volume 238

A Dekker Series of Lecture Notes in Pure and Applied Mathematics

Stochastic Processes and Functional Analysis

A Volume of Recent Advances in Honor of M. M. Rao

edited by
Alan C. Krinik
Randall J. Swift

Stochastic Processes and Functional Analysis

A Volume of Recent Advances in Honor of M. M. Rao

edited by

Alan C. Krinik

Randall J. Swift

California State Polytechnic University

Pomona, California, U.S.A.

MARCEL DEKKER, INC. NEW YORK • BASEL

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress.

ISBN: 0-8247-5404-2

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016, U.S.A.
tel: 212-696-9000; fax: 212-685-4540

Distribution and Customer Service
Marcel Dekker, Inc.
Cimarron Road, Monticello, New York 12701, U.S.A.
tel: 800-228-1160; fax: 845-796-1772

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-260-6300; fax: 41-61-260-6333

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Current printing (last digit):

10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

For M.M. Rao

Professor M.M. Rao has had a long and distinguished research career. His research spans the areas of probability, statistics, stochastic processes, Banach space theory, measure theory and differential equations - both deterministic and stochastic. The prolific published research of M.M. Rao impacts each of these broad areas of mathematics

During M.M.'s career, he has had eighteen Ph.D. students. Many of his students have gone on to very successful careers in mathematics and are recognized experts in their field of study. Six of his former students have written tribute essays about M.M. and each are affectionately dedicated to him. These essays were written by J.A. Goldstein, M.L. Green, N.E. Gretsky, A.C. Krinik, R.J. Swift and J.J. Uhl.

Jerry Goldstein is a Professor of Mathematics at the University of Memphis. He is a internationally known for his outstanding work in semigroup theory, functional analysis and differential equations.

Mike Green is an Assistant Professor of Mathematics at California State Polytechnic University, Pomona. His research is in the area of multi-parameter manifold valued semi-martingales and aspects of applied probability.

Neil Gretsky is an Associate Professor of Mathematics at the University of California, Riverside. He is recognized for his research in the geometry of Banach spaces and recently his work in game-theoretic applications in economics.

Alan Krinik is a Professor of Mathematics at California State Polytechnic University, Pomona. He is noted for his work in lattice path combinatorics and its application to queueing theory and birth-death processes. He is also the co-editor of this volume.

Randy Swift is an Associate Professor of Mathematics at California State Polytechnic University, Pomona. He is well-known for his work in harmonizable processes, mathematical modeling and differential equations. He is also the co-editor of this volume.

Jerry Uhl is a Professor of Mathematics at the University of Illinois, Urbana-Champaign. He is known for his work in vector measures and Banach space theory. He is also noted for his work in calculus reform.

An Appreciation of my teacher, M.M. Rao

I want to record my thoughts about M. M. Rao as a teacher. He was a really great teacher and his teaching continues to have a major impact on my career.

As a first year graduate student at Carnegie Tech, in 1963-64, 1 took Rao's year long course on Functional Analysis. There were a lot of good students around Tech at that time; included in Rao's class were second year students Neil Gretsky and Jerry Uhl. Rao's ambitious style was to cover one major result in each lecture, or three per week. And all major theorems had descriptive names, some standard ("Dominated Convergence Theorem") and some not ("Law of the Unconscious Statistician"). The use of those names made the results easier to remember; I think Rao got this idea from Michel Loeve's book (from which I learned probability theory). Our text was by Angus E. Taylor, but we didn't use it much. Rao taught mostly out of Dunford & Schwartz (Vol. 1) and Hille & Phillips. His organization of the topics was excellent. An unusually large amount of material was covered per class. So much so that details were often omitted (or, sometimes in our minds, incorrectly given). With great regularity Gretsky, Uhl and I would stay after class and work out the complete details of the arguments we had just seen. Sometimes we realized that Rao really had given all the details; after all we were merely beginners and not yet well versed in mathematics. We always found that all of his results had correct versions, occasionally slightly different from what one of us thought when the discussion began. But by the end of the year, I learned so much that, for the first time, I considered myself a mathematician. Gretsky, Uhl and I were somehow teaching assistants to Rao, helping to teach one another. At the time I didn't give Rao credit for orchestrating this, but I think he did, at least to a substantial extent. He conveyed his love of mathematical depth and understanding and his passion for intense mathematical discussions.

I took many more grad courses from Rao prior to graduating in May 1967. They were all great courses, but none matched that extraordinary course in Functional Analysis. That course had a permanent influence on me, and for the rest of my life I will feel close and grateful to M.M. and to Neil and Jerry as well.

Having gotten my BS at Tech in 1963 and anticipating my MS in 1964, I decided to apply elsewhere (in the fall of 1963) to do my Ph.D. away from my birthplace, Pittsburgh. But my wife had a teaching job in Pittsburgh and her applications elsewhere were unsuccessful. So, despite fellowship offers from more prestigious institutions, I was happy to stay at Tech because I knew (from the Functional Analysis course) that a great thesis advisor was available. Gretsky and Uhl were already doing research reading under Rao, and in the spring of 1964 I told Rao I wished to work with him in Functional Analysis (as Neil and Jerry were doing). He said he would be glad to be my advisor, but he had a problem in probability theory for me. I protested, saying I didn't know any probability theory. He pointed out that I had taken a year long junior level probability course from Morrie DeGroot, an excellent teacher and probabilist. (Of course, he was right, but I was mystified, being so in love with Functional Analysis). But, as my main focus was to work under M.M., I said OK. The first paper he gave me to read was by Dynkin, and it defined a Markov process as a 21 tuple (or something like that). Numbers larger than my combined fingers and toes made me nervous; and I was unable to read Dynkin's paper, Rao suggested I try Loeve's book and work a lot of the problems. This was a great suggestion, and Rao helped me a lot when I got stuck. And, happily, 6 or 8 months later I was able to read and understand Dynkin's paper (which was indeed a toughie).

M.M. ran great seminars and, among other things forced his students to present papers they read and their own work. His ferocious but kind questioning taught us never to give a seminar less than fully prepared. And he taught us to work together and learn from one another. This is a very important point which was evident, but I didn't realize it as the time. Rao's teaching and advising styles were shamelessly adopted by me in my capacity as a teacher and advisor. I have had over 20 Ph.D. students ("children") and at least 8 "grandchildren", most of whom never met Rao and probably are largely unaware of the major hidden role he played in their education.

I love to reread occasionally the article Rao wrote in the Raofest volume, celebrating his 65th birthday. Rao did something special and unusual; he gave his best research problems to his students. I have tried to follow his lead, and I believe our profession would be better of if more thesis advisors did the same thing.

Rao was uncompromising in his high standards, but he was gentle and helpful. Not all of his Ph.D. students had the native brilliance of Gretsky or Uhl, but all of them (that I know) wrote excellent theses. Rao got his students to live up to their potential. I think that is the highest praise one can give to a teacher.

Rao was also an excellent researcher. As a departmental citizen he was pretty feisty. He objected to (mathematical) political issues taking precedence over issues of quality and scholarship. Doing the "right thing" is not always the way to maximize one's popularity. But M.M. never hesitated to stand up and fight for his principles.

I suppose I should tell one anecdote. The enormous length of Rao's first two names reminds me of Dynkin's definition of a Markov process (which took me 6 or 8 months to understand). So once I asked Dick Moore what Rao's nickname was. Dick's response: "He doesn't have one. People call him M.M. But he should have one."- Dick thought about it and hit on the nickname Mmmmmmmmm. But it never stuck.

I feel much affection for M.M. I always respected and admired him, and there were moments during my grad student days, when the term "affection" did not characterize my feelings toward him. But I was young, brash and impatient; some things I could figure out very quickly and some not. Was I lucky to have M.M. as my principal teacher and mentor? Absolutely yes. Could I have done better either at Tech or elsewhere? I don't think so. Rao shaped my passion for mathematics, my desires to understand things fully, my standards, and my teaching and advising techniques. I owe him so much, more that I can usually imagine. Thank you M.M., for being such an inspiration and such a friend.

J.A. Goldstein

1001 words about Rao

My first contact with M.M. Rao was in the fall quarter of 1989 in an Advanced Calculus class. Before this course, like a typical undergraduate, I inquired of other students about him. Most of my information came from the graduate students at UCR, since they were the ones who had taken courses from him. The graduate students generally regarded him as a hard, but fair teacher. This positive tone, however, was laced with an undertone, not unlike the sort one would receive about a blind date, who in all other respects was perfect, except for some peculiar habit. It required only one lecture to discover the peculiarity of M.M. Rao. He is so absorbed into mathematics that where the man ended and the math began was blurred until the separation of the two is unimaginable.

His lectures are wonderful. The students of Rao have coined the phrase "Rao Math" for the rather distinctive style he has when presenting mathematics. He carefully prepares all his lectures, often writing them out in their entirety before the beginning of the quarter. An appropriate motto he has given is "If we do this for the general case, the rest will follow as corollaries." One need only read one of his books to see the verity of this motto. A good example is his text for Real Analysis, *Measure Theory and Integration.* He often immediately began lecture upon entering the room and always went over the allotted time leaving the next class waiting at the door. On more than one occasion, he was writing as he walked out the door! These peculiarities are symptoms of his strength, a single-minded dedication to his profession coupled with a deep interest and curiosity in the subject. In M.M. Rao, I met someone that hit the 35th level of Math™, a true math guru. To be fair, not everyone prospered under Rao. The lack of concrete examples was the typical student complaint about M.M. Rao's instruction. I guess M.M. Rao had been getting some grief about not being concrete enough, for during one class he declared, "I am an applied mathematician! I apply this theorem to prove that one!" This is a typical Raoism.

The beauty of mathematics as presented by him seduced me. I know that I am not the only one to experience this and like others I started taking more courses from Rao after my first introduction to him in advanced calculus. I began learning more analysis and in particular probability as a consequence. A tremendous benefit to my education was the open door I always found at his office and the many conversations I had with Rao about mathematics and his research have enriched my life. As a student, I never felt belittled or talked down to by Rao even when discussing his research. During my seeking of an advisor for my Ph.D. thesis, some of the other professors cautioned me about Rao, concerning his ability to win the best students. A reputation well deserved. I mulled over several individuals, all very capable mathematicians, but the accessibility of M.M. Rao won me over, even though my first interests were in algebra and topology.

His work ethic was intimidating. Sleeping four or fewer hours per night working on mathematics most of the day, seven days a week he labored with "a devil on his back" to complete his projects. He only took a half-day off on Sunday. He once said, "Mathematics is a harsh mistress. Either you love her, or she will leave you."

My thesis topic was to extend stochastic integration to multi-parameter manifold valued semi-martingales using the generalized Bochner boundedness principle. The students of Rao have termed his theses topics as "topics in the clouds". A few of the completed theses are "Orlicz spaces of additive set functions and set martingales", "Integral representations of chains and vector measures" and "Multi-parameter semi-martingales integrals and boundedness principles." The last being my own, coming up short on the manifold part. The theses that Rao has guided tend to be on the long side, my own was 138 pages, not the longest.

This propensity to generalize has worked well for M.M. Rao. Take for instance what Rao has done with ideas from S. Bochner. In 1956, Bochner wrote "Stationarity, boundedness, almost periodicity of random valued functions" in the Proceedings of the Third Berkeley Symposium. In this paper Bochner defined V-bounded processes and noted that these processes were an extension Loeve's harmonizable processes. Rao's idea was to define two classes of processes, the V-bounded being called weakly harmonizable processes which includes the processes of Loeve, now called strongly harmonizable. This definition opened up a whole new area of research in harmonizability being still actively pursued. Another idea Rao gleaned from this paper is to define stochastic integration via a boundedness principle. His generalized Bochner boundedness principle provides a unified approach to stochastic integration including all known stochastic integrals under one umbrella. This principle would still be unknown if M.M. Rao had not pursued mathematics in his own distinctive manner. For the Young functions from Orlicz space theory were necessary for the result. Rao met with Bochner three times. Bochner must have been impressed, since he communicated three of Rao's papers to the National Academy of Sciences. Rao still has not been entered as a member of the Academy.

Rao still lies dear in my heart as he does in the hearts of many others who have come across his path. M.M. Rao asked me to not compare him as an equal to Bochner, his modesty is showing, but in my eyes, he is a great mathematician and a great man. He still shows his faith in me and has many expectations for my work, encouraging me to continue my labors. My wish is that he finds a satisfaction in his life and work that brings him peace. I look forward to the years that come to see what new worlds he will open in mathematics.

M.L. Green

A Guide to Life, Mathematical and Otherwise

When I went to graduate school in the early '60s I started in the Systems & Communications Sciences interdisciplinary program at Carnegie Tech. I knew that I wanted to study and work in Numerical Analysis and Computing. In my second year I decided to take a Functional Analysis course because I had some half-formed idea that this would be a valuable tool for Numerical Analysis. I did not in any way anticipate the ensuing life-changing event of meeting M.M. Rao. The course became an almost-religious epiphany for me: this was truly the way, the truth, and the light! M.M.'s lectures were magnificent; the material was spell-binding; the problem sets were *really* challenging. Several of my fellow students felt the same way – especially Jerry Uhl and Jerry Goldstein. We took more courses and seminars with M.M. and we chose to write our theses under his direction. The three of us spent a lot of time challenging each other and guiding each other under M.M.'s firm but insightful hand. In the last vestiges of the medieval guild system, we apprenticed ourselves to a guild master – a true master.

There was certainly a deep love for Mathematics and a lifelong friendship and bond that we developed together under M.M.'s direction, but there was much more to M.M.'s influence. M.M. had a deep concern for, and loyalty to, his students. No matter how busy he was, he always had time and energy for us in all aspects of our development. When my wife left me in the final year of my thesis work, M.M. was there to counsel and comfort me. Unbeknownst to me at the time, he had also spoken with my wife to see if there were any possible solution. When I mistakenly thought that one of my thesis results was contained in an earlier paper, he brought me out of my depression and led me to see the positive differences in the work. When I succumbed to procrastination and other earthly temptations, he was there to inspire me with his example. He was never accusatory or condemning, just exemplary and inspiring. When a new department chair took a personal dislike to me, M.M. was there to defend his student. Of course, this was the same M.M. who liked to put an occasional (unannounced) unsolved problem on his take-home exams in the advanced topics courses.

When I received a job offer from the Mathematics Department at UCR, he told me that it was a great opportunity because Howard Tucker was there and he repeatedly advised me that " ... you will really like Professor Tucker ". This stuck in my mind so deeply that when I finally met Howard and he told me to call him Howard, my natural response was "Yes sir, Professor Tucker".

A few years later, my new department wanted to recruit a senior person in Functional Analysis and Probability. M.M. was not in the market for a new job, but I knew that he was not happy with his department chair. Our department managed to interview him twice and convinced him to come. So we wound up in the same department for thirty years. Once again he me led to learn life's great lessons. At first I needed to assert my independence from him. That must have been painful for him, but he never showed it. Then I needed to again succumb to procrastination and earthly temptations. Once again, he was over the space of many years non-accusatory and supportive. Coming back into the fold, I started to drift into areas of applications of Functional Analysis and Measure Theory. He renamed our continuing Functional Analysis Seminar the Functional Analysis and Related Topics Seminar.

It has been a very large feature of my life as well as a remarkable pleasure and privilege to be his student, his colleague, and his friend.

N.E. Gretsky

Rao and the early Riverside years

M.M. Rao first came to the University of California, Riverside Mathematics Department from Carnegie-Melon in 1973. There was much excitement and anticipation of his coming by both new colleagues and graduate students in the UCR Math Department. Neil Gretsky (a former Rao graduate student from Carnegie-Melon) was already on the faculty at UCR. Neil and others had alerted UCR graduate students of M.M. Rao's prominence in probability and functional analysis. M.M. Rao was a welcome addition of a talented research mathematician who was receptive to graduate students. This enhanced an already formidable mathematics department that had F. Burton Jones in topology, Richard Block in algebra and Victor Shapiro in differential equations among other notable faculty members.

As a new graduate student at UCR (coming from UCLA) in 1973, I knew very little of the anticipation surrounding M.M. Rao's first academic year at UCR. However, I became quickly familiar with Rao's teaching style as I took his inaugural graduate sequence in real analysis at UCR: Math 209, 210, 211 starting in September 1973. The course was taught at a high level of abstraction. The first quarter was measure theory developed on general sigma rings using an outer measure approach restricted down to measurable sets via the Caratheodory construction. The second quarter contained the major results of general integration theory. The third quarter included an introduction to Choquet's capacity theory. There was no specific textbook for the course but rather a list of several recommended texts. The course was carefully and clearly presented by M.M. Rao, a man in his early forties (originally from India) with a lively personality, who wore a suit to class. I tried to take careful notes and absorb the material since I knew a comprehensive qualifying exam on real analysis based on this course was waiting for me at the end of the academic year. However, the material was not easy for me. I passed the qualifying exam but considered myself fortunate. As for this introduction to Rao, I found him an animated professor completely engaged in the subject of real analysis. He developed the theory from a modern abstract viewpoint but was concerned about the history of the subject and was careful to credit various mathematicians as we proceeded (Lebesgue, Caratheodory, Vitali, Saks, Fubini, Egoroff, Choquet, etc.).

Several very talented UCR graduate students, including Stephen Noltie and Michael Brennan, were seemingly planning to work with Rao even before he arrived at UCR. By the time, I asked Professor Rao to be my advisor in 1975-76, I was his sixth Ph.D. student at UCR. I was grateful he agreed to take me as his student. From the beginning, Rao had the reputation of being more demanding than most other professors at UCR. Rao would oversee your progress but he would not help you in the writing of your thesis. Rao expected his students to be prepared in many different areas of mathematics. As a consequence, Rao students routinely took additional coursework past the qualifiers. For example, I took graduate sequences in functional analysis, advanced statistical inference and probability theory after my qualifying exams. The idea was to be prepared to solve our dissertation problem from a variety of different possible perspectives.

Another important aspect of being a Rao graduate student in the 1970's was an ongoing quarterly seminar on functional analysis or stochastic processes. This seminar (which still meets) consisted of Rao, his students and any other interested parties. Everyone attending talked sooner or later. When the discussion became very specialized, the seminar often reduced down to Rao and his graduate students. For me, I recall having to present material that originated from a seminal 1969 article written by D.W. Stroock and S.R.S. Varadhan on solutions of diffusion processes in d-dimensions using the martingale problem approach. I vividly recall preparing this challenging material for what seemed like an endless number of consecutive weeks. It was stressful but very helpful in forcing me to understand this paper which I eventually generalized into my dissertation. Understanding came slowly (and in phases). I learned how to present material when there were holes or unresolved problem areas and how to talk around topics until I was able to make complete arguments. The whole experience also brought the Rao students together in a common misery and made me appreciative of the mathematical abilities of my fellow grad student, Michael Brennan, who kindly helped me understand the more incomprehensible parts of this paper. This seminar experience was a common learning experience for Rao students in the 1970's. It is an activity that I still do today on a modified basis with my own graduate students.

At UCR, M.M. Rao was primarily known among graduate students as a consummate researcher in mathematics-a man whose research interests connected functional analysis and integration theory with probability theory and stochastic processes. He was also recognized as an engaging professor who attracted some of the stronger graduate students to work with him and take a wide range of graduate classes. From a work ethic point of view role, no one worked harder than Rao. In the 70's, Rao occupied the (eastward) end office of a string of about twelve windowed offices on the third floor of Sproul

Hall that faced south overlooking University drive. Any passer-by, looking up at these offices in the evening would customarily see only two or three lights on after dark. Sometimes if the hour was late only one light burned. M.M. Rao's window was almost always lit. He was up there doing research, reviewing articles and in the 70's writing his first book. His colleagues and graduate students knew he was there. They also knew that he would be back in his office at least one day over the weekend as well. Rao displayed a commitment to his profession that was hard to match. From a graduate student's point of view, no one could complain that Rao was inaccessible or difficult to find.

M.M. Rao of the 1970's was a confident, forceful and demanding advisor. As an outstanding mathematician, Rao had expectations or intuition of how the solution of mathematical problems should turn out. Whenever graduate student progress did not fit his long range view, he expected to be consulted or convinced as to why these mathematical objectives were not possible. He also expected graduate students to make a dedicated effort and work hard. Finally, Rao expected his stronger graduate students to make significant contributions by doing future mathematical research. After all, Rao himself lived according to these standards. These expectations sometimes caused tensions between Rao and his students. For example, graduate students, myself included, would at times "disappear" for weeks or even months. When this happened there could be many possible reasonable explanations (and some unreasonable explanations as well)-including outside life factors effecting the unreal graduate school existence. Sometimes, a graduate student just rather "lay low" while trying to achieve progress rather than share their "failed attempts" at solving a problem. I can remember Rao asking "Where is ___?-I haven't seen him in weeks!" These incidents had both good and bad consequences. Rao students developed an independence and self reliance in doing mathematical research and also provided more opportunities for Rao students to bond together. Rao stories, like war stories, were swapped over lunch or over a few beers. Sometimes even an old Rao story from the legendary days of Gretsky, Uhl and Goldstein would be recycled when pertinent. In the end, Rao's forceful personality and expectations played differently among his graduate students (some of whom also had strong personalities and different goals).

M.M. Rao views his graduate students as one big family. Certainly, there are many of his former students who have flourished in mathematical careers engaged in many of the same aspects of the profession that have occupied Rao. There are also former, highly capable, graduate students who presently have little interest in mathematical research and have chosen exciting alternative career paths. M.M. Rao is interested and always enjoys hearing (and talking at length about) how each of his graduate students is doing. However, make no mistake about it, Rao is a true believer. M.M. Rao's career in mathematics is distinguished by his talent, passion and energy in doing mathematics. There has never been any doubt in his mind that (if one has the ability) being a research mathematician is the best way to go. I think that even today, Rao would not understand how a graduate student in mathematics with outstanding potential in research could choose to do anything else. It is also very difficult to imagine M.M. not being engaged in mathematical activities. Rao is a lifer. Currently, at age 74, he is going strong. Rao is busy writing books with plans for additional books in the future.

M.M. Rao did a wonderful job of protecting and promoting his graduate students. He was influential and resourceful in securing teaching assistantships and research assistantships to support his graduate students throughout graduate school. During the 1970's, Rao was preparing his first book, Stochastic Processes and Integration. I, along with several other grad students, worked as a research assistant, proof-reading this monograph. Professor Rao was very receptive to student reaction to his writing. At first, I was hesitant to mention where I had difficulty in understanding his text but it became very clear that he was sincerely interested in both my mathematical and stylistic comments. Rao, in discussion, would often tell me of the history of various portions of the text and what different mathematicians had contributed. These were good times for me. I was seeing mathematics from an insider's perspective. Sometimes, Rao would go off describing some current mathematicians. For example, he knew I was studying a paper of the Russian mathematician Daletskii and Rao would tell me of his personal meeting with Daletskii on a visit to Russia and how nice a man he was even presenting M.M. a bottle of Vodka (or Cognac) as a gift. Rao still had the gift somewhere in his office. These exchanges were memorable.

The academic environment and spirit for faculty and graduate students in the UCR Mathematics Department of the 1970's was very good. The Department was a friendly place and a good place to study mathematics. Al Stralka chaired the Department. I recall colloquia given by Erdos, Halmos, Bing, Stein and Uhl. I recall the excitement of the four color problem being solved at that time. There was an entertaining talk on properties of Fibonocci numbers as well. The colloquia were preceded by a reception that usually included cookies-a sure way to attract graduate students. For at least two years, the math graduate students participated in intramural basketball games. Our team names of "Zorn's Dilemma" and "The Hardy-Haar Measure" accurately reflected our team's abilities. We had measure 1 of going the whole season without a victory. Except, there was one anomalous game where we actually nipped the lowly and equally winless Physics team on a last second miracle shot-which demonstrated once and for all that events of measure 0 can indeed happen! We had fun with basketball but actually looked forward to the pizza and beer get-togethers after the game more than the game itself.

During the mid 1970's, Rao students were united by the knowledge that we were committed to a challenging route working under M.M. Rao and hopeful of his influence to secure us academic employment at a notoriously difficult time period for new Ph.D.'s to find jobs as professors in mathematics. We were also united by having taken an unusually large number of courses from Rao. The following pet phrases (and situations) were often repeated (or experienced) many times in class and today serve to help us recapture, with affection, his unique personality and style:

"We make the following definition with complete 'malice of forethought'."

"Did you think it was going to be easy? No! That is why his name is on the theorem."

"Be careful whenever you see that word 'consider' for what follows is a new idea..."

"From there he went on to develop (pronounced 'devil-up') the whole theory..."

"You ask me if I can change the order of integration? I DID IT!"

"That's the one, that's the condition you need..."

"You work and work and work and that is what comes out..."

"Now we have proved the Dynkin-Doob Lemma which is also used by statisticians who have no idea why it is true, so we call this result the Theorem of the 'Unconscious Statistician'..."

"If you wish to avoid making any mistakes, do nothing at all and that, of course, would be the biggest mistake of all..."

"What a loss...that is the death of his career as research mathematician", (Rao's reaction upon hearing a local mathematics professor had become dean.)

"You can take this book and throw it in the ditch..."

Many times Rao would smile and laugh as he repeated these sayings in different settings. Occasionally, M.M. would re-tell a joke or funny story and break out laughing aloud before reaching his own punch-line.

And finally the signature (literally and figuratively) of most Rao chalk talks was the amazing amount of mathematical material he was able to cram into the lower right hand corner of the board as class time expired. His writing became a space filling curve as he adjusted by writing smaller and smaller-working several minutes past when the class was scheduled to stop-leaving students dazed and hopelessly trying to decipher his final scribbling.

Rao could push and posture. During my last year in graduate school Rao had monitored all my dissertation work. I had passed my oral exams and was in the process of writing up my final results. We were 3 months away from being done. He looked at my folder of dissertation results and then back at me and announced, "It's not enough". I felt my heart sink and had nothing to say. I went home wondering what more I could do. There was no more but he was still seeing if he could squeeze some new results out of me. I did not like the pressure but I understood his intent. It soon became clear to him (if it wasn't already) that there was nothing else to do on my problem. It never came up again and I finished my Ph.D. as originally scheduled three months later.

In 1978, Rao still had three of his six graduate students (Brennan, Kelsh and Krinik) anxious to get out. Rao was leaving on a sabbatical (I believe to France) starting Fall 1978 and the realization dawned on us that Rao's sabbatical was our best chance of finally finishing up. Otherwise, we would have to postpone our graduation until Rao's return to UCR a year later and, of course, no one wanted to wait. In a furious finish, we all made it. I was the last of the three to finish and remember happily driving M.M. to the airport.

After graduation, my relationship and appreciation of M.M. Rao grew and matured. As a graduate student, I was always appreciative of his financial support for all his students and his academic support for me in particular. After graduation, Rao became a major player in my career. He was always in my corner, helping me. From key letters of recommendation to help me secure positions at JPL, University of Nevada, Reno and Cal Poly Pomona, to advising me where to try to publish my results, to being supportive when my efforts were not always successful, to providing me with opportunities to resume research activities and to finally just being there as a good friend. His encouragement and assistance in developing my professional activities has been and remains a constant.

In 1985, I invited M.M. to give a colloquium at Cal Poly Pomona. M. M. did his usual super job and in the audience sat a talented graduating senior who would not forget the talk nor the speaker. That senior was Randy Swift who eventually went on to earn his Ph.D. under M.M. Rao and who today is a good friend and valuable colleague at Cal Poly Pomona. Randy is also the real editor of this volume which we both affectionately dedicate to M.M., our mutual mentor. In tribute to M.M. Rao's stellar career, Randy compiled this volume of research articles.

The eighteen Rao students share a special bond and understanding of what it means to earn your doctorate degree in mathematics under the direction of M.M. This collection of essays and articles in honor of M.M. illustrates this bond crosses four decades and bridges his Carnegie-Mellon University students of the 60's with the University of California, Riverside students of the 70's , 80's and 90's. It's been a pleasure to have the opportunity to celebrate M.M. Rao's many contributions and to be "one of Rao's students".

Alan Krinik

On M.M. Rao

I first met Professor M.M. Rao in 1985, when I was an undergraduate attending California State Polytechnic University, Pomona. Alan Krinik had invited M.M. to give a colloquium talk in the department.

At the time, I was a senior math major and one of the department's promising students. I had attended departmental colloquium talks, but never had I been exposed to a mathematician of the caliber of M.M. Rao.

His talk began in a very elementary fashion, but the breadth and depth of the mathematics it spanned greatly impressed me. I was struck by the passion for mathematics that he displayed. I had not been in the presence of someone totally devoted to his discipline.

After I completed my Masters degree, I worked for a while in the aerospace industry. I found myself desiring to pursue a PhD. My interest in probability theory and my strong recollection of M.M. led me to apply to the University of California, Riverside.

As fate would have it, and in my good fortune, I took M.M. for a graduate course in Probability, his lectures were absolutely beautiful. Spanning the subject with depth and presented with crystal clarity. Of course, he used his text *Probability Theory with Applications*, perhaps the finest graduate text written on the subject.

This course, and indeed, this text, set the tone for what working with M.M. would be like. M.M. believes that homework should challenge his students. During his courses, he assigns a problem or two per week. These problems are not routine homework problems, rather they are problems from the research literature. They are not mere exercises. Indeed, his students spend vast amounts of time working on them. To this end, he is preparing his students for research. Many of these problems aid his students in their future works.

M. M. greatly respects effort. If he sees that a student is working, he will guide the student gently down the appropriate path. He has an incredible memory for details, often if a student was stuck on an idea, he would say, go see this page of a particular paper or text. On that page, you would find what you needed to get going again on the problem. From these interactions, M.M. seems to gauge a student's ability. I became a student of his after I had completed the course in probability theory, a seminar on random fields and a course in stochastic processes. I never asked him to be my advisor; rather, it seemed to be a natural evolution.

The first problem that he asked me to study involved the sample path behavior of harmonizable processes. I spent a large part of that first summer developing my facility with these processes. By the end of the summer, I obtained my first minor result; it was on the analyticity of the sample paths. However, the goal was to consider the almost periodic behavior of these sample paths and I was stuck.

I toiled in vain for the next few months on this problem. One day, late in November, I went to talk with M.M. about the problem, he listened intently and then said, "let it rest there, for now, I would like you to look at this calculation I have been working upon with harmonizable isotropic random fields."

He had obtained a representation for the covariance that involved some rather complicated special functions. He asked me to see what I could do with it, in particular, could it be made to look similar to the representation obtained by Bochner for the stationary isotropic case.

I told him that I would try, and he said "There is no try, there is only do, and I *know* that you can."

By Sunday afternoon of that weekend, I had simplified the expression and had showed that it reduced to Bochner's representation in a very natural way.

That Monday, I gave him the result. He, in a very delighted manner, then said to me, "See if you can push it. Look at Yadrenko's book and use this representation to extend his results."

This began a glorious 3-month stretch of research production. I obtained several major results for harmonizable isotropic random fields.

Riding the tide of this success, he said to me "And what of the almost periodicity?"

With the confidence I had obtained, I went back to the problem. Within a month or so, I had obtained the results I had long sought. This experience gave me great confidence in my ability to do research. It also gave me a very broad research program to pursue. The confidence that M.M. showed in my abilities as a mathematician remains with me today. It has allowed me to flourish.

Many years later, M.M. told me that the first string of results I had obtained after I had obtained the representation was likely enough for the PhD. However, he saw that I was on track and he was going to have me do as much as I could.

This story is very typical of the relationships he has with his students. He works them very hard, always encouraging, and yet unyielding in his determination that they do their very best work.

In this setting, many of his students have become mathematicians completely devoted to the discipline. Whether this devotion is shown through excellence in teaching, or excellence in research, or both, for each of us, it is likely attributable to the role M.M. played, and continues to play, in our lives.

It is no wonder then, that for some of us, M.M. holds a place in our hearts, and we remain devoted to him as former students and now colleagues.

R.J. Swift

Reflections on M.M. Rao

Jerry Goldstein, Neil Gretsky, Bill Kraynek and I all fell under the charms and passions of M.M. at Carnegie Tech at the same time during the mid-1960's. It was a glorious time for us and it was a glorious time for math at Carnegie. We had close access to mathematicians such as Dick Duffin, Nehari (who claimed not to have a first name), Alan Perlis, Dick MacCamy, Roger Pederson, Dick Moore, Vic Mizel, Charlie Coffman and of course M.M.. Not only did we study under these fellows but we also socialized with them too. Some drank beer with us, some drank scotch with us and they all drank coffee with us. They welcomed us into a flourishing community of math research. They made us feel like mathematicians. We all owe so much to them.

With all these mathematicians (most of whom were in their prime) available to us, why did Goldstein, Gretsky and I choose to work under M.M.? I believe the reason was the sheer passion with which M.M. taught his courses. I'll never forget the M.M.'s closing lecture on the Bochner and Pettis integrals. Looking directly in our eyes, he said, "And now you know more about the Bochner and Pettis integrals than anyone in the department." Then he made a dramatic exit. This lecture must have grabbed me in a big way because I spent the next thirty years on research matters centering around the Bochner and Pettis integrals.

There was another reason we chose M.M. The word was out that M.M. really cared about his students. We found this to be true in spades. M.M.'s door was always open and he devoted unlimited time to us. When we had personal problems, M.M. always tried to help. When Gretsky didn't show up for a week, M.M. took to the math department hallways asking everyone, "Where is Gretsky? Where is Gretsky?" Another time when I was finishing my thesis, M.M. was hospitalized for an eye condition under doctor's orders to rest his eyes. But when I visited him in the hospital, there he was going over my thesis in detail.

M.M. did not baby us and certainly did not do our work for us. Somehow he was able to extract good work from us. One of his favorite techniques was to ask why we were stuck on a point. Quite often, we later found how to get around the apparent obstacle. All of our theses contained important new work and opened up wide opportunities for future research. I thank M.M. for preparing me so well to do what I love to do.

There was a lighter side to M.M. One day department chair Ignace Kolodner ordered all advanced grad students to attend a seminar meeting Wednesdays at 4:00. Gretsky exclaimed:"But our intramural team plays Wednesdays at 4:00!" Kolodner said: "What's more important: Sports or Mathematics? Professor Rao what do you say?" In a completely straight face, M.M. said, "I'm on the team; but I don't play." The seminar was rescheduled.

Just after I took my final Ph. D. oral fellow grad students threw a surprise party for me outside Pittsburgh in a bar near Harmarville. M.M. came in Roger Pederson's car and made it clear he was delighted to be there. I went over to him and asked him what he wanted to drink. In true M.M. form, he said, "Ginger ale." I want to the bar and ordered a double bourbon and ginger ale highball. He took the glass, took a swallow and remarked, "This is very good ginger ale." Needless to say, M.M. and the rest of us had a very good time that evening. The next day, M.M. said," I don't understand. On the way home I became dizzy and nauseous - so much so that I had to ask Roger to stop the car for a while." To this day I don't know whether M.M. knew what he had been drinking that night. I prefer to believe he did.

In my long career, I have never met a Ph.D. advisor who was as respected and as loved by his Ph.D. students as M.M. I hope the reasons are now clear.

Jerry Uhl

Stochastic Analysis and Function Spaces

M. M. Rao

Department of Mathematics,
University of California
Riverside, CA 92521

Abstract

In this paper some interesting and nontrivial relations between certain key areas of stochastic processes and some classical and other function spaces connected with exponential Orlicz spaces are shown. The intimate relationship between these two areas, and several resulting problems for investigation in both areas are pointed out. The connection between the theory of large deviations and exponential as well as vector Orlicz, Fenchel-Orlicz, and Besov-Orlicz spaces are presented. These lead to new problems for solution. Relations between certain Hölder spaces and the range of stochastic flows as well as stochastic Sobolev spaces for SPDEs are also pointed out.

1. Introduction

To motivate the problem, consider a real random variable X on (Ω, Σ, P), a probability triple, with its Laplace transform $M_X(\cdot)$, or its moment generating function, existing so that

$$M_X(t) = \int_\Omega e^{tX}\, dP,\ t \in \mathbb{R}, \tag{1}$$

is finite. Since $M_X(t) \geq 0$, consider its (natural) logarithm also called the cumulant (or semi-invariant) function $\Lambda : t \mapsto \log M_X(t)$. Then $\Lambda(0) = 0$ and has the remarkable property that it is convex. In fact, if $0 < \alpha = 1 - \beta < 1$, then one has

$$\begin{aligned}\Lambda(\alpha s + \beta t) &= \log(\int_\Omega e^{(s\alpha + t\beta)X}\, dP) \\ &\leq \log((\int_\Omega e^{sX}\, dP)^\alpha (\int_\Omega e^{tX}\, dP)^\beta), \quad \text{by Hölder's inequality,} \\ &= \alpha\Lambda(s) + \beta\Lambda(t). \end{aligned} \tag{2}$$

So as $t \uparrow \infty, 0 = \Lambda(0) \leq \Lambda(t) \uparrow \infty$, and the convexity of $\Lambda(\cdot)$ plays a fundamental role in connecting the probabilistic behavior of X and the continuouity properties of Λ. First let us note that by the well-known integral representation, one has

$$\Lambda(t) = \Lambda(a) + \int_a^t \Lambda'(u)\, du, \tag{3}$$

where $\Lambda'(\cdot)$ is the left derivative of Λ which exists everywhere and is nondecreasing. Taking $a = 0$, consider the (generalized) inverse of Λ', say $\tilde{\Lambda}'$. It is given by $\tilde{\Lambda}' : t \mapsto \inf\{t > a = 0 : \Lambda'(u) > t\}$, which, if Λ' is strictly increasing, is the usual inverse function $\tilde{\Lambda}' = (\Lambda')^{-1}$. Then $\tilde{\Lambda}'$ is also nondecreasing and left continuous. Let $\tilde{\Lambda}$ be its indefinite integral:

$$\tilde{\Lambda}(t) = \int_a^t \tilde{\Lambda}'(v)\, dv. \tag{4}$$

A problem of fundamental importance in Probability Theory is the rate of convergence in a limit theorem for its application in practical situations. It will be very desirable if the decay to the limit is exponentially fast. The class of problems for which this occurs constitute a central part of the large deviation analysis. Its relation with Orlicz spaces and related function spaces is of interest here. Let us illustrate this with a simple, but nontrivial, problem which also serves as a suitable motivation for the subject to follow.

Consider a sequence of independent random variables $X_1, X_2, \ldots$ on a probability space (Ω, Σ, P) with a common distribution F for which the Laplace transform (or the moment generating function) exists. Then the classical Kolmogorov law of large numbers states that the averages converge with probability 1 to their mean, i.e.,

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to E(X_1) = m. \quad (a.e.)$$

Expressed differently, one has for each $\varepsilon > 0, h_n(\varepsilon) \to 0$ as $n \to \infty$ in:

$$P(A_n) = P[|\frac{1}{n}\sum_{i=1}^{n} X_i - E(X_1)| > \varepsilon] = h_n(\varepsilon) + o(n); \tag{5}$$

and later it was found that $h_n(\varepsilon) = e^{-\tilde{\Lambda}(\varepsilon)}$, with $\tilde{\Lambda}$ as the *Legendre transform* of Λ, the latter being the cumulant function of F, or $\Lambda(t) = \log M_F(t), t \in I\!R$, and $\tilde{\Lambda}$ is given by

$$\tilde{\Lambda}(t) = \sup\{st - \Lambda(t) : t \in I\!R\}. \tag{6}$$

The function $\tilde{\Lambda}$ defined differently by (4) and (6) can be shown to be the same so that there is no conflict in notation. The following example illustrates and leads to further work. [$\tilde{\Lambda}$ of (6) is also termed the *complementary* or *conjugate* function of Λ.]

Let the X_n above be Bernoulli variables so that $P[X_n = 1] = p$ and $P[X_n = 0] = q(= 1 - p), 0 < p < 1$. Then the cumulant function Λ is given by $\Lambda(t) = \log(q + pe^t)$ and hence $\Lambda'(t) = \frac{pe^t}{1-p(1-e^t)}$. One finds its complementary function to be, since $m = E(X_1) = p$ and $\tilde{\Lambda}(m) = 0$,

$$\begin{aligned}\tilde{\Lambda}(t) &= \int_m^t [\Lambda'(u)]^{-1}\, du, \quad t \geq m, \\ &= \int_p^t \log\frac{u(1-p)}{p(1-u)}\, du, \quad 0 < p < u < t < 1, \\ &= t\log\frac{tq}{p(1-t)} - \log\frac{q}{1-t}, \; 0 < t < 1. \end{aligned} \tag{7}$$

and for other values $\tilde{\Lambda}(t) = \infty$. If the X_n describe a fair coin, so that $p = q = \frac{1}{2}$, one gets $\tilde{\Lambda}(t) = t\log\frac{t}{1-t} + \log 2(1-t), 0 < t < 1, = \infty$, otherwise. The complementary function, written in a more symmetric form can be expressed as:

$$\tilde{\Lambda}(t) = t\log 2t + (1-t)\log 2(1-t), \; 0 < t < 1; = \infty, \; t \in I\!R - \{t : 0 < t < 1\}. \tag{8}$$

A direct computation using exact values of (5) with the binomial probability argument shows that for any Borel set $A \subset I\!R$ and $A_n = \{k : |\frac{k}{n} - \frac{1}{2}| \geq \varepsilon\} \cap A$:

$$Q_n(A) = P(A_n) = \sum_{k \in A_n} \binom{n}{k} \frac{1}{2^n}. \tag{9}$$

One can simplify this with Stirling's approximation for the factorials in $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, and get

$$\lim_{n\to\infty} \log Q_n(A) = -\lim_{n\to\infty} \min_{k\in A_n} \tilde{\Lambda}(\frac{k}{n}) = -\inf_{x\in A} \tilde{\Lambda}(x). \tag{10}$$

The result (8) (or (7)) for $\tilde{\Lambda}(\cdot)$ shows how the convexity of Λ may be utilized along with the classical Young complementary function calculation to arrive at (10) directly. The latter observation gives a connection with certain important function spaces. This is better identified if the random variables are also symmetric, i.e.,

$$P[X_n = 1] = \frac{1}{2} = P[X_n = -1], \quad n \geq 1,$$

and $Y_n = \frac{1}{n}\sum_{i=1}^n X_i$, $Q_n = P \circ Y_n^{-1}$. Then it is seen immediately that $\Lambda(t) = \Lambda(-t)$, $\Lambda(0) = 0$ and $\Lambda(t) = \infty, |t| > 1$, so that Λ is a classical Young function as used in functional anlysis (cf., e.g., Krasnoselski and Rutickii (1961)), and its complementary function is seen at once from (7) to be:

$$\tilde{\Lambda}(t) = \frac{1+t}{2}\log(1+t) - \frac{t-1}{2}\log(1-t), \quad |t| \leq 1; \ = \infty, \ |t| > 1. \tag{11}$$

In this case (10) can be obtained directly and it is simpler than the probabilistic calculation.

Similar computations can be performed for random variables with continuous distributions. For instance, let the X_n above be exponentially distributed, i.e., $P[X_n < x] = 1 - e^{-x}, x > 0$ so that its cumulant function is given by $\Lambda(t) = -\log(1-t), t < 1$, and since $E(X_n) = 1$, one has as in (7), $\tilde{\Lambda}(t) = \int_1^t (1 - \frac{1}{y})dy = (t - 1 - \log t), t > 0; = \infty$, otherwise. More general (but not all) distributions, with cumulant functions such as a Gamma, can be illustrated.

The basic problem here is the evaluation of the complementary function explicitly from the (convex) cumulant function of the random variable X_1. However, for a general Young function $\phi : \mathbb{R} \to \mathbb{R}^+$ this is not always simple (or even possible). For instance if $\phi(t) = e^{t^2} - 1$, which is a symmetric nonnegative convex function vanishing at the origin, its complementary function $\psi : t \mapsto \psi(t) = \sup\{|t|s - \phi(s) : s \geq 0\}$ exists uniquely but cannot be computed explicitly, although many properties can be studied in great detail. This implies that one may find a suitable *subclass* of Young functions for which the complementary function may be explicitly calculated, but now even dropping the previous symmetry condition. This is motivated by the large deviation study, and will be considered in the next section. Then in Section 3 the regularity of stochastic processes and exponential Orlicz spaces will be treated. The Fenchel-Legendre transforms and their study on infinite dimensional spaces are discussed in Section 4. Also the perturbation of stochastic differential equations and the sample path behavior leads to Besov-Orlicz spaces. Stochastic flows take values in certain Hölder spaces, and all of this is considered in Section 5. Finally the analysis of SPDE and stochastic Sobolev speces are briefly noted in Section 6 to explain the close relations. The work shows a clear usefulness of analyzing the problems with the ideas and results of one area in the other, and the mutual benefits for both subjects.

2. Large deviations and function spaces of Orlicz type

The first substatial result, subsuming the Bernoulli case discussed in the preceding section, is due to H. Cramér (1938). It will be presented to focus the general problem of large deviations, and to explain the connection more concretely.

1. Theorem. *Let $\{X_n, n \geq 1\}$ be a sequence of independent identically distributed random variables on a probability space (Ω, Σ, P), for which the cumulant generating function $\Lambda(\cdot)$ exists on $\mathbb{R}$. Let Y_n be the sample average, i.e., $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. If $Q_n = P \circ Y_n^{-1}$ is the image measure of Y_n on $\mathbb{R}$, and $A \subset \mathbb{R}$ is a Borel set such that $P \circ X_1^{-1}(\partial A) = 0$ where ∂A is the boundary of A which is clearly measurable, then one has*

$$\lim_{n \to \infty} \frac{1}{n} \log Q_n(A) = -\inf_{x \in A} \tilde{\Lambda}(A), \tag{12}$$

where $\tilde{\Lambda}$ is the complementary function of the Young function Λ, i.e., the one defined by (6) of the last section.

It should be noted that $\tilde{\Lambda} \geq 0$, convex, and not symmetric, which is reflected by (similar) properties of Λ. The proof of this result is nontrivial, compared to the Bernoulli case, and the details are spelled out, e.g., in Rao and Ren [(2002), Chapter VIII]. The point to note is the fundamental role played by the complementary function $\tilde{\Lambda}$ for the assertion of (12). Let us abstract this conclusion to proceed further. Motivated by the special Bernoulli problem, one may term $Q_n(\cdot)$ the 'large deviation probability' and (12) implies that it decreases to zero exponentially fast when A does not contain $E(X_1)$, the pointwise limit of the Y_n being a consequence of the Kolmogorov law of large numbers.

Since for vector or $\mathbb{R}^k$-valued random variables X_n, if the cumulant function $\Lambda : \mathbb{R}^k \to \bar{\mathbb{R}}^+$ exists, $\tilde{\Lambda}$ defined by (6), with st replaced by $\langle s, t \rangle$, (the inner product) is also nonnegative, in this case the mapping is called the *Fenchel-Legendre transform*, alternatively to the term 'complementary function', of

Λ. So to take into account of all the cases seen in the solutions above, one may introduce directly a mapping $I : \mathbb{R}^n \to \bar{\mathbb{R}}^+$, called the rate function, as follows:

Definition 2. Let $I : \mathcal{X} \to \bar{\mathbb{R}}^+$ be a mapping, $\mathcal{X}$ being a complete separable metric (also termed a *Polish*) space, and $Q_n : \mathcal{B}(\mathcal{X}) \to [0, 1]$ (an image probability before) a measure on the Borel σ-algebra of $\mathcal{X}$. Then the sequence $\{Q_n, n \geq 1\}$ is said to have the *large deviation property* relative to I (called the *rate or entropy function*) if the following properties hold:

(i) $I(\cdot)$ is lower semi-continuous,

(ii) $\{x : I(x) \leq a\}, a \geq 0$, is compact in $\mathcal{X}$ for each $a \in \mathbb{R}^+$,

(iii) for each $A \in \mathcal{B}(\mathcal{X})$ one has

$$-\inf_{x \in A^o} I(x) \leq \liminf_{n\to\infty} \frac{1}{a_n} \log Q_n(A) \leq \limsup_{n\to\infty} \frac{1}{a_n} \log Q_n(A) \leq -\inf_{x \in \bar{A}} I(x), \quad (13)$$

where A^o is the interior and $\bar{A}$ the closure of A and $0 \leq a_n \uparrow \infty$. [If $\mathcal{X}$ is infinite dimensional, I is also termed *action functional.* Different names originating in diverse applications.]

It is seen above that if Λ is the cumulant function of P, $\tilde{\Lambda}$ its complementary function (or the Fenchel-Legendre transform), then the latter qualifies to be a rate function of the large deviation problem for Bernoulli variables; but it is also true when $\mathcal{X} = \mathbb{R}^k, k \geq 1$ and $a_n = n$. A detailed argument of this fact is in Ellis ((1985),Chapter II, Section 4) where the rate function is called an *entropy function,* and as noted above it is also termed action functional. If $\mathcal{X}$ is infinite dimensional, for instance a separable Hilbert space, then the Young function and the Fenchel-Legendre transform and its conjugate have to be considered with some care, and this will be discussed further below.

To understand the significance of the close relationship between the (multidimensional) Young complementary pair and the corresponding rate or action function a version of the multidimensional Cramér theorem will be presented, first in $\mathbb{R}^k$ and then for infinite dimensions, both of which actually arise in important applications in stochastic analysis. It will be seen that the connection, not yet fully exploited, shows various aspects of function space results. They are needed here but are also of interest in different parts of (abstract) analysis. This makes it clear that a comparative study of these areas is beneficial mutually.

Theorem 3. *Let* $\{X_n, n \geq 1\}$ *be independent identically distributed ran-*

dom variables on a probability space (Ω, Σ, P) with values in $\mathbb{R}^k, k > 1$, having a well-defined (moment or) cumulant generating function $\Lambda : t \mapsto \log(\int_\Omega e^{\langle t, X_1(\omega)\rangle} dP(\omega)), t \in \mathbb{R}^k$. Then the rate function $I(= \tilde{\Lambda}$, the complementary function of $\Lambda)$ is given by

$$\tilde{\Lambda}(s) = I(s) = \sup\{\langle s,t\rangle - \Lambda(t) : t \in \mathbb{R}^k\} \geq 0, \tag{14}$$

$\langle\cdot,\cdot\rangle$ being the inner product of $\mathbb{R}^k$. Moreover $\tilde{\Lambda}$ is convex, continuous, and has its minimum at $a = E(X_1)$.

A direct proof of this result was first sketched by Sanov (1957), and a full argument is considerably more involved than its one-dimensional counterpart.

The classical study of Young's complementary function suggests that (3) may be considered as a definition of the conjugate function $\tilde{\Lambda}$. This was verified in a basic paper by Fenchel (1949) in $\mathbb{R}^k, k > 1$, but a careful reinterpretation of it gives a useful generalization as follows. Indeed, the complementary Young pair satisfying (2) symmetrically in the sense that if Λ is such a convex (continuous) function, with $\tilde{\Lambda}$ as its conjugate, then

$$\begin{aligned}\tilde{\tilde{\Lambda}}(t) &= \sup\{\langle s,t\rangle - \tilde{\Lambda}(s) : s \in \mathbb{R}^k\} \\ &= -\inf\{\tilde{\Lambda}(s) + \langle s,t\rangle : s \in \mathbb{R}^k\},\end{aligned} \tag{15}$$

since the mapping $s \mapsto -s$ is a homeomorphism on $\mathbb{R}^k$ onto itself. It can be shown that one always has $\tilde{\tilde{\Lambda}} = \Lambda$ if $\mathcal{X}$ is a reflexive separable Banach space, while the same is not true for a general (nonreflexive) $\mathcal{X}$. But this is true for $\mathcal{X} = \mathbb{R}^k$, since all finite dimensional Banach spaces are reflexive. Thus if $\Lambda : \mathcal{X} \to \bar{\mathbb{R}}^+$ is a Young functional and $\tilde{\tilde{\Lambda}}$ is defined with $\mathbb{R}^k$ replaced by $\mathcal{X}$, using the duality mapping $(\cdot,\cdot) : \mathcal{X} \times \mathcal{X}^* \to \mathbb{R}$ so that $\tilde{\Lambda} : \mathcal{X}^* \to \bar{\mathbb{R}}^+$, one has $\tilde{\tilde{\Lambda}} : \mathcal{X}^{**} \to \bar{\mathbb{R}}^+$. Hence $\Lambda = \tilde{\tilde{\Lambda}}|\mathcal{X}$ when $\mathcal{X}$ is identified as a subspace of $\mathcal{X}^{**}$ by the standard natural embedding. [For this extension procedure, see Levin (1975).] These relations can be generalized as follows.

To begin, let us rewrite (4), using the fact that $\tilde{\tilde{\Lambda}} = \Lambda$ on $\mathbb{R}^k$, a finite dimensional Banach space. Thus one has

$$\Lambda(t) = \log(E(e^{\langle t, X_1\rangle})) = -\inf\{\langle s,t\rangle + I(s) : s \in \mathcal{X}^* = \mathbb{R}^k\},\ s \in \mathbb{R}^k. \tag{16}$$

Since the X_n are independent and identically distributed, (13) and (16) give on replacing X by $Y_n = \frac{1}{n}\sum_{i=1}^n X_i$:

$$\Lambda_n(t) = \log(E(e^{\langle t, Y_n\rangle})) = n\log(E(e^{\langle \frac{t}{n}, X_i\rangle})) = n\tilde{\Lambda}(\frac{t}{n}),$$

and (15) becomes

$$\Lambda_n(t) = -\inf\{n\tilde{\Lambda}(s) + n\langle\frac{t}{n}, s\rangle : s \in I\!R^k\}.$$

Replacing t by $n\tau$, this becomes on cancelling the factor n:

$$\frac{1}{n}\log(E(e^{n\langle\tau, Y_n\rangle})) = -\inf\{\tilde{\Lambda}(s) + \langle s, \tau\rangle : s \in I\!R^k\}, n \geq 1.$$

Here letting $n \to \infty$, the limit exists, and one has on setting $h_t(\cdot) = \langle t, \cdot\rangle$, so that h_t is a bounded linear function on $I\!R^k$,

$$\lim_{n\to\infty} \frac{1}{n}\log(E(e^{nh_t(Y_n)}) = -\inf\{\tilde{\Lambda}(s) + h_t(s) : s \in I\!R^k\}. \tag{17}$$

This result admits an extension if $h : \mathcal{X} \to I\!R$ is not necessarily linear but simply bounded and continuous and $I\!R^k$ is replaced by a (separable) Banach space when the left side limit is assumed to exist. This is also motivated by a classical evaluation, due to Laplace, of the integral by a Taylor series expansion of the smooth function $h : [0, 1] \to I\!R$ to obtain

$$\lim_{n\to\infty} \frac{1}{n}\log(\int_0^1 e^{-nh(x)}\, dx) = \min_{x\in[0,1]} h(x). \tag{18}$$

Combining (17) and (18), one gets the following generalization, proved directly by Varadhan (1966) for an arbitrary random sequence $\{X_n, n \geq 1\}$ with values in a complete separable metric (or a Polish) space $\mathcal{X}$, whose distributions (i.e., the image measures on $\mathcal{X}$) satisfy the large deviation principle in the sense of Definition 2 above. This generalizes the result of Theorem 2 above with cumulant function Λ and its conjugate $\tilde{\Lambda}$ as the rate function. More precisely one has:

Theorem 4. *Let $\{X_n, n \geq 1\}$ be a sequence of $\mathcal{X}$ valued random variables obeying the large deviation principle with an action functional $I : \mathcal{X} \to \bar{I\!R}^+$ where $\mathcal{X}$ is a Polish space. Then for any $h \in C_b(\mathcal{X})$, the space of real bounded continuous functions on $\mathcal{X}$, equation (17) holds, i.e., one has:*

$$\lim_{n\to\infty} \frac{1}{n}\log(E(e^{-nh(X_n)}) = -\inf\{h(x) + I(x) : x \in \mathcal{X}\}. \tag{19}$$

Conversely, if (19) holds for any $h \in C_b(\mathcal{X})$, for an $\mathcal{X}$-valued sequence $\{X_n, n \geq 1\}$ of random variables, then the latter satisfies the large deviation principle for a unique action functional I in the sense of Definition 2.

If $\mathcal{X}^*$ is a separable adjoint Banach space, $C_b(\mathcal{X})$ is replaced by $\mathcal{X}^*$ in the above, and Λ_t is the cumulant functional of X_t for each $t \in J$, an index set, then one can extend the above result to the case of a continuous parameter process or field. Let us illustrate this with the case of Brownian Motion (or BM), and then give an immediate extension to all centered Gaussian processes with continuous covariance functions. In these cases admitting infinite dimensional $\mathcal{X}$ is essential as will now be seen. If $\varepsilon = \frac{1}{n} \downarrow 0$ then (17) becomes formally

$$\lim_{\varepsilon \searrow 0} \log(E(e^{-\frac{1}{\varepsilon}h(Y^\varepsilon)})) = -\inf\{\tilde{\Lambda}(s) + h(s) : s \in I\!R^k\},$$

where Y^ε is defined in terms of the X-process, so that $Y^\varepsilon \to 0$ in probability as $\varepsilon \searrow 0$. Indeed, if $Y^\varepsilon(t) = \sqrt{\varepsilon}X(t)$ for a process $X(t)$ with zero mean and covariance r, then the Čebyshev inequality gives for each $\delta > 0$

$$P[|Y^\varepsilon(t)| > \delta] \leq \frac{\varepsilon}{\delta^2}E(|X(t)|^2) = \frac{\varepsilon r(t,t)}{\delta^2} \to 0, \tag{20}$$

as $\varepsilon \searrow 0$ for each t. If X_t is the BM which has continuous paths, i.e., $X_{(\cdot)}(\omega) : [0,1] \to C_0([0,1]) = \mathcal{X}$, then the problem is to find the (exponential) rate of convergence $Y^\varepsilon \to 0$ (for all t) as $\varepsilon \searrow 0$, and this means to find the rate or action functional related to the BM on the (Polish) space $C_0([0,1])$, the space of real continuous functions on $[0,1]$ vanishing at 0 under the uniform norm. To get out of the t-points here one considers the probability measure determined by the entire process, and then the problem is solved from it, as follows.

The canonical representation of a real process is obtained, using the classical Kolmogorov existence theorem, with $\Omega = I\!R^T, \Sigma =$ the cylinder σ-algebra of Ω, and P determined by the compatible family of all finite dimensional distributions. Then the process has the coordinate (or function space) representation $X_t(\omega) = \omega(t), \omega \in \Omega, t \in T$, and whose finite dimensional distributions are the given ones, i.e.,

$$P[\omega : X_{t_i}(\omega) < x_i, i = 1, \ldots, n] = F_{t_1,\ldots,t_n}(x_1, \ldots, x_n),\ x_i \in I\!R,\ t_i \in T.$$

The Fourier transform of the process related to P is defined uniquely as:

$$\int_\Omega e^{i\sum_{j=1}^n X_{t_j}(\omega)x_j}\, dP(\omega) = \int_\Omega e^{i\langle X,x\rangle(\omega)}\, dP(\omega), \tag{21}$$

where $\langle X, x\rangle(\omega) = \sum_{j=1}^n X_{t_j}(\omega)x_j(= \int_T X_t\mu(dt))(\omega)$ with μ as a signed (Stieltjes) measure on T having values x_j at t_j. Thus it may be written as $\langle X, x\rangle = \langle\omega(\cdot), \mu\rangle$, $\omega \in \Omega$. In the case that P is a Gaussian measure on $\mathcal{X} = I\!R^{[0,1]}$, (20) can be evaluated to obtain:

$$\int_\Omega e^{i\langle X,x\rangle(\omega)}\, dP(\omega) = e^{-Q(x,x)},\ x \in \mathcal{X}^*, \tag{22}$$

where $Q : \mathcal{X}^* \times \mathcal{X}^* \to I\!R^+$ is a positive definite bilinear form, $\mathcal{X}^*$ being the adjoint space of the vector space $\mathcal{X}$, so that $Q(x,x) = V_x$, the variance, and $Q(x,y) = C_{x,y}$ the covariance of the process X. For a BM which has continuous paths, $\mathcal{X} = C_0([0,1])$ and so $\mathcal{X}^* = M[0,1]$, the space of regular signed Borel measures, and $\langle X, x\rangle = \langle X, \mu\rangle (= \int_0^1 X_t \mu(dt))$, a standard Bochner (or abstract vector Lebesgue) integral. In this case (21) is expressible as:

$$\begin{aligned}\int_\Omega e^{i\langle X,\mu\rangle(\omega)} dP(\omega) &= \int_\Omega e^{i\langle \omega,\mu\rangle}\, dP(\omega)\\ &= e^{-\Lambda(\mu)}\\ &= e^{-\frac{1}{2}\int_0^1\int_0^1 r(s,t)\mu(ds)\mu(dt)},\end{aligned} \tag{23}$$

where $r(s,t) = s \wedge t$, and $\Lambda(\cdot)$ is the cumulant function as before. It can be verified (nontrivially) that (cf., e.g., Deuschel and Stroock (1989),p.85)

$$\Lambda(\mu) = \frac{1}{2}Q(\mu,\mu) = \int_{\mathcal{X}} \langle x,\mu\rangle^2\, \tilde{P}(dx),\ \ \mu \in \mathcal{X}^*, \tag{24}$$

where $\tilde{P}$ is the image measure of the process on its range space $\mathcal{X}$. It follows by the duality pairing (and the CBS inequality) that ($\|\cdot\|$ denoting norms in both $\mathcal{X}, \mathcal{X}^*$)

$$\Lambda(\mu) \le \|\mu\|^2 \int_{\mathcal{X}} \|x\|^2\, d\tilde{P}(x). \tag{25}$$

Consequently the conjugate function $\tilde{\Lambda}$ of Λ (or the generalized complementary Young function) is given by

$$\tilde{\Lambda}(x) = \sup\{\langle \mu, x\rangle - \Lambda(\mu) : \mu \in \mathcal{X}^*\},\ \ x \in \mathcal{X}. \tag{26}$$

It is now necessary to have explicit expressions for $\tilde{\Lambda}$, and this is now considered.

Recalling (19) and (20), let P be the probability measure of the BM process, and consider the measure P_ε that of $Y_t^\varepsilon = \sqrt{\varepsilon}X_t$ so that it has values again in $\mathcal{X} = C_0([0,1])$ and $\lim_{\varepsilon\searrow 0} P_\varepsilon \Rightarrow \delta_0$, the point measure at 0 (vague convergence). Then one has the following explicit evaluation of $\tilde{\Lambda}$ due to Schilder (1966), with a different (and shorter) proof in Varadhan (1984):

Theorem 5. *The family of probability measures $\{P_\varepsilon, \varepsilon > 0\}$ satisfies the large deviation principle with the action functional $\tilde{\Lambda}$ given on $\mathcal{X}$ by:*

$$\tilde{\Lambda}(f) = \frac{1}{2}\int_0^1 |f'(t)|^2\, dy, \tag{27}$$

for all absolutely continuous $f \in C_0([0,1])$ whose derivatives are square integrable, and $\tilde{\Lambda}(f) = \infty$ otherwise. Thus the domain of $\tilde{\Lambda}$ in which it is finite *is the subset of* $C_0([0,1])$ *given by* $\mathcal{X}_0 = \{f \in \mathcal{X} : \int_0^1 |f'(t)|^2\, dt = \|f'\|^2 < \infty\}$, *which is a Hilbert space, a subspace of* $\mathcal{X}$ *but with a stronger norm topology.*

It will next be shown how this admits an extension to general Gaussian processes, with continuous covariance and mean functions. Thus let $X = \{X_t, t \in T = [a,b]\}$ be a centered Gaussian process with a continuous covariance function $r : T \times T \to \mathbb{R}$. Consider the integral operator R with kernel r, so that

$$(Rf)(t) = \int_T r(s,t) f(s)\, ds, \tag{28}$$

which is compact (even Hilbert-Schmidt) on $\mathcal{X}_0$ and is positive definite. Also R is nonsingular and has a unique square root $R^{\frac{1}{2}}$ which has an inverse. Then the range space $M = R^{\frac{1}{2}}(L^2(T,dt))$ can be described precisely as follows. For each $f \in M \subset \Omega = \mathbb{R}^T$, consider a new process defined by $Y_t^f(\omega) = X_t(\omega) + f(t)$ whose probability measure P_f has the property that $P_f \ll P$. Such an f is called an *admissible mean* of the X_t-process. Then M can be provided with a new inner product as follows. Since $f_i \in M \Rightarrow f_i = R^{\frac{1}{2}} h_i, i = 1,2$, let $(f_1, f_2)_M = (Rh_1, h_2)_{L^2(T,dt)}$, the last symbol being the scalar product of $L^2(T,dt)$. Then M is a Hilbert space (cf., e.g., Rao (1975) for a proof of this fact), and $R^{\frac{1}{2}}$ also has a kernel representation (cf. Dunford-Schwartz (1958),VI.9.59) as:

$$f(t) = (R^{\frac{1}{2}} h)(t) = \int_T G(s,t) h(s)\, ds, \quad h \in L^2(T,dt), \tag{29}$$

so that the covariance r can be expressed as:

$$r(s,t) = \int_T G(s,u) G(t,u)\, du. \tag{30}$$

Next define a new process on the same canonical space: $\tilde{X}_t = \int_T G(s,t)\, dB_s$, where $\{B_s, s \in T\}$ is the standard BM on the same space. This being a linear transformation, $\{\tilde{X}_t, t \in T\}$ is also a centered Gaussian process having the same covariance function r as the X_t-process. Since a Gaussian process is determined by its mean and covariance functions, these two processes can be identified, $X_t = \tilde{X}_t, a.e., t \in T$. Now $f = R^{\frac{1}{2}} h, h \in L^2(T,dt)$ and $R^{-\frac{1}{2}}$exists, and so by Theorem 5, one has the following exact form of the conjugate function $\tilde{\Lambda}$ of Λ:

$$\tilde{\Lambda}(f) = \frac{1}{2}\|R^{-\frac{1}{2}} f\|^2 = \frac{1}{2}(R^{-1} f, f)_{L^2(T,dt)},\ f \in M = R^{\frac{1}{2}}(L^2(T,dt)), \tag{31}$$

and $\tilde{\Lambda}(f) = \infty$, otherwise. This may be summerized as:

Theorem 6. *Let $\{X_t^\varepsilon = \sqrt{\varepsilon} X_t, t \in T, \varepsilon > 0\}$ be a centered Gaussian process with a continuous covariance function. Then X_t^ε-process obeys the large deviation principle with the action or rate functional $\tilde{\Lambda}$ given by (20).*

This was first established by Freidlin with a slightly different argument (cf., Freidlin and Wentzell (1998), Sec. 3.4). The point of these cases here is that the Fenchel-Young (or cumulant) functions Λ are defined on infinite dimensional (separable) Hilbert or Banach spaces, and their action functionals $\tilde{\Lambda}$ are explicitly calculable. They are of primary interest in the large deviation problems. Because of this circumstance these spaces will be analyzed from the function space point of view in Section 4 below.

It is also useful to discuss another class of fast growing Young functions to be included in that study since it is not usually detailed in general abstract analysis.

3. Exponential Orlicz spaces and stochastic processes

The following familiar problem leads to a consideration of exponential Orlicz spaces, as path spaces of the associated process whose structure provides precise description of the growth of sample functions of interest in applications. Thus consider a Poisson process $\{X_t, t \geq 0\}$ used to describe telephone traffic, so that the process has independent and stationary increments. Then $X_0 = 0, a.e.$ and for $0 < t_1 < t_2$ are any time points,

$$P[X_{t_2} - X_{t_1} = k] = e^{-c(t_2-t_1)} \frac{(c(t_2 - t_1))^k}{k!}, \quad k = 0, 1, 2, \ldots. \tag{32}$$

Here $c > 0$ is a constant, called the intensity parameter. Let Y be a symmetric Bernoulli random variable, independent of the X_t-process, so that $P[Y = -1] = P[Y = +1] = \frac{1}{2}$. Let $\{Z_t = Y(-1)^{X_t}, t \geq 0\}$, a process which has practical interest and has bounded paths. It is desired to describe the growth behavior of $t \mapsto Z_t(\omega)$, for almost all sample points ω. Note that the Z_t-process is also stationary and has moving discontinuities. This will follow from the computation, given below, for the growth rate of the increments of the process in terms of an exponential Orlicz norm. Indeed, since X_t and Y are independent $E(Y) = 0$, one has $E(Z_t) = 0$ and using the independent increment property of the X_t-process, it is seen that for $s < t$

$$\begin{aligned} E(Z_s Z_t) &= E(Y^2) E((-1)^{X_t - X_s}) E((-1)^{2X_s}) \\ &= e^{-2c(t-s)}. \end{aligned} \tag{33}$$

So for $0 \leq s, t < \infty$, one has $Cov(Z_s, Z_t) = e^{-2c|t-s|}$, implying the (Khintchine) stationarity of the Z_t-process. Also by definition $Z_t - Z_s$ takes only the three values $0, \pm 1$, and then

$$P[Z_t - Z_s = 0] = P[X_t - X_s = 2k,\ k = 0, 1, 2, \ldots]$$
$$= e^{-c|t-s|} \cosh(c(t-s)) = a(s,t), \text{ (say)}, \tag{34}$$

and similarly that

$$P[Z_t - Z_s = 1] = P[Z_t - Z_s = -1] = \frac{1}{2}[1 - a(s,t)]. \tag{35}$$

If $\Phi(\cdot)$ is an N-function (i.e., a symmetric nonnegative convex function, $\Phi(0) = 0, \frac{\Phi(x)}{x} \to 0(\infty)$ as $x \to 0(\infty)$), and $\Lambda(x) = e^{\Phi(x)} - 1$, which is again an N-function, then the space $L^{\Lambda}(P)$ on a probability space (Ω, Σ, P) of functions $f : \Omega \to \mathbb{R}$ for which $\int_\Omega \Lambda(kf)\, dP < \infty$ for some $k > 0$ is an *exponential Orlicz space* which is a Banach space under the norm:

$$\|f\|_{(\Lambda)} = \inf\{k > 0 : \int_\Omega \Lambda(\frac{f}{k})\, dP \leq 1\}, \tag{36}$$

when equivalent functions are identified. This is well known (cf., e.g., Rao and Ren (2002)). Such Λ satisfy the so-called Δ^2 condition, i.e., there exist constants $K > 0, x_0 > 0$ such that $\Lambda^2(x) \leq \Lambda(Kx)$, $x \geq x_0$, and the above function satisfies this, since $\Lambda^2(x) = e^{2\Phi(x)} + 1 - 2e^{\Phi(x)} \leq e^{\Phi(2x)} - 1 = \Lambda(2x), x \geq 0$. Moreover, any Δ^2 function has the growth condition satisfying (as seen from Rao and Ren, *loc. cit.*):

$$e^{x^\alpha} \leq \Lambda(x), \quad x \geq x_1 \geq 0,$$

for some $\alpha > 0$ and an x_1. It is also easy to see that

$$L^\infty(P) \subset L^\Lambda(P) \subset \cap_{p>1} L^p(P) \subset \cup_{p>1} L^p(P) \subset L^{\tilde{\Lambda}}(P) \subset L^1(P), \tag{37}$$

where $\tilde{\Lambda}$ is the complementary function of Λ. Since Z_t is bounded, it is in $L^\Lambda(P)$. Using (36) one can compute $\|Z_t - Z_s\|_{(\Lambda)}$ for any exponential N-function Λ and get:

$$\|Z_t - Z_s\|_{(\Lambda)} = \inf\{k > 0 : [\Lambda(-\frac{1}{k}) + \Lambda(\frac{1}{k})]\frac{1 - a(s,t)}{2} \leq 1\}, \ (\Lambda(0) = 0)$$
$$= \inf\{k > 0 : \Lambda(\frac{1}{k})(1 - a(s,t)) \leq 1\}, \text{ by symmetry of } \Lambda,$$
$$= [\Lambda^{-1}(\frac{1}{1 - a(s,t)})]^{-1}.$$

Also from the fact that for any $t_0 > 0$

$$P[|Z_{t_0} - Z_0| > 0] = 1 - a(0, t_0) \geq 1 - e^{-ct_0}, \tag{38}$$

it follows that $P[|Z_{t_0+\varepsilon} - Z_{t_0-\varepsilon}| > 0] \to 0$ as $\varepsilon \searrow 0$, since $a(0,\varepsilon) \to 1$ as $\varepsilon \to 0$. Thus $t_0 > 0$ is not a fixed discontinuity point, but is a moving one.

This example shows that exponential Orlicz spaces $L^\Lambda(P)$ for $\Lambda(x) = e^{\Phi(x)} - 1$, with Φ as an N- (or a continuous Young-) function, are of interest in applications. The spaces $L^\Lambda(P)$ lie between $L^p(P)$ and $L^\infty(P)$ for all $p \geq 1$, as seen from (37). They also play a key role in pre-Gaussian random variables. Several applications may be found in Budygin and Kozachenko (2000).

Recall that a random variable X is pre- (or sub-) Gaussian if its moment generating function exists, in a nondegenerate neighborhood of the origin (or $\mathbb{R}$) and is dominated by that of a centered Gaussian variable in that neighborhood (in all $\mathbb{R}$), i.e., there are an $\alpha > 0$ and a $\beta \geq 0$ such that

$$E(e^{tX}) \leq e^{\frac{t^2\beta^2}{2}}, \quad \forall t \in (-\alpha, \alpha), \ (t \in \mathbb{R}). \tag{39}$$

Thus the cumulant function of X is dominated by a quadratic form in that interval. If $\mathcal{G}_0$ denotes the class of pre-Gaussian random variables, then $\mathcal{G}_0 \subset L^\Lambda(P)$ for a suitable exponential Orlicz space, and is a closed subspace, hence a Banach space. This is a nontrivial fact, depending on properties of the latter function spaces, somewhat similar to the preceding sections. This point is a consequence of the following result.

Theorem 1. *Let $\Lambda : x \mapsto e^{\Phi(x)} - 1$ where Φ is a continuous Young function such that $\Phi(x) > 0$ for $x > 0$, and consider the (exponential) Orlicz space $L^\Lambda(P)$, as above. Then:*

(i) $f \in L^\Lambda(P)$ iff there exist constanta $C(= C_f) > 0$, $D(= D_f) > 0$ such that

$$P[\omega : |f(\omega)| > x] \leq Ce^{-\Phi(\frac{x}{D})}, \quad x \geq 0, \tag{40}$$

and then

$$\|f\|_{(\Lambda)} \leq (1 + \frac{C}{2})D; \tag{41}$$

(ii) if $\Phi(x) = |x|^p, p \geq 1$, and $N(f) = \sup_{n\geq 1} \frac{\|f\|_{np}}{\sqrt{n}}$, where $\|\cdot\|_{np}$ is the Lebesgue norm of $L^{np}(P)$, then $N(\cdot)$ is a norm equivalent to $\|\cdot\|_{(\Lambda)}$.

Here (i) is from Buldyagin and Kozachenko (2000) whose bound for the norm in (10) is slightly improved ($(1 + \frac{C}{2})$ instead of $(1 + C)$ there). This whole result is detailed in Rao and Ren (2002, Sec. 8.3). For (ii) which is essentially due to Fernique (1971), and which is also in the latter reference, there is unfortunately a typographical slip at a key point which may confuse the reader. So it will be sketched here for convenience. As a consequence of this part, one obtains that the space of pre-Gaussian variables $\mathcal{G}_0$ is a Banach space.

Sketch of Proof of (ii). Since $\Phi(x) = |x|^p, \Lambda(x) = e^{|x|^p} - 1 = \sum_{n\geq 1} \frac{|x|^{np}}{n!}$, define the functional $M : f \mapsto \sup_{n\geq 1} \frac{\|f\|_{np}}{(n!)^{\frac{1}{np}}}$ and an $a_n > 0$, by the equation $a_n(n!)^{\frac{1}{np}} = \|f\|_{np}$. Then one has

$$1 \geq \int_\Omega \Lambda(\frac{|f|}{\|f\|_{(\Lambda)}})\, dP = \sum_{n\geq 1} \frac{1}{n!}(\frac{\|f\|_{np}}{\|f\|_{(\Lambda)}})^{np}$$
$$= \sum_{n\geq 1}(\frac{a_n}{\|f\|_{(\Lambda)}})^{np}, \tag{42}$$

Thus $a_n \leq \|f\|_{(\Lambda)}, n \geq 1$, and so $M(f) \leq \|f\|_{(\Lambda)}$. However

$$\int_\Omega \Lambda(\frac{|f|}{2^{\frac{1}{p}} M(f)})\, dP = \sum_{n\geq 1}(\frac{a_n^p}{2M(f)^p})^n$$
$$\leq \sum_{n\geq 1} \frac{1}{2^n} = 1,$$

since by definition of a_n one has $a_n \leq M(f), n \geq 1$. Consequently, $\|f\|_{(\Lambda)} \leq 2^{\frac{1}{p}} M(f)$, using the definition of the gauge norm. Hence

$$M(f) \leq \|f\|_{(\Lambda)} \leq 2^{\frac{1}{p}} M(f) \leq 2M(f),\ p \geq 1, f \in L^\Lambda(P), \tag{43}$$

so that the functionals $M(\cdot)$ and $\|\cdot\|_{(\Lambda)}$ are equivalent.

Now observe that $\|f\|_n \uparrow \|f\|_\infty$, and by Stirling's approximation for $n!$, one gets easily $\sup_{n\geq 1}(\frac{1}{n!})^{\frac{1}{np}} \sim \sup_{n\geq 1} \frac{1}{\sqrt{n}}$, and hence $N : f \mapsto N(f) = \sup_{n\geq 1} \frac{\|f\|_{np}}{\sqrt{n}}$ is a norm functional and is equivalent to $M(\cdot)$. Thus $N(f) \sim \|f\|_{(\Lambda)}$, or the norms $N(\cdot)$ and $\|\cdot\|_{(\Lambda)}$ are equivalent. □

Remark. If f is a pre-Gaussian variable, then one finds that $\tilde{N} : f \mapsto \sup_{n\geq 1}(\frac{\|f\|_n}{n!})^{\frac{1}{n}}$ is a norm and it is equivalent to N and hence to $\|\cdot\|_{(\Lambda)}$. Thus the set $\mathcal{G}_0$ of pre-Gaussian variables f for which $\tilde{N}(f) < \infty$ is a subspace of $L^\Lambda(P)$, by part (ii) of the above theorem. By the equivalence of norms and the completeness of an (exponetial) Orlicz space, one concludes that $(\mathcal{G}_0, \tilde{N}(\cdot))$ is a Banach space. A direct proof of this fact without the Orlicz space theory is considerably more difficult.

It is natural to seek a common abstract function space analysis that combines both this and the preceding multidimensional Young function versions to give an overview of the underlying functional structure. This is discussed in the next section.

4. Fenchel-Orlicz spaces for stochastic applications

In the problems of large deviations of BM and the general Gaussian processes, it was seen that the Fenchel-Young function $\Lambda : \mathcal{X} \to \bar{\mathbb{R}}^+$ where $\mathcal{X}$ is an infinite dimensional Polish space, and in the preceding section the classical Young functions leading to exponential Orlicz spaces. These two facts motivate the following functional analysis study. Thus Λ must be convex but not necessarily symmetric and not bounded, i.e., $\Lambda(tx) \to \infty$ as $t \to \infty$ for each $0 \neq x \in \mathcal{X}$. Simple examples show that even when $\mathcal{X} = \mathbb{R}^2$ the last condition need not hold for a convex Λ, and has to be assumed. In this extension, even if $\Lambda(-x) = \Lambda(x)$, i.e., symmetric, its conjugate $\tilde{\Lambda}$ need not be. A first step in this direction may be formulated as follows.

Definition 1. Let $(\mathcal{X}, \|\cdot\|)$ be a Banach space, $\Lambda : \mathcal{X} \to \mathbb{R}^+$, a convex function such that $\Lambda(0) = 0, \Lambda(-x) = \Lambda(x)$ and $\{x : \Lambda(tx) < \infty$, for some $t > 0\} = \mathcal{X}$. Then the *Fenchel-Orlicz space* on a measure space (Ω, Σ, μ) denoted $L^\Lambda(\mu, \mathcal{X})$ is the class of strongly measurable $f : \Omega \to \mathcal{X}$ such that $\int_\Omega \Lambda(kf)\, d\mu < \infty$ for some $k > 0$ (Bochner integral). The (gauge) norm, with which equivalent classes are identified, is given by

$$\|f\|_{(\Lambda)} = \inf\{k > 0 : \int_\Omega \Lambda(\frac{f}{k})\, d\mu \leq 1\}. \tag{44}$$

The basic statement about the space can be presented as:

Proposition 2. *The normed linear space $(L^\Lambda(\mu, \mathcal{X}), \|\cdot\|_{(\Lambda)})$ is a Banach space if either $\mathcal{X}$ is finite dimensional, the complementary (or conjugate) function $\tilde{\Lambda}$ is continuous at 0, or Λ satisfies the equivalent condition that $\Lambda(x) \geq \alpha\|x\| + \beta$ for some $\alpha, \beta > 0$ and all $x \in \mathcal{X} - \{0\}$.*

A proof of this result, and several of the properties discussed below can be obtained from the work of Turett (1980). After stating some geometrical aspects of $L^\Lambda(\mu, \mathcal{X})$, the problem of interest in stochastic analysis will be highlighted.

The standard growth condition used for Young function is also meaningful for the general case. Thus Λ is Δ_2-regular if there are constants $K_i > 0, i = 1, 2,$ such that $\Lambda(2x) \leq K_1\Lambda(x)$ for $\|x\| \geq K_2$. A similar condition is meaningful for the conjugate function $\tilde{\Lambda}$ of Λ, as in the classical case. With these concepts the following exemplifies some geometric facts of these spaces.

Theorem 3. *Let $\mathcal{X}$ be a Banach space and $\Lambda : \mathcal{X} \to \bar{\mathbb{R}}^+$ be a convex function, as in Definition 1, and $L^\Lambda(\mu, \mathcal{X})$ be the corresponding Fenchel-Orlicz space. Then $(L^\Lambda(\mu, \mathcal{X}), \|\cdot\|_{(\Lambda)})$ is reflexive whenever $\mathcal{X}$ is reflexive and both $(\Lambda, \tilde{\Lambda})$ satisfy a Δ_2-condition.*

For these spaces, the geometric structure as well as other related properties have been detailed by Turett (1980). The analysis of the corresponding spaces if Λ is not symmetric, as desired by the work of the preceding two sections, is not yet available but desirable.

This will also be of interest even in the classical convex analysis, as discussed, for instance, by Brøndsted (1964), and in fact much of the work in Rockafellar's (1970) book, to have extensions to the general case with $\mathcal{X}$ infinite dimensional. This is at present largely unexplored.

In the next section, another set of spaces that are needed for an analysis of the sample paths of solutions of stochastic differential equations and related perturbation theory will be considered.

5. Function spaces for stochastic differential equations

Suppose that a process $\{X_t, t \geq 0\}$ is a solution of a first order stochastic differential equation

$$dX_t = \varepsilon\sigma(X_t, t)dB_t + b_\varepsilon(X_t, t)dt, \quad t \geq 0 \tag{45}$$

where $b_\varepsilon : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$, called the 'drift' and $\sigma : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^+$, the 'diffusion' coefficients, and $\varepsilon > 0$ is a parameter, B_t being the BM-process. If b_ε is independent of ε, and then $\varepsilon = 0$ making the first term disappear, one has an ordinary differential equation. In the general case when $\varepsilon > 0$ is fixed, and b_ε, σ satisfy a standard Lipschitz condition, and a given initial value $X_0^\varepsilon = x$ (or $X_0 = x$ in the nonstochastic case) the Itô theory implies that there is a unique solution of (45). [The matter is discussed even for higher order equations in Rao (1997).] A solution of (45) is called a diffusion process. The problem is to find conditions on b_ε and σ and the range space in order that the solution of the perturbed equation tends to the unperturbed one, and find conditions that the deviations decrease exponentially. Thus the problem becomes a continuous parameter analog of the large deviation questions considered in the earlier sections. For simplicity b_ε, σ will be assumed to be defined just on $\mathbb{R}^d$ so that the diffusion process will have stationary transitions, but the extension to the general case is then also possible and useful, as discussed in the last reference.

Since the solution should be "smooth", one has to find a proper space in which the process takes its values. Naturally these are Orlicz type spaces whose elements are "smooth". The appropriate spaces here are the Besov-Orlicz and Orlicz-Sobolev spaces, the latter being more suitable for the stochastic partial differential equations. First a brief recall of these spaces is given to make the discription intelligible, and then to state the results for the solution of (45).

Consider the probability space (Ω, Σ, P) where $\Omega = [0,1], \Sigma =$ Borel σ-algebra. For $f \in L^{\Lambda_p}(P), \Lambda_p(x) = e^{|x|^p} - 1, p \geq 1$, the modulous of continuity is given as:

$$\omega_{\Lambda_p}(f,t) = \sup_{0 \leq h \leq t} \|\Delta_h f\|_{(\Lambda_p)}, \ 0 \leq t \leq 1, \tag{46}$$

the norm $\|\cdot\|_{(\Lambda_p)}$ was defined earlier in Section 3, and $(\Delta_h f)(x) = f(x+h) - f(x)$ for $0 \leq x \leq 1-h$. This is scaled by $\omega_{\alpha,\beta} : t \mapsto t^\alpha(\log \frac{1}{t})^\beta, 0 < t < 1$, and set

$$\|f\|_{(\Lambda_p),\omega_{\alpha,\beta},\infty} = \|f\|_{(\Lambda_p)} + \|\frac{\omega_{\Lambda_p}(f,\cdot)}{\omega_{\alpha,\beta}}\|_\infty. \tag{47}$$

Then (47) defines a norm and the subset of the exponential Orlicz space $L^{\Lambda_p}(P)$ defined by

$$\mathcal{B} = B^{\Lambda_p}_{\omega_{\alpha,\beta},\infty} = \{f \in L^{\Lambda_p}(P) : \|f\|_{(\Lambda_p),\omega_{\alpha,\beta},\infty} < \infty\}, \tag{48}$$

is the desired vector space, and it can be verified that $\mathcal{B}$ is a Banach space, called the *Besov-Orlicz space.* The diffusion process, a solution of (45), can be shown to take values actually in the separable subspace $\mathcal{B}^0$ of $\mathcal{B}$, determined by the Schauder functions in it, and it can be defined precisely.

Moreover, the early work of Schilder (1966) has already shown that the BM process verifies the Large Deviation Principle (LDP), and this property is reflected for diffusion processes, the solutions of equations such as (47) driven by the BM with coefficients satisfying the usual Lipschitz conditions. Thus the conjugate (or complementary) function of Λ_p is definable again on subspaces such as $\mathcal{X}_0$ given in (26) of Section 2 above. In the present context, it is the space:

$$\mathcal{Y}_0 = \{h \in C_0([0,1]) : \int_0^1 |h'(t)|^2\, dt < \infty\}$$

relative to which the rate function can be defined. This is given by the following:

Theorem 1. *Let X_t^x and $X_t^{\varepsilon,x}$, $t \in [0,1]$ be the (unique) solutions of (45) in the original and perturbed cases with the initial condition $X_0^x = x = X_0^{\varepsilon,x}$ where X_t denotes the solution with $\varepsilon = 1, b_1 = b$, and the Lipschitz conditions are assumed. Then $X_t^x, X_t^{\varepsilon,x}$ take values in the separable subspace $\mathcal{B}^0$ of the Besov-Orlicz space $\mathcal{B}$ given by (48), and $X_t^{\varepsilon,x}$ satisfies the LDP, i.e., for each Borel set $A \subset \mathcal{B}^0$ one has (with the process in canonical form so that $X_t(\omega) = \omega(t), t \in [0,1]$):*

$$-I(A^o) \leq \liminf_{\varepsilon \searrow 0} \varepsilon^2 \log P_x[X_{(\cdot)}^{\varepsilon,x} \in A] \leq \limsup_{\varepsilon \searrow 0} \varepsilon^2 \log P_x[X_{(\cdot)}^{\varepsilon,x} \in A] \leq -I(\bar{A}), \tag{49}$$

where $I : \mathcal{B} \to \bar{\mathbb{R}}^+$ is the rate (or conjugate of Λ_2) function, written $I(A) = \inf\{I(f) : f \in A\}$, $A^o, \bar{A}$ are the interior and closure of A, with

$$I(f) = \inf\{\frac{1}{2}\int_0^1 |h'(t)|^2\, dt : f \in \mathcal{Y}_0, s(h)(x) = f(x)\}$$

and $s : \mathcal{Y}_0 \to \mathcal{B}^0$ is a mapping such that $s(h)$ is continuous on balls of $\mathcal{B}^0$, and satisfies the ordinary differential equation, associated with (also called a 'skeleton' of) (45) as:

$$\frac{ds(h)}{dt}(t) = \sigma(s(h)(t)h'(t) + b(s(h))(t), \quad s(h)(0) = x. \tag{50}$$

The details of proof, as one can expect, involve several estimates that are tedious but somewhat standard in this area, and are given by Eddahbi and Ouknine (1997). A two parameter extension (i.e., for fields $X_t, t = (t_1, t_2) \in [0,1] \times [0,1]$) is also available with an exact extension, but involves more work. It is recently obtained by Boufoussi, Eddahbi and N'zi (2000). The role of Besov-Orlicz spaces and exponential Orlicz spaces here need no further emphasis, being crucial for the work.

Let us also note briefly the spaces that appear in the study of stochastic flows, which are continuous function spaces but are of a different type. Here nonlinear operations appear from the start, and the analysis is localized. A brief discussion will now be included.

Consider the mappings $\phi_{st} : \mathbb{R}^d \times \Omega \to B(\mathbb{R}^d)$, the space of $d \times d$ matrices on $\mathbb{R}$ and suppose the following three conditions hold on (Ω, Σ, P) for $0 \leq s \leq t \leq u < \infty$:

about the book . . .

An expansion of the recent American Mathematical Society Special Session celebrating M. M. Rao's distinguished career, this extraordinary compilation includes most of the presented papers as well as ancillary contributions from session invitees. This book decisively shows the effectiveness of abstract analysis for solving fundamental problems of stochastic theory—specifically the use of functional analytic methods for elucidating stochastic processes, as made manifest in M. M. Rao's prolific research achievements.

Features a biography of M. M. Rao, a complete bibliography of his published works, and meditations from former students.

Over 30 notable researchers imbue ***Stochastic Processes and Functional Analysis*** with new insight into generalized distributions...nonlinear operators and doubly stochastic operators...Orlicz, weighted Sobolev, and reproducing kernel Hilbert spaces...stochastic integration...the monomer–dimer problem...birth–death processes...geographically-uniform coevolution...optimal asset allocation...van der Pol oscillators...asymptotically stationary processes and cyclostationarity...nonlinear filtering...harmonizable processes and fields...pseudoergodicity...information channels...approximations of scale mixtures...and numerous other active areas of research.

about the editors . . .

Alan C. Krinik is Professor and Associate Chair of the Department of Mathematics and Statistics at California State Polytechnic University, Pomona. His general research interest is applied probability, specializing in queueing theory, an area in which he delights in coauthoring articles with colleagues and graduate students. Dr. Krinik received the B.A. and M.A. degrees from the University of California, Los Angeles, and the Ph.D. degree from the University of California, Riverside.

Randall J. Swift is Associate Professor of Mathematics and Statistics at California State Polytechnic University, Pomona. Dr. Swift's research interests are stochastic processes and probability theory, differential equations, and mathematical modeling. He is the author or coauthor of three mathematical texts and more than 40 journal articles. He received the B.S. degree from California State Polytechnic University, Pomona, the M.A. degree from the University of California, Santa Barbara, and the Ph.D. degree from the University of California, Riverside.

Printed in the United States of America

Marcel Dekker, Inc.
New York • Basel

ISBN 0-8247-5404-2

90000

9 780824 754044