A Math Theory for Why People Hallucinate

Psychedelic drugs can trigger characteristic hallucinations, which have long been thought to hold clues about the brain's circuitry. After nearly a century of study, a possible explanation is crystallizing.

By Jennifer Ouellette

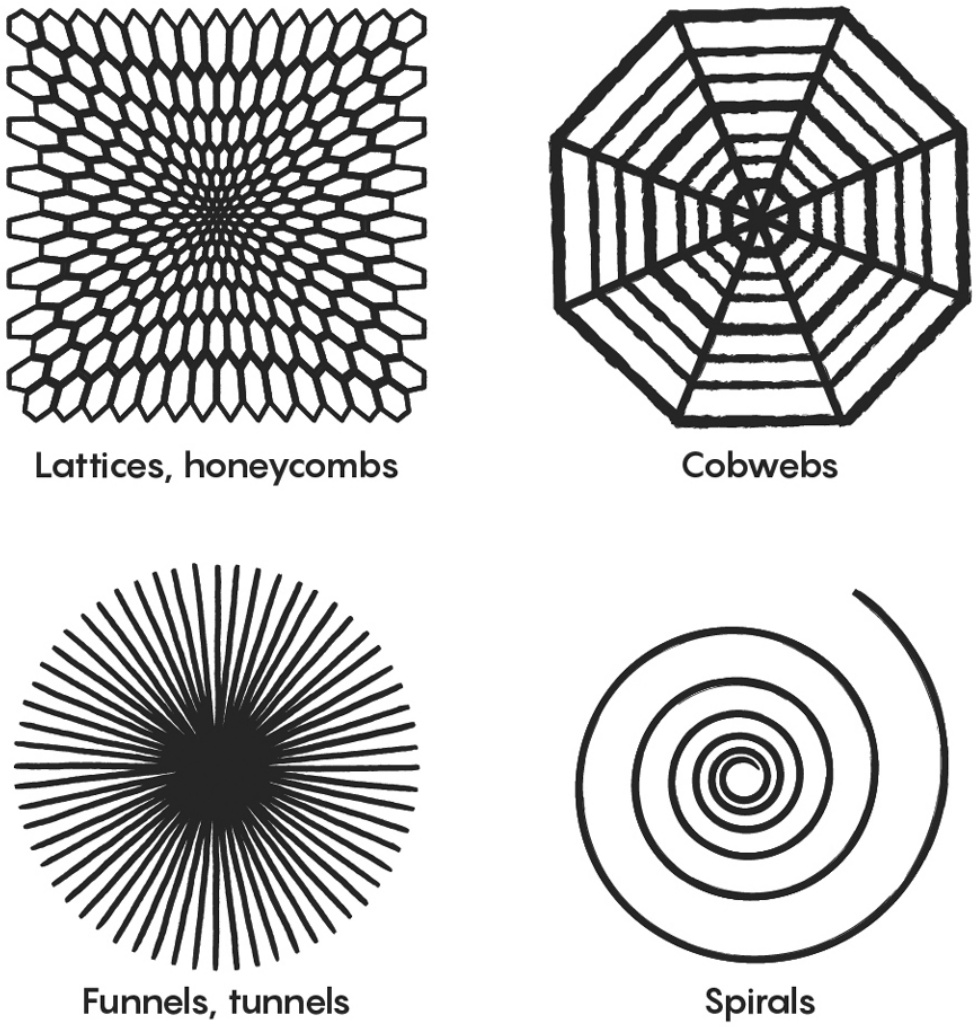
aeforia and Olena Shmahalo/Quanta Magazine

In the 1920s, decades before counterculture guru Timothy Leary made waves self-experimenting with LSD and other psychedelic drugs at Harvard University, a young perceptual psychologist named Heinrich Klüver used himself as a guinea pig in an ongoing study into visual hallucinations. One day in his laboratory at the University of Minnesota, he ingested a peyote button, the dried top of the cactus *Lophophora williamsii*, and carefully documented how his visual field changed under its influence. He noted recurring patterns that bore a striking resemblance to shapes commonly found in ancient cave drawings and in the paintings of Joan Miró, and he speculated that perhaps they were innate to human vision. He classified the patterns into four distinct types that he dubbed "form

constants": lattices (including checkerboards, honeycombs and triangles), tunnels, spirals and cobwebs.

Some 50 years later, Jack Cowan of the University of Chicago set out to reproduce those hallucinatory form constants mathematically, in the belief that they could provide clues to the brain's circuitry. In a seminal 1979 paper, Cowan and his graduate student Bard Ermentrout reported that the electrical activity of neurons in the first layer of the visual cortex could be directly translated into the geometric shapes people typically see when under the influence of psychedelics. "The math of the way the cortex is wired, it produces only these kinds of patterns," Cowan explained recently. In that sense, what we see when we hallucinate reflects the architecture of the brain's neural network.

But no one could figure out precisely how the intrinsic circuitry of the brain's visual cortex generates the patterns of activity that underlie the hallucinations.



Lucy Reading-Ikkanda/Quanta Magazine

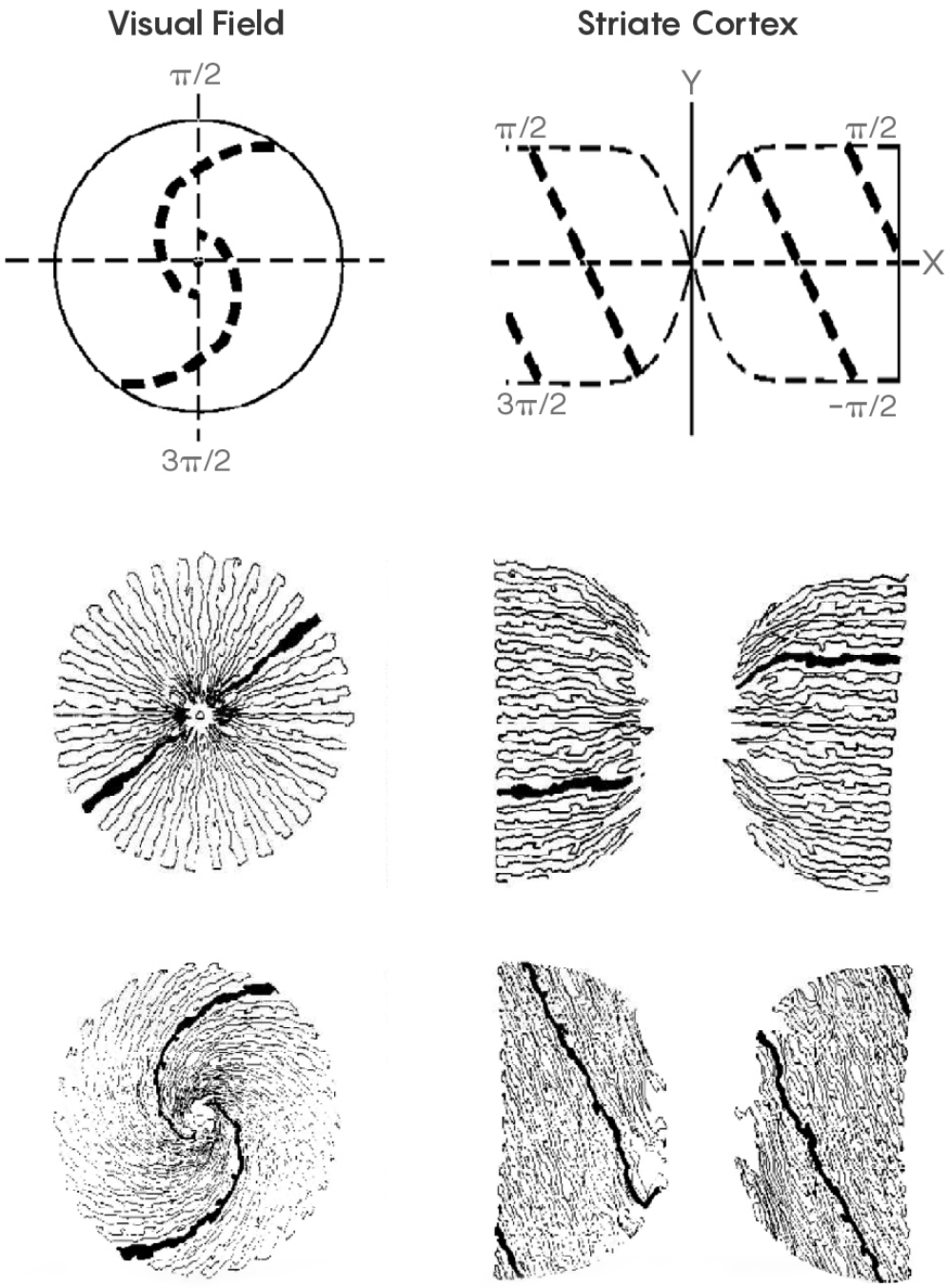
Heinrich Klüver classified the shapes he saw while under the influence of hallucinogenic drugs into four categories, known as "form constants."

An emerging hypothesis points to a variation of the mechanism that produces so-called "Turing patterns." In a 1952 paper, the British mathematician and code-breaker Alan Turing proposed a mathematical mechanism for generating many of the repeating patterns commonly seen in biology — the stripes of tigers or zebra fish, for example, or a leopard's spots. Scientists have known for some time that the classic Turing mechanism probably can't occur in a system as noisy and complicated as the brain. But a collaborator of Cowan's, the physicist Nigel Goldenfeld of the University of Illinois, Urbana-Champaign, has proposed a twist on the original idea that factors in noise. Experimental evidence reported in two recent papers has bolstered the theory that this "stochastic Turing mechanism" is behind the geometric form constants people see when they hallucinate.

Sweaty Grasshoppers

Images we "see" are essentially the patterns of excited neurons in the visual cortex. Light reflecting off the objects in our field of view enters the eye and comes to a focus on the retina, which is lined with photoreceptor cells that convert that light into electrochemical signals. These signals travel to the brain and stimulate neurons in the visual cortex in patterns that, under normal circumstances, mimic the patterns of light reflecting off objects in your field of view. But sometimes patterns can arise spontaneously from the random firing of neurons in the cortex — internal background noise, as opposed to external stimuli — or when a psychoactive drug or other influencing factor disrupts normal brain function and boosts the random firing of neurons. This is believed to be what happens when we hallucinate.

But why do we see the particular shapes that Klüver so meticulously classified? The widely accepted explanation proposed by Cowan, Ermentrout and their collaborators is that these patterns result from how the visual field is represented in the first visual area of the visual cortex. "If you opened up someone's head and looked at the activity of the nerve cells, you would not see an image of the world as through a lens," said Peter Thomas, a collaborator of Cowan's who is now at Case Western Reserve University. Instead, Thomas explained, the image undergoes a transformation of coordinates as it is mapped onto the cortex. If neuronal activity takes the form of alternating stripes of firing and non-firing neurons, you perceive different things depending on the stripes' orientation. You see concentric rings if the stripes are oriented one way. You see rays or funnel shapes emanating from a central point — the proverbial light at the end of the tunnel common in near-death experiences — if the stripes are perpendicular to that. And you see spiral patterns if the stripes have a diagonal orientation.



Bressloff, P.C., Cowan, J.D., Golubitsky, M., Thomas, P.J., & Wiener, M.C. (2002). What Geometric Visual Hallucinations Tell Us about the Visual Cortex. *Neural Computation*, 14, 473-491.

Diagram showing how lines in the visual field (circular regions on the left) map to lines in the striate cortex, part of the visual cortex involved in directly processing visual information.

But if geometric visual hallucinations like Klüver's form constants are a direct consequence of neural activity in the visual cortex, the question is why this activity spontaneously occurs — and why, in that case, it doesn't cause us to hallucinate all the time. The stochastic Turing mechanism potentially addresses both questions.

Alan Turing's original paper suggested that patterns like spots result from the interactions between two chemicals spreading through a system. Instead of diffusing evenly like a gas in a room until the density is uniform throughout, the two chemicals diffuse at different rates, which causes them to form distinct patches with differing chemical compositions. One of the chemicals serves as an activator that expresses a unique characteristic, such as the pigmentation of a spot or stripe, while the other acts as an inhibitor, disrupting the activator's expression. Imagine, for example, a field of dry grass dotted with grasshoppers. If you start a fire at several random points, with no moisture present, the entire field will burn. But if the heat from the flames causes the fleeing grasshoppers to sweat, and that sweat dampens the grass around them, you'll be left with periodic spots of unburned grass throughout the otherwise charred field. This fanciful analogy, invented by the mathematical biologist James Murray, illustrates the classic Turing mechanism.

Turing acknowledged that this was a greatly simplified toy model for how actual patterns arise, and he never applied it to a real biological problem. But it offers a framework to build on. In the case of the brain, Cowan and Ermentrout pointed out in their 1979 paper that neurons can be described as activators or inhibitors. Activator neurons encourage nearby cells to also fire, amplifying electrical signals, while inhibitory neurons shut down their nearest neighbors, dampening signals. The researchers noticed that activator neurons in the visual cortex were mostly connected to nearby activator neurons, while inhibitory neurons tended to connect to inhibitory neurons farther away, forming a wider network. This is reminiscent of the two different chemical diffusion rates required in the classic Turing mechanism, and in theory, it could spontaneously give rise to stripes or spots of active neurons scattered throughout a sea of low neuronal activity. These stripes or spots, depending on their orientation, could be what generates perceptions of lattices, tunnels, spirals and cobwebs.

While Cowan recognized that there could be some kind of Turing mechanism at work in the visual cortex, his model didn't account for noise — the random, bursty firing of neurons — which seemed likely to interfere with the formation of Turing patterns. Meanwhile, Goldenfeld and other researchers had been applying Turing's ideas in ecology, as a model for predator-prey dynamics. In that scenario, the prey serve as activators, seeking to reproduce and increase their numbers, while predators serve as inhibitors, keeping the prey population in check with their kills. Thus, together they form Turing-like spatial patterns. Goldenfeld was studying how random fluctuations in predator and prey populations affect these patterns. He knew about Cowan's work in neuroscience and soon realized his insights could apply there as well.

Houses With Eyes and Jaws

A condensed matter physicist by training, Goldenfeld gravitates toward interdisciplinary research, applying concepts and techniques from physics and math to biology and evolutionary ecology. Roughly 10 years ago, he and his then graduate student Tom Butler were pondering how the spatial distribution of predators and prey changes in response to random local fluctuations in their populations, for instance if a herd of sheep is attacked by wolves. Goldenfeld and Butler found that when a herd's population is relatively low, random fluctuations can have big effects, even leading to extinction. It became clear that ecological models need to take random fluctuations into account rather than just describe the average behavior of populations. "Once I knew how to do the fluctuation calculation for pattern formation," Goldenfeld said, "it was an obvious next step to apply this to the hallucination problem."

In the brain, it's the number of neurons that are on or off that randomly fluctuates rather than sheep and wolf populations. If an activator neuron randomly switches on, it can cause other nearby neurons to also switch on. Conversely, when an inhibitory neuron randomly switches on, it causes

nearby neurons to switch off. Because the connections between inhibitory neurons are long-range, any inhibitory signals that randomly arise spread faster than random excitatory signals — exactly what's needed for a Turing-like mechanism. Goldenfeld's models suggested that stripes of active and inactive neurons will form in a Turing-like pattern. He dubbed these stochastic Turing patterns.

However, to function properly, the visual cortex must be primarily driven by external stimuli, not by its own internal noisy fluctuations. What keeps stochastic Turing patterns from constantly forming and causing us to constantly hallucinate? Goldenfeld and colleagues argue that even though the firing of neurons can be random, their connections are not. Whereas short-range connections between excitatory neurons are common, long-range connections between inhibitory neurons are sparse, and Goldenfeld thinks this helps suppress the spread of random signals. He and his cohorts tested this hypothesis by creating two separate neural network models. One was based on the actual wiring of the visual cortex, and the other was a generic network with random connections. In the generic model, normal visual function was substantially degraded because the random firing of neurons served to amplify the Turing effect. "A generically wired visual cortex would be contaminated by hallucinations," Goldenfeld said. In the realistic model of the cortex, however, internal noise was effectively dampened.

Seth Lowe for Quanta Magazine

Nigel Goldenfeld, a physicist at the University of Illinois, Urbana-Champaign, hypothesizes that the stochastic Turing mechanism underlies visual hallucinations.

Goldenfeld suggests that evolution has selected for a particular network structure that inhibits hallucinatory patterns: The sparseness of connections between inhibitory neurons prevents inhibitory signals from traveling long distances, disrupting the stochastic Turing mechanism and the perception of funnels, cobwebs, spirals and so forth. The dominant patterns that spread through the network will be based on external stimuli — a very good thing for survival, since you want to be able to spot a snake and not be distracted by a pretty spiral shape.

"If the cortex had been built with these long-range inhibitory connections all over the place, then the tendency to form these patterns would be stronger than the tendency to process the visual input coming in. It would be a disaster and we would never have survived," Thomas said. Because long-range inhibitory connections are sparse, "the models don't produce spontaneous patterns unless you force them to, by simulating the effects of hallucinogenic drugs."

Experiments have shown that hallucinogens like LSD appear to disrupt the normal filtering mechanisms the brain employs, perhaps boosting long-range inhibitory connections and therefore permitting random signals to amplify in a stochastic Turing effect.

Goldenfeld and collaborators have not yet tested their theory of visual hallucinations experimentally, but hard evidence that stochastic Turing patterns do arise in biological systems has emerged in the last few years. Around 2010, Goldenfeld heard about work done by Ronald Weiss, a synthetic biologist at the Massachusetts Institute of Technology who had been struggling for years to find the appropriate theoretical framework to explain some intriguing experimental results.

Years earlier, Weiss and his team had grown bacterial biofilms that were genetically engineered to express one of two different signaling molecules. In an effort to demonstrate the growth of a classic Turing pattern, they tagged the signaling molecules with fluorescent markers so that the activators glowed red and the inhibitors glowed green. Although the experiment started out with a homogenous biofilm, over time a Turing-like pattern emerged, with red polka dots scattered throughout a swath of green. However, the red dots were much more haphazardly located than, say, leopards' spots. Additional experiments also failed to yield the desired results.

When Goldenfeld heard about these experiments, he suspected that Weiss' data could be viewed from a stochastic point of view. "Rather than trying to make the patterns more regular and less noisy," Weiss said, "we realized through our collaboration with Nigel that these are really stochastic Turing patterns." Weiss, Goldenfeld and collaborators finally published their paper in the *Proceedings of the National Academy of Sciences* last month, 17 years after the research began.

The biofilms formed stochastic Turing patterns because gene expression is a noisy process. According to Joel Stavans of the Weizmann Institute of Science in Israel, that noise is responsible for disparities among cells, which can have the same genetic information yet behave differently. In recently published work, Stavans and his colleagues investigated how noise in gene expression can lead to stochastic Turing patterns in cyanobacteria, ancient organisms that produce a large proportion of the oxygen on Earth. The researchers studied anabaena, a type of cyanobacteria with a simple structure of cells attached to one another in a long train. An anabaena's cells can specialize to perform one of two activities: photosynthesis, or converting nitrogen in the atmosphere into proteins. An anabaena might have, for instance, one nitrogen-fixing cell, then 10 or 15 photosynthesis cells, then another nitrogen-fixing cell, and so on, in what appears to be a stochastic Turing pattern. The activator, in this case, is a protein that creates a positive feedback loop to produce more such proteins. At the same time, the protein may also produce other proteins that diffuse to neighboring cells and inhibit the first protein's production. This is the primary feature of a Turing mechanism: an activator and an inhibitor fighting against each other. In anabaena, noise drives the competition.

Researchers say the fact that stochastic Turing processes appear to be at work in these two biological contexts adds plausibility to the theory that the same mechanism occurs in the visual cortex. The findings also demonstrate how noise plays a pivotal role in biological organisms. "There is not a direct correlation between how we program computers" and how biological systems work, Weiss said. "Biology requires different frameworks and design principles. Noise is one of them."

There is still much more to understand about hallucinations. Jean-Paul Sartre experimented with mescaline in Paris in 1935 and found it distorted his visual perception for weeks. Houses appeared to have “leering faces, all eyes and jaws,” clock faces looked like owls, and he saw crabs following him around all the time. These are much higher-level hallucinations than Klüver’s simple form constants. “The early stages of visual hallucination are very simple — these geometric patterns,” Ermentrout said. But when higher cognitive functions kick in, such as memory, he said, “you start to see more complex hallucinations and you try and make sense of them. I believe that all you’re seeing is the spontaneous emergence of [stored memories] as the higher brain areas become more excited.”

Back in the ’20s, Klüver also worked with subjects who reported tactile hallucinations, such as cobwebs crawling across their skin. Ermentrout thinks this is consistent with a cobweb-like form constant mapped onto the somatosensory cortex. Similar processes might play out in the auditory cortex, which could account not only for auditory hallucinations but for phenomena like tinnitus. Cowan agrees, noting that the brain has similar wiring throughout, so if a theory of hallucinations “works for vision, it’s going to work for all the other senses.”