these methods again allow rigorous constructions of diffusions on nonlinear path spaces and lead to rigorous treatment of such things as operators and Sobolev structures on these spaces for the first time.

## 6 The Support Theorem—Understanding the Itô Differential Equation

**6.1. On Extending the Wong–Zakai Theorem.** One of the key results in [12] and [13] associates a unique diffusion process to any PDE of type (5.1), providing that  $a$  is strictly positive as a quadratic form and continuous and  $b$  is bounded and measurable.

If Stroock and Varadhan had assumed their equations had Lipschitz continuous coefficients then (as was remarked in [12]) Itô's theory would have given them all the information they would have needed.

A vector field  $V$  on  $\mathbb{R}^d$  defines, for each  $x \in \mathbb{R}^d$ , a vector or direction  $V(x) = (v_1(x), \dots, v_d(x))$  in  $\mathbb{R}^d$ . Vector fields are useful for describing how systems evolve. For example, one might look for paths  $t \rightarrow \gamma_t \in \mathbb{R}^d$  that have velocity that matches the vector field or, in more mathematical language, solve the differential equation

$$\dot{\gamma}_t = V(\gamma_t).$$

For Lipschitz vector fields there is always a unique solution with given starting point. In the example  $V = (-y, x) \in \mathbb{R}^2$  the trajectories are circles centred at the origin. As we vary the initial state  $\gamma_0$ , the map from initial state to state at time  $s$  defines an invertible map or flow  $\pi_s$  by

$$\pi_s(\gamma_0) := \gamma_s.$$

In our example it rotates the plane by an angle  $s$ .

Now suppose that one places more interest on the interactions of the system  $\gamma$  with other evolving systems  $t$  and  $\tau$ . There are two very simple interactions, the first is replication where every increment of  $\tau$  is replicated by an increment of  $\gamma$

$$d\gamma_t = d\tau_t$$

and the second is the ordinary differential equation reflecting the effect of a vector field, which we now write as

$$d\gamma_t = V(\gamma_t)dt$$

for consistency.

A much richer range of interactions becomes immediately available if we mix these approaches. Suppose that  $V^0, \dots, V^n$  are a family of vector fields on  $\mathbb{R}^d$  and that  $\tau = (\tau_1, \dots, \tau_n)$  is a (smooth enough) path in  $\mathbb{R}^n$  then consider

$$d\gamma_t = \sum_{j=0}^n V^j(\gamma_t) d\tau_{j,t}$$

or

$$d\gamma_t = V(\gamma_t) d\tau_t \tag{6.1}$$

for short. It is helpful to view  $\gamma$  as the response of the system to the control  $\tau$ ; the map  $\tau \mapsto \gamma$  is today known as the Itô map.

This equation includes the previous two via a judicious choice of  $\tau$  and  $V$ . By adjoining variables the framework also includes the apparently more general case

$$d\gamma_t = V(\gamma_t, \tau_t) d\tau_t$$

through replicating the variable  $\tau$  to  $\tilde{\tau}$  and setting  $\tilde{\gamma} = (\gamma, \tilde{\tau})$ :

$$d\tilde{\tau}_t = d\tau_t, \quad \tilde{\tau}_0 = \tau_0$$

$$d\gamma_t = V(\gamma_t, \tilde{\tau}_t) d\tau_t$$

is a differential equation of the prescribed type. In this way, geometers might note that one class of such systems is formed by the connections on a manifold. In this case, the path on the manifold is the control and the horizontal lift to the bundle is the response. Another class of examples appears in Cartan development—in this case the control is naturally the path in the Lie algebra and the response is the path developed onto the Lie group.

In classical control theory the language is usually a little different. One can affect the evolution of a system by applying a control  $\sum c_j(t) V^j$ , mixing the different effects and considering the response

$$d\gamma_t = \sum_{j=0}^n c_j(t) V^j(\gamma_t) dt. \tag{6.2}$$

In optimal control, the discussion usually constrains  $(c_j(t))$  to lie in some fixed compact convex set  $K$  and optimises the choice to maximise some coordinate of the solution at the terminal time. By introducing

$$C_t = \left( \int_0^t c_1(s)ds, \dots, \int_0^t c_n(s)ds \right)$$

one sees that the classical control theory problem is also easily articulated in this language

$$d\gamma_t = V(\gamma_t)dC_t$$

where the standard optimal control problem can then be seen to maximise some coordinate of the solution (the benefit) at the terminal time subject to a constraint (the cost) on the length of the path  $C_t$  in an appropriate semi-norm and parametrises this control at unit speed.

To make sense of (6.1) requires some regularity between  $V$  and  $\tau$ . If  $V$  is Lipschitz and so is  $t \rightarrow \tau_t$ , then, by rewriting the equation in a form similar to (6.2), one can use the classical theory of ordinary differential equations to deduce the existence of unique solutions to the initial value problem and the existence of flows and so on. However, as the replication example makes clear, (6.1) should have meaning in some contexts where  $\tau$  is not differentiable and there are no  $c_j(t)$ . It was the remarkable achievement of Itô, now taught in most major mathematics and engineering schools world-wide, to give this equation meaning as  $\gamma$  ranges over almost every Brownian or semi-martingale path. Because Brownian motion  $W_t \in \mathbb{R}^n$  has independent identically distributed increments with the correct homogeneity, it is at least intuitively obvious that the solutions to the Itô–Stratonovich equation

$$dX_t = V(X_t) \circ dW_t + V^0(X_t)dt, \quad X_0 = a$$

would evolve in a way that was independent of their history, given their current position. That is to say  $X_t$  is a Markov diffusion process. There are subtleties to get the exact theorems, but broadly it is easy to see from Itô’s theory that if

$$Lu = \left( \sum_j v_k^j \frac{\partial}{\partial x_k} v_l^j \frac{\partial}{\partial x_l} + v_k^0 \frac{\partial}{\partial x_k} \right) u \tag{6.3}$$

and  $f(x)$  is a bounded continuous function, then starting  $X$  at  $x$  at time  $t$  and allowing  $X$  to run until time  $T$ , and then evaluating  $f$  at that location and averaging:

$$u(x, t) = \mathbb{E}[f(X_T) \mid X_t = x]$$

gives the solution to the boundary value problem

$$\begin{aligned} \frac{\partial u}{\partial t} &= Lu \\ u(x, T) &= f(x) \end{aligned}$$

so expressing the solution  $u$  as an integral of  $f$  (Itô functional) over Wiener Space. One of the striking points is that the equation did not have to be elliptic or satisfy Hörmander's condition. As the example

$$V = (1, 0, \dots, 0)$$

$$L = \frac{\partial^2}{\partial x_1 \partial x_1}$$

shows clearly, even though  $X$  diffuses it is not forced into any sort of spreading out in all directions; the diffusion, if it starts at a point, might well stay in a sub-manifold going through the starting point (a diffusive Hamiltonian system might still conserve momentum or energy). This explains the difficulty that one has in the subelliptic case in interpreting (6.3) in any naive way. If the process stays on a submanifold, e.g. on a circle centred on the origin, then the solutions will have no natural smoothness as one traverses from one sphere to adjacent ones. In general the curved nature of these surfaces means that  $\frac{\partial^2}{\partial x_i \partial x_j} u$  will have no meaning, even though  $u$  will—under reasonable hypotheses—be smooth on the circle. The fundamental solutions to (5.1) are evolving measures that, in general, do not have a density to Lebesgue measure that can be differentiated twice.

It is obviously an utterly basic question to understand, in terms of equation (5.1), the dependency of the solution at a point  $(x_0, t_0)$  on this boundary data and identify the conservation laws. In probabilistic language: which sets support and contain the evolution? how does the evolution leave the domain?

“The support theorem” is the decisive answer, and today extends even to stochastic PDEs. It was motivated not only by the desire to extend Nirenberg's maximum principle to elliptic  $L$  with continuous coefficients, but also to understand how to generalise Nirenberg's result to the degenerate setting where  $a$  is only non-negative. The solution is a story of paths and control theory. The solution, in [15] and [14], is both intuitive, radical and technically deep. It is a story of a fundamental kind about approximation of solutions to stochastic differential equations by ordinary differential equations. The entire long last chapter of Ikeda and Watanabe's book (another ground breaking contribution to Stochastic Analysis) is dedicated to giving a second proof.

The ramifications of the methods have implications that go far beyond the fundamental problem solved. For example, the ideas underpin the standard construction of stochastic flows—a crucial concept for Bismut in [3] who gave a new proof of the index theorem.

**One-Dimension of Rough Control.** Today, the first steps in understanding this work come from two later and, in some sense, deeply misleading papers, by Doss and Sussman. Consider the controlled differential equation

$$d\tau_t = V(\tau_t)d\gamma_t + V_0(\tau_t)dt \tag{6.4}$$

$$\tau_0 = a$$

where  $\gamma$  is a smooth *real* (i.e., in  $\mathbb{R}^n$ ,  $n = 1$ ) valued path and  $V, V_0$  are Lipschitz vector fields, then the path  $\tau$  is uniquely defined using classical tools. Independently, each showed that the map  $\gamma \mapsto \tau$  extends to all continuous  $\gamma$  and is *continuous in  $\gamma$  in the uniform topology*. It is clearly the unique extension with this property. We can thus give meaning to (6.4) for any continuous real valued path  $\gamma$ . In the case where  $\gamma$  is a realisation of a 1-dimensional Brownian path, this definition almost surely coincides with the Itô–Stratonovich solution.

Now, if  $\gamma_t \in \mathbb{R}$  is a continuous path for  $t \in [0, 1]$ , and  $\gamma_t^{(n)}$  is its dyadic piecewise linear approximation (i.e.  $\gamma_t^{(n)} = \gamma_t$  if  $t \in 2^{-n}\mathbb{Z} \cap [0, 1]$ ,  $\gamma^{(n)}$  is continuous on  $[0, 1]$  and linear on each interval in  $[0, 1] \setminus 2^{-n}\mathbb{Z}$ ) then  $\gamma^{(n)} \rightarrow \gamma$  uniformly. In particular, no matter how rough  $\gamma$ , if one defines  $\tau^{(n)}$  to be the solution of (6.4) obtained by replacing the control  $\gamma$  with its dyadic piecewise linear approximation  $\gamma^{(n)}$  and solving the classical equation, then the  $\tau^{(n)}$  converge to  $\tau$ . This gives a strong intuitive interpretation of (6.4) for rough one-dimensional  $\gamma$ . Restricting to the (Itô–Stratonovich) case where  $\gamma$  is a randomly chosen Brownian path, one recovers the earlier result by Wong and Zakai that gave insight into the meaning and support of one-dimensional SDEs (stochastic differential equations).

**Multiple Dimensions of Rough Control.** The results of Doss and Sussman are misleading—not because they are wrong, but because something rather miraculous happens if  $\gamma$  takes its values in a one-dimensional space. In the multi-dimensional case it is simply and spectacularly false. The Itô functional  $\gamma \mapsto \tau$

$$d\tau_t = \sum_i V^i(\tau_t) d\gamma_t^i + V_0(\tau_t) dt, \quad \tau_0 = a$$

is not at all continuous in  $\gamma$  in the uniform norm if the vector fields  $V^i$  do not commute. It is not even closable in the space of continuous paths.

To get their support theorem, Stroock and Varadhan proved two conceptually fundamental results.

Stroock and Varadhan proved that if one approximates a Brownian path  $X_t$  through its dyadic piecewise linear approximations  $\gamma^{(n)}$  and solves the resulting classical ordinary differential equations then, almost surely, the  $\tau_t^{(n)}$  again converge in law to the Itô–Stratonovich solution to the SDE driven by  $X$ . In this way they proved that the closed support of the law of  $\tau_t$  is contained in the closure of the points where  $\tau$  could reach when controlled by a piecewise smooth path  $\gamma$ .

Stroock and Varadhan also proved that if  $\gamma$  was smooth, and if one conditions a Brownian path to be close to it in the uniform norm, then the solutions to the stochastic equation driven by the conditioned process converge to the solutions to the deterministic equation. From here one can see that the closed support of the  $\tau_t$  controlled by piecewise smooth  $\gamma$  is not bigger than the stochastic support.

From a more modern perspective, this extension of the Wong–Zakai theorem<sup>2</sup> also follows from work of Sipilainen [11] who proved that the dyadic polygonal

---

<sup>2</sup>At around the same time, and independently, J.M.C. Clark also tackled this problem in his doctoral thesis.

approximations to a Brownian path converge almost surely in the rough path metric.

The extension to Wong–Zakai, and the weak continuity theorem are far, far deeper results than the analogous one dimensional results. The difficulty of these results can be gauged by the fact that the statement that a path be close to  $\gamma$  in the uniform norm on  $[0, 1]$  implicitly depended on a notion of distance in  $\mathbb{R}^n$ . Brownian motion conditioned to lie within  $\varepsilon$  of  $\gamma$  for one norm on  $\mathbb{R}^n$  bears little statistical resemblance to the same process conditioned to lie within  $\varepsilon$  of  $\gamma$  for a second norm on  $\mathbb{R}^n$ , because these events correspond, are both rare and essentially distinct. The Stroock–Varadhan argument held for the box norm; later, Ikeda and Watanabe used the last chapter of their book to prove a similar result for the spherical norm. More recent work, presented in *St Flour*, allows the same result for any norm but depends on many more recent and deep developments, such as the boundary Harnack principle of Burdzy and Bass.

The conceptual content of these results is easy to take for granted today because it is so natural that the results should be true; they are basic and absorbed into our folklore. But they are also difficult. Had they not been true, then of course the whole subject would have developed differently. As an illustration of how the techniques go further than the basic results, we briefly discuss the existence of flows. Flows are solutions defined for every starting point  $a$ . Because there are uncountably many  $a$ , the Itô approach does not in itself allow such simultaneous solution. On the other hand the approximations to the solution using piecewise linear paths automatically provide flows for each approximation, and getting this sequence of diffeomorphisms to converge is possible and easier than trying to obtain the flow in other ways directly. Such flows were important to Bismut in his ground-breaking work on the Index theorem.

## 7 The Donsker–Varadhan Theory of Large Deviations

Introducing new mathematical abstraction can be dangerous and sometimes even an intellectual black hole. But it is also the life blood that allows mathematics to move forward and contribute to the wider world. At its most positive, one could think of “zero”—who needs a name for nothing? Closer to home, the introduction of filtrations of sigma algebras (adapted from measure theory) as a way of describing an evolving but partial knowledge of a system seems so dry that it could not have application.

My own score chart is positive about abstraction when it leads to simplification of arguments, and the extension of existing results to a broader class of significant examples. I get very excited when the approach leads to a completely new application. Successful innovations distinguish themselves from the black holes because they tend to be accompanied by an equally serious new approach which, although naturally proved in this level of abstraction, applies widely in settings where the abstract theory that justifies it is of little consequence. Leonardo Pisano brought us zero, but he also gave us Liber Abaci, brought the Hindu-Arabic algorithms

of arithmetic to Europe, demonstrated the value of positional notation in denoting numbers, and made a non-trivial contribution to the development of (Merchant) Banking. Kolmogorov set out the axioms of probability in terms of measure theory, but immediately considered filtrations (sequences of nested  $\sigma$ -algebras), gave us a proper definition of independence and conditional expectation, and proved the strong law of large numbers in this context. The ideas are very abstract—it took me weeks to see any connection between the definitions and the intuitive concepts when I first came across the ideas. But the solid foundations it gave to probability, and the new theorems that came out of it, were a fundamental prerequisite to the massive development of Stochastic Analysis and probability as the toolset we have today.

In developing the theory of Large Deviations, Varadhan made a landmark achievement—he certainly proved hard theorems, demonstrating the value of the framework. Today, whether one is explaining the superconductivity associated to Josephson junctions, modelling stochastic resonance, managing communication networks, understanding the spectra of the large random matrices that occur in the analysis of mobile wireless communication channels, or understanding random surfaces, one gains substantial insight from Large Deviation theory. Rare events can be the crucial element in understanding the evolution of high dimensional (they happen) non-linear (their effects persist) systems.

**7.1. Short Time Behaviour—Geodesics.** Large Deviation theory provides the language and the basic results needed to describe and analyse these types of events, and even to model the systems when they are conditioned to have exceptional behaviour. In the papers [17] and [18] one can see the picture of large deviations theory being shaped as he proves a hard theorem. As before, Varadhan is proving theorems about non-constant coefficient elliptic and parabolic PDEs under relatively weak conditions on the coefficients. For example, consider  $L$  where  $a$  is Hölder and uniformly elliptic:

$$Lu = \sum_{j,k} a_{jk} \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} u.$$

Then one can construct (Miranda, Friedman) a unique minimal solution on  $\mathbb{R}^d$ :

$$\frac{\partial}{\partial t} p_\varepsilon = \frac{1}{2} L p_\varepsilon,$$

which is positive  $p_t(x, y)$  on  $D$  and has boundary data  $\delta_x$  at  $t = 0$ .

In the customary way at that time, (Stroock and Varadhan came later) this is associated to a semi-group and a unique diffusion. Suppose that  $D$  is open and that all boundary points are externally accessible; then we might be interested in the distribution of the time  $T_D$  as the associated diffusion  $X_t$ , started at  $x$ , hits  $\partial D$ . From the Itô perspective it is easy to see that the Laplace transform of this random

time solves the eigenfunction problem

$$\begin{aligned} \phi(x, \lambda) &:= \mathbb{E}_x[e^{\lambda T}] \\ L\phi &= \lambda\phi, \quad \phi|_{\partial D} = 1. \end{aligned}$$

Varadhan gives a bare handed proof that

$$\lim_{\lambda \rightarrow \infty} \left[ -\frac{1}{\sqrt{2\lambda}} \log \phi(x, \lambda) \right] = d(x, \partial D)$$

and uses this to prove that the fundamental solution  $p(t, x, y)$  has the property

$$\lim_{t \searrow 0} -2t \log p(t, x, y) = d(x, y)^2$$

where the convergence is uniform on bounded regions. It is only for solutions defined on the whole of  $\mathbb{R}^d$ . This result is obvious if the coefficients are constant, but far deeper in this generality. It is the core technical bound that, when combined with the Large Deviations framework, gives a series of beautiful results in the second paper and provides a rich immediate justification for the framework.

The rigorous probability starts here. It was well known by the time the paper was written that the conditions on the operator  $L$  were strong enough to associate to it a unique strong Markov diffusion  $(X_t)_{t \in [0, T]}$  whose Feller transition semigroup was  $p_t(x, y)$ . We can use this reference process to define a family of measures on  $C[0, T]$  by rescaling time

$$\mathbb{P}_x^\varepsilon(A) = \mathbb{P}_x(X_{\varepsilon t} |_{t \in [0, T]} \in A).$$

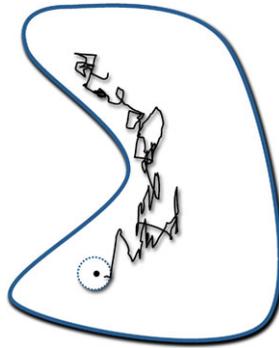
We can fix  $A$  and consider in general the behaviour as  $\varepsilon \rightarrow 0$  of  $\mathbb{P}_x^\varepsilon(A)$ . Note that  $A$  is a generic set of paths and so this is truly an infinite-dimensional question. If the set of paths  $A$  is separated, in the uniform metric, from the path that stays at  $x$  for all time, then the probability that the reparameterised path  $X_{\varepsilon t} |_{t \in [0, T]}$  started at  $x$  is in  $A$  becomes increasingly unlikely as  $\varepsilon \rightarrow 0$ . In fact one can often get a really quite precise asymptotic picture. Suppose that  $x \in D$ , an open set in  $\mathbb{R}^d$ , and  $A_D$  is the set of paths  $\gamma$  in  $C([0, T], \mathbb{R}^d)$  that remain in  $D$  until at least time  $T$  and which, at time  $T$ , are in a ball of radius  $\delta > 0$  centred on  $y \neq x$

$$A_D = \{\gamma \mid \gamma([0, T]) \subset D, d(\gamma(T), y) < \delta\}.$$

The set  $A_D$  is open in the uniform topology and, because  $y \neq x$ , is separated from the trivial path at  $x$ . It is probably best to give an example.

Can one be more precise? For example, how does the probability change if we change the shape of the region  $D$  (in particular if we let it be the whole of  $\mathbb{R}^d$ )? How does it depend on the diffusion coefficients  $a$  that determine  $L$  and  $X$ ?

We have already mentioned that, although rare events are rare, if there are several of them then it is very unusual that they are equally rare. There is an intuition that, if we aggregate a family of rare events  $A_i$  into a single rare event  $\cup A_i$ , one of the rare



events in the aggregate will be much less rare than the others, and so dominate and determine the rareness of the aggregate. This would certainly be true if  $\mathbb{P}_\varepsilon(A_i) = r_i e^{s_i/\varepsilon}$ ; if  $r_j > 0$  and the  $s_j$  are distinct then

$$\max_{j < N} \lim_{\varepsilon \rightarrow 0} \frac{\sum_{i < N} r_i e^{s_i/\varepsilon}}{r_j e^{s_j/\varepsilon}} = 1.$$

At the heart of any Large Deviations argument one has to prove that an appropriate extension of this relationship holds in a setting where the  $i$  range over a continuum instead of a finite set.

Varadhan introduced a functional on paths  $\gamma$  that would determine their local rareness for the diffusion process over short times. It now has a natural feel to it as the energy of the path

$$I(\gamma) := \frac{1}{2} \int_0^T \dot{\gamma}_t a(\gamma_t)^{-1} \dot{\gamma}_t dt$$

(or  $+\infty$  where  $\dot{\gamma}$  is not defined almost everywhere) determined by the diffusion coefficient  $a$ . He could then use his estimate of the probability of the reparameterised process being close to  $\gamma$  to prove the following result.

**Theorem** *Let  $G \subset C([0, T], \mathbb{R}^d)$  be open and  $C \subset C([0, T], \mathbb{R}^d)$  be closed. Then*

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{P}_x^\varepsilon[C] &\leq - \inf_{\gamma \in C} I(\gamma), \\ \liminf_{\varepsilon \rightarrow 0} \varepsilon \log \mathbb{P}_x^\varepsilon[G] &\geq - \inf_{\gamma \in C} I(\gamma). \end{aligned}$$

In this way Varadhan established asymptotic lower bounds on the probability that a random path  $X_{\varepsilon t}$ , run over the interval  $[0, T]$ , would be in a given open set of paths. At the same time he had an asymptotic upper bound on the equivalent inclusion for closed sets of paths. Varadhan also allowed the sets  $G$  and  $C$ , on the left hand side, to vary.

We can now look again at the example of the paths that stay in the domain  $D$ . As  $\varepsilon \rightarrow 0$ , almost all paths from  $x$  to  $y$  will stay close to a minimising geodesic if there

is one. There is an obvious phase transition according to whether  $A_D$  contains the shortest path (in the  $a$  metric) from  $x$  to the ball centred on  $y$ . In our picture it does not.

Varadhan made this all precise and, in a beautiful application, related the fundamental solutions  $p_D, p$  to the PDE restricted to the domain and on the whole of  $\mathbb{R}^d$ . An immediate consequence of Theorem 4.9 in that paper is that if  $p_D$  is the fundamental solution for the equation restricted to a domain  $D$  with Dirichlet boundary conditions then

$$\lim_{t \searrow 0} \frac{p(t, x, y)}{p_D(t, x, y)} = 1$$

for every pair  $(x, y) \in D \times D$  if and only if  $D$  is convex in the Riemannian sense for the metric  $a^{-1}$ . This is a very fine remark which is considerably more delicate than the basic large deviations result above. It symbolises how, even in 1967, Varadhan had very serious examples to test the technology and perfect the theory of Large Deviations as well as beautiful applications.<sup>3</sup>

**7.2. Long Time Behaviour—Occupation Measures.** Suppose a visitor to a finite neighbourhood is forced by his job to move around the neighbourhood in a Markovian way. As he moves around he interacts with the local inhabitants and would like to make friends. But they are cautious and only make friends if the visitor spends a considerable amount of time at their site and gives them significantly more than average attention. On average the visitor will spend equal time at each site, but even over long time intervals then time will not exactly equilibrate. How many friends can the visitor expect to make? How many good friends?

Mathematically, the problem is to look at a recurrent diffusion over long time periods. For example  $X$  might be a Brownian motion on a compact multi-dimensional manifold. Then, from the ergodic theorem, one knows that for almost every  $\omega$  the empirical occupation measure  $\mu_{(T, \omega)}$  defined by

$$\mu_{(T, \omega)}(f) := \frac{1}{T} \int_0^T f(X_t(\omega)) dt$$

will converge to the normalised volume measure. But in general it will never get there. One would like to describe the fluctuations and exceptional behaviour of this random(ly evolving) measure.

I suppose (but do not know) that the theory really became confirmed when, with Donsker [4, 5, 7, 8], Varadhan looked at this second intrinsically important and

---

<sup>3</sup>As matter of history, Varadhan was coming back after a seminar at Rockefeller University in a taxi with colleagues when someone remarked about a theorem of Cieselski that says the ratio  $p_G(t, x, y)$  of the Dirichlet fundamental solution of Brownian Motion to the Gaussian  $p(t, x, y)$  tends to 1 for all  $(x, y) \in G$  if and only if  $G$  is convex (modulo sets of capacity 0). Varadhan wondered what was the Diffusion analog. Varadhan tells the author “I thought the answer should be that diffusions if they go some place in a short time interval will go along geodesics. That was what led me to the short time asymptotics of diffusions.”

clearly infinite-dimensional problem of convergence. Again, the dimensionality and the specificity of the problem ensured that there were hard issues to be resolved, but the Large Deviations framework pointed out the direction and focused the considerable effort required allowing much better results than a more ad hoc approach.

They identified a rate function

$$I(\mu) = \inf_{\substack{u>0 \\ u \in \mathcal{D}(L)}} \int_{x \in M} \frac{Lu}{u}(x)\mu(dx),$$

and established the Large Deviation principle, so providing a powerful family of estimates for exceptional behaviour of the occupation measure and its deviations from the invariant measure. This gives insight into what happens when one gets a large fluctuation and gives valuable information on the dynamics of the system conditional on this exceptional behaviour by minimising  $I(\mu)$ , subject to the constraint of the exceptional behaviour for  $\mu$ .

In other words, if we condition on our traveller making an unusually large number of friends, then we can expect and predict a lot of structure in how these contacts are built up. This work on occupation measures is highly applicable and recaptures Kac’s use of exponential functionals to capture solutions to PDEs. As before there is always at least one basic concrete result that is needed to make it all work (no free lunch!). In this case one could cite [6], which proves under appropriate hypotheses that

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \mathbb{E}_x [e^{-t\Phi(\hat{\mu})}] = - \inf_{\mu} \frac{1}{8} \left( \int \left| \nabla \left( \log \frac{d\mu}{dv} \right) \right|^2 d\mu + \Phi(\mu) \right).$$

This is determined using the diffusion metric, where  $\nu$  is the invariant measure,  $\hat{\mu}$  is the normalised general random occupation measure, and  $\Phi$  is a lsc functional on probability measures with compact sets  $\Phi \leq \lambda$ .

The ramifications of the methodology established in these early papers have spread so widely that it would be impossible to do the whole subject any sort of credit and in any case others are better equipped to do this than this author.

## 8 Hydrodynamical Limits, Interacting Particles and Other Questions

We have seen that large scale non-linear phenomena can arise because of the influence of large deviations. The effect is exactly a balanced trade off between the rareness of the event, the magnitude of the event, and the influence it has on the system as a whole.

However, one needs caution. For the fact that large deviations matter throws doubt onto the framework in which they are studied, on the non-linear PDEs they justify, and on modelling in general. Most real world models have a granularity, but, like the simple random walk, sometimes there are central limit type theorems,

and they can be modelled very effectively by Markov diffusion processes. However, as the introductory section in this article shows, the large deviations of the granular process may be very different to those of the limiting Markov diffusion. The macroscopic large deviations of the granular process are a function of the microstructure of the process. Microstructure that is lost on taking the diffusion limit. This is important, and also a challenge—one cannot interchange orders of limits without justification.

So, there is a very real challenge to take reasonable microscopic models of non-linear physical systems (such as fluids), systems where the non-linear and large deviation effects are likely to be highly relevant, and prove that the large scale behaviour can be successfully modelled (with PDEs or by other means). This is a major programme, proving there is a fluid limit for models with non-trivial microstructure, of current interest. Varadhan, with younger colleagues and students Kipnis, Olla (Rome), Quastel (Toronto) and H.T. Yau (Harvard) did as much as anyone to initiate this very important programme.

Almost by definition, the current state of understanding on the hydrodynamical limits and on particle systems has a huge amount of detail, and would require another long chapter to do it justice. Moreover, and sadly, I do not have enough expertise to summarise it accurately.

**8.1. Homogenised Hamilton–Jacobi Equations.** Sometimes homogenisation works, you can forget local structure, and central limit phenomena dominate! In this case, one might replace the rapidly fluctuating coefficients in the basic diffusion with a much more smoothly varying model, but still have a diffusion which, on the appropriate scales, was indistinguishable from the original. Varadhan had early work on this homogenising the coefficients of PDEs. He was also a co-author of the very influential unpublished—but widely circulated and referenced—contribution on homogenisation of the Hamilton–Jacobi equation, with Lions and Papanicolou which provided the weak viscosity solutions to the HJE equation that underpin recent work on the construction of Mather invariant sets, reconnecting PDEs and Lagrangian dynamics.

The impact of this work persists to the current time. KAM theory provides a suitable framework for understanding the dynamics of attracting sets for a system subject to a small periodic potential. Aubry and Mather, and then Mather alone, constructed invariant sets in the non-KAM non-perturbative framework. The work of Fathi and Evans, using weak viscosity solutions to the HJE developed in the seminal work with Lions and Papanicolou, constructs Mather invariants in the non-KAM non-perturbative framework.

## 9 Conclusion

The Fields Medals at Madrid illustrate how probability has come of age. Okounkov used random surfaces to connect work of Gromov and Witten with that of Donaldson

and Thomas. Werner's work is motivated by statistical physics and links the evolving geometry of two-dimensional Brownian motion to conformal field theory. Tao used probabilistic motivations for many aspects of his plenary lecture.

It is inconceivable that Okounkov and Werner have not been substantially, even subconsciously, influenced by the work of Varadhan and his collaborators. Varadhan's work has influenced applications at the same level and his results and perspectives infuse over a wide horizon. Much of my work would have little interest without his contributions.

Probabilists around the world were delighted when Varadhan was awarded the Abel Prize; his deep and decisive contributions to this significant and young area of mathematics ensure that the Abel prize will get distinction from choosing Varadhan!

The author of this short article would like to offer his thanks to Professor Varadhan for his inspiration, to the editors for their patience, and to the reader for getting this far. At the same time as he hopes to have given a flavour of some of Varadhan's work, he is very conscious of, almost by the same token, the simplifications and omissions that have distorted the accuracy of the presentation; for this he asks the readers' forgiveness.

## References

1. Barlow, M.T.: Heat kernels and sets with fractal structure. *Contemp. Math.* **338**, 11–40 (2003)
2. Barlow, M.T., Bass, R.F., Kumagai, T., Teplyaev, A.: Uniqueness of Brownian motion on Sierpinski carpets. [arXiv:0812.1802v1](https://arxiv.org/abs/0812.1802v1) (2008)
3. Bismut, J.-M.: *Mécanique aléatoire*. Lecture Notes in Mathematics, vol. 866. Springer, Berlin (1981). With an English summary
4. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. I. *Commun. Pure Appl. Math.* **28**, 1–47 (1975)
5. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. II. *Commun. Pure Appl. Math.* **28**, 279–301 (1975)
6. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Wiener integrals for large time. In: *Functional Integration and Its Applications*, Proc. Internat. Conf., London, 1974, pp. 15–33. Clarendon Press, Oxford (1975)
7. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. III. *Commun. Pure Appl. Math.* **29**(4), 389–461 (1976)
8. Donsker, M.D., Varadhan, S.R.S.: Asymptotic evaluation of certain Markov process expectations for large time. IV. *Commun. Pure Appl. Math.* **36**(2), 183–212 (1983)
9. Nadirashvili, N.: Nonuniqueness in the martingale problem and the Dirichlet problem for uniformly elliptic operators. *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (4)* **24**(3), 537–549 (1997)
10. Nirenberg, L.: A strong maximum principle for parabolic equations. *Commun. Pure Appl. Math.* **6**, 167–177 (1953)
11. Sipilainen, E.-M.: *A pathwise view of solutions of stochastic differential equations*. PhD thesis, University of Edinburgh (1993)
12. Stroock, D.W., Varadhan, S.R.S.: Diffusion processes with continuous coefficients. I. *Commun. Pure Appl. Math.* **22**, 345–400 (1969)
13. Stroock, D.W., Varadhan, S.R.S.: Diffusion processes with continuous coefficients. II. *Commun. Pure Appl. Math.* **22**, 479–530 (1969)
14. Stroock, D.W., Varadhan, S.R.S.: Diffusion processes with boundary conditions. *Commun. Pure Appl. Math.* **24**, 147–225 (1971)

15. Stroock, D.W., Varadhan, S.R.S.: On the support of diffusion processes with applications to the strong maximum principle. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Univ. California, Berkeley, CA, 1970/1971. Probability Theory, vol. III, pp. 333–359. Univ. California Press, Berkeley (1972)
16. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. *Commun. Pure Appl. Math.* **19**, 261–286 (1966)
17. Varadhan, S.R.S.: Diffusion processes in a small time interval. *Commun. Pure Appl. Math.* **20**, 659–685 (1967)
18. Varadhan, S.R.S.: On the behavior of the fundamental solution of the heat equation with variable coefficients. *Commun. Pure Appl. Math.* **20**, 431–455 (1967)
19. Varadhan, S.R.S.: Stochastic processes. Notes based on a course given at New York University during the year 1967/68. Courant Institute of Mathematical Sciences, New York University, New York (1968)
20. Varadhan, S.R.S.: Boundary value problems with rapidly oscillating random coefficients. In: *Random Fields*, Vols. I, II, Esztergom, 1979. *Colloq. Math. Soc. János Bolyai*, vol. 27, pp. 835–873. North-Holland, Amsterdam (1981)

# List of Publications for S.R. Srinivasa Varadhan

## 1962

- [1] Limit theorems for sums of independent random variables with values in a Hilbert space. *Sankhyā Ser. A*, 24:213–238.
- [2] (with K.R. Parthasarathy and R. Ranga Rao). On the category of indecomposable distributions on topological groups. *Trans. Amer. Math. Soc.*, 102:200–217.

## 1963

- [3] (with K.R. Parthasarathy and R. Ranga Rao). Probability distributions on locally compact abelian groups. *Illinois J. Math.*, 7:337–369.

## 1964

- [4] Limit theorems in probability. *Math. Student*, 32:17–21.
- [5] (with K.R. Parthasarathy). Extension of stationary stochastic processes. *Theor. Probab. Appl.*, 9:65–71. Also available in *Teor. Veroyatnost. i Primenen.*, 9:72–78 (in Russian).

## 1966

- [6] Asymptotic probabilities and differential equations. *Comm. Pure Appl. Math.*, 19:261–286.

## 1967

- [7] Diffusion processes in a small time interval. *Comm. Pure Appl. Math.*, 20:659–685.
- [8] On the behavior of the fundamental solution of the heat equation with variable coefficients. *Comm. Pure Appl. Math.*, 20:431–455.

## 1968

- [9] *Stochastic processes*. Courant Institute of Mathematical Sciences, New York University, New York.

## 1969

- [10] (with D.W. Stroock). Diffusion processes with continuous coefficients. I. *Comm. Pure Appl. Math.*, 22:345–400. Also in *Matematika: Periodicheskii sbornik perevodov inostrannykh statei. Tom 15 (1971)*, vyp. 6, pages 66–113. Izdat. “Mir”, Moscow, 1971 (in Russian).

- [11] (with D.W. Stroock). Diffusion processes with continuous coefficients. II. *Comm. Pure Appl. Math.*, 22:479–530. Also in *Matematika: Periodicheskii sbornik perevodov inostrannykh statei. Tom 16 (1972), vyp. 1*, pages 100–142. Izdat. “Mir”, Moscow, 1972 (in Russian).

## 1971

- [12] (with D.W. Stroock). Diffusion processes and martingales. I. In *Martingales (Report Meeting, Oberwolfach, 1970)*, pages 60–66. Lecture Notes in Math., Vol. 190. Springer, Berlin.
- [13] (with D.W. Stroock). Diffusion processes and martingales. II. In *Martingales (Report Meeting, Oberwolfach, 1970)*, pages 67–75. Lecture Notes in Math., Vol. 190. Springer, Berlin.
- [14] (with D.W. Stroock). Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.*, 24:147–225.

## 1972

- [15] (with D.W. Stroock). On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, pages 333–359. Berkeley, Calif., 1972. Univ. California Press.
- [16] (with D.W. Stroock). Diffusion processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability Theory*, pages 361–368. Berkeley, Calif., 1972. Univ. California Press.
- [17] (with D.W. Stroock). On degenerate elliptic-parabolic operators of second order and their associated diffusions. *Comm. Pure Appl. Math.*, 25:651–713, 1972.

## 1973

- [18] (with D.W. Stroock). Limit theorems for random walks on Lie groups. *Sankhyā Ser. A*, 35(3):277–294.
- [19] Strassen’s version of the law of the iterated logarithm. In *Topics in probability theory (Sem., Courant Inst. Math. Sci., New York, 1971–1972)*, pages 15–31. Courant Institute of Mathematical Sciences, New York University.
- [20] (with D.W. Stroock). Martingales. I, II, III. In *Topics in probability theory (Sem., Courant Inst. Math. Sci., New York, 1971–1972)*, pages 113–161. Courant Institute of Mathematical Sciences, New York University.
- [21] (with G.C. Papanicolaou). A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations. *Comm. Pure Appl. Math.*, 26:497–524.
- [22] (with D.W. Stroock). Probability theory and the strong maximum principle. In *Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971)*, pages 215–220. Amer. Math. Soc., Providence, R.I., 1973.
- [23] D.W. Stroock and S.R.S. Varadhan, editors. *Topics in probability theory*. Courant Institute of Mathematical Sciences, New York University, New York.

## 1974

- [24] (with D. Stroock). A probabilistic approach to  $H^p(R^d)$ . *Trans. Amer. Math. Soc.*, 192:245–260.
- [25] *Mathematical statistics*. Courant Institute of Mathematical Sciences, New York University, New York.

## 1975

- [26] E. Phillips and S. Varadhan, editors. *Ergodic theory*. Courant Institute of Mathematical Sciences, New York University, New York. With contributions by R. Adler, S. Alpern, N. Bitzenhofer, E. Phillips and S. Varadhan.

- [27] (with M.D. Donsker). Asymptotic evaluation of certain Wiener integrals for large time. In *Functional integration and its applications (Proc. Internat. Conf., London, 1974)*, pages 15–33. Clarendon Press, Oxford.
- [28] (with M.D. Donsker). Large deviations for Markov processes and the asymptotic evaluation of certain Markov process expectations for large times. In *Probabilistic methods in differential equations (Proc. Conf., Univ. Victoria, Victoria, B.C., 1974)*, pages 82–88. Lecture Notes in Math., Vol. 451. Springer, Berlin.
- [29] (with M.D. Donsker). Asymptotics for the Wiener sausage. *Comm. Pure Appl. Math.*, 28(4):525–565. Erratum, *Comm. Pure Appl. Math.*, 28(5):677.
- [30] (with M.D. Donsker). Asymptotic evaluation of certain Markov process expectations for large time. I. II. *Comm. Pure Appl. Math.*, 28:1–47; *ibid.* 28:279–301.
- [31] (with M.D. Donsker). On a variational formula for the principal eigenvalue for operators with maximum principle. *Proc. Nat. Acad. Sci. U.S.A.*, 72:780–783.

## 1976

- [32] (with M.D. Donsker). On some problems of large deviations for Markov processes. In *Proceedings of the 40th Session of the International Statistical Institute (Warsaw, 1975)*, Vol. 1. *Invited papers*, volume 46, pages 409–416, 417–418.
- [33] (with M.D. Donsker). Asymptotic evaluation of certain Markov process expectations for large time. III. *Comm. Pure Appl. Math.*, 29(4):389–461.
- [34] (with M.D. Donsker). On the principal eigenvalue of second-order elliptic differential operators. *Comm. Pure Appl. Math.*, 29(6):595–621.

## 1977

- [35] (with M.D. Donsker). Some problems of large deviations. In *Symposia Mathematica, Vol. XXI (Convegno sulle Misure su Gruppi e su Spazi Vettoriali, Convegno sui Gruppi e Anelli Ordinati, INDAM, Rome, 1975)*, pages 313–318. Academic Press, London.
- [36] (with G.C. Papanicolaou and D. Stroock). Martingale approach to some limit theorems. In *Papers from the Duke Turbulence Conference (Duke Univ., Durham, N.C., 1976)*, Paper No. 6, pages ii+120 pp. Duke Univ. Math. Ser., Vol. III. Duke Univ., Durham, N.C.
- [37] (with M.D. Donsker). On laws of the iterated logarithm for local times. *Comm. Pure Appl. Math.*, 30(6):707–753.

## 1978

- [38] (with M.D. Donsker). On the number of distinct sites visited by a random walk. In *Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978)*, pages 57–62. Academic Press, New York.

## 1979

- [39] (with M.D. Donsker). On the number of distinct sites visited by a random walk. *Comm. Pure Appl. Math.*, 32(6):721–747.
- [40] (with D.W. Stroock). *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin. Reprint, *Classics in Mathematics*, Springer-Verlag, Berlin, 2006.
- [41] (with M.D. Donsker). On the principal eigenvalue of elliptic second order differential operators. In *Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976)*, pages 41–47, New York. Wiley.

**1980**

- [42] (with G.C. Papanicolaou). Diffusion in regions with many small holes. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, pages 190–206. Springer, Berlin.
- [43] *Lectures on diffusion problems and partial differential equations*, volume 64 of *Tata Institute of Fundamental Research Lectures on Mathematics and Physics*. Tata Institute of Fundamental Research, Bombay.
- [44] (with K.L. Chung). Kac functional and Schrödinger equation. *Studia Math.*, 68(3):249–260.
- [45] Some problems of large deviations. In *Proceedings of the International Congress of Mathematicians (Helsinki, 1978)*, pages 755–762, Helsinki. Acad. Sci. Fennica.
- [46] (with M.D. Donsker). A law of the iterated logarithm for total occupation times of transient Brownian motion. *Comm. Pure Appl. Math.*, 33(3):365–393.

**1981**

- [47] Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloq. Math. Soc. János Bolyai*, pages 835–873. North-Holland, Amsterdam.
- [48] (with M.D. Donsker). Some problems of large deviations. In *Stochastic differential systems (Visegrád, 1980)*, volume 36 of *Lecture Notes in Control and Information Sci.*, pages 41–46. Springer, Berlin.
- [49] (with M.D. Donsker). The polaron problem and large deviations. *Phys. Rep.*, 77(3):235–237.

**1982**

- [50] Large deviations. In *Advances in filtering and optimal stochastic control (Cocoyoc, 1982)*, volume 42 of *Lecture Notes in Control and Inform. Sci.*, pages 382–392. Springer, Berlin.
- [51] (with G.C. Papanicolaou). Diffusions with random coefficients. In *Statistics and probability: essays in honor of C. R. Rao*, pages 547–552. North-Holland, Amsterdam.

**1983**

- [52] Random walks among random scatterers. In *Theory and application of random fields (Bangalore, 1982)*, volume 49 of *Lecture Notes in Control and Inform. Sci.*, pages 282–283. Springer, Berlin, 1983.
- [53] (with M.D. Donsker). Asymptotic evaluation of certain Markov process expectations for large time. IV. *Comm. Pure Appl. Math.*, 36(2):183–212.
- [54] (with M.D. Donsker). Asymptotics for the polaron. *Comm. Pure Appl. Math.*, 36(4):505–528.

**1984**

- [55] *Large deviations and applications*, volume 46 of *CBMS–NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

**1985**

- [56] Large deviations and applications. *Exposition. Math.*, 3(3):251–272.
- [57] (with C. Kipnis). Central limit theorems for additive functionals of reversible Markov chains and applications. *Astérisque*, (132):65–70.
- [58] (with G. Papanicolaou). Ornstein–Uhlenbeck process in a random potential. *Comm. Pure Appl. Math.*, 38(6):819–834.

- [59] (with M.D. Donsker). Large deviations for stationary Gaussian processes. In *Stochastic differential systems (Marseille–Luminy, 1984)*, volume 69 of *Lecture Notes in Control and Inform. Sci.*, pages 108–112. Springer, Berlin, 1985.
- [60] (with R.J. Williams). Brownian motion in a wedge with oblique reflection. *Comm. Pure Appl. Math.*, 38(4):405–443.
- [61] (with M.D. Donsker). Large deviations for stationary Gaussian processes. *Comm. Math. Phys.*, 97(1–2):187–210.

## 1986

- [62] Stochastic differential equations—large deviations. In *Phénomènes critiques, systèmes aléatoires, théories de jauge, Part I, II (Les Houches, 1984)*, pages 625–678. North-Holland, Amsterdam.
- [63] V.F. Kolchin. *Random mappings*. Translation Series in Mathematics and Engineering. Optimization Software Inc. Publications Division, New York. Translated from the Russian, With a foreword by S.R.S. Varadhan.
- [64] (with C. Kipnis). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19.
- [65] (with A.-S. Sznitman). A multidimensional process involving local time. *Probab. Theory Relat. Fields*, 71(4):553–579.

## 1987

- [66] (with M.D. Donsker). Large deviations for noninteracting infinite-particle systems. *J. Statist. Phys.*, 46(5–6):1195–1232.
- [67] K. Itō. *Selected papers*. Springer-Verlag, New York. Edited and with an introduction by S.R.S. Varadhan and D.W. Stroock.

## 1988

- [68] Large deviations and applications. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 1–49. Springer, Berlin.
- [69] M. Métivier, J. Neveu, and S.R.S. Varadhan, editors. *Colloque Paul Lévy sur les Processus Stochastiques*. Société Mathématique de France, Paris.
- [70] (with M.Z. Guo and G.C. Papanicolaou). Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.*, 118(1):31–59.

## 1989

- [71] (with B. Gopinath and J.A. Gubner). Bounding functions of Markov processes and the shortest queue problem. *Adv. Appl. Probab.*, 21(4):842–860.
- [72] (with M.D. Donsker). Large deviations from a hydrodynamic scaling limit. *Comm. Pure Appl. Math.*, 42(3):243–270.
- [73] (with C. Kipnis and S. Olla). Hydrodynamics and large deviation for simple exclusion processes. *Comm. Pure Appl. Math.*, 42(2):115–137.

## 1990

- [74] On the derivation of conservation laws for stochastic dynamics. In *Analysis, et cetera*, pages 677–694. Academic Press, Boston, MA.

## 1991

- [75] (with J.R. Baxter and N.C. Jain). Some familiar examples for which the large deviation principle does not hold. *Comm. Pure Appl. Math.*, 44(8–9):911–923.
- [76] (with S. Olla). Scaling limit for interacting Ornstein-Uhlenbeck processes. *Comm. Math. Phys.*, 135(2):355–378.
- [77] Scaling limits for interacting diffusions. *Comm. Math. Phys.*, 135(2):313–353.

**1992**

- [78] Entropy methods in hydrodynamical scaling. In *Mathematical physics, X (Leipzig, 1991)*, pages 103–112. Springer, Berlin.

**1993**

- [79] Relative entropy and hydrodynamic limits. In *Stochastic processes*, pages 329–336. Springer, New York.
- [80] Nonlinear diffusion limit for a system with nearest neighbor interactions. II. In *Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990)*, volume 283 of *Pitman Res. Notes Math. Ser.*, pages 75–128. Longman Sci. Tech., Harlow.
- [81] Entropy methods in hydrodynamic scaling. In *Nonequilibrium problems in many-particle systems (Montecatini, 1992)*, volume 1551 of *Lecture Notes in Math.*, pages 112–145. Springer, Berlin.
- [82] (with S. Olla and H.-T. Yau). Hydrodynamical limit for a Hamiltonian system with weak noise. *Comm. Math. Phys.*, 155(3):523–560.
- [83] (with H. Berestycki and L. Nirenberg). État fondamental et principe du maximum pour les opérateurs elliptiques du second ordre dans des domaines généraux. *C. R. Acad. Sci. Paris Sér. I Math.*, 317(1):51–56.

**1994**

- [84] Regularity of self-diffusion coefficient. In *The Dynkin Festschrift*, volume 34 of *Progr. Probab.*, pages 387–397. Birkhäuser Boston, Boston, MA.
- [85] (with T. Digernes and V.S. Varadarajan). Finite approximations to quantum systems. *Rev. Math. Phys.*, 6(4):621–648.
- [86] (with H. Berestycki and L. Nirenberg). The principal eigenvalue and maximum principle for second-order elliptic operators in general domains. *Comm. Pure Appl. Math.*, 47(1):47–92.

**1995**

- [87] Entropy methods in hydrodynamic scaling. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 196–208, Birkhäuser, Basel.
- [88] The work of Pierre-Louis Lions. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pages 6–10, Birkhäuser, Basel. Also in *Fields Medalists' lectures*, volume 5 of *World Sci. Ser. 20th Century Math.*, pages 555–559. World Sci. Publ., River Edge, NJ, 1997.
- [89] Self-diffusion of a tagged particle in equilibrium for asymmetric mean zero random walk with simple exclusion. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(1):273–285.

**1996**

- [90] Nongradient models in hydrodynamic scaling. In *Analysis, geometry and probability*, volume 10 of *Texts Read. Math.*, pages 397–416. Hindustan Book Agency, Delhi.
- [91] (with A.F. Ramirez). Relative entropy and mixing properties of interacting particle systems. *J. Math. Kyoto Univ.*, 36(4):869–875.
- [92] The complex story of simple exclusion. In *Itô's stochastic calculus and probability theory*, pages 385–400. Springer, Tokyo.
- [93] (with C. Landim and S. Sethuraman). Spectral gap for zero-range dynamics. *Ann. Probab.*, 24(4):1871–1902.

**1997**

- [94] (with H.-T. Yau). Diffusive limit of lattice gas with mixing conditions. *Asian J. Math.*, 1(4):623–678.
- [95] (with J. Quastel). Diffusion semigroups and diffusion processes corresponding to degenerate divergence form operators. *Comm. Pure Appl. Math.*, 50(7):667–706.
- [96] (with H.T. Yau). Scaling limits for lattice gas models. In *Stochastic processes and functional analysis (Riverside, CA, 1994)*, volume 186 of *Lecture Notes in Pure and Appl. Math.*, pages 39–43. Dekker, New York.

**1999**

- [97] Large deviations for interacting particle systems. In *Perplexing problems in probability*, volume 44 of *Progr. Probab.*, pages 373–383. Birkhäuser Boston, Boston, MA.
- [98] E.P. Hsu and S.R.S. Varadhan, editors. *Probability theory and applications*, volume 6 of *IAS/Park City Mathematics Series*. American Mathematical Society, Providence, RI.
- [99] (with J. Quastel and F. Rezakhanlou). Large deviations for the symmetric simple exclusion process in dimensions  $d \geq 3$ . *Probab. Theory Related Fields*, 113(1):1–84.
- [100] M. Giaquinta, J. Shatah, and S.R.S. Varadhan, editors. *Differential equations: La Pietra 1996*, volume 65 of *Proceedings of Symposia in Pure Mathematics*. American Mathematical Society, Providence, RI.

**2000**

- [101] (with C. Landim and S. Olla). Asymptotic behavior of a tagged particle in simple exclusion processes. *Bol. Soc. Brasil. Mat. (N.S.)*, 31(3):241–275.
- [102] Scaling limits of large interacting systems. In *The mathematical sciences after the year 2000 (Beirut, 1999)*, pages 144–158. World Sci. Publ., River Edge, NJ.
- [103] Lectures on hydrodynamic scaling. In *Hydrodynamic limits and related topics (Toronto, ON, 1998)*, volume 27 of *Fields Inst. Commun.*, pages 3–40. Amer. Math. Soc., Providence, RI.
- [104] S. Feng, A.T. Lawniczak, and S.R.S. Varadhan, editors. *Hydrodynamic limits and related topics*, volume 27 of *Fields Institute Communications*. American Mathematical Society, Providence, RI.
- [105] Infinite particle systems and their scaling limits. In *Mathematical physics 2000*, pages 306–315. Imp. Coll. Press, London.
- [106] (with S. Sethuraman and H.-T. Yau). Diffusive limit of a tagged particle in asymmetric simple exclusion processes. *Comm. Pure Appl. Math.*, 53(8):972–1006.

**2001**

- [107] (with C. Landim and S. Olla). Symmetric simple exclusion process: regularity of the self-diffusion coefficient. *Comm. Math. Phys.*, 224(1):307–321.
- [108] Large deviation and hydrodynamic scaling. In *Taniguchi Conference on Mathematics Nara '98*, volume 31 of *Adv. Stud. Pure Math.*, pages 265–286. Math. Soc. Japan, Tokyo.
- [109] Diffusion processes. In *Stochastic processes: theory and methods*, volume 19 of *Handbook of Statist.*, pages 853–872. North-Holland, Amsterdam.
- [110] *Probability theory*, volume 7 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York University.

**2002**

- [111] Rare events, large deviations. In *Mathematical finance—Bachelier Congress, 2000 (Paris)*, Springer Finance, pages 85–92. Springer, Berlin.
- [112] (with C. Landim and S. Olla). Finite-dimensional approximation of the self-diffusion coefficient for the exclusion process. *Ann. Probab.*, 30(2):483–508.

**2003**

- [113] Particle systems and partial differential equations. In *Lectures on partial differential equations*, volume 2 of *New Stud. Adv. Math.*, pages 217–222. Int. Press, Somerville, MA.
- [114] Large deviations and entropy. In *Entropy*, Princeton Ser. Appl. Math., pages 199–214. Princeton Univ. Press, Princeton, NJ.
- [115] Large deviations for random walks in a random environment. *Comm. Pure Appl. Math.*, 56(8):1222–1245.
- [116] Stochastic analysis and applications. *Bull. Amer. Math. Soc. (N.S.)*, 40(1):89–97 (electronic).

**2004**

- [117] (with E. Dynkin., A.D. Venttsel', et al.). Mark Iosifovich Freĭdlin. *Russian Math. Surveys*, 59(3):593–597. Also available in *Uspekhi Mat. Nauk*, 59(3(357)):185–188 (in Russian).
- [118] Random walks in a random environment. *Proc. Indian Acad. Sci. Math. Sci.*, 114(4):309–318.
- [119] (with C. Landim and S. Olla). On viscosity and fluctuation-dissipation in exclusion processes. *J. Statist. Phys.*, 115(1–2):323–363.
- [120] (with C. Landim and S. Olla). Diffusive behaviour of the equilibrium fluctuations in the asymmetric exclusion processes. In *Stochastic analysis on large scale interacting systems*, volume 39 of *Adv. Stud. Pure Math.*, pages 307–324. Math. Soc. Japan, Tokyo.
- [121] Large deviations for the asymmetric simple exclusion process. In *Stochastic analysis on large scale interacting systems*, volume 39 of *Adv. Stud. Pure Math.*, pages 1–27. Math. Soc. Japan, Tokyo.

**2005**

- [122] (with S. Sethuraman). A martingale proof of Dobrushin's theorem for non-homogeneous Markov chains. *Electron. J. Probab.*, 10(36):1221–1235 (electronic).

**2006**

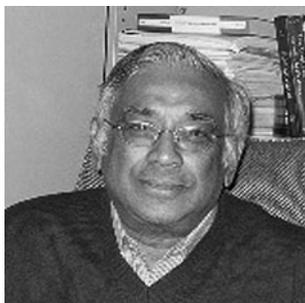
- [123] (with E. Kosygina and F. Rezakhanlou). Stochastic homogenization of Hamilton–Jacobi–Bellman equations. *Comm. Pure Appl. Math.*, 59(10):1489–1521.

**2007**

- [124] *Stochastic processes*, volume 16 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York University.
- [125] Homogenization. *Math. Student* 76:129–136.

**2008**

- [126] Homogenization of random Hamilton–Jacobi–Bellman equations. In *Probability, geometry and integrable systems*, volume 55 of *Math. Sci. Res. Inst. Publ.*, pages 397–403. Cambridge Univ. Press, Cambridge, 2008.
- [127] (with N. Zygouras). Behavior of the solution of a random semilinear heat equation. *Comm. Pure Appl. Math.*, 61(9):1298–1329, 2008.
- [128] (with E. Kosygina). Homogenization of Hamilton–Jacobi–Bellman equations with respect to time-space shifts in a stationary ergodic medium. *Comm. Pure Appl. Math.*, 61(6):816–847, 2008.
- [129] Large deviations. *Ann. Probab.*, 36(2):397–419, 2008.
- [130] Large deviations and scaling limit. *Lett. Math. Phys.* 88:175–185.



## Curriculum Vitae for Sathamangalam Rangaiyenga Srinivasa Varadhan

- Born:** January 2, 1940, Madras (Chennai), India
- Degrees/education:** BSc Madras University, India, 1959  
MA Madras University, India, 1969  
PhD Indian Statistical Institute, Calcutta, India, 1963
- Positions:** Postdoctoral Fellow, New York University, 1963–66  
Assistant Professor, New York University, 1966–68  
Associate Professor, New York University, 1968–72  
Professor, New York University, 1972–  
Frank J. Gould Professor of Science, New York University, 1999–
- Visiting positions:** Guest professor, Mittag-Leffler Institute, 1972  
Guest professor, Stanford, 1976–77  
Visiting member, Institute for Advanced Study, Princeton, 1991–92
- Memberships:** American Academy of Arts and Sciences, USA, 1988  
Third World Academy of Sciences, 1988  
Institute of Mathematical Statistics, 1991  
National Academy of Sciences, USA, 1995  
Royal Society, UK, 1998  
Indian Academy of Sciences, 2004  
Norwegian Academy of Science and Letters, 2009
- Awards and prizes:** Alfred P. Sloan Fellowship, 1970–72  
Guggenheim Fellowship, 1984–85  
Birkhoff Prize, 1994  
Margaret and Herman Sokol Prize, New York University, 1995  
Steele Prize, American Mathematical Society, 1996  
Abel Prize, 2007

**Honorary degrees:** Université Pierre et Marie Curie, Paris, 2003  
Indian Statistical Institute, Kolkata, India, 2004

**Presidencies:** Director, Courant Institute, 1980–84, 1992–94  
Institute of Mathematical Statistics, 2003