Postdoc Paper Talk

Welcome to the Postdoc Paper Talk of the Cluster of Excellence Mathematics Münster. In this series of talks, postdocs from the Cluster of Excellence talk about their research. We would like to give an insight into the research of our more than 150 early career scientists.

In this interview from May 2024, we will learn about the research of Dr. Markus Tempelmayr. Markus studied Applied Mathematics at the TU Vienna then did his doctorate at International Max Planck Research School Leipzig and since 2023 has been working as a postdoc in Professor Hendrik Weber's workgroup for Stochastic Analysis in Münster.

Professor Christopher Deninger, professor of Arithmetic Geometry and Representation Theory at the University of Münster, talks with Markus about this paper recently published in Inventiones mathematicae.

P. Linares, F. Otto, M. Tempelmayr, P. Tsatsoulis (2024): A diagram-free approach to the stochastic estimates in regularity structures. Inventiones mathematicae. 2024
DOI: `https://doi.org/10.1007/s00222-024-01275-z`

Listen to the audio podcast online at `https://uni.ms/3sr50`

Introduction to Tempelmayr's paper

Christopher Deninger: I have heard of some important paper by you and I am curious what you are doing. I had a look at the paper and I did not understand it. So what is it all about?

Markus Tempelmayr: Maybe we can start with a little bit of context about the rough field we are in: What we consider is what is called "stochastic partial differential equations" - those equations that are supposed to describe some physical systems, for example how temperature evolves over space and time. And so "stochastic" in this context means that there are some random terms in this equation that typically describe fluctuations. They are a modelling assumption because, for example, we are not able to exactly prescribe a physical system and detect every molecule in which direction –

CD: – that's clear. We always have some noise.

MT: As you say, well there is some background noise in your system. And what we typically consider are then systems on a scale where exactly these fluctuations may have an impact on the effective behaviour. For example, I guess that's called a mesoscopic system. We are not interested in one litre of water and also not in a single particle but rather a very thin liquid film where these fluctuations that arise from the thermal noise may have an impact on the effective behaviour. That's a typical situation where such stochastic partial differential equations come up.

CD: Is this a general question in the field? If you see the effect of the randomness?

MT: That's definitely one big question. How does this randomness affect the behaviour?

CD: I know ordinary stochastic differential equations, the Itô calculus, and for example, if it's driven by Brownian motion, would that be an example of what you are saying?

MT: Yes. That's exactly what we're looking at, but for systems that do not only depend on time but also on space.

CD: Is this still the mesoscopic range?

MT: Yeah, exactly.

CD: What would be a range which is less than this where it doesn't affect it directly?

MT: I guess if you consider a single particle and you want to track that then this equation is probably not the correct framework. Then you would have to go to a quantum mechanical system or so.

CD: Okay, so now we have partial differential equations.

MT: Yes.

(Introduction to Tempelmayr's paper: Analogies 00:03:00)

CD: I actually read something about if you generalise Brownian motion to two dimensions, they coined the term "Wiener sausage". Isn't that funny?

MT: I haven't heard that. We also consider some higher dimensional analogue of Brownian motion or derivatives of Brownian motion, and this is what we call "white noise" or "spacetime white noise". And that is exactly this "independent, in all directions, at every point" randomness that you inject in your system. Such equations can also come up by describing some fluctuations of particle systems that can model magnetisation of materials.

CD: Is it true that the injection of randomness into some classical equation leads to more stability?

MT: Yeah, but that's not exactly what we are looking for. But that's also not far away from our field. Having some randomness, can actually improve the stability properties of your system. If you try to put a pencil on the tip that's a highly unstable phenomenon. But in a deterministic equation, the equation would tell you, you could observe this pencil in this position. Physically, you will never do that. If you insert randomness, then this position is not a stable point anymore.

CD: If you want a tidy apartment, it is necessary to have one box where you put all the random stuff, that doesn't have a good place. Adding a little bit of randomness keeps the apartment more tidy.

MT: Yes. That's a good analogy.

CD: Or these examples of people marching over a bridge, if they go in a certain rhythm, you may get these resonances and then the bridge may collapse. Okay anyhow. But this is not what you're doing.

From the physical idea to the mathematical approach and the problem of ill-defined non-linearities 00:05:19

MT: Exactly. Now, we would like to study such equations. And typically, the first question that you ask is: Ok, a physicist derives some equation and tells you this is the correct equation to look at. As a mathematician we would then like to make this a bit rigorous and say, in what sense can you solve this equation? Do solutions exist? Are they unique? And once we know there exist solutions, we would like to study their qualitative properties. What is the long-time behaviour? Does the theoretical long-time behaviour we then compute actually coincide with some experiments or with the model the physicists try to describe? And so, the problem in the equations we look at is already "what does it mean to be a solution of this equation?". And the problem is the following: This randomness that you add to the equation, is typically very irregular. As maybe the example of Brownian motion already suggests, the Brownian motion is a continuous curve, but it looks quite rough.

CD: It is nowhere differentiable.

MT: Exactly. It's not differentiable. But what you write in these equations is basically the derivative of Brownian motion. One can take derivatives of non-differentiable functions. The result is not a function anymore, but a distribution. And the theory of distributions is quite rich and general, and one can do a lot of things with these distributions. For example, one can take derivatives as often as one wants and in general one can do linear operations with them. The problem is now if the equation is non-linear then at some point you would have to do a non-linear operation with your distribution, and this is forbidden. At least it's mathematically ill-defined. You try to solve an equation that just doesn't have a sense.

CD: You can perform certain manipulations. Non-linear distributions, multiplications and so on, but not in your cases?

MT: Exactly. The distributions we would like to square are not allowed to be squared. That has to do with the roughness of these stochastic terms that we put in and that these non-linearities can then create resonances. And if you try to simulate such equations, things blow up and you cannot, not even numerically, try to find solutions.

CD: You mean if your notion of solution is not the correct one? Then any numerical experiments will fail?

MT: Yeah, exactly.

CD: Even finding the correct notion of what a solution is?

MT: That's already the problem. What is the good notion of solution there? A typical answer is, you regularise your noise term, you make it somehow smoother by a certain approximation. Then the equation is well-defined, you can solve it by some classical theory. You get a solution to this regularised equation, and then you try to remove this regularisation parameter and track what happens with this sequence.

CD: You really make it smooth. Does that mean that it makes sense as an ordinary partial differential equation?

MT: Exactly. You can then treat it pathwise as a deterministic equation. You forget there is randomness. It's a forcing in your equation, which is maybe even smooth. It doesn't have to be C^∞, but so smooth that the non-linearity makes sense. Then you can solve your equation and you get a solution for every regularisation parameter. What you, at first, try to do is: What happens if you now slowly remove this parameter? What happens with your sequence of solutions? And the typical situation is either it converges to zero or to infinity. And both are not satisfying.

Martin Hairer's ansatz: Renormalisation in the context of stochastic PDEs 00:12:30

(Relevance of non-linearity; Invariance under rescaling / zooming in; Subcritical, Critical and Supercritical dimensions)

MT: Somehow the solution to this whole problem is what is called "renormalisation", which is around in physics since almost a hundred years now. Maybe not hundred, but 80 or so, and the big breakthrough in the field I'm working on came by Martin Hairer in around 2014, when he managed to apply similar ideas in this context of stochastic PDEs and develop a well-posedness theory for such equations.

CD: And how does he do it?

MT: The idea is that the equations that he can consider are equations, where if you try to zoom into your equation, meaning that you look on very small scales, the non-linearity drops out. This means that on small scales, the linear equation dominates the behaviour and only on larger scales the non-linearity matters.

CD: Can you write it down what it means to zoom in? In a simple example?

MT: Sure, let's write down a typical example. One example is

$$(\partial_t - \Delta)\varphi = -\varphi^3 + \xi, \tag{1}$$

where the left-hand side is called the heat operator and ξ is a noise, which due to my bad handwriting already looks like a noise. This function $\varphi(x,t)$ will be a function of space and time, so of a parameter $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$. And zooming in means you rescale your function. It basically means you take a parameter $\lambda > 0$ and you consider the rescaled

$$\lambda^{\frac{2-d}{2}} \varphi\left(\frac{t}{\lambda^2}, \frac{x}{\lambda}\right) =: \varphi^\lambda(t,x). \tag{2}$$

That means now if you tune the parameter either to zero or to infinity, that what you were looking at, say a graph of a function at scale one, gets zoomed out (you squeeze your graph) or zoomed in.

CD: And can we see this phenomenon here, that the non-linearity now becomes less important?

MT: Yes. You can now assume you have a solution φ to this equation. If you plug the rescaled object φ^λ in, you get again a solution,

$$(\partial_t - \Delta)\,\varphi^\lambda = -(\varphi^\lambda)^3 + \xi^\lambda, \tag{3}$$

where $\xi^\lambda(t,x) = \lambda^{-(d+2)/2}$ is somehow the same noise as ξ. The noise is somehow invariant under this scaling that I do here. That makes basically a noise a noise, that zooming in it looks exactly the same as before.

CD: But is there something missing here (in the equation)?

MT: Exactly, there is something missing now. And what is missing is that if you do that correctly, which probably now I can't do, should be something like λ^{d-4}, i.e.

$$(\partial_t - \Delta)\,\varphi^\lambda = -\lambda^{d-4}(\varphi^\lambda)^3 + \xi^\lambda. \tag{4}$$

You see if you do this correctly, that this parameter λ comes in front of the non-linearity with a certain exponent - and the exponent is $d-4$.

MT: So now under this rescaling the noise somehow looks the same. That's the same as for Brownian motion. Under a specific rescaling, it looks again exactly as a Brownian motion. And now you see here this exponent $d-4$ pops up. Now if you send $\lambda \to \infty$ and the dimension is less than four, the non-linear term drops out. This regime is called "subcritical", if it happens that on small scales the non-linear term drops out.

CD: Oh, and if $d = 4$, then it's critical?

MT: Yes, this is critical. And that's a special case where the equation is somehow invariant under this scaling. Or you could even go to supercritical dimensions, which in this case means $d > 4$. Then the non-linearity becomes more and more important on small scales. In this case nothing is known for this equations. Also not for the critical one. What we can study is the subcritical case.

CD: The case where you can make it less important. Of course, this kind of renormalisation is a very old idea.

MT: In physics it is, in mathematics, at least in this context of SPDEs (Stochastic partial differential equations) it's only since 2014, with one exception in the early 2000's, which we may come back to later.

CD: This is very funny. When I was a student, renormalising partial differential equations – yeah, it's true, I learned it in a physics course. I mean they did this and then they talked a lot about it, but I don't know actually why. I forgot.

MT: Yeah. In their context, it comes up in these Feynman diagrams when you have.

CD: We didn't have this. But I mean, to make these kind of substitutions and check for invariance, this is of course very old. I'm surprised that it was used here only recently.

Solving the problem of ill-defined non-linearities: Generalised Taylor expansion (still Hairer's ansatz) 00:15:50

MT: I mean, this is basically just the starting point that gives you a hint of what could be the correct idea. And the idea is now the following: The problem in this equation is exactly the small-scale behaviour, that there is some roughness, that the non-linear term doesn't make sense. But on the other hand, you see that on the small scales, the non-linear term is maybe not that important, because it goes anyway, it stands away. The idea is that if you have a good enough understanding of the linear equation, then you can pass from knowledge of the linear equation also to the non-linear equation. That is somewhat the basic idea.

CD: By scaling back or how?

MT: It's more like a generalisation of a Taylor expansion. You try to describe the solution to the non-linear problem as a linear combination of generalised monomials. In this case, you don't try to do it just with standard monomials, as a Taylor expansion would be, because monomials are smooth and would approximate very well smooth objects. But here we are in a rough situation. But if you allow yourself, for example, also to take the solution of the linear stochastic heat equation into this expansion, then you can hope that this thing approximates your solution, locally at least, very well. And the idea is that you locally, at every point, try to expand your solution and then glue together all your local expansions and build one global object, which is called in this context "reconstruction". For this reconstruction to work one needs a certain consistency condition. Because if I give you an arbitrary sequence of functions, it's not said that this family is the Taylor expansion of a certain function at every point. So, there must be some consistency condition and if that is satisfied, you can find a global object that at every point looks like your local description. That was the basic idea and now the question is: What are these generalised monomials that you have to use for your equation? This will heavily depend on your equation: For every equation you develop your own calculus and, so to say, your own notion of Taylor expansion with which you describe then your solution.

CD: Why do you call it Taylor? Couldn't you also say call it Fourier or whatever? Is it closer to Taylor?

MT: It is some sense.

CD: Because it's local?

MT: It is done in physical space. You really try to expand locally at every point, and one takes in this expansion all the polynomials into account, and you add a couple of more terms. It's somehow really a generalisation of that, yes.

CD: Oh okay, right. This is the general strategy of Hairer's program?

MT: This is exactly what he came up with, yes, and somehow that's now a big industry. A lot of people work on that and try to push it further in several different directions.

CD: For example, in this equation (1), what would be the additional functions?

MT: So that depends now on the dimension. In dimension two it's quite simple. There you take exactly the solution to the stochastic heat equation. So that's

$$(\partial_t - \Delta)v = \xi. \tag{5}$$

CD: The linear one, right?

MT: Exactly. That's this object v and you take its square and its cube, so you also take v^2 and v^3. And these three objects v, v^2, and v^3 are sufficient in dimension two. The problem is that this v^2 and v^3 is somehow still analytically ill-defined for the same reason as the equation before was ill-defined. The solution to this linear problem (5) is well defined. You can solve this equation, but the solution will still be a distribution. And so again you cannot say what is v^2 and v^3. But now you can use probabilistic tools to give a sense to v^2 and v^3 by doing again a modification: You replace ξ by a smooth version ξ_ε and you get a solution $v^{(\varepsilon)}$, which depends on ε:

$$(\partial_t - \Delta)\, v^{(\varepsilon)} = \xi_\varepsilon. \tag{6}$$

CD: Aha. So now you use this regularisation procedure, which fails in the non-linear case?

MT: The idea is to replace now $(v^{(\varepsilon)})^2$ by $(v^{(\varepsilon)})^2$ minus a certain constant. And the constant in this case is just its expectation:

$$:\left(v^{(\varepsilon)}\right)^2: \; = \left(v^{(\varepsilon)}\right)^2 - \mathbb{E}\left(v^{(\varepsilon)}\right)^2. \tag{7}$$

And the strange thing is that this expectation $\mathbb{E}(v^{(\varepsilon)})^2$ is roughly a logarithm - it behaves like $\log(\varepsilon)$. In particular, it diverges, as you said when I am thinking of $\varepsilon \to 0$. If you send epsilon to zero, then ξ_ε converges to ξ. And in that context, this constant that you subtract from $(v^{(\varepsilon)})^2$, it diverges. But the nice insight is that as a whole object, this thing, if you call it $: v^2 :$ (typically it's denoted with a colon and it's called Wick product) converges to a certain distribution. And this is exactly why it is called renormalisation: You subtract the divergent term to make this whole expression convergent. You do this for the square and for the cube. And with these three terms v, $: v^2 :$, and $: v^3 :$ you can then describe the solution to this equation, at least in two dimensions. The observation one has to make is that by changing this $(v^{(\varepsilon)})^2$ to $: (v^{(\varepsilon)})^2 :$, my local description that I made, even for positive ε, even in the setting where everything is smooth, is not the description anymore of the original equation we started with. Instead, it's now a description of a modified equation. It turns out that in this two dimensional case, the modified equation is

$$(\partial_t - \Delta)\,\varphi^{(\varepsilon)} = -(\varphi^{(\varepsilon)})^3 + \xi_\varepsilon - C_\varepsilon \varphi^{(\varepsilon)}. \tag{8}$$

So there appears an additional term $C_\varepsilon \varphi^{(\varepsilon)}$ in the equation, which in this case is called a mass counter term. It has some physical interpretation. And you just see in the equation for $\varepsilon > 0$, you get a divergent term, because C_ε blows up as $\varepsilon \to 0$. For fixed ε it is still a smooth equation, and you can consider its solution. And if you now pass to the limit $\varepsilon \to 0$, you will converge to something non-trivial. Neither to zero, nor to infinity, whereas before – if you didn't have this counter term, if you dropped it and you tried to pass to the limit – you will just converge to zero.

CD: I thought that by renormalisation, by looking at small scales, we were getting rid of the cube, the non-linearity? But now you again have this cube?

MT: Yes. This is rather a heuristic that tells you, which objects you have to consider, to see, why we consider v, v^2 and v^3, so you can do a certain iteration.

CD: But these guys $: v^2 :$ and $: v^3 :$, if you define them by (7), they have nothing to do with the cube in (3)?

MT: Yes exactly, at first they have nothing to do with it.

CD: Then probably I misunderstand what the $\varphi^{(\varepsilon)}$ is. How is it?

MT: Okay, I mean, I didn't explain. The equation for φ had a problem. What you try to do is take $w := \varphi - v$. Now, let's forget the ε for a second. You subtract the linear equation. And since you expect that on small scales anyway φ looks similar than v, you take the roughness away, it cancels somehow out. So now what you can do is, you plug this difference in the equation and see which equation does this difference satisfy. We can try to do that:

$$(\partial_t - \Delta)\, w = (\partial_t - \Delta)\,(\varphi - v) = -\varphi^3 + \xi - \xi = (*)\,.$$

$(\partial_t - \Delta)\,\varphi$ gave this $-\varphi^3 + \xi$ and $(\partial_t - \Delta)\,v$ just gives the ξ. You see already the ξ cancels. And now we try to re-express φ by w and v and continue the equation from above to get

$$(*) = -\,(w+v)^3 = -w^3 - 3wv^2 - 3w^2v - v^3.$$

One can now check the regularities of all these objects that come up. Somehow the surprising fact is that now this equation for w is well-defined. The only problematic terms in this expression is v^2 and v^3. The product w^3 is fine, wv^2 is fine, provided v^2 makes sense. And the same for the remaining terms. Now by modifying v^2 and v^3, by renormalising these two expressions, we get

$$(\partial_t - \Delta)\,w = -w^3 - 3w\!:v^2\!:\, - 3w^2v \,-\!:v^3\!:. \tag{9}$$

Here we replace v^2 and v^3 by the renormalised objects $:v^2:$ and $:v^3:$. Now, this w makes sense, and I just postulate my solution φ to be $w + v$ from this ansatz. This is now like a definition of a solution. I call this my solution.

CD: You call this your solution?

MT: Exactly. As long as you have some ε parameters flying around, you can now check which equation does this object solve. And it's exactly this equation (8). It's the original one, just slightly modified. And in the limit, as epsilon fades away, as the regularisation goes, there is no equation that this object satisfies. Because in the equation things explode. This $(\varphi^{(\varepsilon)})^3$ doesn't make sense in the limit, this C_ε doesn't make sense in the limit. Somehow it both creates some infinities, and they cancel each other exactly out so that you get a well defined limiting expression φ, but no equation it solves.

CD: This is a little bit like black magic, like a cooking book: Here is your recipe. These terms don't make sense, you replace them by regularized objects. And now our limits exist.

MT: Yes, that's somehow the philosophy.

Physical interpretation of the equations 00:28:30

CD: I mean, the physicists they do similar things, and then in the end, they check with experiments and find in quantum electrodynamics, it fits up to 16 places after the decimal. So that confirms that they were doing the right thing. But mathematically...

MT: Probably there are equations without a physical interpretation. But most of the equations we look at, they have a physical interpretation. And the counter term that one derives usually has a physical interpretation. And it even makes sense to have it there. It's dependent on the equation: For this equation we looked at, it has a relation to self-interactions in quantum field theory. Another very famous example is the KPZ equation, which models interface growth. What you do there is somehow that you look at a growing interface and you hold the frame, at which you look, fixed. What you will see is that your interface grows and grows and grows and diverges to infinity. And in some limit, you don't see anything. It's just infinity. What you instead have to do is, you have to move your frame with the same speed as your interface is growing. And this is somehow the interpretation of the counter term. So, there is typically an interpretation, why it is valid to do that.

CD: So, in retrospect you do this and then you get new terms, and the physicists tell you, "oh, they have to be there!"

CD: I mean, in the case of Feynman diagrams and Quantum electrodynamics there's this approach by Connes and Kreimer where they use Hopf algebras to make sense of these renormalisations. So, it has a meaning in those terms, what the physicists have been doing all the time, a mathematical meaning. Is there anything like this here?

MT: There is a very close connection, exactly. In two dimensions, we could do everything by hand and it's not so deep. In three dimensions it already gets much more complicated. One could still do it by hand and that's what Hairer actually did in his original paper. But then one can look at equations that are much closer to the criticality threshold – here it would be four dimensions. What one can do is, one can mimic fractional dimensions and approach the critical dimension. And what will happen is that more and more of these generalised monomials will enter. I think now here it was three terms. In three dimensions it's already maybe ten terms or so. And it approaches infinitely many. In the critical dimension it would be infinitely many. You need an automated machinery that keeps track of everything. For example in your renormalisation procedure, you cannot always read off by hand from all the terms, how you have to modify them to get well-defined limits. And that's where Hopf algebras enter the picture. They systematise exactly which terms you renormalise.

CD: They are relevant for you?

MT: Yes, they are. They come up in this business and I think what Hairer did was inspired by what physicists did with these Hopf algebras. It's the analogue that comes up for SPDEs.

CD: Maybe we can return to this in the office later, but maybe now focus a little bit on your work.

Back to Tempelmayr's paper: Alternative approach replacing diagrams by the Malliavin derivative 00:32:32

MT: Our work answers basically: Given an equation, what are these generalised monomials that you take? First of all, which ones do you take? Second, how do you modify them to make them converge? And do they converge? The question we ask and answer is: How do you modify these generalized monomials to make them converge?

CD: What type of stochastic partial differential equations are you looking at?

MT: We looked at a specific, quasi linear equation. That's a certain type of equation that's of interest for some applications. Before that, people already looked at semi-linear equations, which is slightly less non-linear.

CD: Do you also have these heat equations?

MT: Yes, typically we look at parabolic or elliptic equations. One could also drop the time derivative. But what pretty much changes the picture is if one looks at wave equations or Schrödinger equations. Then the techniques developed do not work so well. Or rather, other techniques work much better.

MT: There was much progress for semi-linear equations. And what we did is that we basically redid an argument of the convergence of these objects for quasi-linear equations. And now the proof we gave also applies for semi-linear equations and is very different from the proof that was earlier given. While the earlier proof was very much inspired from this physics point of view of diagrams and contractions and Hopf algebras, the point of view we take is much less combinatorial. So there do not appear any diagrams and it's a much more analytic approach and direct to the equation, I would say.

CD: These Hopf algebras, they are out again?

MT: They are still somewhere in the background, but not so crucial for us.

CD: But I mean they are a formalism. What do you replace it with?

MT: Basically, we replace it by taking derivatives with respect to the noise. Which in the end means, by the product rule (also called Leibniz rule). This also has some combinatorics in it. And that somehow for us replaces all the combinatorics of Hopf algebras, that comes up.

CD: Ah, by the combinatorics of binomial coefficients and by repeated applications?

MT: Yeah, exactly. Just, how you take the derivative of a product. Iteratively, which is much simpler.

CD: Yes. And the derivative with respect to the noise? Is this a new notion?

MT: This is not a new idea. It's typically called "Malliavin derivative".

CD: Oh terrible, I tried to learn them in calculus at one stage.

MT: Exactly. So that's quite old also.

CD: It was developed to give another probabilistic proof of Hörmander's theorem about hyper elliptic equations.

MT: Yes, so it's already old and has been used in many contexts, but not in this context to show the convergence of these monomials. Somehow the new idea is that you can use these tools in this context.

Application of the Malliavin derivative and spectral gap assumption 00:36:39

CD: And how do you do this?

MT: The philosophical idea is the following: What comes up in these equations are multi-linear expressions of the noise. Something like a square or a cube of the noise term. And what you do if you take a derivative with respect to the noise of such a multi-linear expression is that you get rid of one of the noise terms, it's like taking the derivative of x^2. You get rid of one of the instances of x, just one x remains. And this is good in this context, because the noise is what creates this irregularity and the troubles in your non-linearity. If you, in your non-linear term, get rid of one of the instances, then somehow your problem becomes more regular, which is the opposite of what one usually has in mind if one takes derivatives. If your function is $\mathcal{C}^2$, then its derivative is only $\mathcal{C}^1$, it's less regular. Now in this context, because we remove some irregular noise, things get more regular actually. And you can use this in a systematic way to show convergence of these objects. Should we go in more details?

CD: Yes. It sounds like you take the stochastic differential equations, you differentiate with respect to the noise, and you turn this into an ordinary differential equation.

MT: There is still some instances of the noise typically left because you do not just have linear expressions in the noise in your equation. For linear expressions if you take a derivative, you get rid of all the randomness and all the noise. But we always have this multi-linear expressions, like for example, the square. So, if you take a derivative, there is still some noise somewhere left, but not so much anymore. This is one ingredient.

It comes in combination with a certain assumption, which is called a "spectral gap assumption". So that's already on the technical side the crucial ingredient. Sometimes it's called "Poincare inequality", maybe people are more familiar with that. What it basically tells you is that if you control the average of a function, then you can also control the fluctuations around this average if you also control the derivative of a function – which somehow makes sense, right? If you know you have a function, you fix its average, so you know roughly where in space you are. If you then can also control the derivative so you know, how far can you at most go away from your average, then you somehow know everywhere roughly where you can go. And now we use this in a stochastic context.

CD: And how do you know that you will have this spectral gap property?

MT: This is an assumption, and we can consider noise terms that satisfy this spectral gap assumption. A lot of physically interesting noise terms satisfy that. For example, spacetime white noise, which is the analogue of the Brownian motion, satisfies it. Any Gaussian noise is covered. But certainly, there are situations which are not covered by this approach.

CD: Right. But for example, if we look at equation (1) that we discussed. Now you may apply your technique of differentiating. How would you go on?

MT: We're rather not taking the derivative of equation (1), but of these expressions v, v^2, and v^3 built from the linearised equation (5). A first step is to linearise your problem as before. (00:46:33-9)

CD: Hairer's ideas are still necessary, yeah?

MT: They are still there, yes. The big philosophy is exactly the same. You linearise your problem and then of these generalised monomials, of these objects that you introduce, there you apply your derivative with respect to the noise. And with that you show that they converge, and one can easily see this for example again with this squared term. Let's write down what would be the assumption. It would be that we can control this object v^2 minus its expectation by its derivative:

$$\mathbb{E}\left|v^2 - \mathbb{E}v^2\right|^2 \leq \mathbb{E}\left|\frac{\partial}{\partial \xi} v^2\right|^2. \tag{10}$$

That's the spectral gap assumption, and it says the following: If we control the average of this object v^2, and we control also its derivative, then we also control what is called the variance, a quantified version of the fluctuations. You see again how this expectation of the mollified thing comes in: That's what we have to subtract so things stay under control. But provided we subtract the average, the only thing we have to control is the derivative. If we control the derivative and the average, we control the fluctuations. And now why can you control the derivative? This is a derivative in an infinite dimensional setting. It's a bit delicate object. Basically, the same rules of calculus apply, so we have

$$\frac{\partial}{\partial \xi} v^2 = 2v \frac{\partial}{\partial \xi} v. \tag{11}$$

You can see that this object here, v, was the solution to the stochastic heat equation (5)

$$(\partial_t - \Delta)\, v = \xi.$$

Now, simplified said if you take the derivative here w.r.t ξ, i.e.

$$(\partial_t - \Delta)\, \frac{\partial}{\partial \xi} v = \frac{\partial}{\partial \xi} \xi \tag{12}$$

you replace $\frac{\partial}{\partial \xi}\xi$ by a 1. So, the derivative with respect to the noise satisfies an equation where the solution is smooth. This $\frac{\partial}{\partial \xi} v$ is actually a smooth object. Suddenly this problem, that products were ill-defined and didn't make sense, disappeared. The right hand side of (11) is now a product that perfectly makes sense, analytically.

CD: Why again was this derivative now smooth? Because it just satisfies the ordinary heat equation?

MT: Exactly. You see, $\frac{\partial}{\partial \xi} v$ satisfies equation (12). And it is therefore much better behaved, even smooth. This product on the right hand side of (11) is now a product of a distribution with a smooth function, and that's okay. The only thing that's not okay is a product of two honest distributions. You somehow got rid of the problem that certain multi-linear expressions didn't make sense by taking the derivative. And now one could also do that with this modification parameter ε around and show that things actually converge.

CD: This is a completely basic idea, I think. I mean I understand why it works.

MT: It gets, of course, much more difficult in the full general setting. Because also the terms, you usually have to come up with, are very complicated, not just so simple terms as v^2. But this philosophy somehow always works for all the terms that you need.

CD: Yeah, you differentiate with respect to the noise and then suddenly things become smooth and now you can multiply or whatever non-linear operations you want to make. That's super. Can you remind me how do you differentiate with respect to the noise?

Differentiating w.r.t. ξ using the Fréchet derivative 00:45:40

MT: Yes, sure. The correct term in a mathematical context is also a Fréchet derivative.

CD: This would be between Banach spaces also?

MT: Yeah, exactly. There is a notion of derivatives also in infinite dimensional spaces.

CD: I know Gateaux and Fréchet and somehow Fréchet was more uniform?

MT: Yeah. A Gateaux derivative is a directional derivative.

CD: Ah and so Fréchet is like in analysis the total differential. Okay, this I understand, but what would be the spaces?

MT: It's similar to Sobolev spaces. Also in this context, you can start with a certain notion of functions that are like the analogue in infinite dimensions of compactly supported and C^∞. In infinite dimensions you don't take compactly supported. Instead, one takes cylindrical functions, whatever that means. It's some notion of smooth functions where a derivative makes sense. And then you take completion with respect to a certain norm as you can define Sobolev spaces as the completion of compactly supported smooth functions with respect to a certain norm. And one can do exactly the same here. And I think this space probably has a name. Maybe "Malliavin Sobolev space".

CD: I actually remember, I tried to understand Malliavin's proof, and I think to some extent I understood it, but I forgot it completely. But now that you are saying this.

MT: I also don't know this proof.

CD: Yeah, but now it comes back. I remember such spaces there.

MT: What I wrote earlier in (10) is not precisely true since this is an infinite dimensional object. It's like an infinite dimensional gradient. What you have to put in (11) is really a norm:

$$\mathbb{E}\left|v^2 - \mathbb{E}v^2\right|^2 \le \mathbb{E} \parallel \frac{\partial}{\partial \xi} v^2 \parallel_{\dot{H}^s}^2 .$$

And it matters which norm because it's an infinite dimensional space. We take here some Sobolev space, say $\dot{H}^s$, and so this is now the assumption on the noise. And what you put here somehow determines the behaviour of ξ.

CD: You said in the Gaussian case or white noise case, these spectral gap inequalities were known?

MT: Yes. More generally, I think this is like a research field in its own to show which stochastic processes satisfy a spectral gap inequality, because it also tells you something about mixing properties and decay of correlations. It is a very strong concept.

CD: And you could lean on this on their results and use them.

MT: I think it's still current research proving such kind of estimates for various stochastic processes.

CD: Aha! So now I am pretty confident that if I look at your paper again, I will understand much better.

MT: Maybe the introduction makes more sense now.

CD: Yes okay, I will certainly look at it again. Thanks a lot for these enlightening explanations. And I have the feeling that I got the gist of the basic idea. And hopefully any readers or listeners too. Okay, thanks a lot!

MT: Thanks!