Interview with Abel Laureate 2020 Gregory Margulis

Bjørn Ian Dundas (University of Bergen. Norway) and Christian Skau (Norwegian University of Science and Technology, Trondheim, Norway)

Professor Margulis, first and foremost we would like to congratulate you on being awarded the Abel Prize 2020, together with Professor Furstenberg, for your, and we quote the Abel Committee, "pioneering the use of methods from probability and dynamics in group theory, number theory and combinatorics."
It's a great honour.

We will come back to the background story in a moment, but just to place your mathematics: you received the Wolf Prize mostly for your contributions to algebra. Your NSF grants emphasise analysis, and the Abel Prize focuses on probability and dynamics. So where should we place you? What sort of problems attract you?
I mostly consider myself to be a geometer. Once I talked to Jacques Tits, and he said "I am a geometer and you are a geometer". Somehow, I am a geometer in the sense that my mathematical thinking is mostly based on imagination and intuition. There are different types of geometers.

Let's see if we can get back to geometry as we go along. Jacques Tits, of course, is a previous winner – in 2008 actually – of the Abel Prize. You were born in Moscow in 1946. When did your interest in mathematics begin?
Probably from an early age. My father was also a mathematician, but he was mostly interested in mathematical education and didactics, and he wrote his PhD, or candidate thesis, under the direction of Aleksandr Khinchin, who was a famous probabilist. But my father's dissertation was in didactics. He was, how can I put it, a candidate of pedagogical sciences.

You played chess, and you were included in the so-called mathematical circles run by the Moscow State University. What was life like for you as a kid in Moscow in the 1950s and 1960s?
It was okay. My family was relatively well off. My father worked at an educational institute, and his salary was relatively high. From the time I was seven or eight, this was in 1954, we had our own apartment, which was quite exceptional at that time. People usually lived in so-called communal apartments. So one family occupied a room, and several families lived in one apartment. My family actually encouraged me to do mathematics. My father probably realised that I had some mathematical talent. From an early age, I was able to multiply double-digit numbers.

So it was natural that you were included in these mathematical circles?
Included is not the right word. Somehow it was no competition. The mathematical circles probably started for me around seventh grade when I was 12 or 13 years old.

Was it hard work being in the circles?
No, the circles were okay. It was quite informal. I don't remember many details, but it was run by students, sometimes by graduate students, from Moscow University, and we discussed various problems. And we were encouraged to solve them. There was also supervision by some more senior mathematicians. At that time, the mathematical Olympiads were connected to these mathematical circles.

So it was expected that one participate in Olympiads?
Yes, it was. It was in Moscow, I think it was in the so-called City Olympiad. A little later it was for the entire country. Somehow, there was no competition to enter, but it was quite challenging for many.

We can imagine that you actually did quite well in these Olympiads, didn't you?
Yes.

And later, in 1962, at the age of 16, you participated in the International Mathematical Olympiad, where you won a silver medal.
Yes.

Let us talk about your early career as a research mathematician. You wrote your first mathematical paper before you were 20. And in 1968 you and your fellow student Kazhdan published a very influential paper. How did that play a role in your later life, for example, toward arithmeticity of lattices and so forth.
Actually, my first paper was written while I was attending the Dynkin seminar. It was about positive harmonic functions on nilpotent groups. Probably it has some influence even now. My joint paper with Kazhdan was about the existence of unipotent elements in non-uniform lattices, or what is also called the non-cocompact case. This was a conjecture by Selberg. But our paper was also directed towards a proof of the arithmeticity of non-uniform lattices in semisimple Lie groups. So, actually, this paper with Kazhdan was the starting point towards the proof of arithmeticity.

We will talk more about arithmeticity later. However, we will remark that the joint paper you wrote with Kazhdan caused quite some excitement, and Armand Borel talked about it in a Bourbaki seminar. You were in the Dynkin seminar when you wrote the paper on positive harmonic functions on nilpotent groups. However, your undergraduate and PhD advisor was Yakov Sinai, who, by the way, was the Abel Prize Recipient in 2014.
Yes, it was Sinai. I attended the Sinai seminar on dynamical systems. I did actually then publish two works which are related to dynamical systems. One of them was about Anosov systems on compact 3-manifolds and the conditions on the fundamental groups this entailed. The other was about the counting of closed geodesics on compact manifolds of negative curvature. And this work was done because I attended and was inspired by this seminar. Sinai became my advisor in the third year of my undergraduate studies, so around 1965. I finished my PhD thesis in 1970.

After you finished your PhD, you began to work at the Institute for Problems in Information Transmission. According to Dynkin, you had the luxury of spending most of your time there on your own research. Is that a fair assessment?
Yes, that's quite a fair assessment. I didn't get a position at the Moscow University or the Steklov Institute, but the Institute for Problems in Information Transmission was one of the institutes of the Soviet Academy of Sciences. By Soviet standards at that time, the Institute was relatively small, only 200 researchers, but by Western standards it would probably be considered a huge institution. My immediate boss was Roland Dobrushin, who was a famous probabilist and a mathematical physicist. There was also another group there, which was headed by Mark Pinsker. He was famous for his work on information theory. In a sense, I was lucky to be there, because my most well-known and widely cited paper on expander graphs, published in 1973, was written under Pinsker's influence. Expander graphs, which incidentally have many applications in computer science, were first defined by Pinsker. Their existence was first proved by Pinsker in the early 1970s. My paper gave the first explicit construction of an infinite family of expander graphs. The

work on expanders was in some sense done under pressure or from a sense of duty. But I was mostly working on discrete subgroups of Lie groups, on arithmeticity and super-rigidity, and so on. It was done in parallel. All my own independent research was not, to put it mildly, quite related to the main direction of the Institute.

We will get back in more detail to your work on lattices in Lie groups. But before that, if you were to rank the people in Moscow who influenced you the most when you embarked on your research career, who would you name?
Sinai, of course, and then Piatetski-Shapiro, Kazhdan and Vinberg.

Did Piatetski-Shapiro make you aware of Selberg's and his own conjectures about lattices in semisimple Lie groups? Did you get the problem from them?
Yes, in a sense. It started with Selberg, and then with Piatetski-Shapiro. Selberg only stated the conjecture for non-uniform lattices. For uniform lattices he didn't actually have the definition of an arithmetic subgroup. That was done by Piatetski-Shapiro. So this problem somehow circulated, mostly thanks to Piatetski-Shapiro. I first proved the arithmeticity for non-uniform lattices by following the strategy which is essentially due to Selberg and Piatetski-Shapiro. You start with the unipotent elements and do various pieces of quite intricate work. For some special cases, the arithmeticity of non-uniform lattices had been proved by Selberg. But in the general case, the proof is much more complicated and requires a lot of additional, non-trivial arguments. For me, it took quite a long time, something like two or three years, to write a detailed proof. Regarding the case of uniform lattices, there was no strategy before my work.

So it is fair to say that your strategies for proving the non-uniform case and the uniform case are vastly different?
Yes, they are vastly different. For the non-uniform case there are these unipotent elements, and there are some kinds of building blocks which allow you to get the structure of arithmetic groups. For uniform lattices there are no such building blocks. For non-uniform lattices the method of proof is of algebraic and geometric nature, but for uniform lattices you have to use transcendental methods.

Before we go on, could you explain to us what it means to say that a lattice is arithmetic?
First, in Lie group theory a lattice Γ is a discrete subgroup of a Lie group G with the property that the quotient space G/Γ has finite invariant measure. Or, as we say, Γ has finite covolume. A lattice Γ is uniform (or cocompact) if the quotient G/Γ is compact, and non-uniform (or non-cocompact) otherwise. Consider the group $\mathrm{SL}_n(\mathbb{R})$ of real invertible $n\times n$-matrices with determinant 1, and the subgroup $\mathrm{SL}_n(\mathbb{Z})$ of matrices with integral coefficients. This is the standard example of an arithmetic subgroup. It's a classical result that $\mathrm{SL}_n(\mathbb{Z})$ has finite covolume in $\mathrm{SL}_n(\mathbb{R})$. It probably goes back to Hermite and Minkowski.

Full studio at the Norwegian Academy for the interview which was conducted remotely due to the covid-19 pandemic, with Professor Margulis on line from Yale.

Armand Borel and Harish-Chandra generalised this to semisimple Lie groups: $G(\mathbb{Z})$ is a discrete subgroup which has finite covolume in $G(\mathbb{R})$, where G is a semisimple Lie group. I proved that, under certain conditions, any lattice in a semisimple (algebraic) Lie group G is arithmetic (precisely, the lattice must be irreducible and the real rank of the group must be greater than 1). However, one needs to extend the definition of arithmetic subgroups. One extension is that the subgroup Γ is commensurable with $G(\mathbb{Z})$, i.e. the intersection of Γ and $G(\mathbb{Z})$ has finite index in both Γ and $G(\mathbb{Z})$. Another extension, which was actually due to Piatetski-Shapiro, is that, vaguely speaking, there is some construction which comes from maybe a bigger group that maps onto the original group with compact kernel. Selberg probably did not know about this definition.

Can you give us the timeline of your proof of the arithmeticity of lattices in higher rank semisimple Lie groups?
For non-uniform lattices, the crucial step was an announcement in 1969. I wrote quite a long paper about this crucial step, which was finished in 1971, but because of some difficulties in getting it published, it did not appear before 1975. As for getting the arithmeticity result from this crucial step, that was finished in 1973.

As for uniform lattices, the initial inspiration came in 1969 or 1970, when I learned about Mostow's fundamental work on strong rigidity. Thinking about it, I realised at some point that it would be possible to prove the arithmeticity of uniform higher rank lattices if one could prove a statement which is now called superrigidity. I believe, and this was confirmed by Mostow, that superrigidity was a new phenomenon which had not been discovered before. The first proof of superrigidity was based on a combination of methods from ergodic theory and algebraic group theory. One of the important ingredients was Oseledec's multiplicative ergodic theorem. It involves methods which are actually quite far from the original formulation

of arithmeticity. I was invited to give an address at the ICM Congress in Vancouver in 1974, but I was prevented from attending. Instead, I sent a report, later published in the Proceedings of that Congress, where I outlined a proof of arithmeticity in the uniform case.

At the next ICM Congress in Helsinki in 1978 you received the Fields Medal, but you were not allowed to attend, mostly due to the opposition of the top Soviet mathematical establishment at that time. Jacques Tits said in his citation of your work, and we quote: "Margulis has completely, or almost completely, solved a number of important problems in the theory of discrete subgroups of Lie groups, problems whose roots lie deep in the past and whose relevance goes far beyond that theory itself. It is not exaggerated to say that, on several occasions, he has bewildered the experts by solving questions which appeared to be completely out of reach at the time. He managed that through his mastery of a great variety of techniques used with extraordinary resources of skill and ingenuity."

I am perhaps not the right person to comment on that. However, Dennis Sullivan told me that when Jacques Tits gave a presentation at Collège de France – or maybe it was at IHES at Bures – of my proof of the arithmeticity of cocompact lattices, Armand Borel was extremely surprised that ergodic theory was a crucial ingredient in the proof. After all, the theorem was stated in arithmetic terms.

Could you comment on a later and quite different proof of the superrigidity theorem and its application to arithmeticity – both in the uniform and non-uniform case – using the work on boundaries by Furstenberg, with whom you share the Abel Prize?

It seems strange now, but when I worked on superrigidity I was not influenced by Furstenberg's work, because I was essentially not familiar with it. It is indeed strange, because many ideas and methods introduced by Furstenberg are very similar in style to what I used.

As I mentioned before, my proof of the uniform case is vastly different from my proof in the non-uniform case. For the uniform case, there are actually two parts in the proof. The first part is to prove the existence of equivariant measurable maps and the second part is to show the rationality of these equivariant measurable maps. I did the first part for the uniform case, and using certain integrability estimates my argument could be extended to the non-uniform case. Actually, for that case I had to use arithmeticity, or at least the crucial statement in my proof of arithmeticity, to obtain these estimates. For the existence of equivariant measurable maps, Furstenberg gave a different proof that is not based on the multiplicative ergodic theorem, but is based on his boundary theory, which is, in a sense, quite related to the multiplicative ergodic theorem. One of the origins of the multiplicative ergodic theory of Oseledec was previous work by Furstenberg and Kesten on products of random matrices dating back to 1960.

Incidentally, let me relate a story which has some bearing on your question. Around 1970, Furstenberg visited Yale. At that time, Mostow was working on strong rigidity, and Furstenberg was working on applications of boundary theory to the theory of discrete subgroups of Lie groups. Furstenberg and Mostow were good friends, but somehow at that time they did not pay much attention to each other's work. In retrospect it looks strange, because their works looked very closely related. Mostow told me later on several occasions that Furstenberg probably could have proved superrigidity if he had been aware of the problem. I mentioned that in my talk during one of the workshops in honour of Furstenberg. Furstenberg was present and he immediately said something like: "I would never have been able to do that".

You are on record saying that you consider the proof of the so-called normal subgroup theorem as your best proof. Could you tell us what the normal subgroup theorem says, and also tell us why you think this proof is so good?

The normal subgroup theorem says that if G is a connected semisimple Lie group of rank at least 2 with no compact factors and with finite centre, and if Γ is an irreducible lattice in G, then any normal subgroup N of Γ either belongs to the centre of Γ or has finite index in Γ. So a special case is, for example, $\mathrm{SL}_3(\mathbb{R})$ of 3×3 matrices of determinant 1. Take a discrete subgroup Γ which has finite covolume. Then any normal subgroup of the lattice Γ is either central or has finite index in Γ.

The general proof is divided into two parts. You have this irreducible lattice Γ and a normal subgroup N. Consider the quotient Γ/N. First consider the case that Γ/N is an amenable group. This case can be treated using representation theory arguments, in particular Kazhdan's property T. Then consider the case when Γ/N is a non-amenable group. Somehow I realised that one can use algebras of measurable sets. A crucial tool was one of the initial lemmas occurring in Furstenberg's paper on his boundary theory. I cannot explain how I came upon the idea behind the second part of the proof – it was some sort of intuition. Also, the idea of subdividing the proof into two parts was quite new. Anyway, I consider it the best proof I have done because it is mostly based on intuition.

Were there any precursors, or did the statement of the normal subgroup theorem come out of the blue?

The statement of the normal subgroup theorem was known to be true in certain cases, for example for $\mathrm{SL}_n(\mathbb{Z})$. The proofs were done by algebraic methods. Maybe it was natural to assume that the statement was true in general. However, the proof in the cocompact – or uniform – case was obtained partially by using measure theory, as I alluded to above. For example, one of the ingredients in the proof is the density point theorem in measure theory. So even though the theorem is stated in purely algebraic terms, the proof in the general case is mostly non-algebraic.

Another example where you prove results in algebra and Diophantine approximation by your sort of methods, which, to us at least, seem very surprising, is when you prove the Oppenheim conjecture. Could you explain to us what that is, and how we should think about it?
The Oppenheim conjecture is actually a quite natural conjecture. There is a classical theorem called Meyer's theorem which says that if you have an indefinite rational form Q in at least five variables, then it nontrivially represents zero over rational numbers. That is, you have a nonzero integral vector x such that $Q(x)=0$.

Oppenheim – he was British and a student of Dickson's in Chicago – worked on these rational forms in four variables. He published a paper in the Proceedings of the National Academy of Sciences in 1929, and there was a footnote where he formulated the conjecture. It can be considered as an analogue of Meyer's theorem for irrational forms: if you have an indefinite irrational form Q in at least five variables, then for every positive ϵ there exists an integral vector x, such that the absolute value of $Q(x)$ is less than ϵ.

So the image of the set of all integral vectors is dense in $\mathbb{R}$, is that what you say?
More or less. It was later realised that the conjecture could be strengthened. There was a lot of work on using analytic number theory methods. I think that the main progress had been done by Davenport and his coauthors, starting from 1946 up until 1959, where they proved this when the number of variables is at least 21. But it was mostly analytic number theory methods.

Davenport realised that the conjecture could be stated not just for dimension at least 5; for irrational forms it could be stated for dimension at least 3. For dimension 3 and 4 the statement of Meyer's theorem is not true for rational forms, but for its analog for irrational forms it *is* true. Later, in the mid 70s, Raghunathan realised that the Oppenheim conjecture can be reformulated as a statement in dynamical systems. Let's say G is $\mathrm{SL}_3(\mathbb{R})$ and Γ is $\mathrm{SL}_3(\mathbb{Z})$, and if you take H to be SO(2,1), then any bounded orbit of H in G/Γ should be closed. Actually, the Oppenheim conjecture can be stated for a very special case: if you take the form $Q(x,y,z)=x^2+y^2-\sqrt{2}\,z^2$, then $Q(\mathbb{Z}^3)$ should be dense in $\mathbb{R}$. And the proof of this special case is not any simpler than the general case.

You learned about this problem in Bonn, as we understand?
Yes, I visited Bonn for three months in '79 from the beginning of July. I met Gopal Prasad, who was a student of Raghunathan's, and he is now a well-known mathematician.

The proof was published in 1986, is that right?
Yes, the proof was published in '86. But I remember that I already gave some kind of oral presentation in '84.

Let's move on to what we briefly touched upon before, namely expander graphs. You said that it was Pinsker who introduced you to this topic. Could you be more specific? But first tell us what an expander graph is!

The Abel Prize Laureate 2020 Gregory Margulis. © Dan Rezetti.

Intuitively, an expander graph is a finite, undirected graph in which every subset of the vertices that is not "too large" has a "large" boundary. This notion was first introduced by Mark Pinsker in his work on so-called concentrators. Pinsker's original definition was not the same as the standard definition one can find in textbooks today. Vaguely speaking, a regular and undirected finite graph with n vertices is an expander graph if for any subset A of m vertices, where m is less than $n/2$, A has $m(1+\epsilon)$ neighbours for some (small) ϵ. Actually, the definition is not just for one graph, but for an infinite family of graphs. That the graph is d-regular means that each vertex has exactly d neighbours.

As I said earlier, this started with the work of Pinsker, who proved the existence of expander graphs by probabilistic methods. In fact, almost all graphs which come from his construction are expander graphs. But there were no explicit constructions. I realised that by using some group theory, especially involving property (T), I could explicitly construct an infinite family of expander graphs. This was probably unexpected for people working in computer science.

Later on you constructed even more examples of expander graphs, isn't that correct?
Yes, later on there were these graphs that came from using quaternions. This was in 1984 and at that time I was mostly interested in studying the girths of regular graphs, and finding upper estimates of the girth size. In graph theory, the girth of a graph is the length of the shortest cycle contained in the graph. There was some probabilistic construction due to Erdős and Sachs that gave an upper asymptotic estimate $2\log_p n$ for the girth of a $(p+1)$-regular graph with n vertices (this is simple), while the asymptotic lower estimate was $\log_p n$. Quite surprisingly, my explicit construction gave an asymptotic lower estimate $4/3\log_p n$. I believe that up until now there hasn't been any probabilistic construction which goes beyond $\log_p n$.

At the same time as I studied these explicitly constructed graphs I realised, based on some deep work by Deligne, that they were also expander graphs. Slightly later and completely independently, Lubotzky, Phillips and Sarnak basically gave the same construction, but with some variations. They also used the work of Deligne, and they called these graphs which come from this construction Ramanujan graphs, because it is related to some Ramanujan conjecture.

We have to talk about another problem that you solved, this was at the very beginning of the 80s, and we quote Mathematical Reviews: "Although it is not explicitly stated, this paper settles a long-standing problem of Banach on the uniqueness of invariant means on the n-sphere". We understand that this came to you in a flash, and it involves property (T) again. Could you tell us a wee bit about the story behind this?
I believe I was at a conference in Poland in 1980, in May or June. I met Rindler, an Austrian mathematician, who mentioned this problem by Banach and Ruziewicz about invariant means for algebras for measurable sets and the reformulation of the problem due to Rosenblatt, who explained that it would follow from the statement about small almost invariant sets. I immediately realised that this statement can be deduced from property (T) for certain subgroups.

Okay, if you start in dimension 5, say SO(5), then it contains certain arithmetic lattices, or *S*-arithmetic lattices, and it has property (T). I think – essentially around the same time – Dennis Sullivan gave another proof using a slightly different subgroup, but also using work by Rosenblatt.

Actually, the proof came to me almost immediately.

Is this typical, does this happen to you from time to time that all of a sudden you see the answer?
Maybe I had an answer before I had a question…

Fair enough.
This work became quite famous, but actually I didn't spend much time on it.

Can we ask you about your working style, because there is one thing that strikes us. Early in your career, you were the single author of nearly all of your papers, and then in the past thirty years, almost all of them have been joint papers. How do you explain that?
When I was in Moscow, essentially all my papers were written by myself. One notable exception was the joint paper with Kazhdan. Actually, for me it was quite challenging to write up papers, so it took a lot of time. But when I moved to the United States and to Yale in 1991, it was a completely different environment. I started to work with many mathematicians, mostly younger than me. I also started to work with my graduate students.

Was that fun, did you enjoy it?
Yes, working with graduate students was fun. I did not have graduate students in Moscow, but after I moved to Yale I had several graduate students. It was a rewarding experience to work with them.

We noticed that even last year, in 2019, you had a paper together with A. Mohammadi. So you are still working with younger fellows?
A. Mohammadi was my graduate student, who finished more than ten years ago. I had joint papers with him when he was a graduate student, and then we had several joint papers after that. Another graduate student of mine that I would like to mention is Dmitry Kleinbock. We have several joint papers. In 1995 I noticed a book in the maths department library by Sprindzuk, where he presents his proof of the Mahler Conjecture on transcendental numbers and Diophantine approximations. I asked Dmitry to look at it and see if there are relations between the subject of Sprindzuk's book and dynamics. Soon after, Dmitry gave a reformulation of the Baker–Sprindzuk conjectures in dynamical terms. After that we realised that a modification of methods used in the proof of non-divergency for unipotent flows can be applied to prove the dynamical reformulation of the Baker–Sprindzuk conjectures.

Now I am retired, and I had my last student finish this current year.

A final question: How do you rank yourself on a scale from theory builder to problem solver?
Probably I am more a problem solver than a theory builder. But I find this division rather artificial.

That brings us to the end of the interview. We want to thank you on behalf of the Norwegian Mathematical Society and the European Mathematical Society. Also, the two of us would like to thank you personally for this very interesting interview. Thank you very much!
Thank you!

Bjørn Ian Dundas is a professor of mathematics at the University of Bergen. His research interests are within algebraic K-theory, homotopy type theory and algebraic topology.

Christian Skau is a professor emeritus of mathematics at the Norwegian University of Science and Technology (NTNU) at Trondheim. His research interests are within C-algebras and their interplay with symbolic dynamical systems. He is also keenly interested in Abel's mathematical works, having published several papers on this subject.*