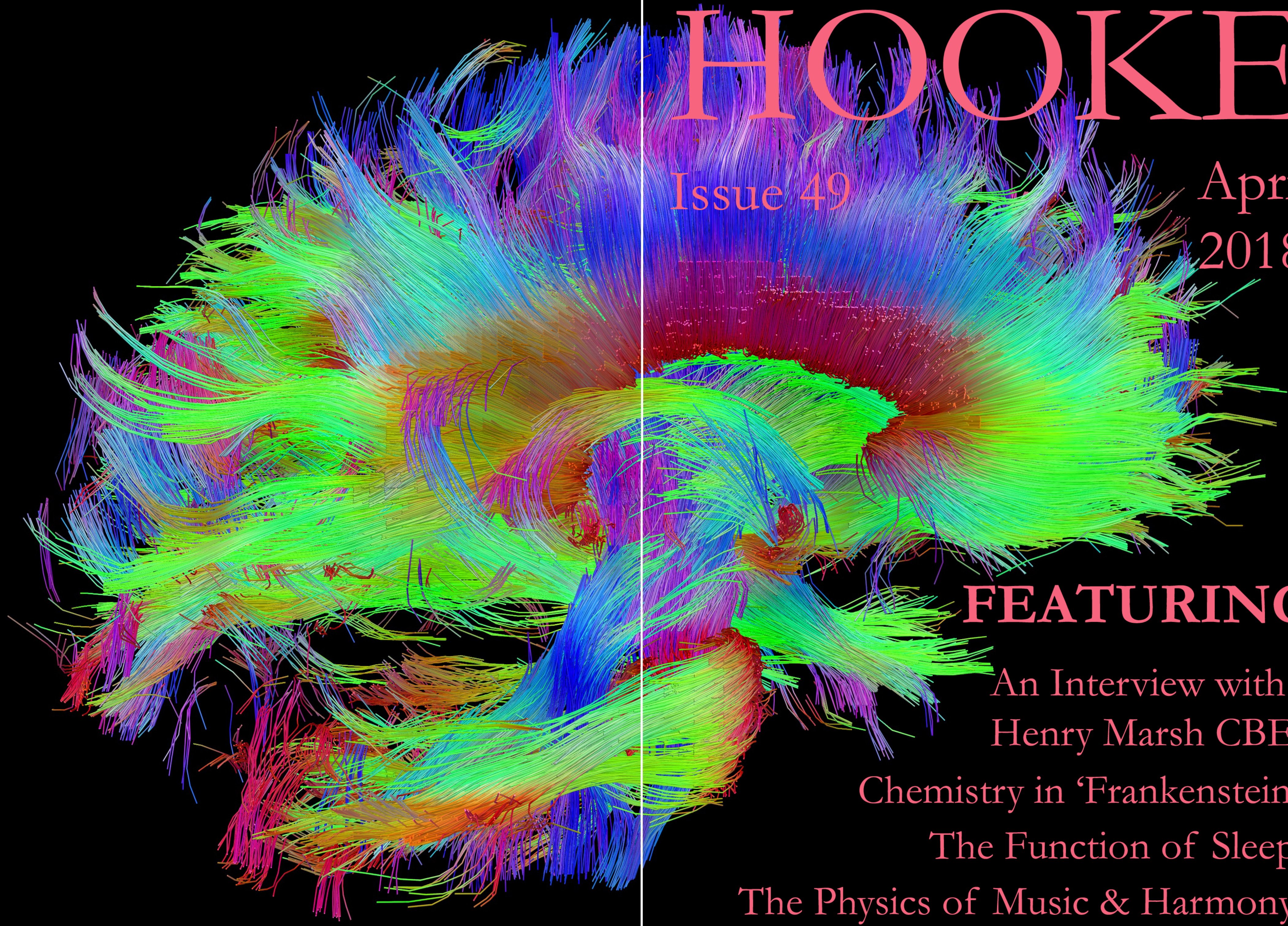


HOOKE

Issue 49

April 2018

FEATURING

An Interview with Henry Marsh CBE

Chemistry in 'Frankenstein'

The Function of Sleep

The Physics of Music & Harmony

Editors' Note

For many, scientific research remains shrouded in a veil of off-putting technicality. Few of us can claim, for example, to understand the intricacies of quantum mechanics, the networks of the brain, or the algorithms that program a computer. Yet in this edition, we feature a wealth of articles about such topics, such as those focussed on neuroscience (featuring an exclusive interview with neurosurgeon/author Henry Marsh), computing, and their synthesis, neural mapping - which inspired our choice of cover from the Human Connectome Project. Our writers also consider applications of science outside of its theory, from its use in industry, to its interpretations in literature, to its significance in achieving political goals.

By bringing the inherent intricacy of scientific research to light, Hooke aims to help people appreciate its value. The need for nuance and the scientific method is often forgotten outside the bubble of research, especially when it comes to matters of media literacy and political rhetoric. Science, therefore, isn't just synonymous with the eight floors of Hooke, but is a subject that remains a cornerstone of any questioning, open mindset. It isn't a field of study limited to prescriptive GCSE and A-Level specifications, but rather a tool to confront issues with rational evidence-based argument, and the fuel that enables collective discovery and progress. We may not be able to lift the veil, but we can definitely respect all it stands for.

The articles compiled here have, once more, showcased the curiosity, creativity, and attention to detail that goes into Westminster's science writing, and stand alongside all the countless academic achievements and extra-curricular feats all of you have gathered. We would like to take this opportunity to thank our contributors for this edition's submissions, and you, for taking the time to read them.

The Editors

April 2018

Editors: Lorna Bo, Alessandro Esciua-Blanco, Ryan Kang, Navyaa Mathur, Isky Mathews, Raghav Nayak, Brandon Tang & Hannah Virji

The editors would like to express their gratitude to Matthew Bradshaw and Abigail Farr.

Submissions for the next issue of Hooke are warmly invited at any time: please email **hooke.editors@westminster.org.uk** to send in any science-related essays or articles (preferably between 800-2500 words).

Contents

An Interview with Henry Marsh CBE

Lorna Bo

LB: What prompted you to choose neurosurgery as a career?

HM: I didn't see any neurosurgery as a medical student, and only came across an aneurysm operation more or less by chance. I was working as an intensive care doctor at the time, and was helping out with the anaesthetic. I just decided then, while I was watching: this is what I want to do. I was already interested in surgery, but I was always disillusioned by general surgery - it was relatively crude and macroscopic at the time. It was the combination of the fiddling done on microscopes, and the drama, excitement, and seriousness, which I found irresistible. The fact that my son had had a brain tumour operation a year before then influenced me too, and probably helped me transition into paediatric neurosurgery as well as adult neurosurgery in my consultant career.

LB: You also talked quite a lot about how difficult it is making mistakes in medicine.

HM: Well it's easy to make mistakes. It's just very hard to admit to them – indeed, to admit to *yourself* that you've made a mistake, let alone to anybody else.

You're making endless decisions which might seem quite trivial, but a misread blood test or a forgotten scan can have catastrophic consequences. In my experience, when there have been disastrous problems, it's not been one big mistake, it's been a series of things which compound.

LB: Do you feel as though there's less psychological support available to junior doctors now?

HM: Yes, there's not really a sense of belonging anymore. My generation of doctors worked very long hours and had no life outside medicine - but you were at the same place for months on end, and so you felt part of a family. You got to know all the hospital staff on a face-to-face basis, thanks to the very long hours.

Junior doctors now work shifts, and so the disruption to their circadian rhythms is probably worse than it was for me. They don't get to know people, and they don't have a structure of senior doctors looking after them. Before, there were only a small number of senior doctors who had clear authority, and they ran the hospitals. But now the whole hospital hierarchy has changed, in the sense that now there's a whole *series* of hierarchies. You've got the managers, you've got the nurses, you've got the paramedics, you've got the doctors. There's nobody really in charge, and the managers who *are* there interfere with the daily running of one's clinical workload, and are basically there to save money. They never see patients, exist on a sort of alternate universe, and ultimately are answerable to the politicians in the Department of Health, who pretend that things are better than they actually are.

LB: You've said in a previous interview that healthcare is a 'bottomless pit'. What did you mean by that?

HM: Well, it is. If you spent the entire national income on healthcare, everybody would still die. It's a question of when to stop, of when somebody is too old to treat. In America, however, you're never too old. There, they say 'death is optional' - provided your insurance is good enough!

LB: But how do you sit someone down and tell them that they shouldn't have any more treatment?

HM: It's very difficult, and ultimately, it's all about judging probabilities. But since most of us are frightened of dying, probabilities get very skewed. If you know the chemotherapy has no chance of working you obviously won't use it, but if there's even a 5% chance of success, then the patient will want it. We all say we'll be the lucky one, like soldiers in warfare.

LB: What do you do if a patient keeps trying to push for more and more treatment?

HM: There's a story in my first book ['Do No Harm'] where I was pushed by the patient's family to continue with surgery. I agreed to operate, but the operation ended up only making things worse. So, it can be very difficult. It's always hard to say 'go away and die' to a patient, and it's always easier to do something than do nothing. I can see it in my junior consultants at St George's, for example. But you do get better at it as you get older – because you can look back and see how you've made things worse in the past.

JS (Jake Swann): There's this culture of thinking that if you've done nothing, then you must be accountable.

HM: Well, that's right. You worry if you do nothing because you're worried you might be wrong. But if you operate on every single patient who wants it, you cause a lot of suffering for the sake of one or two patients who do well. Where you strike that balance is a very difficult question.

LB: Do you think medicine is more of an art or a science, then?

HM: I don't really like that choice between art and science. The original phrase, I think, was said by William Osler - a great doctor of the turn of the last century. He said, 'medicine is a science of uncertainty, and an art of probability'. It's profoundly true, because the future's uncertain and so we deal in probabilities. That's frightening to patients, who want certainties rather than gambles. Sometimes it's obvious when to stop, but often it's not. It's one of the reasons why healthcare costs are going up and up - because doctors are finding it hard to stop.

LB: You've talked quite a lot before about the limitations of bureaucracy, but also that power corrupts, and doctors must be held accountable – what do you think about the meeting of those two things in medical litigation?

HM: Everything in life is a question of balance; almost everything is multifactorial. We tend to think in unitary causation: if A, then B. In fact, most things are enormously complex, and most things interreact. Obviously, doctors need regulating, but there are many different ways of doing it. In New Zealand, they have no medical negligence litigation. If a mistake is made you're automatically compensated by the state, which means pulling up income tax a bit. But in England, the NHS litigation bill costs billions of pounds and is still increasing - 40% of any settlement is purely in legal costs. In a sense, that's money wasted. So, there are different ways of doing things, and there's nothing inevitable about bad design. Nothing is so good that it can't be done better.

LB: You're often called a 'maverick' by the media. Is that an accurate portrayal of yourself?

HM: No…well, yes *and* no. Because I went into medicine late, I've always been a little bit detached. I've never really identified with the medical profession, which gave me a slightly more objective, detached view that helped me write my books. I was lucky that I also had a proper education [due to studying PPE before switching to medicine] – most doctors don't. They do their science A-Levels and multiple-choice exams, and they never learn to write or think. It's all done in a rather dogmatic, fact-learning way. I think I was immensely lucky that I had this sort of separate education and didn't actually become a doctor until I was 30 years old. Yes, I'm slightly eccentric and rather outspoken. But actually, I'm very thin-skinned. I get terribly frightened and upset if I get criticised, and I get this compulsion to stand up, argue, and cause problems.

LB: So, do you think medical training should be changed so doctors are taught resilience, empathy and other soft skills involved in medicine?

HM: Students and junior doctors are hugely influenced by the senior doctors who train and work with them. Setting an example is the most effective way of changing people's behaviour - rather than ordering them to do things, or having all these GMC regulations saying how you should behave. Professional standards are all rather weak and woolly. The idea of public service and professional standards is something done for purely ethical, moral reasons.

LB: Do you think the NHS will end up being a two-tiered system?

HM: I fear it will, yes. There's many different ways it could pan out, and it also depends where you live. If you live in central London and go to UCH, you'll get the best treatment - as good as anywhere else in the world - free of charge. If you have the same disease and you live in Barrow-in-Ferness, you won't. There are big provincial differences, and that's inevitable to some extent. But the worry is that care will become segregated to the point of a two-tiered system. But for now, if you have cancer or a trauma in London, you're actually better off going to an NHS hospital, not a private one, because the backup is not as good in a private hospital as it is in the NHS.

Competition in hospitals has never worked. There was one hospital in Hinchingbrooke which was taken over by a private healthcare company called Circle Health, who had a lot of grand ideas about running it as a commercial business getting paid per case by the NHS. It bombed. It didn't work.

Obviously, you can't let hospitals go the wall. What typically happens if there's a problem in a business is that you'll sack half the managers and appoint new ones to save the business. But in the NHS, you appoint *more* managers to discuss the problem. I met one person who'd been appointed to a new committee to discuss how to reduce the number of committees in their hospital!

LB: Final question now. Why do you write?

HM: It's a compulsion. I've written all my life. Both books ['Do No Harm' and 'Admissions'] are basically a version of the diary I've kept all my life. In fact, ['Do No Harm'] only became a book because I read bits of my diary to my second wife – the anthropologist Kate Fox – and she said to me, 'that's got to be a book'. Writing a diary is a good habit, because you realise when you look back at it how wrong and how stupid you were – particularly at your age! I regret I destroyed the diary I kept as a teenager, I'd love to read it now. When I was young I had lots of bright ideas, and I suspect many of the things that trouble me now as I get older were already present when I was a teenager. It's a period of fantastic intellectual and emotional creativity. I don't want to be young again because I was so silly, but as Wordsworth said, 'Bliss was it in that dawn to be alive, but to be young was very heaven!'

My diaries did become more serious as the years went by, because I knew that as a doctor I was living a very interesting life and seeing extraordinary things. I didn't want to lose those experiences; I didn't want to feel like they were just sand slipping through my fingers. Luckily, I'd had the benefit of having been taught English extremely well by one or two very inspiring teachers. I did 'King Lear' and 'The Tempest' for English A-Level which left me with a deep love of the sound of language. So, I do owe Westminster for some things.

The Physics of Music and Harmony

Isaac Adni

In this essay I will look at sound itself – what it is, how we can represent it and how we can store it, as well as at hearing, harmony, and the production of sound in musical instruments.

The Nature of Sound

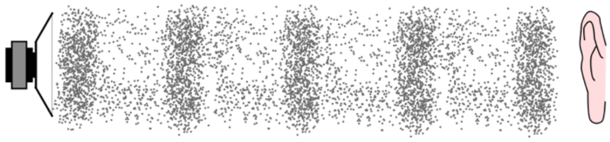


Fig. 1. A sound wave moving to the right.

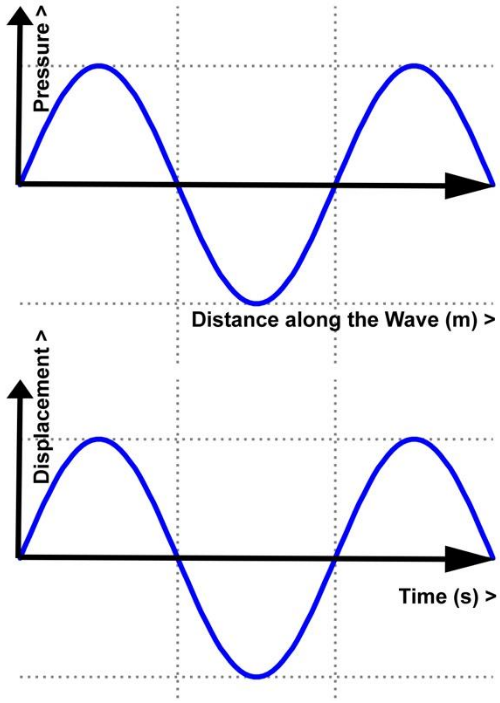


Fig. 2. The two ways of graphing a sound wave.

Sound is a wave. This means it transfers energy with no overall displacement of particles – it travels through the air, and the air molecules vibrate around fixed positions. These vibrations are in the same direction as the direction of travel as the sound wave, making the sound wave a longitudinal wave, or a 'pressure wave'. Thinking of the wave in terms of pressure is very useful, as we shall soon see. We call the areas of high pressure 'compressions', and the areas of low pressure 'rarefactions', seen in Fig. 1 where a speaker creates a sound wave that travels due to particles hitting their neighbours. The number of compressions produced per second is called the frequency (measured in Hertz - Hz), and the amplitude is the maximum displacement of a particle from equilibrium.

Fig. 2 shows the two ways in which we can graph a sound wave. We can take a 'photo' of the wave at a moment in time, graphing distance along the wave on the horizontal axis, and pressure on the vertical axis. Therefore compressions and rarefactions are shown as the extremes of the graph.

Alternatively, we can graph the displacement of a single particle from its original position over time. When storing sound on computers, we are essentially storing the graph of displacement at the microphone against time. This is because when a microphone is recording sound, its electrical signal represents the change in position of its diaphragm over time as it is pushed by the changing displacement of the particles at the point where the microphone is, and this motion will be mimicked by the speaker when it reproduces the sound.

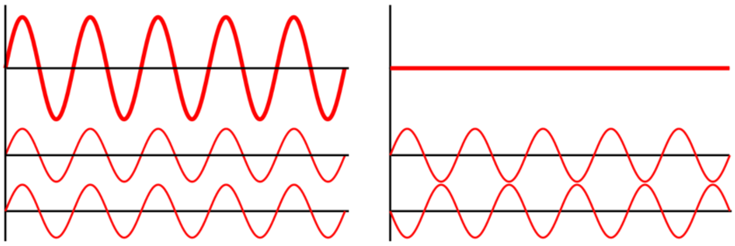


Fig. 3. The two waves on the bottom combine into the wave on the top.

However, some electronic storage formats are slightly different, but before I go on to explain them, it is useful to look at how waves add together. When waves meet, the resultant effect is a single wave. Therefore, when we hear complex sounds composed of multiple waves, we perceive the result of the overlap, rather than the individual waves themselves. When waves are added together, the resultant wave can be represented by adding together the individual displacements, as shown in Fig. 3. This fact can now be used to produce other waves from sine waves we have seen in the earlier diagrams. Fig. 4 shows a square wave being created by adding harmonics together: waves with frequencies that are integer multiples of the fundamental, the lowest frequency present. To make a Square wave we add only odd harmonics, with amplitudes of $\frac{1}{n}$ where n is the ratio of its frequency to the fundamental. If we add infinitely many of these odd harmonics, the wave that results is perfectly square. To make a sawtooth wave, as in Fig. 5, we add all harmonics, also with amplitudes of $\frac{1}{n}$ from 1 to infinity, creating a perfectly steep sawtooth wave. Square and sawtooth waves are used very frequently in music synthesisers.

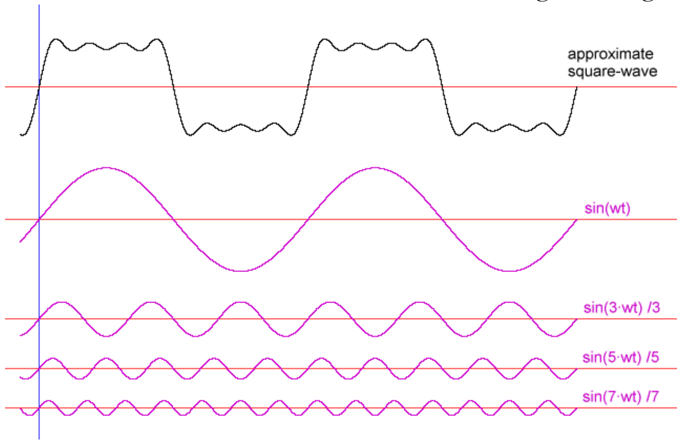


Fig. 4. Synthesis of a Square Wave (the purple waves add to give the black wave).

In fact, it turns out we can generate any wave from a sum of sine waves, and mathematically decompose any periodic function into its constituent sine waves using the ‘Fourier Transform’. If we do not have an infinitely repeating wave, we can perform the transform on a section of it.

This is the basis of mp3 compression: a sound file (effectively a digital graph of displacement against time) is split up into small ‘frames’, each of which is Fourier transformed. The sine waves outside the human hearing range are removed, and the rest stored by the encoder in an mp3 file, reducing the data required to store the file.

The Production of Sounds in Musical Instruments

Musical instruments can essentially be split into two categories, using either columns of air or strings. When a sound wave is sent down a pipe, it reflects from the open end of a pipe as well as a closed end. This is because when a rarefaction reaches the open end of a pipe, the air outside the pipe is now at higher pressure than the air just inside the pipe, so the outside air rushes in as a compression. Therefore, at an open end of a tube, a rarefaction reflects as a compression and a compression reflects as a rarefaction. At a closed end of a tube, a compression is squeezed up against the end and then reflects back as a compression, and a rarefaction reflects as a rarefaction. Of course, some energy is lost into the surroundings when a wave reflects from an open end – that is how we hear the sound.

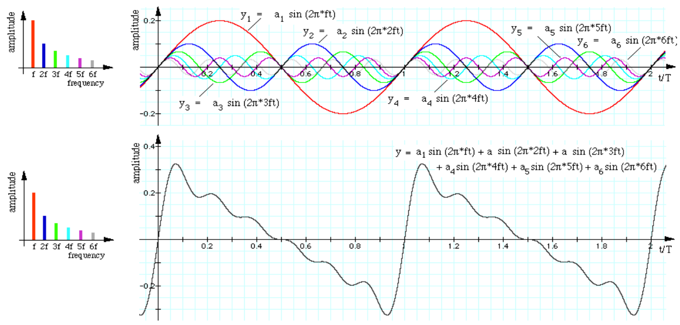


Fig. 5. Synthesis of a Sawtooth Wave (the waves in the top graph add to give the wave in the bottom graph).

However, when a continuous wave is sent down a tube, the resultant effect of the original wave and its reflection will be a single wave. At the open end of a tube, these waves will always cancel out to give no pressure change at the open end, as expected by its proximity to the atmosphere (though in the real world, a small amount of pressure change requires an 'end correction'). By contrast, at a closed end of a pipe, the pressure change is very large, so the overall pressure at a closed end varies between this large compression (high pressure) and large rarefaction (low pressure).

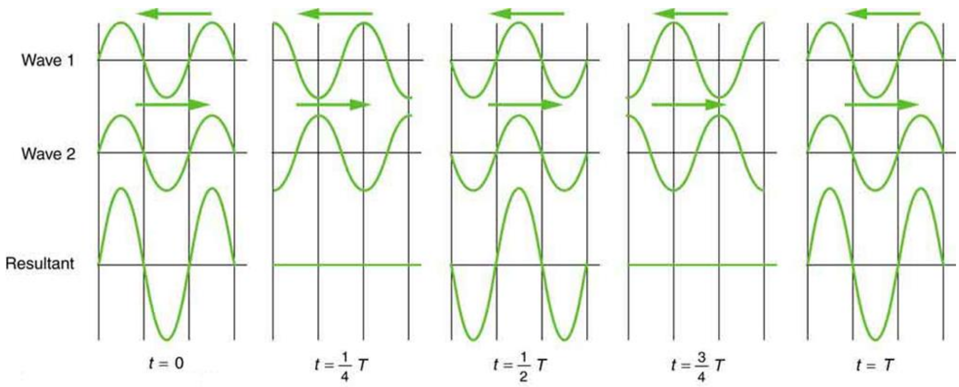


Fig. 6. The creation of a standing wave in a pipe open at both ends.

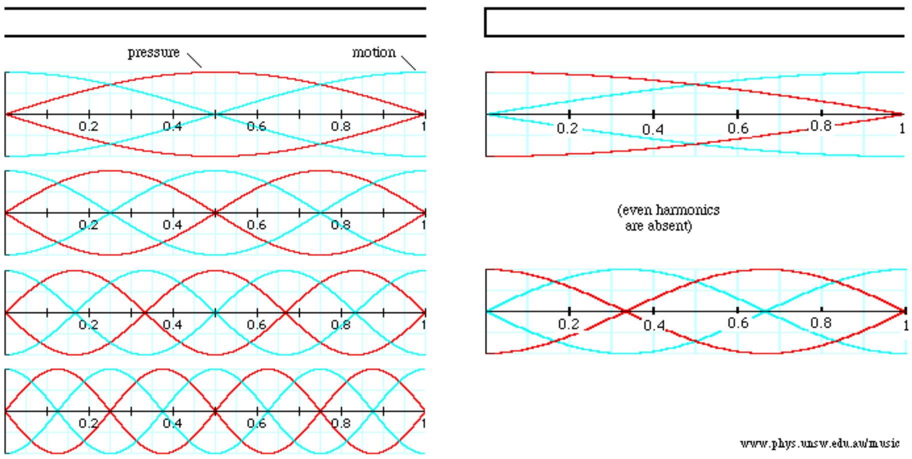


Fig. 7. Modes of resonance in a pipe open at both ends and open at one end.

All in all, the combination of the original wave and the reflected wave create what is called a 'standing wave' (see Fig. 6). At a node, the pressure does not change because the particles are free to move. At an antinode, the pressure varies between high and low with the frequency of the wave, because the particles don't move at the antinodes. Fig. 7 shows the possible frequencies (or 'resonant modes') for pipes open at both ends, or closed at one end. The difference between the types of pipe shows why a clarinet (open) and a flute (closed) sound different in tone and pitch, despite being a similar length.

The resonant mode of such an instrument can be controlled using fingering and by how the player controls their breath.

Resonant modes also occur on strings. If you hold the end of an elastic string, and tie the other end to a solid object, it is possible to shake one end of the string so that the first few modes of vibration shown in Fig. 8 are created: these are the harmonics of the string.

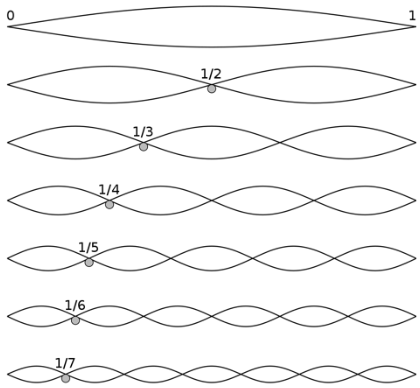


Fig. 8. Modes of vibration of a string.

The harmonics present, and their volumes relative to each other through time, are what make instruments sound different to one another when playing the same frequency. After all, if notes had no harmonics, all instruments would sound like a tuning fork, playing a pure sine wave.

Hearing

There are two major theories of hearing: Place or Resonance Theory and Temporal Theory. Place theory is considered more robust for the vast majority of frequencies, so I will concentrate on it here. Inside the ear are hair-like cells which resonate to different frequencies, and the specific hairs that resonate tell the brain what frequency is being heard.

Researchers have proved this theory of hearing by damaging part of the cochlea in animals, which then become deaf to only those frequencies which those hairs resonate to.

The interaction of harmonics means that we can often hear sounds that aren't actually present. If two notes with frequencies $3f$ and $2f$ are played together, the overtones match most of those of frequency f, so we perceive f as also being played. This is why we can hear the bass through a phone speaker, which cannot physically play those notes.

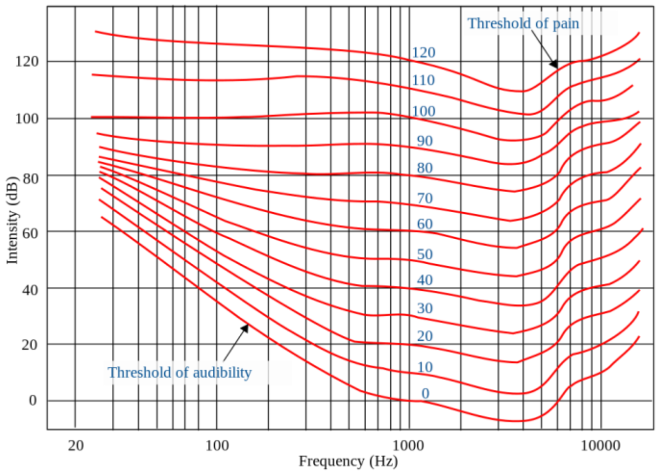


Fig. 9. A comparison of decibels and phons

Our ears also hear sounds at different volumes depending on the frequency. Fig. 9 represents this information on a graph, plotting pressure variation in decibels against frequency. 0 decibels is defined to be 20µPa (0.00002 Pascals) of pressure variation. Every increase of 10dB represents a 10x increase in pressure. The red lines on the graph represent values in phons. Phons are a scale that represents how loud our (the average person's) ears perceive sounds to be, and unlike decibels, phons are a linear scale.

Therefore, we can draw two conclusions from the graph. Firstly, that our hearing is actually logarithmic – a sound with twice as much pressure variation as another sound, if the original sound is already loud, will sound far less than twice as loud. Secondly, the volume at which we perceive sounds varies tremendously based on the frequency – the volume at which we perceive very low notes increases hugely with an increase of a very small number of decibels.

When two notes that are very close to each other in frequency/pitch are played together, we hear beats, because the notes are too close to each other in frequency for us to distinguish the two as separate from each other. The frequency of the beat is given by the difference in frequency between the two notes, as shown in Fig. 10. When the difference in frequency between the two notes is greater than ~25Hz, we no longer hear beats and perceive the notes as separate.

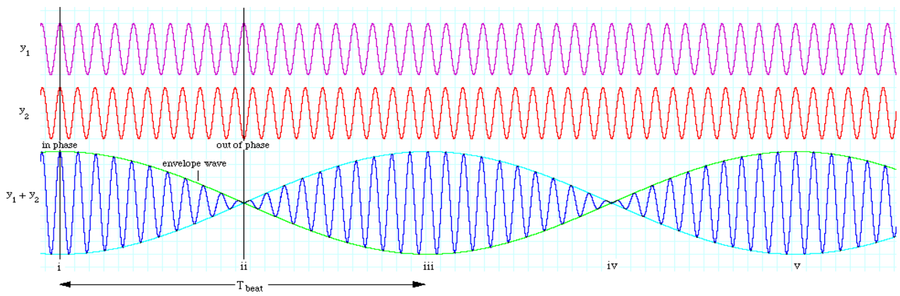


Fig. 10. Beats.

Harmony

Legend has it that Pythagoras was the first to notice what notes sounded harmonious to one another when he walked into a blacksmith's shop and found that hammers with weights in small ratios gave the most harmonious tones when sounded together. Hammers do not actually behave like this, and Pythagoras most likely used a stringed instrument.

The 19th century physicist Helmholtz developed a theory of harmony that is still used today. He stipulated that notes sound bad together when their harmonics make unpleasant beats with each other, and used to draw a version of the graph shown in Fig. 11. The graph shows the dissonance perceived with intervals between a unison and an octave. As Pythagoras had found, the least dissonant intervals are those of small frequency ratios. The least dissonant frequency ratio is $\frac{2}{1}$, an octave, the second is $\frac{3}{2}$, a fifth, the third is $\frac{4}{3}$, a fourth, and the fourth is $\frac{6}{5}$, a major third. It is useful to note that frequency is inversely proportional to the length of a string.

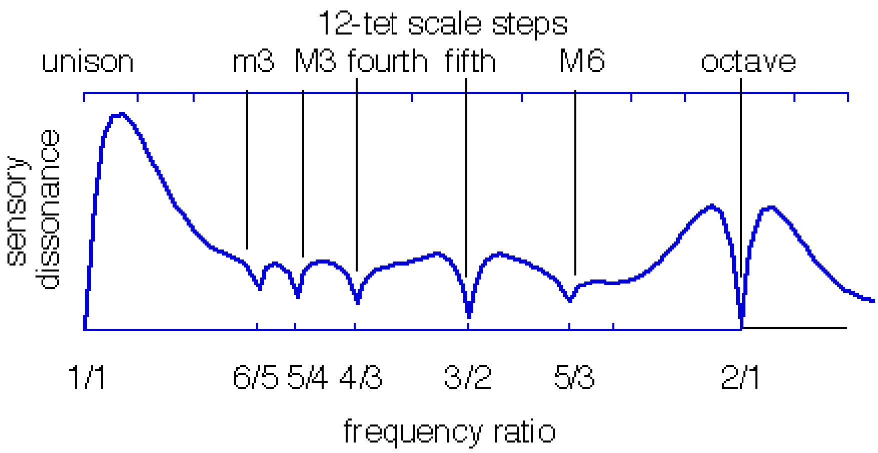


Fig. 11. A Helmholtz dissonance graph

To build a scale, we keep tuning up fifths from our starting note. We might assume that we should eventually get to the same note a number of octaves up. However, the note that we finally reach is in fact slightly sharp – because $1.5^{12} \approx 129.7$ whereas $2^7 = 128$. This ratio (129.7:128) is known as the 'Pythagorean Comma'.

This is very troubling, because we cannot extend our scale infinitely without it getting more and more out of tune with itself. So, in practice, what is done is a starting note is picked, and we tune a few fifths upwards and a few fifths downwards from that note. Then all the octaves are taken of these notes in order to extend the scale as much as is necessary.

In the Pythagorean scale, the perfect fifths actually mean that the major and minor thirds are rather dissonant, so instead a scale called the mean-tone scale prevailed for centuries. Creating the notes in the mean tone scale is very similar to the Pythagorean scale, except all the fifths are flattened $\frac{1}{4}$ a Pythagorean comma to ensure that the major thirds are in tune. However, using the last note created by going upwards is usually very dissonant with the last note created by going downwards, because this interval is greater than any other fifth. It is known as the wolf fifth, because it is said to sound like the howling of wolves.

The wolf fifth, for many centuries, prevented music being written in many keys, because keys with lots of sharps or flats were bound to contain the wolf fifth. This explains why Bach's organ music uses keys with few sharps or flats while his piano music uses more adventurous keys – the organs of the day were tuned to the mean-tone scale, whereas the harpsichords and pianos were beginning to use well-tempered scales.

Today, our instruments use equal temperament. This tuning system works by spreading the Pythagorean Comma equally across all the notes that make up an octave. No intervals apart from the octave have the perfect frequency ratios shown in Fig. 11, but this does allow for music to be played in any key. Contrary to popular teaching, Bach's Well-Tempered Clavier was not actually intended for equal temperament – the word for equal temperament existed in German, but Bach did not call it the Equal-Tempered Clavier. Well temperament simply refers to any tuning system that allows music to sound good in any key, and although it would have to be close to equal temperament, it is important to note that the two are not the same.

The debate about whether equal temperament is good rages fiercely because of the beats frequently heard, but it is important to note that equal temperament is not always used – unaccompanied string players play with perfect Pythagorean intonation and even woodwind and brass instruments can slightly bend pitches to play in different temperaments to a limited extent. However, keyboard instruments can only play in the temperament that they have been tuned to.

Conclusion

The study of the science of music can take a lifetime, but I hope this has been a brief insight into that world - I have not discussed acoustics or the psychological effect of music. Behind even the very simplest music is very complex science.

Chemistry in 'Frankenstein'

Hannah Virji

There is a clear dichotomy within the portrayal of the chemist in literature. In his 'Introduction to a Course of Lectures on Chemistry', Davy describes the figure of the chemist. On one hand, he is an inspired genius, deeply spiritual and possessed of a "sublime philosophy" which lets him transform "the elements that surround him". The creative power of the chemist here is seen in a more positive light, allowing new ideas to flourish. This interpretation of the role of the chemist in society is cohesive with the Promethean figure seen throughout the Romantic literary canon, produced by many of Davy's contemporaries and friends. For example in Prometheus Bound, Percy Shelley writes about the torments of the Greek mythical figure Prometheus. On the other hand, the chemist is a hubristic man obsessed with gaining power over his surroundings and subverting natural order by "vainly and presumptuously" attempting to "tear" apart Nature's secrets. Davy touches upon the dichotomous nature of chemistry through his language, describing chemical theories as both "visionary and seductive". Whilst contemporary developments in chemistry (many made by Davy himself) were forward thinking and helped to develop society, the unique power gifted to chemists was "seductive" because of its unattainability to those who had not learnt about the science and because of the huge amount of power and control it gave humans.

The best text to illustrate this dichotomy is 'Frankenstein' by Mary Shelley, as it shows the gradual breaking down of the Promethean figure of the chemist in favour of a more traditional, darker "mad alchemist" image. First published in 1818, Frankenstein is a science fiction and Gothic horror novel. It tells the story of Victor Frankenstein, a promising young chemist who discovers the secret to creating life. Frankenstein creates a being from human body parts which eventually rises to destroy everyone who Frankenstein holds dear; both ultimately go to their death alone. Victor Frankenstein and his work at first appear to conform to the Promethean myth. Frankenstein is enticed by an inspiring vision of chemistry in a talk given by his university lecturer, who describes chemistry as the only science with the true ability to alter the world and to transform. Whilst this vision of science encourages Frankenstein to pursue his own scientific discoveries, the religious, almost cult-like language used throughout the talk by Shelley suggests a darker side to the subject. Frankenstein is a "disciple" of his professors, and as such makes his own discoveries in chemical instruments, allowing Shelley to integrate him into the 'proper' science epitomised by his university studies. However, Frankenstein originally studied alchemy as a replacement for chemistry and whilst he apparently casts these disproved theories aside, he retains many key aspects of the stereotypical alchemist in popular culture. Firstly, his ultimate pursuit, and successful discovery is of the creation of life. This achieves the highest aim of an alchemist - the secret to immortality, the secret to resurrection and most importantly the secret to true creation. Shelley combines this aspect of Frankenstein's work with the intensely modern method he uses to create his "monster": electricity. Shelley's scientific contemporaries were using electricity to create the illusion of life in recently executed criminals' corpses, and this realisation of modern scientific ambitions intrinsically links ancient alchemical pursuits with modern chemistry in the eyes of the reader. By making these links within a wildly popular form of literature - the Gothic ghost story - Shelley both acknowledges and reinforces the popular image of the mad scientist. Furthermore, Frankenstein does actually go mad. He obsessively works on his creation, retreats into his room away from other people and eventually succumbs to mental breakdown and psychosis at the horrific realisation of his monster. Shelley was writing as a female author, without a pseudonym, at a time when almost all her prominent literary and scientific contemporaries were men. Another enduring stereotype of the century was the hysterical woman, and by transferring this madness and obsession onto a male figure, Shelley interestingly subverts this stereotype, illuminating a legitimate deeper, feminist reading of her novel. Indeed, the entire novel is notably devoid of female characters - Frankenstein's female relation dies and his monster's request for a wife is forcibly denied.

So, Frankenstein is indisputably the epitome of the "mad scientist" and the enduring popularity of both that image and this story are testament to the power of Shelley's vision. However, for a ghost story and arguably the first popular science fiction novel, this conclusion is very predictable. Shelley therefore integrates the Promethean myth into the novel to reinforce her portrayal of chemistry. She even refers to Frankenstein as the "modern Prometheus" in the subtitle of the book. The Greek myth of Prometheus is one of the most famous. It describes the Titan Prometheus who repeatedly defied Zeus, the King of the gods, to create mankind from clay and then to bring fire to humans, ultimately providing us with the power to survive independently of the gods. Prometheus' loyalty to humans over his immortal comrades leads to his eventual punishment by Zeus and he is condemned to have his organs pecked out by a giant eagle every day. In the nineteenth century Prometheus remained a representation of the enlightenment of humankind and the ultimate creative figure. At first sight, Shelley creates an impression of Frankenstein which conforms to this figure. Frankenstein is the star of the university Chemistry department, universally praised by his professors and innocently pursuing his ambition to create life itself. Shelley's notable use of electricity and light to bring his creation to life echoes Prometheus' gift of fire to the world. However, whilst Frankenstein's experiment conforms to the Greek myth of creation, it directly contradicts Shelley's contemporary ideas. For example, the scene where Frankenstein illegally digs up corpses from charnel houses and lies in graves overnight to observe the decay of human flesh serves as the moment which cements the depravity of his actions in the mind of the reader. Death and the sanctity of human life is an idea which has persisted across centuries and around the world, and this sanctity is brutally disrupted by Frankenstein's actions. Furthermore, Frankenstein subverts not only universal human ideas of life and death but also key Christian teachings. Firstly, the Christian idea of life after death is ruined by Shelley, who portrays the corpses as very much dead, with no hope of reincarnation as the body decays. Moreover, the ultimate Christian belief of God as the single Creator is destroyed here as well, as Frankenstein takes on the power to create life using modern scientific techniques which are not spiritual. As mentioned above, Shelley's use of spiritual and religious language when talking about science suggests that Frankenstein has become a "disciple" of his own perverted branch of science, rejecting Christianity. The creation of Frankenstein's monster is reminiscent of the creation of Adam by God except that Frankenstein never creates a comparable 'Eve'. Frankenstein's oversight here implies that he does not have the wisdom of God which he thinks he has, and emphasises the sanctity of God's creation over a human's. In addition, the idea of Frankenstein rejecting God is emphasised by his creation of a living being without a soul. Whilst his "monster" is alive, it is inherently evil, destructive and not at all human, which reinforces the religious ideal of an eternal soul being the possession of only humans as a gift from God. When Frankenstein overlooks his omission of a soul, he overlooks the ultimate inability of humans to create in the way that God can. By creating Frankenstein as a figure who breaks with both natural and religious views of death and human existence, Shelley creates an evil, hubristic scientist whose creation is far from the purity of Prometheus', due to Frankenstein's rejection of human values.

Frankenstein differs from the Promethean myth in Shelley's novel through his relationship with his creation. Frankenstein's "monster" is horrific, a hideous whole which cannot be improved by the selection of beautiful parts. Whilst Prometheus remains distinct from his creation, Frankenstein's monster becomes a part of him. The fact that most people today believe that 'Frankenstein' is the name of the monster rather than the scientist reflects this idea of Shelley's remarkably well, even though she presumably did not intend that! The monster is the embodiment of the darkest side of humanity, and the darkest side of Frankenstein himself, murdering and rampaging through what was once his creator's peaceful happy world. In contrast to this, Prometheus is never human and his immortality ensures that he remains separate from the people he creates. Whilst Prometheus is great enough to create an entire human race which can exist independently of himself, Frankenstein does not possess this power, although he does aspire to it. Frankenstein can only create a single being, which is not even a whole human and which is intrinsically linked to himself for the rest of his own life. The events in both Frankenstein's and his monster's lives are linked throughout the novel: when Frankenstein destroys the monster's prospective wife, the monster murders Frankenstein's wife Elizabeth, and both Frankenstein and his creation die alone and miserable.

Ironically, the monster ultimately rises to a position of physical power far greater than that of his creator, which humans can never do in our relationship with the immortal Titan Prometheus, despite his defeat by Zeus. Finally, the reaction of the creator to the final creation differs dramatically between Shelley's story and the myth of Prometheus. Whilst Prometheus loves and cherishes his human race, giving us more power and knowledge, Frankenstein is repulsed by his monster, attempting at all costs to destroy it as soon as he has created it. Whilst the differences in the perfection of the creations is arguably a result of the differences in creative power between Frankenstein and Prometheus, the difference in reaction is also indicative of a difference in intentions. Whilst Prometheus intends only to bring happiness to humankind, Frankenstein creates purely for his own pursuit of knowledge and science, which is ultimately portrayed by Shelley as destructive and inhumane. The view of creation presented in the novel is also interesting due to its basis in only male characters. Again, only a man is responsible for the creation of the monster, and by removing the woman entirely from the act of creation, Shelley emphasises the essential role of the female in creation, as Frankenstein inevitably fails to produce a recognisable human. This reading enhances the overall message of Shelley's novel, that true creation by scientists is ultimately impossible, and Frankenstein remains constricted by the limits of human scientific endeavour. Indeed, whilst Shelley's chemist contemporaries created a vision of chemistry that seemed creative, the fact remains that even today, the so-called creations of chemists can only ever be different combinations of known elements and particles: chemistry is still the science of transformations rather than creations.

Through the destruction of the Promethean figure of the chemist and the reinforcement of the stereotypical figure of a mad scientist obsessed with a single unobtainable goal, Shelley in 'Frankenstein' presents an overwhelmingly negative view of chemistry. Chemistry is portrayed as a science which subverts natural order and religious convention, and by extension is no longer a 'natural' science. This destruction of scientific and intellectual ideals has been reiterated throughout modern science fiction and is no longer restricted to chemistry. So why was chemistry Shelley's specific choice to be vilified in this way? Percy Shelley, Mary's husband, was a keen chemist and enjoyed performing amateur chemical experiments at their home. Through him, and his close friend Lord Byron, Mary Shelley gained a wide, if not academically rigorous, knowledge of chemistry and her understanding in this field may have led her to include it in her writing. However, another interpretation is that Shelley's portrayal simply arose from a mistrust of new scientific discoveries in society – a common theme even now. The portrayal of morally impure, avaricious, gold-obsessed alchemists has morphed into contemporary sensationalist news stories about the lack of research behind global warming theories or medical advances. Science has always been perceived as a threat to everyday life, just as any radical change is a threat. Because scientists who make great discoveries tend to be those who, isolated, devote a lot of time to their studies, science in general is seen as a discipline which remains very much separate from 'normal' people. It is the domain of the intelligent elite and so is not trusted. Chemistry in the nineteenth century was making huge leaps in human knowledge with discoveries of the first elements and exciting experiments. However, for the first time, science was beginning to move far beyond the realm of public understanding and so the situation that remains today emerged. This essay essentially discusses how chemistry and the figure of the chemist interacted with popular culture in Shelley's time - through ghost stories, ideas of religion and common Romantic themes - and it remains clear that the same phenomenon will reappear in literature now and in the future unless a remarkable shift in public attitudes towards science, and a less insular scientific world emerges. In a world where science and crucially a trust in science is increasingly relevant and important again in our everyday lives, this shift seems now more essential than ever.

Biomimicking Superhydrophobicity

Navyaa Mathur

Superhydrophobicity, in layman's terms, is the inability of a surface to be wetted by water. This is due to its chemical composition and surface architecture, which together cause water falling upon it to form spherical droplets that roll or may even bounce off.

This enables superhydrophobic surfaces to remain clean and dry, in a phenomenon known as the 'lotus effect': water falling on the surface of a lotus leaf rolls to the base of the leaf, taking along dirt particles, and denying water-borne bacteria the chance to infect. This is especially important in water-surface plants like the lotus, as the stomata are on the upper surface, so it must stay dry in order to keep them unclogged and allow the free passage of respiratory gases.

Superhydrophobicity is also seen in the animal kingdom where it keeps butterflies' wings dry, enables insects such as the water strider to walk on water, and inspires the idiom 'water off a duck's back' - as the feathers of water birds tend to be superhydrophobic to avoid waterlogging.

This could potentially be mimicked to significant effect, from self-cleaning antibacterial surfaces to literally waterproof fabrics, from sieves that easily filter oil out of water to precise fluid flow control in lab-on-a-chip devices. There are, however, several issues that must be worked through before they can be used in the wider world, such as the fragility and stability of their surface architecture.

How can superhydrophobicity be defined quantitatively?

To understand the technical definition of superhydrophobicity, we first need to be able to quantify the interactions of a liquid with the surface it is on. When describing this, the surrounding gas is also relevant to the system's behaviour. This is due to the interfacial tensions between the phases at their surfaces, especially at the line of contact. Interfacial tension is the energy required to increase the area of the contact surface, and can be used as a broad term to describe interactions at the junction of two phases. Surface tension is therefore a category of interfacial tension, as it is the interfacial tension between a liquid and its own vapour. It is caused by the attraction of particles in a liquid to each other; those at the surface, as they are not surrounded completely by other particles, cohere more strongly to those next to and below them, creating a 'film' at the surface which can hold the inner liquid together, as in a raindrop. The shape of a mass of liquid is a playoff between gravity (and any other forces acting on it) and the action of surface tension to minimise the surface area to volume ratio. When the liquid is not in a container, this generally leads to the formation of a pool with a perceptible thickness and a curve to the surface. Surface tension is also responsible for the spherical shape of droplets, as spheres are the most efficient solids with regards to conserving surface area.

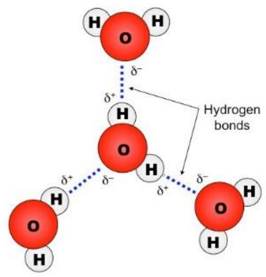


Fig. 1.

Water has an exceptionally high surface tension, because of the hydrogen bonding that takes place due to its chemistry. Hydrogen bonding is the unexpectedly high attractive force between molecules which have a hydrogen atom covalently bonded to a highly electronegative atom (giving the hydrogen atom a high partial positive charge), as well as a lone pair of electrons on another electronegative atom (giving the electronegative atom a high partial negative charge). These partial charges attract each other, giving rise to intermolecular forces as represented by dotted lines in Fig. 1. Every water molecule has two hydrogen atoms bonded to an electronegative oxygen atom, and therefore hydrogen bonding occurs.

This causes strong cohesion in water (and is also responsible for its unexpectedly high melting and boiling points). Intuitively, one would assume that a liquid with stronger cohesion and therefore higher surface tension forms droplets more easily, and this is in fact the case. This is a bonus for any attempt at hydrophobicity, which occurs when water molecules are more attracted to each other than they are to the surface, and therefore the intermolecular forces involved work against gravity to minimise the area of contact between the water and the surface.

The interaction between a liquid and a surface can be quantified by calculating the contact angle: this is the angle between the flat surface and a tangent drawn to the liquid at the point of intersection, as shown by Fig. 2. It can be calculated theoretically if the relevant interfacial tensions between the surface, the liquid and the surrounding vapour are known.

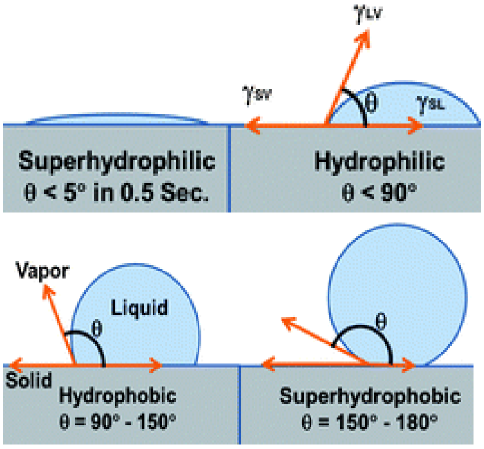


Fig. 2.

The static contact angle for a homogeneous surface is given by Young's contact angle equation:

$$\cos \theta = (\delta_{sv} - \delta_{sl})/\delta_{lv}$$

where θ is the contact angle,
δ_{sv} is the solid-vapour interfacial tension, or the surface free energy of the solid,
δ_{sl} is the solid-liquid interfacial tension, and
δ_{lv} is the liquid-vapour interfacial tension, or the surface tension of the liquid.

The static contact angle is achieved when a thermodynamic equilibrium is reached in the system. It provides a technical definition of superhydrophobicity, as a surface is classified as such when it exceeds a static contact angle of 150° with water.

Dynamic contact angles follow a similar principle but are measured when the liquid is rolling, as shown by Fig. 3, where *a* is the advancing contact angle and *r* is the receding contact angle. The rolling angle *t* is the tilt of the surface at which the liquid rolls off under gravity – significant when designing a surface which is intended to self-clean by allowing water to roll off it. The rolling angle for water is generally lower than 5° for superhydrophobic surfaces, and may even reach around 1°.

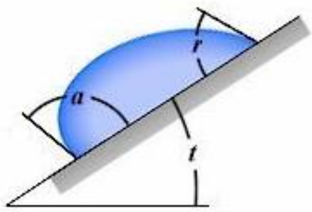


Fig. 3.

How is superhydrophobicity manifested in lotus leaves?

As mentioned before, Young's static contact angle formula works under the assumption that the surface is homogenous: flat and without a micro- or nano-structure that interacts significantly with the liquid. However, a lotus leaf has a multi-level surface architecture – hierarchical roughness – that contributes very significantly to its superhydrophobic ability. It was a serendipitous discovery made in the 1970s by Wilhelm Barthlott, a botanist who observed it under a scanning electron microscope while studying evolution.

The surface consists of tube-shaped wax crystals a few nanometers across, growing out of projections known as papillae that are a few microns across: essentially waxy bumps upon bumps that are excellent at keeping off the rain, as shown in Fig. 4.

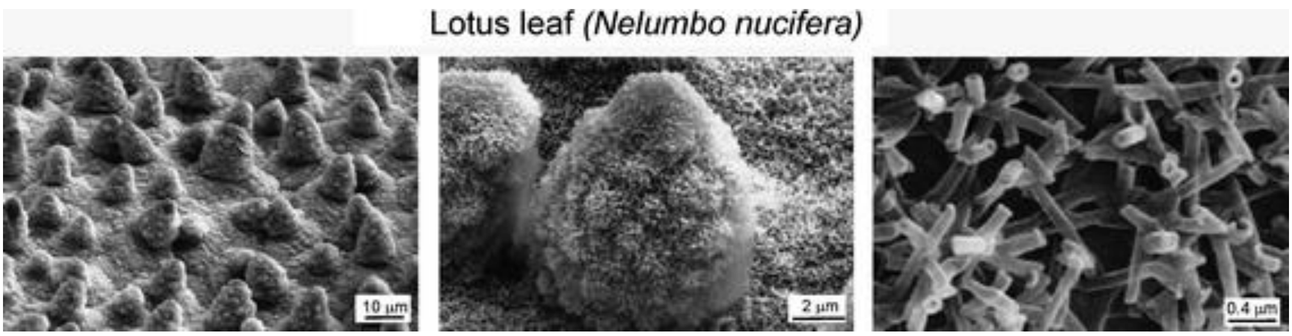


Fig. 4. Electron microscope images of the surface of a lotus leaf at different magnifications, showing the multilayer surface architecture

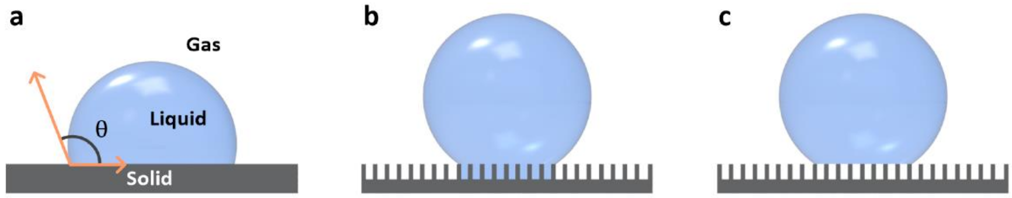


Fig. 5. a) Young regime b) Wenzel regime c) Cassie-Baxter regime

To describe this sort of surface, the Wenzel equation *or* the Cassie-Baxter equation may apply; both situations are illustrated by Fig. 5. The Wenzel equation is something of a modification of Young's original equation, and describes what happens when the liquid fills in the gaps created by the roughness. The equation demonstrates that the original tendency of the surface, either towards hydrophobicity or hydrophilicity, is amplified, as might be expected from the increased surface-liquid interaction due to the significantly greater surface area. This equation does not, however, apply to lotus leaves and similar surfaces, as it does not account for the effect of a trapped layer of air, meaning that the liquid interacts with a patchwork of wax and trapped air. In this situation the Cassie-Baxter equation is applied; it is a further extension of Young's equation which acknowledges the different surface-liquid interfacial tensions as well as the proportion of the overall area claimed by each surface. It therefore describes what happens when the liquid 'floats' on top of the roughness, which, due to the greatly reduced contact area, leads to a much lower adhesive effect regardless of the surface chemistry. In a real-life situation, a system might transition between the Wenzel and Cassie-Baxter regimes depending on the other forces acting upon it.

Papillae on the lotus leaf grow to different heights so that the tiniest droplets get caught on the tallest papillae, while larger drops nestle between them while still hovering above the shorter papillae, causing a deformation to the smooth surface of the droplet. This causes a further repulsive force due to the 'bumpiness' of the surface film: an irregular shape is far more inefficient and so the drop retreats to form a smooth curve, rather than wicking down.

These tubules are composed of epicuticular wax; its slippery texture is endowed by the long-chain (C_{29}) diols of which it is largely composed, and the long carbon chain is what causes the hydrophobicity (a diol is shown in Fig. 6). In all molecules, the electron density constantly shifts, resulting in the formation of short-lived partial charges which are far weaker than permanent dipoles such as those in water (which are already exceptionally strong, as mentioned earlier). These temporary dipoles nevertheless attract each other, resulting in cohesion even in entirely nonpolar substances such as elements that exist in a diatomic state. The effect increases with the number of electrons, as the temporary dipoles are more pronounced and slower to ebb. This explains why melting and boiling points increase with increasing carbon chain length: methane, with one carbon atom, is a gas at room temperature, while n-pentane, with five carbon atoms in its chain, is the lightest straight-chain alkane which is a liquid at room temperature, due to the increased number of electrons forming stronger temporary dipoles. However, temporary dipole attractions are nowhere near strong enough to compete with the hydrogen bonding between water molecules, so solubility in water decreases along with hydrocarbon chain length, and creating a surface out of such molecules means that water coheres far more than it adheres to the surface, causing an even higher contact angle.

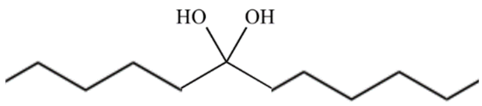


Fig. 6. A diol; the vertices of the hydrocarbon skeleton represent carbon atoms

Could nature have missed a trick?

The lotus leaf is undoubtedly a marvel of evolution, remaining permanently clean and dry in its watery habitat in monsoon-prone areas of the world such as India and Malaysia. However, the highest measured lotus leaf static contact angle of 163° has been improved upon in the lab: the record is around 179°, very close to the theoretical limit of 180°, at which the droplet would be a perfect sphere instantaneously in contact with the surface it rests upon.

The lotus leaf's hierarchical roughness does appear to be the optimal surface architecture for superhydrophobicity. Synthetic surfaces tend to mimic it closely, mainly tinkering with the surface chemistry to make it more hydrophobic. However, the most superhydrophobic surface in the world has uniform glass spike-like protrusions, unlike the lotus leaf with its varied papillae heights; the loss of the ability to deal with different sizes of droplets does seem amply compensated for.

In terms of surface chemistry, the lotus leaf's surface is largely composed of carbon, as might reasonably be expected of a waxy, nonpolar surface. More surprisingly, that there is also significant oxygen content (as hinted at earlier, when discussing the diols that make up epicuticular wax). This induces permanent dipoles in the surface – and therefore makes the surface more hydrophilic, as the partial charges on the water molecules are attracted to it. However, the hydrogen bonding effect does contribute to the high melting point of the wax, which is necessary in the tropical regions which are the lotus's natural habitat. Most of the alcohol groups are buried below the surface of the wax as they are not at the ends of the molecule, so do not necessarily hydrogen bond with the water – the hydrophobic $-CH_3$ groups at the ends are exposed.

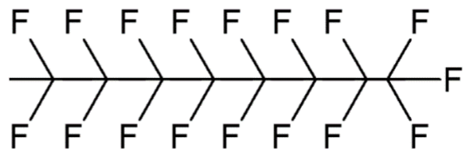


Fig. 7. A fluoroalkyl strand; the left edge of the strand might be bonded directly to the microstructure.

To increase the hydrophobicity of the surface, these groups can be replaced through a condensation reaction with a fluorosilane, which results in extremely hydrophobic fluoroalkyl nanostrands ($-C_nF_{2n+1}$) at the surface (as shown in Fig. 7). Fluoroalkanes form temporary dipoles even weaker than those in alkanes due to fluorine's electronegativity, which is the highest on the periodic table; the electron density around it is strongly held and is less liable to shift and form dipoles. While the lotus leaf does not have many alcohol groups at the surface, this reaction is important when synthesizing superhydrophobic surfaces from scratch, as will be explained later.

How could superhydrophobicity be used, and why isn't it widely in commercial use yet?

An environmentally significant potential use is the development of sponges and meshes which, due to being superhydrophobic but relatively oleophilic (attractive to oil), can separate oil and water. This may allow oil spills to be cleaned up without the need to burn the oil or to add even more foreign chemicals to the local ecosystem – it would be an extension of a technique already in use, which involves skimming the oil layer off the surface of the water.

Coating only some areas of a surface with a superhydrophobic layer could be used to precisely control fluid flow, which is important in lab-on-a-chip devices: scaled-down laboratory processes housed on a chip which can be used for synthesis and analysis of very low amounts of fluid. This could also be useful to minimise wastage in large-scale industry processes.

One issue is that the cost of fabrication for large flat superhydrophobic surfaces can be quite high as they need fairly extensive processing; a primary method of creating the glass microstructure is using a technique known as photolithography, which uses light to transfer a pattern onto the substrate. This technique is itself expensive, and only covers small areas of substrate, which then must be joined together, further increasing the complexity of the process. Another processing issue is that the reagent which forms the self-assembling nanostructure may be damaged during the reaction which joins it to the glass, reducing the superhydrophobicity.

The all-important nanoscale component of hierarchical roughness also tends not to be very stable physically as the nanostrands mat down and may then expose the more hydrophilic areas of the surface, whereas the rigid wax crystals on a lotus leaf do not present this issue. Rigidity in the nanostrands is therefore essential but potentially hard to achieve, when also considering chemical stability: filling the chain up with rigid double bonds would lead to a very reactive surface, highly vulnerable to addition reactions which might significantly reduce the superhydrophobicity (depending on what the attacking agent was); this would entirely defeat the purpose of the double bonds.

In a similar vein to this, localised high water pressure may eliminate the trapped air layer, which would cause the surface to transition from the Cassie-Baxter regime to the less hydrophobic Wenzel regime. Pressure due to rubbing may damage the surface architecture, which, unlike a plant, is not self-healing – the wax that makes up the epicuticular layer passes through the cuticle along with respiratory gases, and so is constantly replenished. Oils and surfactants would have the effect of reducing the surface tension in the water, making the water less likely to cohere into a sphere.

An innovative way to solve this problem is through the use of superhydrophobic nanoparticles. The superhydrophobic reagent itself is not very expensive or rare: part of it can be produced in large quantities from seawater, or more precisely from the skeletons of unicellular algae called diatoms which are made of hydrated silica. The surface architecture is the desired one for superhydrophobicity, but the silica is hydrophilic and is treated with a fluorosilane reagent to make it superhydrophobic. This forms a powder known as super-hydrophobic diatomaceous earth.

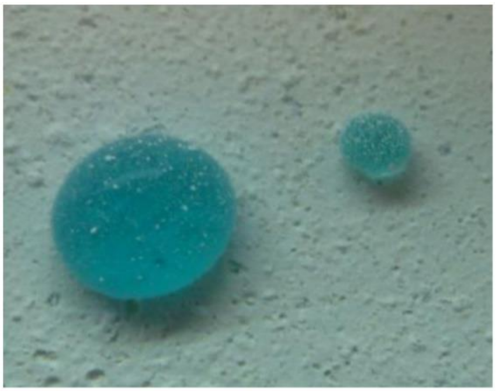


Fig. 8.

A variation on this theme is to make only one side of the nanoparticle superhydrophobic and react the other side with a superhydrophilic reagent instead. The resulting nanoparticles can be used to create water marbles, shown in Fig. 8. These are balls of water that are self-contained and are themselves superhydrophobic due to being held together by the superhydrophilic side and further surrounded by the superhydrophobic side. They undoubtedly have a wow factor but practical applications, beyond demonstration of the concept, do not spring to mind. However, the use of nanoparticles solves the issue of durability, as they are 'volumetric' – a layer of nanoparticles is superhydrophobic all the way through, so slight abrasion or corrosion is not an issue.

Another way to alleviate the manufacturing problems is by applying the superhydrophobic surface as a paint or similar product which does not need fabrication in the same way that a pre-prepared large surface does, and is also volumetric but less vulnerable to movement. Example of these are the commercially available StoLotusan render, intended to be used over normal paint on buildings in order to make them self-cleaning, and the Nanex Always Dry spray for shoes. These work by using a superhydrophobic powder in a solvent which, when it evaporates, leaves behind a layer with micro- and nano-porosity, which is the required finish for hierarchical roughness – mixing them directly with a paint or dye would not work because the non-evaporating component of the other product would fill in the porosity, rendering it useless.

Products like this could be extremely useful in hospitals, laboratories and any other place where sterile conditions are crucial, as the surfaces are antibacterial and self-cleaning when water is poured over them, saving time and money. This will be made possible when production challenges are fully overcome, allowing us to take advantage of the lotus's miraculous ability.

Increasing Internet Coverage as the Vehicle to a Better World

Senkai Hsia

In a dusty hanger in California, a group of engineers from Google are working on an idea that could revolutionise life for people in rural and remote regions: to give them access to the Internet. Project Loon's aim is to address the 60% of the worldwide population who currently live in areas without cellular infrastructure, in the hope that they can use the Internet as a medium to gain knowledge and improve their lives. The awareness, education and economic benefits that the Internet can universally provide in all languages and in all locations, will be essential in addressing the plethora of challenges that current and future generations will face. Therefore, access to the Internet must, and is being expanded to reach those who do not yet have these opportunities, and to prepare them for the future.

The most critical barrier to a better future is a lack of awareness about the challenges of today. Without knowledge of the contributing causes and detrimental impacts of issues such as climate change and anti-bacterial resistance, bad practices will continue to occur. This is particularly true in developing countries which have a higher proportion of the population without Internet access, and are more likely to follow harmful health and environmental policies. The knowledge that the Internet provides can solve these problems. A pertinent example is India's rural agricultural industry, where poor productivity in wheat and rice has led to perpetual cycles of poverty, increased river pollution from pesticide runoff, and the unsustainable extraction of groundwater for irrigation. One of the main causes has been a lack of awareness from farmers of modern agricultural practices, with high illiteracy rates preventing the acquisition of such knowledge. In response, Ashok Khosla stepped in and founded TARAhaat.com with the aim of educating farmers in rural areas about more efficient cultivation methods online. With access to the Internet provided by cyber-kiosks from local entrepreneurs, rural farmers can use TARAhaat's 'Agri-advisory', which contains an extensive database on crop types, herbicides, information of financial loans for farmers, and a phone helpline. This new information provided by increasing internet coverage to rural areas has empowered farmers to take steps to increase their yields and to mitigate infectious diseases afflicting their crops. This in turn helps to reduce their impact on the environment by decreasing the quantity of natural resources required for a successful harvest. Through resources such as TARAhaat, the Internet can be used as a powerful tool to increase awareness for better practices that are more financially and environmentally sustainable.

A major reason why environmental degradation continues to persist is that the causal factors, like logging and pollution, are often the cheapest available option for consumers or the most lucrative for organisations. It is only when people have more money that they are able to afford more expensive medical supplies like antibiotics or sustainable eco-friendly products. The Internet can promote the social mobility of the poor in two ways: facilitating education and business. Education is essential in responding to future challenges as it allows more people to gain knowledge and the skills to develop solutions to problems. But more importantly, it lifts people out of poverty to better face these challenges by allowing them to gain better employment. In recent years the Internet has become a key tool to enable free, accessible education to people across the world from all social backgrounds. There is no better example than Khan Academy, which reaches 10 million students with 3,400 short instructional online videos in 42 languages on topics from Differential Equations to Gastrointestinal Physiology. This extensive resource allows children to educate themselves up to secondary school level in Maths and Sciences no matter their background, location or language: all they need is access to the Internet. Take the story of Moawia Eldeeb, an immigrant from Egypt who was forced to drop out of high school in New York to provide money for his family. Through being able to access the Internet at a homeless shelter, Eldeed was able to teach himself Mathematics using Khan Academy and this allowed him to gain entry to Columbia University and set up his own company. If his success story is repeated in places that do not currently have Internet access, it will lead to many gaining greater education and being liberated from poverty.

With greater awareness and education provided by Internet access comes greater economic prosperity. Aside from gaining greater global knowledge of problems and higher levels of education, access to the Internet also facilitates business opportunities in rural remote areas. Eduardo who lives in Neltume Chile, has greatly benefited from the 'Todo Chile Comunicado' governmental program to provide low cost 3G wireless connectivity targeting 1,400 rural locations with high poverty levels. He is now able to use the Internet to promote his artisanal woodwork to regions outside the local area by using email to conduct transactions with clients and to present his products at regional fairs. This has led to increased income that, without using the Internet, would not have been possible. What's more, now his children can use the web as part of their classroom learning. The Internet thereby improves the lives of the poorest by improving their awareness of global issues, enabling their education and enhancing their financial opportunities, allowing them to be better prepared for the problems of the future. From climate change to scarcer natural resources, the Internet facilitates more effective and wider responses from everyone worldwide.

With such great benefits, the issue now is how to deliver the Internet to the 60% without access. The problem lies in the remote locations without cellular infrastructure that would not be economically viable for governments or private companies to invest in. The response has been spearheaded by innovative technology companies like Facebook and Google. At Moffatt Airfield, California, Google engineers are working on Project Loon, which derives its name from both the idea itself, and how ludicrous it seems. Loon is using a fleet of special high-altitude helium balloons 20 miles up in the stratosphere to transmit the Internet to users on the ground. What started out as a 'Harmless Science Experiment' in 2012 has now developed into tennis court sized sheets of polyethylene with miniaturised transceivers surviving up to 100 days in 100km/h winds and -90^{o}C temperatures. The idea is to use predictive computational models of the weather to guide the balloons on an airborne conveyer. This enables wireless coverage to be provided over a large area by using different strata of wind speeds at different altitudes to guide the balloons into a constantly cycling network. With the balloons being consistently launched by a custom 'auto launcher' and recycled after recovery, it provides a cheaper alternative than building ground infrastructure to users in isolated mountain regions or over very long distances. Initial testing has been successful, with Google signing contracts with Sri Lanka and Indonesian telecoms companies for deployment within the next two years. Meanwhile, Facebook has opted for both drones and satellites - though both projects are several years from fruition. The prototype for a future fleet of Internet-beaming drones made its maiden flight in July 2016, but Facebook's first Internet satellite was unfortunately destroyed with the SpaceX rocket explosion in September. While their public goal may be virtuous and have many benefits for the recipients, Google and Facebook are pushing for the Internet's spread to new areas primarily to increase the range of their advertising markets. But it is through the innovation of such companies and the co-operation of governments that social mobility can be enabled. This ultimately allows everyone to be in a better position to respond to the global problems of tomorrow.

Therefore, increasing the reach of the Internet to people of rural and remote areas is the most instrumental idea of today that will best prepare us for the challenges that lie ahead. To effectively address any global issue of the scale of climate change, population growth or infectious disease, greater awareness from everyone who could be affected is required. As such, people require education to understand and develop their own novel solutions to these problems. The Internet is the vehicle to enable this change, as it can uniquely provide these benefits for everyone, no matter their background or upbringing. As Project Loon has shown, it is the innovation of today that ultimately enables the innovation of tomorrow.

Neuromorphic Computing: Reaching Superintelligence

Alessandro Esciua-Blanco

Reaching a level of intelligence in computer simulations on par with humans or even less complex animals has long proved difficult. Firstly, replication of the neurological processes that encompass who we are and what we can do is simply impossible, since each of us are in possession of billions of neurones and trillions of synapses between them. We can see neurological activity on a macroscopic scale incredibly simply nowadays by recording electrical activity across one's brain using sensing technology but understanding the actual processes is clearly a task far beyond any sort of comprehension we can yet establish.

Why is this so? Well, one may refer to the stadium analogy: it is easy to tell which parts of a crowd are jeering, cheering, or silent, but, if you want to find out the state of a particular person in the stadium, it's nearly impossible to tell through listening unless you are right next to them. Perhaps it's the same here; we can see which parts of the brain are active on a large scale but cannot find individual interactions between neurones. Knowing which parts of the brain are engaged at a moment in time is not enough to understand it - to do that we must map connections between neurones on an enormous scale. In the past, this was done for the tiny *C. elegans* worm, which has a mere 302 neurones. This worm's connectome (a complete representation of its nervous system) can be seen below in Fig. 1. Take a moment to try to comprehend the distinct relationships between neurones, and see how manageable it is.

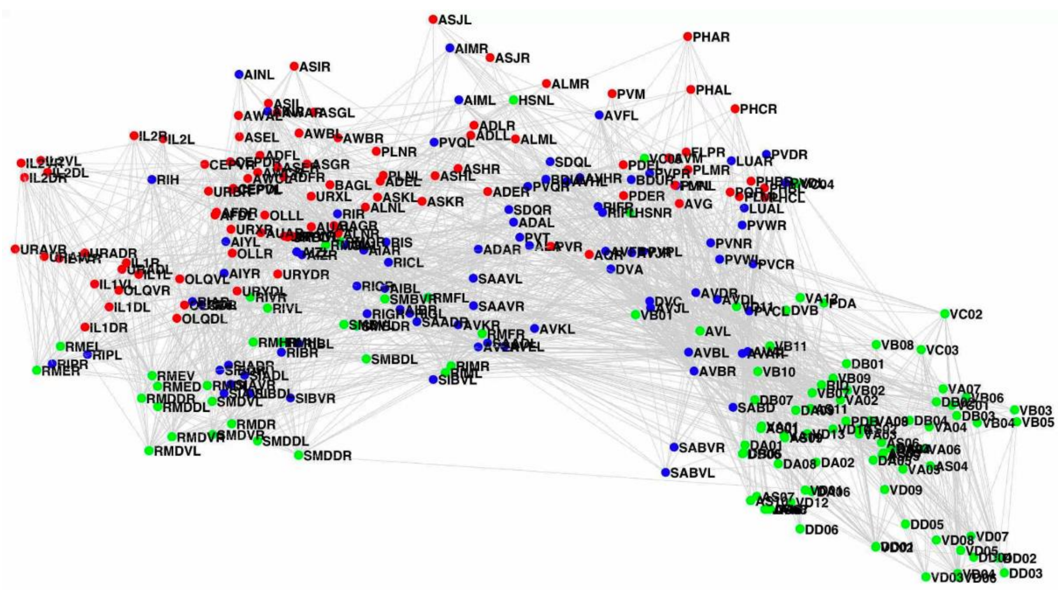


Fig. 1. The *Connectome* of a *C. elegans* worm. Each point represents a neurone, while the lines represent synapses.

Now think about the scale we were talking about earlier. The human brain has *billions* of neurones, with *trillions* of connections. That is at least 10^8 times greater than the near-indecipherable mental map in Fig. 1. Of course, so many more neurones results in a vastly greater number of connections between them. It is perhaps not a matter of theory - as we have a vague idea of how the brain works on the micro level - but rather a question of the scale and power involved.

One way to understand the trillions of neural interactions would be to develop an extremely powerful supercomputer, essentially designed to replicate point-to-point interactions and connections with weightings established for each synapse as a function of previous interactions, and with responses adjusted for each situation.

However, a far more organic way to go about this process would be to map and replicate brains of increasing complexity, from mice to monkeys, with the ultimate goal of simulating the human brain. Either way, extensive research needs to be carried out, not only to find out which neurones are associated with instinctive reactions/reflexes, but also the pathways involved in more sophisticated thinking. Alongside this, we would of course require more advanced neural imaging technology than is currently available. The stadium analogy explains this: until we can find a method to accurately map neurones' interactions without relying on area-based activity, we simply cannot replicate the human brain in detail. Of course, this is a limitation that must be overcome, but it is certainly conceivable.

Through comprehension and, by extension, simulation, we could essentially build an artificial form of the human brain. This alone would be a fantastic achievement, a sort of artificial intelligence which we would have total and unrestricted control over. It could be enhanced, tailored and specialised to a level of intelligence above any previously known to man. Imagine a modern Beethoven for whom we could adjust the connectome to compose countless more beautiful pieces than he ever had the time or even ability to dream up, or a physicist comparable to Einstein - with his preserved brain, one could even bring him back to life in a virtual form.

It sounds like the stuff of sci-fi, yet it is not entirely implausible. With brain simulations, we could develop a mind with a sort of intelligence unachieved in history. On a larger, more human scale, we could upload our minds to a computer. Of course, everything described so far would require huge progress in neurological mapping, but it is simply a matter of scale. Perhaps brain-tailored supercomputers could even analyse data from living beings, enabling them to read neurones and synaptic strength through mapping activity associated with different tasks, creating a neurological clone.

Implanting concepts into virtual minds could eventually become a reality, since everything in the mind comes down to neurone interactions, chain networks and hierarchies. Say one simply enables or implants a neural network into a simulated mind - it could be reprogrammed in little to no time. While this has enormous potential in quickly implanting knowledge into countless people, there is of course the potential for abuse of such systems vulnerable to hacking.

However, the possibilities do not end at enhanced human brains, but rather at super-intelligent brains. Since all brains are structured similarly, super-intelligence could stem from developing neural maps more complex than our own, perhaps even by using the artificial brains described above. Such 'minds' would allow us to reach unparalleled levels of understanding in any topic we wish.

A contemporary development in this emerging field is the invention of neuro-synaptic chips: computer chips developed with the model of a brain in mind, allowing them, within constraints, to process information in the style of a brain. This has been explored by IBM to great length, with the rough connectome mapping of a Macaque monkey's brain (Fig. 2) leading to the development of computer chips based on such interactions.

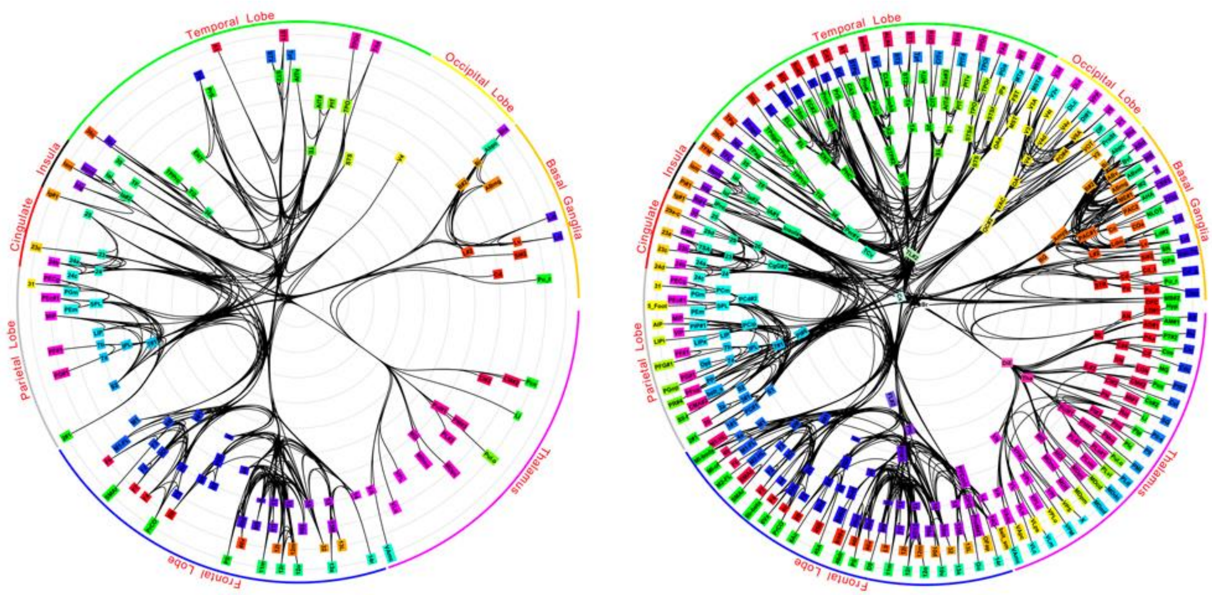


Fig. 2. A rough model of the connections within a Macaque monkey's brain developed by researchers at IBM and DARPA, used in the development of neurotrophic chips such as TrueNorth.

TrueNorth is the most recent chip, developed in 2014 with a total of 1 million neurones and 256 million programmable synapses. It uses far less energy in comparison to traditional computers rendering similar simulations, and can harness greater processing power and speed. This represents a distinct move away from the von Neumann architecture that has dominated computing since its inception.

With the enormous potential of neuromorphic computing, it's easy to get lost in the huge advantages it promises to bring. For example, there are countless ethical issues one must tackle regarding the matter, not only regarding the standard ethical debates on artificial intelligence and synthetic sentience, but also regarding whether demeaning one's brain to a virtual form is morally palatable (alongside other debates). By extension, there are huge dangers that such developments could entail - for example, through hacking.

Despite all this, the promise of brain-simulation computing has limitless potential: with researchers now dedicating huge resources towards it I believe it is inevitable that comprehension of the brain will be achieved by the end of this century (likely earlier), enabling us to develop complex simulations and use neuro-synaptic hardware. The knowledge gained will enable the development of robust AI, likely operating at higher degrees of intelligence and efficiency than humans.

Of course, widespread usage would require greater computing power, a problem that will almost certainly be solved by huge progression in circuitry and processors. That would, by extension, lead to an alternative to traditional silicon-based processing chips. Carbon nanotubes are proving to be a prime candidate (barring their current restrictions with regards to manufacturing costs).

One advancement this carbon medium does bring is greater miniaturisation than silicon; the thinnest semiconductors would be able to be under 1 nanometre, an order of magnitude smaller than what we have today. In addition, CNTs exhibit ballistic conduction, enabling conductivity to surpass that of similar semiconductors such as silicon, and enabling scientists to harness quantum mechanics to their advantage.

In conclusion, brain-focused computing via hardware and implementation has enormous potential, one that I believe far exceeds the most complex neural networks that can be reinforced through gradient descent. While other forms of artificial intelligence loosely model their processing on the neurosynaptic action of our brains (such as neural networking), true AI would be a genuine rendering of the thought processes that encompass our every action, decision and thought. This, I believe, will be an enormous progression not only for computing, through neuromorphic architecture, but also in reaching the ultimate goal of synthetic artificial intelligence - with whatever lies beyond.

A Day of Birdwatching

Lorna Bo & Brandon Tang

10.14am: The first bird of the day is spotted – an inquisitive pigeon (*Columba livia domestica*). It waddles around the bins outside the station, and is eventually scared away by a stray commuter.

10.12am: A trio of *Columba livia domestica* is observed, teaming up to defend their hard-won territory of Upper Crust tables.

10.28am: Another trio assemble, this time consisting of Mr Moore, Mr Law, and Mr Ullathorne (left to right). Having asserted their dominance, they lead the rest of the flock onto a c2c train headed for the dark depths of Essex.

11.14am: The train arrives at Purfleet, and the dominant trio leads the flock to the River Thames.

11.51am: The trio finally lead the flock into Rainham Marshes, expectantly looking about them for signs of birds.

11.52am: The author gets bored and starts photographing leaves.

12.08pm: A large unidentifiable bird is seen to be digging into the earth using powerful claws.

12.31pm: Mr Moore steals a fellow birdwatcher's telescope.

12.38pm: Flocks of hundreds of lapwing (*Vanellinae)* fly overhead.

12.40pm: A cormorant (*Phalacrocoracidae*) poses amongst a crowd of common gulls (*Larus canus*) and shelduck (*Tadorna*).

12.46pm: A cow (*Bos Taurus)* observes the invading birdwatchers, terrified.

12.55pm: The flock takes a brief respite for lunch in the incredibly named 'Shooting Butts' hide.

1.33pm: Mr Moore and his beloved telescope.

1.36pm: A mute swan (*Cygnus olor*) performs strange gyrating head and neck acrobatics.

1.39pm: A coot (*Fulica*) tries to drown itself, to no avail. Another coot looks on.

1.42pm: A wild Brandon photobombs an otherwise perfectly captured Mr Ullathorne.

1.44pm: A cormorant (*Phalacrocoracidae*) rises from a failed hunting dive.

2.05pm: A flock of greylag geese (*Anser anser*) passes overhead, honking aggressively.

2.15pm: A kestrel (*Falco tinnunculus*) pauses overhead, presumably watching its prey.

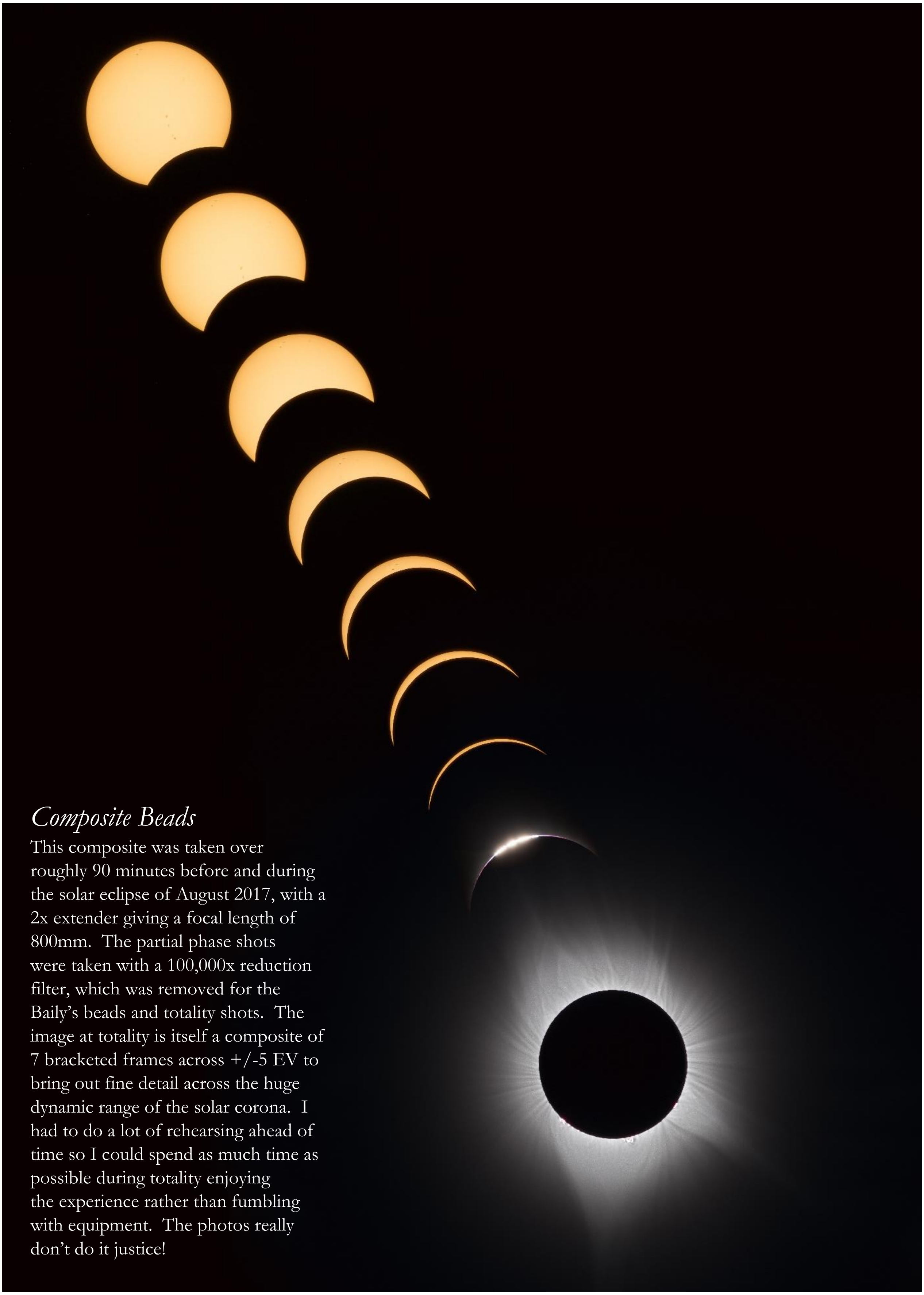
2.18pm: A robin (*Erithacus rubecula*) in the moment before flight.

2.34pm: A pheasant (*Phasianus colchicus*) coyly obscures itself behind a tree.

2.51pm: A group of teal (*Anas crecca*) gather together on an island.

Astrophotography

Jeff Tomasi



Composite Beads

This composite was taken over roughly 90 minutes before and during the solar eclipse of August 2017, with a 2x extender giving a focal length of 800mm. The partial phase shots were taken with a 100,000x reduction filter, which was removed for the Baily's beads and totality shots. The image at totality is itself a composite of 7 bracketed frames across +/-5 EV to bring out fine detail across the huge dynamic range of the solar corona. I had to do a lot of rehearsing ahead of time so I could spend as much time as possible during totality enjoying the experience rather than fumbling with equipment. The photos really don't do it justice!

Moon

The nearly full moon in this shot provides plenty of light, so I could push the focal length out by stacking both a 1.4x and a 2x extender to create a 1120mm f/16 lens. I set up an intervalometer to take 120 exposures over the course of several minutes to allow mechanical vibrations to dampen out between frames, but even tiny puffs of wind or atmospheric distortion can blur the image at such a long focal length. I discarded the 40 worst frames and stacked the remaining 80 to average away any pixel noise, which allowed me to really push the contrast and saturation without the final result looking grainy.

Andromeda

The Andromeda galaxy is a very large object in the sky, about 6 times as wide as the full moon, so a 400mm lens alone has it filling most of the frame. The problem is that it's also very faint, so I needed to use a tracking mount to follow the rotation of the earth over the eight separate 2-minute exposures that make up this image. These frames were combined with dark frames taken with the lens cap on to eliminate sensor noise, and then stacked and heavily stretched to bring out the still faint spiral arms. It's impressive what you can see without a telescope!

Obituary: Stephen Hawking

Raghav Nayak

How did everything originate? What goes on "out there"? How do we reconcile what happens at the microscopic level with the deceptively profound concept of gravity? In the eyes of many scientists, it was Stephen Hawking who put forward the most compelling theories in response to these inexhaustible questions. From his discoveries concerning black hole radiation that have now become synonymous with his name, to assuming the role of one of the world's most pre-eminent ambassadors for science at large, Hawking's impact will forevermore shape discussion in the scientific community. While he may no longer be a part of this world, his legacy is nothing short of, in the truest sense of the word, universal.

Hawking is perhaps best known for his ground-breaking work on general relativity and black holes. Perhaps somewhat counter-intuitively, both of these concepts were nothing especially novel at the time. The former had famously been postulated by Einstein shortly after the turn of the century, while the latter had been explicated in considerable detail by Oppenheimer in 1939. But Hawking's genius was in further unifying these two fields and carrying them through to their inevitable conclusion. While a black hole, on account of the pull of its own weight, is formed from the collapse of a very massive star, whereby all of its mass collapses into a single point of ostensibly infinite density (the singularity), Hawking, along with Penrose in 1970, realised that, in a similar vein, general relativity implied the universe beginning from a point of singularity and expanding thereafter, in a sensational "Big Bang". While Darwin's "natural selection" theory may have explained how we, the human species, came about, Hawking's theory traced the origins of quite literally everything around us, from the atoms in our body to galaxy clusters light years away.

Hawking's potential showed itself once again in his direct discoveries concerning black holes. While the Big Bang could be very superficially described as the "inverse" of the formation of a black hole, Hawking could never content himself with such a sparse association with these bewildering phenomena. In the early 1970s, his "eureka" moment kick-started a series of discoveries that would transform our understanding of the workings of the universe: the observation that entropy, the state of disorder in a system, was inextricably linked to the increasing surface area of a black hole's event horizon (the so-called "point of no return", at which not even light can escape the black hole's clutches). There was just the nagging problem that certain elements of this theory violated the infamous Second Law of Thermodynamics, which consequently filled the entire theory with controversy.

Prompted by Bekenstein, a fellow, albeit markedly junior, physicist, who had taken Hawking's theory further to proclaim that the event horizon was in fact itself a direct measure of the black hole's entropy, thus resolving the Second Law, Hawking's quest to disprove this equally controversial proposition ironically led not only to a confirmation of Bekenstein's results, but also to another colossal discovery from Hawking. Since, via Bekenstein's theory, black holes possessed entropy, this meant they also had had a temperature above absolute zero, which in turn signified that they had to be radiating heat. At first, this would appear an inherent contradiction: if nothing can escape from the event horizon of a black hole, then how can it radiate heat? The answer, discovered by Hawking, transformed cosmology and quantum mechanics forever.

At the quantum level, particle-antiparticle pairs are consistently popping into existence, and almost instantaneously after that, right back out of it. However, within the context of an event horizon, Hawking made the key observation: these two particles become separated, and owing to the strength of the gravitational pull, cannot re-join. A particle is thus able to escape from the black hole, in a process aptly named "Hawking Radiation". Thus, Hawking was able not only to develop a far more substantiated version of his original theory, but also to validate it in relation to the thermodynamic axioms which had originally hindered him. Because this discovery opened the floodgates for a new field of physics, Quantum Gravity, physicists by and large consider this Hawking's greatest achievement, even more so than the infamous Big Bang theory. It is worth noting that there is still no observational evidence to confirm such radiation - but, testament to the immense credibility of Hawking, his theory has become the overwhelming scientific consensus on the matter and virtually unchallenged since.

Of course, it doesn't take a quantum physicist to comprehend his impact for science. His now world-famous book, A Brief History of Time detailed his wide-ranging research to an international layman audience. It has now sold over 10 million copies and been translated into 40 languages. In the public sphere, this is arguably where he will most resonate: his impassioned advocacy of science will have inspired the next generations of scientist to achieve their full potential. At least during his lifetime, science gained a great deal of the much-needed respect it deserves among the media thanks to paragons of the subjects like him. For his sake and the sheer sake of science, we must not let denial become the new asbestos. In 2018, it is people like him we now need more than ever to take on the mantle of safeguarding the principles of scientific research and observation.

Moreover, it doesn't pay to imagine what being told, at the ripe age of 21, that you have no more than a couple years left in your life, feels like. It speaks volumes not only for the genius but also for the character of a man that, after being paralysed and latterly unable to speak without the aid of technology, he was able to accomplish so much with such inhibiting limitations. While his determined and even obstinate persona often meant he was slow to discard his theories that did not work, when the right time came, he was always willing to admit his errors. In 2012, when Hawking conceded an initially heated bet to Peter Higgs that the Higgs boson would never be directly observed, he did so with the utmost grace, before insisting that his colleague and former rival should be awarded the Nobel Prize for Physics, which Higgs promptly was. In the same spirit of this determination, Hawking never let any of the obstacles life threw at him get in his way – and the world of science has benefitted enormously as a result of his courage. How fortunate for us all that he survived ALS for over half a century - his life was certainly one lived to its fullest.

Stephen Hawking's ashes will be interred in Westminster Abbey, beside the graves of Isaac Newton and Charles Darwin, in a thanksgiving ceremony to be held later this year.

Below is the equation Hawking discovered with the assistance of Bekenstein. It was the one of which he was most proud and, in line with his will, it shall be the one which adorns his grave:

$$S = \frac{\pi A k c^3}{2hG}$$

The Hawking-Bekenstein equation, which is now commonly referred to as the Second Law of Black Hole Thermodynamics. This equation made the fundamental link between the change in the entropy (state of disorder) in a black hole and the increase in the area of the event horizon. This equation is especially beautiful scientifically as it combines several of the most important constants in physics into a meaningful result. In particular:

S = Entropy – more specifically, the quantity of entropy required such that the black hole conforms to the laws of thermodynamics relative to external observers.

A = Area of the Event Horizon

k = Boltzmann's constant – more precisely, $1.38064852 \times 10^{-23}\ m^2\ kg\ s^{-2}\ K^{-1}$. This relates the average kinetic energy of the particles in a gas to its temperature (i.e. this is the conversion factor from Joules to Kelvin).

c = The speed of light, $299{,}792{,}458\ ms^{-1}$

h = Planck's constant – more precisely $6.62607004 \times 10^{-34}\ m^2\ kg\ s^{-1}$, which gives the ratio of the energy in one photon of electromagnetic radiation to the frequency of that very radiation.

G = Newton's constant – more precisely $6.67408 \times 10^{-11}\ m^{-3}\ kg^{-1}\ s^{-2}$. It is used both in calculations of the force between two objects owing to gravity and in Einstein's theory of general relativity.

And a few iconic photos of the man himself:

The image that Hawking has become so well-known for. His contraction of motor neurone disease in 1963 left virtually his whole body paralysed. In 1985, he contracted pneumonia and was forced to undergo a tracheotomy to save his life, resulting in the loss of his already limited speech. The photo epitomises what Penrose aptly proclaimed in his eulogy to Hawking: **mind over matter**.

Stephen Hawking, 23, and Jane Wilde, 21, on their wedding day in July 1965

Hawking had stated in a BBC interview in 2007 that one of his greatest yearnings was to experience flight in space. As a result, Richard Branson offered Hawking the opportunity to experience this sensation on one of Virgin Galactic's cutting-edge reduced gravity flights. The expression on his face says it all.

And finally, a short extract from his *A Brief History of Time,* in which Hawking encapsulates the underlying purpose of the work he does:

"The discovery of a complete unified theory may not aid the survival of our species. It may not even affect our lifestyle. But ever since the dawn of civilization, people have not been content to see events as unconnected and inexplicable. They have craved an understanding of the underlying order in the world. Today we still yearn to know why we are here and where we came from. Humanity's deepest desire for knowledge is justification enough for our continuing quest. And our goal is nothing less than a complete description of the universe we live in."

An Exploration of Chronocomputation

Isky Mathews

--the end.

Time travel has been dreamed about for many decades – some, such as H.G. Wells in The Time Machine, have considered what it would be like to travel into the far future to discover the nature of humanity's demise or into the past where we could meet Robert Hooke to discuss biology before breakfast, eat lunch with Richard Feynman over conversation about QED and perhaps dinner with Carl Gauss, if he would have the time or be *interested*. Of equal intrigue have been the many discovered paradoxes an unthinking time traveller could cause, with many arguing that this shows the impossibility of the very notion. This article will examine possibilities for computation that time travel could bring: a rarely considered but intriguing field of research.

To understand how time travel can speed up the computation of problems, we need to first understand something about the limitations of normal, *classical* computers. In modern computer science, computational problems can be classified and analysed based on how long they take and how much space they need to be computed by a classical computer (a *Turing machine*). For example, the problem of detecting whether a given number n is odd or even takes a constant number of steps no matter the magnitude of n since we can use an algorithm that checks whether the last digit is odd or even, while sorting a list of n numbers into increasing size order takes roughly n^2 steps using the "naïve" algorithm

1. *In the kth iteration of this algorithm, find the least element of the remaining $(n-k)$ elements.*
2. *Put that at the end of the new sorted list the algorithm is creating.*
3. *Repeat, with 1 added to k's value.*

This can be seen from the fact that for the 1st step, the algorithm performs $(n-1)$ comparisons of items in the list to find the smallest element of all n elements, then in the 2nd step, the algorithm only has to look at $(n-1)$ items and so performs $(n-2)$ comparisons etc. so that the total number of comparisons (steps) the algorithm performs is

$$(n-1)+(n-2)+(n-3)+\cdots+2+1=\frac{n(n-1)}{2}$$

which, computer scientists say, for larger and larger n essentially grows *on the order of* the simpler function n^2, which is notated $O(n^2)$.

The essential idea behind this is *computational complexity:* when we refer to a *problem* here, we are actually talking about an infinite class of such problems of different sizes where each *instance* of the problem can be referenced using a value n, so that n represents the number of pieces of information needed to describe the situation. The *time complexity* of the problem is a function of *n* that describes the number of steps required to solve an *n*-size instance; another complexity measure is *space complexity*, which describes the number of values that need to be stored by an algorithm.

The 4th ARM article in *The Librarian*, which can be found on its website, explained parts of this in greater detail but overall, there are a few classes within which we can place most problems (the study of these classes is known as *computational complexity theory*):

P – the set of problems whose time complexities are polynomial (or smaller) functions of n; this makes up many everyday processes used in phones, computers etc. and are considered *efficiently solvable problems* due to the speed of modern day CPUs.

$EXPTIME$ – the set of problems whose time complexities are exponential (or smaller) functions of n; this class includes problems such as deciding whether a given move in $n \times n$ (generally, for $n \geq 8$) chess is optimal.

L – the set of problems whose space/memory complexities are logarithmic functions of n.

$PSPACE$ – the set of problems whose space/memory complexities are polynomial functions of n.

NP – the set of problems such that the process of verification for an answer to an n-size instance has polynomial time complexity; this includes large numbers of problems such as the Travelling Salesman Problem, where the goal is to determine whether there is a path through n cities (with defined distances between each pair) that has length smaller than k.

There are many more whose meanings can be guessed from some knowledge of nomenclature: the suffix "X-**hard**" denotes the set of problems whose time complexity is *at least as great* as those in class X, "X-**complete**" denotes those problems in "X" and in "X-hard", and "$XSPACE$" denotes those problems whose space complexity is an X function of n, where X is an abbreviation for some well-known class of functions (such as EXP or P). It should be noted that, definitionally, a problem in a given class will also be in all slower classes.

From this, we can understand the capabilities of modern computers: some things, from multiplying large numbers to primality checking can be done "efficiently", whilst others (especially the elusive NP-complete problems) can take far, far longer. Let's take the example of chess – imagine there's a game going on at a chess camp and a grandmaster looks down at a move taken by a novice and says "Ah. That was a mistake – the game is lost now." To determine truly whether that was a *good* or *bad* move, we would have to look at the game and consider all the possible moves that the novice could have taken (roughly 30 on average) and then, for each of those, consider all the possible retaliatory moves their adversary could take and then for each of those, the moves the novice would in response *and so on* until in each of these possible games on the *game tree* we reach either a stalemate, a checkmate for white or a checkmate for black. Then, we could say for each move the novice could have taken back in the original game whether, with perfect play, the novice could force a win – *these moves would be optimal*. However, the number of games that need to be examined grows by thirty-fold each level down the tree, meaning that from the beginning of a chess game it would be necessary to examine more than 10^{120} possibilities which would take far longer than the lifetime of the universe for even the combined computing power of humanity.

This is where "chronocomputing" might help. The first and perhaps easiest way to do so would be to start your computer running with a normal program and then time travel into the future by an upper bound on the time it will take to finish and collect your results. In fact, you might think, this doesn't need a "forwards-or-backwards" time machine, all we need is to take a trip round the solar system at close to the speed of light and thanks to relativistic time-dilation, time will speed up enough to finish the computation quite soon. However, it turns out that for a problem of a given time complexity, let's say exponential, you would be required to get exponentially close to the speed of light and thus need exponential amounts of fuel to get to that speed – but it would be necessary to burn all the fuel to get to the required speed and that would, in of itself, take time equivalent to some exponential function of n. Thus we would not see any significant speedup.

So, what could we have instead? In General Relativity, there are solutions to the field equations found in spacetime environments really not too dissimilar to ours called *Closed Timelike Curves*, meaning that there could regions of space with the property that if we moved *along them,* we would take some path through spacetime which would eventually arrive back at the *same point in space and in time*. By moving along such closed timelike curves (CTCs for short) we could potentially travel back to some point in the past without any more energy than would otherwise be required to move that distance. Let's suppose that in the future, the creation of CTCs becomes so cheap that we can install *negative delay elements* (NDEs) in circuits just like any other component; these would output what their inputs were/are going to be within some fixed amount of time designated by their manufacturer - for example, I might buy a 5m20s NDE from Rapid Electronics, signifying a negative delay of 320s.

Why would this be useful? There's a computer programmer who wants to find the only real root of the equation $y = x^2$ and is so bad at intuiting maths that she can't see that it's 0, so instead she sets up an approximation system where she can choose a starting guess, x_0, and then she iterates it through some sort of approximator which outputs a guess closer to the answer ($x_n \to x_{n+1}$). In order to get the answer really quickly, she sets up a circuit as shown in fig. 1.

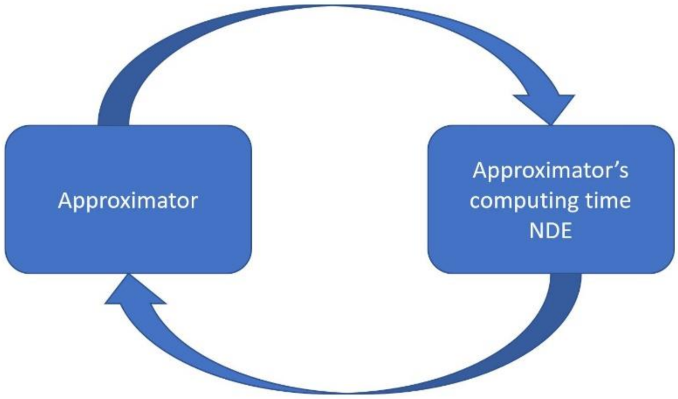


Fig. 1

When the programmer first inputs her bad guess x_0, the approximator outputs x_1. NDEs output *now* what they *will* receive as an input in the future by some amount – so this NDE, whose time value is the length of time it takes for the approximator to compute a single approximation iteration, receives the x_1 and sends that back in time so that *that was the original output* of the NDE. Thus x_1 was the original input to the approximator and when it outputs x_2 into the NDE, once again it makes x_2 the original input etc. This cycle continues until a fixed point is reached, where the input to the approximator is the same as its output – this *fixed point* is the *exact* answer to the problem, since the approximator would output it unchanged! The programmer starts the machine with the bad guess x_0 and immediately sees the exact answer to the problem…

This instant computation is a side effect of *the self-consistency principle* as formulated by computer scientist David Deutsch – he believes that relativity would only allow *self-consistent* causal systems to exist and so would propagate towards such a situation. The first few situations above were not *self-consistent,* since the input of x_0 led to the input being changed to x_1 - a self-consistent situation would be one where the output does not change anything about what the input had to be and the simplest way in which this could be achieved is for the approximator to find the solution. The situation described above, it should be noted, is analogous to a classic bootstrap paradox; from the perspective of the programmer, the system *is* self-consistent in that the input is the same as the output but how did it ever get to this state? It seems to them that there was no point of origin for this information.

We can do even more with these negative delay elements and solve NP-complete problems immediately! I mentioned the Travelling Salesman Problem above: we must decide whether there is a path between n cities whose total length is smaller than or equal to k. Our computer programmer would create a set up as in fig. 2.

Here, the programmer would input a starting path and the setup above essentially tries an arbitrary number of paths through repeatedly making random changes to the initial path and it only stops if it finds a path of length $\leq k$. If there is such a path, then it will immediately display this as the result but if not, then the machine will linger in a sort of "undecided" position. Presumably, if one wished to find the shortest path, one could simply input a smaller and smaller k until any lower k leads to an "undecidable" state. Such a system could be made for any NP-complete problem – a worrying idea would be a system which could attempt to generate proofs for some specific theorem of length smaller than k (in terms of characters) and the machine would examine them and see if each step followed logically according to certain standard rules programmed into it and stop if the theorem had been proved.

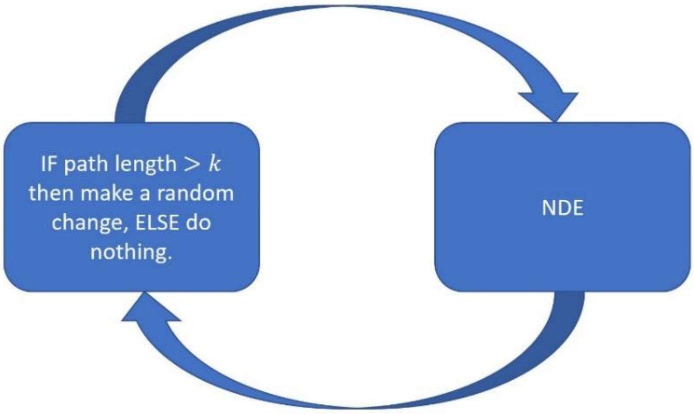


Fig. 2

Can we go further, to problems exponential or harder? Yes, but it requires some trickery. The programmer starts making *chess units*, whose job is to take in a given situation from Input Channel 1 from a chess game, along with knowledge of whose turn it is, and tries all the possible moves for that player; if one of the moves directly leads to a checkmate for the side playing, it would send an output of "YES" to Output 1, if the move directly leads to a stalemate or a checkmate for the side not playing, then it would send a "NO" to Output 1 and if a move does not end a game, it would send to Output 2 the game with that move having been played and await an answer of either "YES" or "NO" from Input 2 (with all previous rules about receiving a "YES" or "NO" applying to this). Then, the programmer sets up a number of these (the number being some upper bound for the longest possible chess game in terms of moves, such as 700) in a chain, with the nth chess unit's Input 1 set up to the $(n-1)$th unit's Output 2 and its Input 2 set up to the $(n+1)$th unit's Output 1, as shown in Fig. 3.

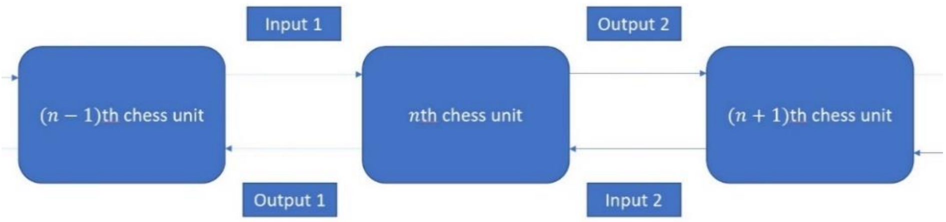


Fig. 3

Now, without time travel, the computer programmer would input into the chain of 700 or so chess units a chess game along with a move she made. This system would eventually (over a period longer than the age of the universe) come back with a "YES" or "NO" saying whether the move was optimal or not. What if every chess unit's Output 1 has an NDE placed between it and the previous unit with a time value equivalent to the average amount of time it takes for such a unit to go through all possible moves of a given situation? Then, when you input a situation to the device, it would immediately output an answer *and then* compute it…

There are some interesting caveats here. The largest one comes about in considering the exponential nature of the problem: the first chess unit is given a game and has to go through all the possible moves the programmer's adversary could take in retaliation to the move she inputted, which is going to be (as we said previously) 30 on average, but in evaluating each move it must query the next chess unit to evaluate the possible retaliatory moves to *those* so, thanks to the NDE setup, it apparently has to consider 30^2 situations at the same time and the next unit has to simultaneously evaluate 30^3 of them etc. So, surely we have to build exponentially larger chess units and therefore our computer would *still* be larger than the observable universe?

Not if we use quantum computers – these can be in the superposition of computing all of these games by seeing them each as separate wavefunctions that it is perform parallel "quantum gate" operations on. Although quantum computers still have a finite number of possible "parallel" things they can be computing, this scales exponentially with the number of "qubits" it has, so we only have to use a *linear amount* of space! However, this means that the computation should not be observed at all costs since to do so would cause quantum decoherence, meaning that the computation would fail.

When the programmer sets up the chess machine to compute whether a given move is optimal or not, *as soon as* she has seen the answer, she must close all openings or windows to the computer and just let it run; the longer it will run without disturbance, the higher the probability that the answer they saw is correct (since all quantum computations are in essence probabilistic ones).

We are in the interesting situation where we can get the answer to a computational problem instantly and have to continue computing it for a somewhat indefinite time *afterwards*. The answer we get would be *probabilistically* consistent, as David Deutsch said – he thought that the universe would resolve the *apparent* grandfather paradox by applying quantum mechanical logic, thus saying that there is a 50% chance of you being born, thus making a

50% chance of you time traveling back, thus making a 50% chance of you killing your grandfather, resulting in a 50% chance of you being born etc.!

In conclusion, a full time-machine, which can send back in time large equipment such as a full computer, would not be necessary for speeding up computing – in fact, just through the manufacturing of small devices which can send a few bits of data back in time we can create marvellous structures to solve problems *now*. Perhaps in the future, if a practical time-reversal process was discovered it wouldn't be banned as many naïve sci-fi writers predict but rather limited to usage in machinery – if so, the nature of algorithm design would be changed forever. But in cycles like closed timelike curves, who is to say they can distinguish the beginning from--

Neuroregeneration and Spinal Cord Injury

Lorna Bo

The origins of spinal cord injury can be traced back more than 4500 years ago, to the oldest known trauma text: the ancient Egyptian Edwin Smith Papyrus. It was here that spinal cord injury was first described by clinicians, who, even then, described it as an injury 'not to be treated' - an attitude that was to last for millennia. Only after breakthroughs in imaging, medicine, and rehabilitation in the 20th century did we begin to develop a greater understanding of the mechanisms behind the debilitating, often paralysing injury, which affects between 250,000 to 500,000 people worldwide each year. We now understand that although the peripheral nervous system (PNS) can regenerate fully after damage, the central nervous system (CNS) cannot, and this is why CNS trauma carries such a poor prognosis. It is only by picking apart the reasons for this dichotomy that researchers are now able to develop therapies incorporating the rapidly expanding fields of gene therapy and stem cell research to stimulate regeneration, finally offering hope of treatment to an injury historically thought to be untreatable.

In order to better understand the inability of the CNS to heal, it is useful to first consider the mechanisms behind successful regeneration in the PNS, as illustrated in Fig. 1. If a PNS neuron's axon is damaged, it will regrow at a rate of around 1mm a day in small neurons and 5mm a day in larger ones (note that if the cell body is damaged, regeneration is impossible). Immediately after axotomy (the severing of the axon), supporting glial cells such as Schwann cells recruit macrophages by releasing cytokines, and accompany them to the site of the injury to clear away debris - for example, by the phagocytosis of myelin. The distal stump (the end of the neuron not attached to the cell body) then undergoes Wallerian degeneration, a process that takes around 24 hours and results in the complete fragmentation of the axon – the endoplasmic reticulum degrades, mitochondria disintegrate and microtubules are depolymerised. The endoneurium (layer of tissue around the myelin sheath) remains intact, however, to provide a conduit to guide growing axons in a later stage of regeneration. This involves the proximal end (the end of the neuron attached to the cell body) sprouting axons with growth cones on their ends, which produce a protease that further digests debris on its journey to reinnervation through the endoneurial, or basal laminar tube, along which Schwann cells assemble in ordered longitudinal columns called 'bands of Büngner' to preserve the channel and help guide the axon to its target. Schwann cells and macrophages also upregulate neurotrophic factors such as nerve growth factor (NGF), while the PNS neuron itself upregulates regeneration-associated genes (RAGs). This intracellular encouragement, structural guidance, and expression of growth factors all create a favourable environment for regeneration and eventual reinnervation.

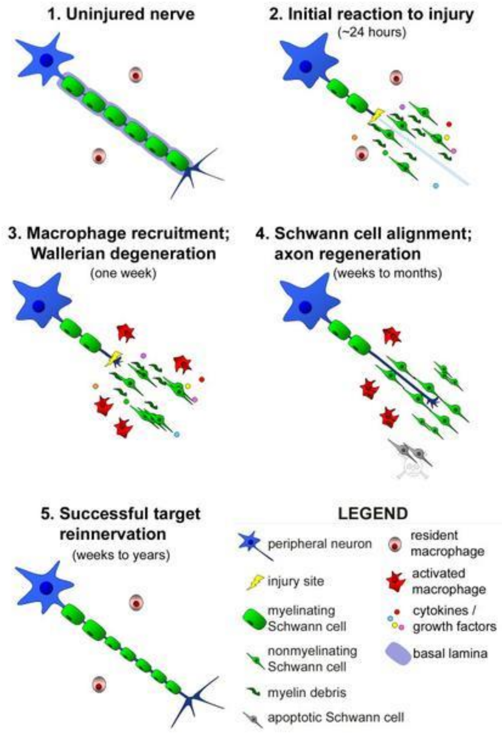


Fig. 1. Diagrammatic overview of PNS neuron regeneration. Adapted from Gaudet, A. D. et al. (2011)

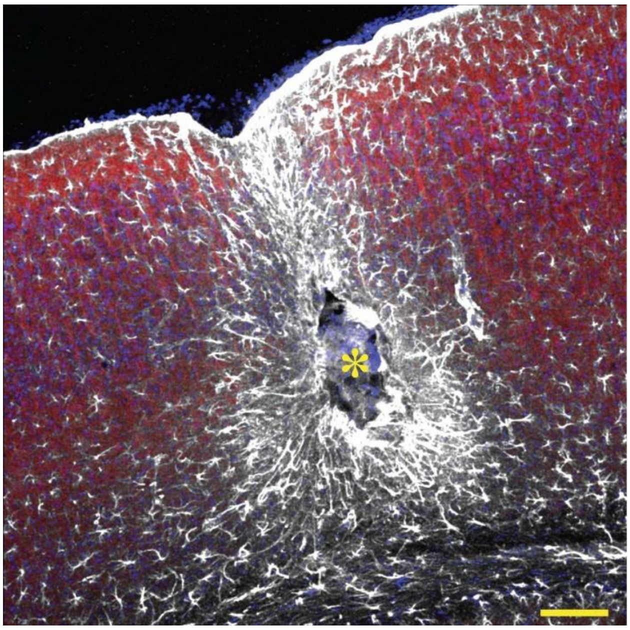


Fig. 2. Cross section of rodent spinal cord following spinal cord injury. Yellow asterisk marks a core of fibroblasts surrounded by reactive astrocytes (shown in white) that form the glial scar. From Nedergaard Laboratory, University of Copenhagen.

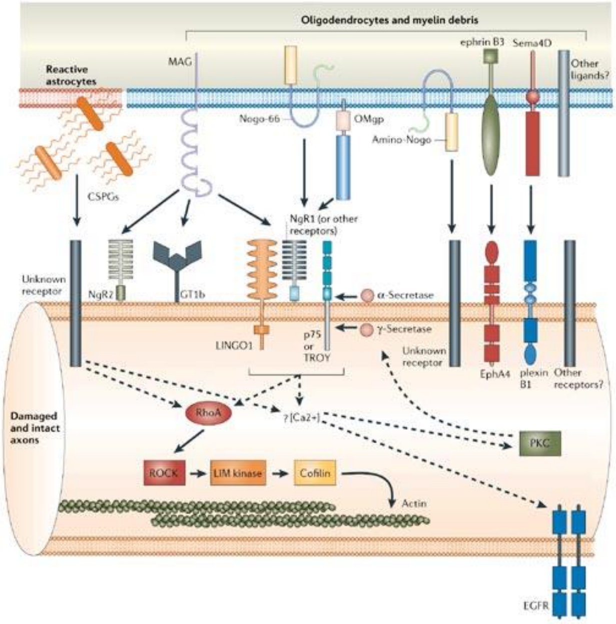


Fig. 3. Inhibitory factors to CNS regeneration. Nogo-A, ephrin-B3 and Sema-4D can be seen on the right, and the RhoA/Rho-kinase pathway at the bottom. From Nature Publishing Group (2006)

However, the general mechanism detailed above does not occur for a damaged CNS neuron. This is not entirely due to the CNS neuron's inherent inability to regenerate – on the contrary, when placed into a permissive environment of a peripheral nerve graft, they are able to grow long distances, thus indicating that the explanation lies in the CNS environment. Indeed, research has shown that the PNS-CNS regeneration dichotomy is likely to be due to their different glial cell populations and their reactions to injury. The CNS's glial cells are not Schwann cells, but oligodendrocytes and astrocytes. Oligodendrocytes require axon signals to survive, and therefore undergo apoptosis (cell death) and fail to recruit macrophages to clear debris after injury, as Schwann cells do in the PNS. Instead, the CNS must rely on the action of microglia (the CNS analogue of macrophages in the peripheral immune system), which are slower than macrophages and may fail to function as well. This causes the distal stump to degenerate slower than in a PNS neuron, resulting in the accumulation of inhibitory myelin debris. This is exacerbated by the formation of a glial scar (see Fig. 2) which axons cannot cross. It consists mainly of reactive astrocytes, which undergo heavy proliferation after CNS injury and form a dense network of gap junctions which act as a physical barrier to regrowth. Compelling proof of this inhibitory environment hypothesis comes from studies in which axotomy of the dorsal root ganglion neurons is followed by regeneration within its peripheral environment, but further growth is arrested at the PNS-CNS interface, also known as the dorsal root entry zone (DREZ). This is a phenomenon which has been known since the early twentieth century, and for which astrocytes in the glial scar have been shown to be responsible, through ultrastructural analysis. Although this barrier serves the beneficial purpose of preventing cytokine release from the injury site causing further damage to surrounding tissue, it is the main physical inhibitor of CNS regeneration. Furthermore, the CNS glia do not produce neurotrophic factors, and instead produce factors that inhibit remyelination - oligodendrocytes express myelin-associated inhibitors (MAI) such as Nogo-A, ephrin-B3 and Semaphorin-4D, and the astrocytes in the astroglial scar produce chondroitin sulfate proteoglycans (CSPGs) such as neurocan (as well as providing a physical barrier in the form of the glial scar) (Fig. 3). CNS neurons themselves are also at fault, as it has been found that they upregulate RAGs less than do PNS neurones. Therefore, a mixture of intracellular poor regenerative response and extracellular hostile environment serve to inhibit axonal regrowth and restoration of function in the CNS.

It is by targeting these factors that combine to prevent regeneration that researchers have begun to make headway in developing treatments to promote it. For example, neurotrophin (NT) treatment showed some success in a study in which rats were injured on dorsal roots, and osmotic mini-pumps were immediately implanted to infuse neurotrophic factors such as NGF, GDNF and NT3 over the next week. This allowed the dorsal roots to overcome the transitional zone between the PNS and CNS that had previously been shown to be insurmountable, restoring some function. Yet delivery by mini-pump is inconvenient for long-term treatment and does not stimulate the production of NTs in the spinal cord itself, and can cause more parenchymal damage, than, for example, viral-vector based gene therapy. This type of therapy involves injecting gene-edited viruses which express NTs along the DREZ, leading to near-normal recovery of function. However, this also causes the sprouting of non-injured neurons, which, although it can enhance recovery, may cause chronic pain. Research that involves strategically using a gradient of both neurotrophic factors and growth inhibitory factors can prevent hyperinnervation, however. Alongside using NGF to encourage growth across the DREZ, the use of increasing concentrations of the inhibitory Semaphorin-3A can prevent the growth of the neuron into unwanted areas, allowing for directed growth. NT treatment, either through protein delivery or gene therapy, has therefore been shown to have potential therapeutic applications.

Another option is to, instead of increasing growth factors, decrease factors inhibitory to growth. The genetic deletion of the aforementioned MAI protein Nogo-A promotes growth and enhances recovery after spinal cord injury (although phenotypic expression varies among individuals), as can its targeting by anti-Nogo-A monoclonal antibodies.

As well as manipulating the extracellular environment, the intracellular hindrances to regeneration can also be manipulated. For example, induced RAG over-expression in CNS neurons has been shown to promote sensory axon regeneration. Inhibitory intracellular signalling pathways can also be manipulated to promote regeneration. Take the RhoA/Rho-kinase pathway, for example. When activated, the protein RhoA activates the protein kinase 2, which, in turn, regulates the dynamics of the cytoskeleton and results in the cessation of neurite growth. C3 transferase, an enzyme that deactivates RhoA, has shown to promote axonal sprouting and motor function in mouse models, and kinase 2 knockout mice (mice in which the gene to produce kinase 2 is deleted) also showed functional recovery after spinal cord injury.

Or perhaps the answer may not lie in changing the extrinsic or intrinsic, but rather replacing the damaged neuron altogether. Stem cell transplants have received much attention from the media and general population, and they do show some promise for morphologically replacing neurons or glial cells in the damaged CNS, sometimes even altered with therapeutic genes (for example, to over-produce neurotrophins). This is due to their multipotency, meaning they can differentiate into the appropriate neuronal or glial subpopulation. When oligodendrocytes were replaced by embryonic stem cells in a rat with spinal cord injury, for example, the rat's locomotion improved. Yet this study is one of very few that have shown a functional use for stem cells in spinal cord injury, and it is unclear whether the stem cells actually differentiated into functional cells that contributed to structural reorganisation, or whether they simply secreted factors that aided the pre-existing cells to recover. In any case, despite the need for more rigorous studies as this new field develops, its preliminary success does show that future therapies may not need to be derived from the nervous system itself.

Thanks to our newfound understanding of multifactorial causes that facilitate PNS regeneration and hinder CNS regeneration, the latter, once an unachievable goal, is now a viable possibility. Targeting the glia that create physical barricades and inhibitory cues, as well as boosting the intrinsic regenerative capacity of the neuron itself, and even perhaps replacing them altogether with stem cells are all therapies that have shown promise in *in vivo* experimentation. Yet extrapolating these results to the inherently complex adult human CNS would be naïve, and for each successful experiment, one must question its implications carefully. What role does neural plasticity – the compensation of undamaged systems for the function of damaged systems – play, for example? To what extent does recovery of function correlate with anatomical and structural recovery? Perhaps most importantly – how applicable is the therapy for clinical use? The answers to these questions will only be found with a more rigorous evaluation of all these different therapies, facilitated by ongoing technological advances. Only then will we be able to bring them together and formulate a solution to a clinical mystery that only now, is being unveiled as the potentially treatable injury it is. We may finally be able to bring new life not only to the damaged spinal cord, but to lives that have been irrevocably impaired by it.

The Function of Sleep

Chloé Pitts

Humans spend on average 33% of their lives asleep. Nowadays, there is no doubt that sleep is a vital part of not only our existence but the existence of all animals on earth. Despite the slight variations in patterns and types of sleep not only within species but across species, the indisputable fact remains that sleep is essential for survival. However, although a large percentage of our lives is spent in this strange unconscious state, scientists still have not uncovered most of the mysteries behind it. Through numerous studies on sleeping disorders and sleep deprivation, they have discovered just how imperative sleep is, and how cutting it off completely would eventually lead to death. Yet there is no clear conclusion on why sleep is actually needed, and what function it serves in the complex web of survival. Despite this, there are many proposed theories by different scientists with varying levels of credibility which can be used to explore this strange phenomenon.

The evolutionary theory is one of the earliest developed concerning the function of sleep. It proposes that inactivity at night, when the environment became potentially more dangerous (due to the lack of light and warmth from the sun), served as an adaptation which provided a key survival function by keeping organisms away from harm when they would be most vulnerable. Therefore, sleep improves an animal's likelihood of survival, and those animals with sleeping habits appropriate to their environment are most likely to survive, which explains the variation observed in sleeping patterns within many species. For example, variation is observable within sleeping habits of humans in different countries and environments, such as people who engage in polyphasic sleep, and countries such as Spain where siestas or biphasic sleep is nationally recognised. This obvious variation supports the idea that sleep is a behavioural strategy which developed into its current state through natural selection. However, this evidence is largely flawed due to the lack of experimental data and application of the scientific method, and could be completely consequential (especially in humans) where cultural differences and traditions influence behaviour and sleep patterns.

Further counterarguments which render this theory almost negligible is the obvious fact that although it might explain a period of quiescence or inactivity during times of danger, it does not explain why sleep should leave animals so vulnerable and defenceless (with greatly decreased sensitivity to external stimuli, and sometimes complete paralysis) at such a critical time. It is always safer to remain conscious to be able to react to an emergency or potential threat. It also fails to explain why carnivores or even apex predators with no predation to fear - for example, lions - sleep for such extended periods of time. Through experiments in sleep deprivation, it can be concluded that sleep is not only a passive removal from the environment, as it is treated as a necessity in species who will alter their behaviours in order to ensure they gain sufficient amounts of it. The evolutionary theory fails to address this behaviour and the distress and hallucinations extended sleep deprivation causes. Overall, the large evolutionary costs associated with sleep are high: sleeping animals don't eat, drink or reproduce and are more vulnerable to predation. This high cost implies a critically important adaptive value. So how can sleep serve this evolutionary purpose?

The energy conservation theory of sleep is closely linked with the evolutionary theory, as one of the strongest factors in natural selection is competition for and effective utilisation of energy resources. The energy conservation theory claims that diurnal, nocturnal and crepuscular species may use sleep as an energy-saving device, reducing energy expenditure and foraging during those times when they cannot exploit their environment and gain easy access to food. Multiple experiments have produced results both for and against this theory. Metabolic rates have been found to be higher in smaller animals, which typically sleep longer hours, and cold-blooded animals like reptiles and fish tend to exhibit less obvious and unequivocal sleep - supporting the theory of energy conservation. Research shows that during sleep, the core temperature of humans is reduced by around 1°C. Besides body temperature, there is also no doubt that caloric demand decreases during sleep, as less energy is needed to maintain body temperature and perform other bodily functions.

Another feature of this theory is that the waking brain uses a great deal of energy, which needs to be recuperated. However, it is actually only during non-REM sleep that the brain uses less energy at all, and brain wave activity during REM sleep is very similar to (or sometimes even higher than) that of the waking state. In fact, sleeping only reduces metabolism and energy use in humans by at most 5-10% overall. This is supported by the fact that the energy saved by sleep is quite modest – the equivalent of a piece of toast (80 to 130 calories) per night in humans. It seems unlikely that such a high behavioural price is paid for such small returns. It appears that it would be much more efficient to stay awake and eat. However, others argue that sleep only occurs when animals cannot easily gain food from their environment (such as at night when it is dark), and that even a small energy saving can have an important impact on evolutionary selection and survival. This argument is easily countered through the obvious process of hibernation in some species, which is a truly energy-saving state. This suggests that species which depend on energy conservation for survival hibernate, which labels hibernation as the answer to this energy-conserving selection pressure - not sleep. Sleep must therefore serve a different primary purpose, although energy conservation could potentially be a secondary one.

Another theory for the function of sleep is the restorative theory. This states that sleep restores what is lost in the body whilst we are awake, and therefore rejuvenates and repairs it. This encompasses physical restoration - the neutralisation of neurotoxins, the renewal of tissues and nerve cells, and the balancing of chemicals within the body - and brain restoration. Scientists argue that REM sleep appears to be devoted to brain repair and NREM sleep devoted to body repair. Experimental evidence of the restorative theory shows that Stage 3 slow-wave sleep has been associated with increased levels of growth hormone levels in the body, which is an important factor in tissue regeneration and repair. Furthermore, athletes' bodies (which are under a lot of physical pressure) spend proportionately more time in slow-wave sleep than the average person, and growing children spend more time in it than older people. Hard physical exercise also typically causes a moderate rise in deep slow-wave sleep the next night, all of which supports the cell restoration and repair theory. It is logical that slow-wave sleep gives the body an opportunity to focus on physical healing and damage repair during quiescence, whilst in REM sleep there are more active brain wave patterns, suggesting a restorative function focused in the brain.

There is a lot of further evidence to support this theory through both human and animal experimentation and studies. The results of these experiments indicate that sleep strengthens the immune system in general, so the physical healing of wounds is expedited by sleep. An experiment carried out on rats showed that when deprived of sleep, they show distinctly inferior healing capacities, develop skin lesions, lose body mass, and are unable to maintain a stable body temperature, ultimately dying of sepsis (bacteria and toxins). Sleep-deprived rats have been shown to exhibit substantially fewer leukocytes (white blood cells), the body's main defence against infection, and sleep-deprived humans show less than half of the protective antibodies compared to people with healthy sleep patterns. Metabolic activity during sleep is mainly anabolic, during which new molecules are constructed and built up (as opposed to catabolic in which molecules are broken down for subsequent re-use), which strongly supports the theory.

However, there are, again, many counterarguments to the restorative theory. It is not apparent that more repair happens during sleep than during wakefulness, and the process of protein synthesis actually decreases during sleep. Although the physical benefits of Stage 3 NREM sleep are very apparent, the value of Stage 2 NREM sleep is not clear and this type of sleep accounts for over 50% of the adult human sleep period. However, Stage 2 NREM sleep is only found in rudimentary form in sub-primate species. It is also worth highlighting that individuals who spend the whole day at rest in bed sleep as long (or even longer) than highly active individuals.

Another more recent theory for the function of sleep is that it evolved as a mechanism for the consolidation of memory and specific neuronal networks, a process which occurs during the part of the rest-activity cycle when there is sufficient energy and minimal extraneous neural interference. 'Sleep dependent memory processing' is supported by evidence from both animal and human studies. Human studies on the effect of sleep deprivation have shed a lot of light on the theory. These experiments demonstrate that sleep deprivation leads to reduced attention span and short-term/working memory which influences what gets saved as long-term episodic memories. Sleep deprivation also impacts high-level cognitive functions such as reasoning and decision making, and seems to play a role not only in memory consolidation after learning but in preparing the memory for encoding before learning. Further explorations suggest that REM sleep benefits procedural memory (memory involved in how to carry out tasks), and slow-wave sleep benefits declarative memory (related to facts and events). Synaptic and neuronal activity in the brain during sleep is also significantly greater in the areas where learning took place during the day, and the content of most dreams frequently revolves around experiences that day, suggesting the replaying of events during dreams is closely linked to consolidating those particular memories. This is supported by the fact that numerous studies have shown that tasks learned on one day are performed better the next day after good sleep versus bad/no sleep. The chance of gaining insight is almost three times higher if the individual is allowed to sleep, and some tasks are never learned if sleep is restricted the night after learning.

Despite the abundance of evidence for the theory that the function of sleep is consolidation of memory and learning, there are some factors which make us question the validity of it. There are three main counter arguments. The first is that the theory only explains sleep related to those species with a complex brain. The second is that memory and learning can occur in the absence of sleep, negating somewhat its importance. The third is that, if sleep evolved to serve this sole function, it would make logical sense that animals with the largest and most complex brains would sleep more - yet there is no correlation between the complexity of a species' nervous system and its length of sleep.

Overall, there is no doubt that sleep is an essential physiological process for humans and most animals, without which survival is impossible. Scientists have developed multiple theories pertaining to the function of sleep and the role it plays, and there is experimental data which both supports and challenges aspects of them. The main theories outlined here are the evolutionary theory, the energy conservation theory, the restorative theory and the consolidation of memory theory. The most flawed are the evolutionary and energy conservation theories, as they can be argued against through pure logical reasoning which is supported through experimental results. The restorative theory and the idea that sleep developed as a means of consolidation of memory hold more promise in uncovering the reasons behind the need for sleep. Even though both have flaws, there is considerable support for the notion that sleep is required for processing short-term memories and strengthening long-term memories through changes in brain plasticity and neuronal circuitry, and it is the theory which boasts the largest quantity of supportive experimental evidence.

However, the lack of consensus regarding a definition of sleep, and by extension the presence or absence of sleep across different species of the animal kingdom addresses a vital problem in trying to discover its true function. Through the exploration of underdeveloped species (even single celled organisms) which may show a primitive form of sleep, such as active and passive cellular activity, a possible precursor to sleep could be studied in order to define the term more accurately. The most probable conclusion we can draw from all these theories is that the function of sleep is most likely to be a combination of these factors. Yet its main function is still up for debate, and it is only through more research that it can be uncovered.

Despite this, scientists have made massive developments in understanding sleep and sleeping disorders. As sleep researcher Robert Stickgold says, scientists that used to ask the question, "Does sleep do any good, or serve any function?" are now forced to ask instead, "Is there anything that isn't improved by sleep, or impaired by sleep deprivation?" – to which he claims the current answer is no. Therefore, we could amend the entire perspective on which we approach the concept of the function of sleep and say that sleep is so beneficial that the question is really why animals ever bother to wake up. Maybe it is actually the harmful state of wakefulness that is an evolutionary mystery, not sleep. Sleep could even be just an alternate state of wakefulness which does not pose any risks if the body is kept in a safe environment. It begs the question that if dreams could be controlled (which already occurs naturally through lucid dreaming) and a safe environment guaranteed, would it not be better for us to remain in a constant state of sleep, living in an alternate reality that poses no real bodily harm?

Sleep: A User's Guide

Alina Nicheperovich

How does sleep work?

According to the Oxford Dictionary, sleep is "a condition of body and mind which typically recurs for several hours every night, in which the nervous system is inactive, the eyes closed, the postural muscles relaxed, and consciousness practically suspended." But what exactly happens in the human brain during our unconscious hours?

The answer is quite simple: no one knows for sure. It is known that the brain orchestrates the daily sleep-wake cycle by responding to external cues such as sunlight, and the body's own rhythms (circadian clock). Then there's sleep drive, or sleep pressure. The longer you are awake, the more energy carrier ATP (adenosine triphosphate) molecules are broken down. This leads to a built-up of a chemical called adenosine which triggers neurone activity in the ventrolateral preoptic nucleus (VLPO) and increases your desire for sleep (Fig. 1).

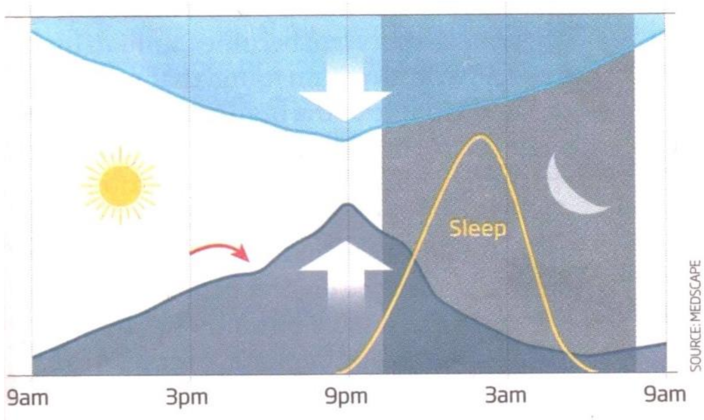


Fig. 1. Sleep pressure builds up the longer you are awake. It is balanced by the force of circadian alertness. As night falls, however, circadian alertness drops off more quickly – and you enter the land of nod.

Brain activity changes continuously throughout the night. This is why there are four stages of sleep. These different phases have been shown to help consolidate different types of memories. How much time you spend in each phase also affects sleep quality (Fig. 2).

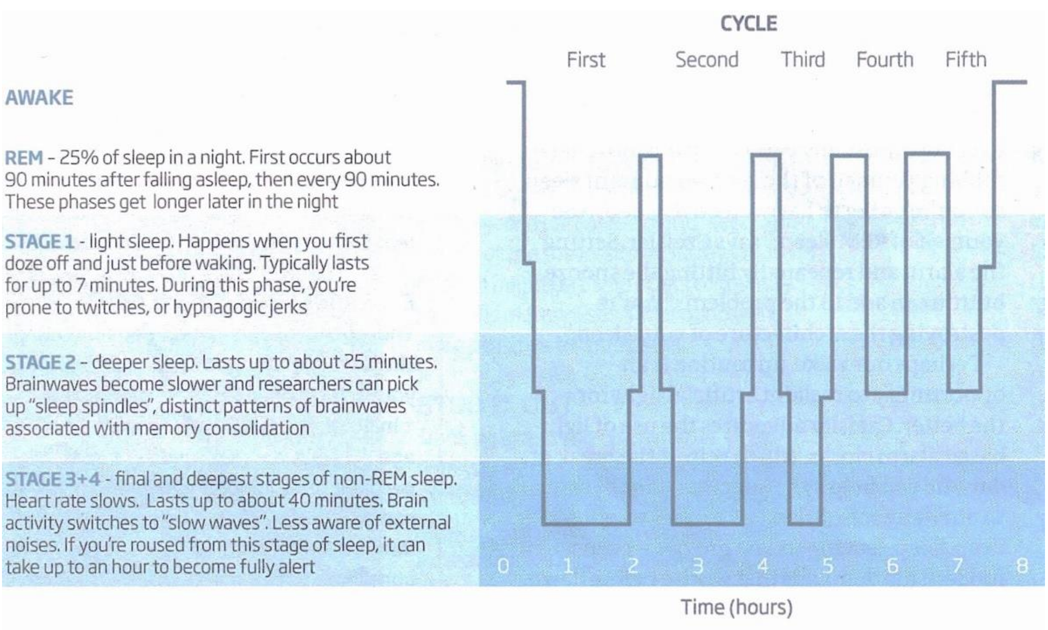


Fig. 2. A typical night's sleep involves several cycles

Why is sleep so important?

Sleep is as vital for life as food or water. Scientists state that too little time in the land of nod messes with your emotions and your ability to make sound decisions. It affects your immune system and appetite, and has been linked to metabolic diseases such as obesity and type 2 diabetes. In addition, a lack of sleep is proved to increase the risk of development of mental health problems including depression, bipolar disorder and schizophrenia.

5

is the number of minutes it takes you to fall asleep if you're sleep deprived. The ideal is 10-15 minutes.

Sleep is important for various aspects of brain functions such as cognition, concentration, productivity and performance. In addition, the process of cell regeneration takes place when we sleep, therefore a lack of sleep can lead to premature aging.

EEG machines monitoring people during the deepest stages of sleep have shown electrical impulses moving between the hippocampus, thalamus and cortex, which serve as relay stations of memory formation. This shows that a good sleep might increase your learning potential.

Also, the next time you feel down, think about how much sleep you had last night. Sleep deprivation decreases the concentration of hormones responsible for good mood, and can lead to a lack of motivation.

What about coffee?

How many stamps have you got on your Grumpy Mule loyalty card yet? Quite a few? Not surprising, because caffeine is the world's most used drug.

In the human body, caffeine acts as a stimulant for the central nervous system. It keeps us awake by blocking one of the body's key sleep-inducing molecules, a substance already mentioned above: adenosine. When adenosine binds to receptors in our brain, it slows down the release of ATP molecules which we use as an energy source. In other words, high concentrations of adenosine makes us sleepy. Caffeine is an adenosine receptor antagonist. That means it blocks adenosine receptors from binding to adenosine molecules. This happens because adenosine and caffeine have similar double-ring structures (Fig. 3).

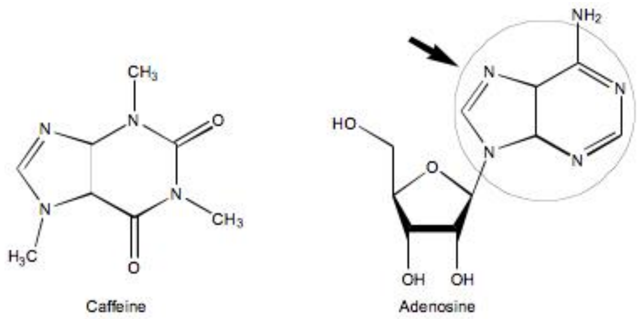


Fig. 3. Structural formulae of caffeine and adenosine

Caffeine can also boost positive feelings, which is why so many of us are coffee addicts. In some neurones, adenosine receptors are linked to receptors for another chemical called dopamine, which promotes feelings of pleasure. When adenosine binds to one of these paired receptors, it makes it harder for dopamine to fit in its own spot, interrupting its mood-lifting work. But when caffeine takes adenosine's place, it doesn't have the same effect and dopamine can attach to its receptor, making us happier.

The question "Is coffee good or bad for our health?", however, has no answer. On the one hand, there's evidence that caffeine's effects on adenosine and dopamine receptors can have long-term benefits, reducing the risk of diseases like Parkinson's, Alzheimer's, and some types of cancer. Caffeine can also increase the body's ability to burn fat. On the other hand, it can raise your heart rate and blood pressure, cause urination or diarrhoea, and contribute to insomnia and anxiety.

How can I nap like a pro?

Napping is known to be very beneficial for focus. A short nap, lasting only 10-15 minutes, can boost alertness, concentration and attention for as long as 4 hours. Take 20 minutes and you increase your powers of memory and recall, too. Research on NASA pilots showed that a 26 minute nap enhanced their performance by 34% and overall concentration by 54%.

Still, deep sleep provides the biggest boost to learning. If that's your aim, go for a nap of between 60 and 90 minutes, says Matthew Walker of University of California, Berkeley. Unlike a short nap, it will improve your alertness for up to 10 hours. As well as helping you to retain factual information, longer naps can increase motor memory, which is useful for training skills such as sport or playing a musical instrument.

What's the best way to get to sleep?

Everyone has different sleep patterns, but there are still some small things that can help us all have a good rest at night:

1. **Switch off.**
   Tablets, laptops and smartphones emit lots of short wavelength blue light, which interferes with production of our sleep hormone, melatonin. Using screens for 2 hours before bed reduces melatonin concentrations by 22%. Screen time before bed can also mean it takes longer to fall asleep, and seems to cause a decrease in how much REM sleep you can get.

2. **Temperature control.**
   It is a common knowledge that we sleep better when it is a little chilly (but not too cold). Melatonin cools the body by a couple of degrees while we sleep, and an overheated bedroom can interfere with this process. A general rule is to keep the bedroom between 18 and 21°C, with a window open if it's not too noisy.

3. **Develop a relaxing bedtime routine.**
   It can be a chapter of a book, 15 minutes of yoga or a hot bath, but avoid doing any intense physical exercise 2 hours before sleep. Sure, it gets your blood flowing, but it will wear you out mentally. Studies have also shown that memorising information three hours before going to sleep is ideal for consolidation during the night.

4. **Sniff some herbs.**
   Finally, if you want to get really serious about good sleep, you can incorporate herbs into your night-time routine. You can either breathe in lavender oil before bed or put a few drops on your pillow.

Hopefully, this article answered the questions keeping you up at night. Sleep occupies nearly a third of our lives, but many of us give surprisingly little care and attention to it. Sleep deprivation is a major cause of many serious diseases, so if you want to be as healthy and productive as possible, make sure your sleep routine is one of your top priorities. Otherwise, a long-term habit of catching up on class late at night may well catch up with you in the end.

The Possibilities of Teleportation: Quantum Entanglement and the Tale of Two Particles

Rohan Dhir

'Beam me up, Scotty' is the catchphrase from Star Trek that has become instantly recognisable around the world. It precedes an almost magical transportation in the Star Trek universe where a character is moved across a distance ranging from 40,000km to light years in a matter of seconds, more commonly known as teleportation. Scientists have worked up an obsession of making teleportation a reality, but as of yet we have not found a way. However, this could soon change with the emergence of the most researched and fastest growing quantum mechanical phenomenon - that of 'Quantum Entanglement'. This area has vast potential and its concept is explained below as well as its potential uses to society.

What is Quantum Entanglement?

Quantum Entanglement is not a 'property' that particles possess, but rather a 'state' in which they can be. When two particles are in the state of Quantum Entanglement, the quantum state of one particle cannot be described without mention of the other i.e. the quantum state thus must be described as a whole. This state can persist even if the particles are large distances away from each other- even light years. The distance aspect of the phenomenon has exciting implications, especially in some form of future communication.

The physical properties of the entangled particles remain intricately linked. In particular, the measurements of properties such as spin and polarization of the two particles are correlated - for a measured property of one linked particle, the same property measured on the other will be the opposite.

Let us take an example; let's assume we begin with two entangled photons each of which can have a spin of -1 or +1. If we measure the spin of one of the photons as +1, we instantly know that the spin of the other photon is -1. The photons can be separated by light years but we would still know the spin state of one instantly as we measure the other. This phenomenon can be mathematically described using the full quantum mechanical framework.

Let us take two separate quantum mechanical 'physical systems' Y and Z. For such a 'physical system' we define it as a system such that if you applied classical wave theory would give you the completely wrong answer - it is a notion of states, observables and dynamical law. The term 'physical system' cannot be applied to a single particle either, as objects that act as quantum mechanical systems can be as large as stars.

Having taken the systems Y and Z, we can define the *Hilbert space* of the combined system as

$$HY \otimes HZ$$

The Hilbert space of the products is simply an infinite stretch of the 3D coordinates of Euclidean space. First, we must consider the systems in a non- entangled form, in 'separable states'. This can be denoted by the equation:

$$|\Psi\rangle\ Y \otimes |\Phi\rangle\ Z$$

If the two systems are linked however, the equation changes to:

$$|\Psi\rangle\ YZ = \sum_{l,m} clm\, |l\rangle\ Y \otimes |m\rangle\ Z$$

In the equation above we have denoted the set of vectors describing system Y as l and the set of vectors describing system Z as m. This is also known as setting a basis for the Hilbert spaces of Y and Z. With this equation, under certain vector conditions, the systems become entangled and inseparable, such as here:

$$\frac{1}{\sqrt{2}}(|0\rangle\ \mathrm{Y} \otimes\ |1\rangle\ \mathrm{Z} - |1\rangle\ \mathrm{Y} \otimes\ |0\rangle\ \mathrm{Z})$$

This entangled state equation above returns us to the measurement situation mentioned earlier, where if you measure the value of the properties of one of the states, the other state would be in the opposite form.

We can add a bit more detail to the situation by demonstrating how this might work. If a person measuring system Y can observe {0,1} then two equally likely possibilities arise. The first is if the person observing Y recorded a 0, which will result in the person observing Z recording a 1 and the second is those switched around. Either way the system has been changed by the result of Y and so the system is in a state of quantum entanglement.

However, this is where the paradox of quantum entanglement really begins. At any one moment two states of Systems Y and Z as described above are linked to each other, in an entangled state, but when we take a measurement of a property of one of the particles the state of entanglement is broken. This can result in any sort of a random change in the quantum (and other) properties of the system. In the case of the entangled pair of particles, it may appear that the system is acting completely as normal - this is untrue and any quantum entanglement-like property imitation is entirely by chance. The sheer luck of the particles continuing to act as entangled particles can change at any moment and is not guaranteed after the entangled state is broken.

Pairs of entangled photons have proved to be highly durable. In a recent experiment conducted by the University of Vienna and the Austrian Academy of Sciences, an entangled pair of photons retained its entangled state after a series of stress experiments. The pair was dropped 12m and rotated with an acceleration of 30g (30 times the acceleration of earth). This unlocks massive potential for the usage of Quantum Entanglement for communication in space.

Albert Einstein, Boris Podolsky and Nathan Rosen first considered the phenomenon of entanglement in their 1935 paper, which later became known as the EPR paper (named after its 3 authors). They believed that such behaviour was impossible and that the current theory surrounding Quantum Mechanics at the time was flawed or incomplete. The three scientists proposed the phenomenon of 'hidden variables' which must have some effect on the entangled particles in order for properties to be shared faster than the speed of light.

Quantum Entanglement experiments and proofs:

There have been a series of experiments to test the Quantum Entanglement theory and also different methods to produce two Quantum-Entangled particles. Each experiment have required different machinery, interesting and complex in and of themselves.

How do you entangle two particles? There are two main methods of doing so:

The first is to entangle photons as they are emitted. This is usually done using a cascade source - calcium atoms are elevated to a highly excitable state and forced to emit two photons. When the emitted photons are released in opposite directions, they become entangled because their polarizations have to be correlated due to the law of the conservation of charge. This process of generating entangled photons is extremely slow as they are emitted in random directions. However, a new method has been discovered recently using beta barium borate crystals to separate a beam of photons. It has reduced randomness, as one is able to predict exactly where the photons will be emitted. The normal wavelength of light used to produce the beam of photons is violet as the process produces photons that are nearly infrared. This new method is known as a 'Parametric Downconversion' and has massively increased the count rate of entangled particles obtained, allowing them to be of much larger statistical significance than the previous method.

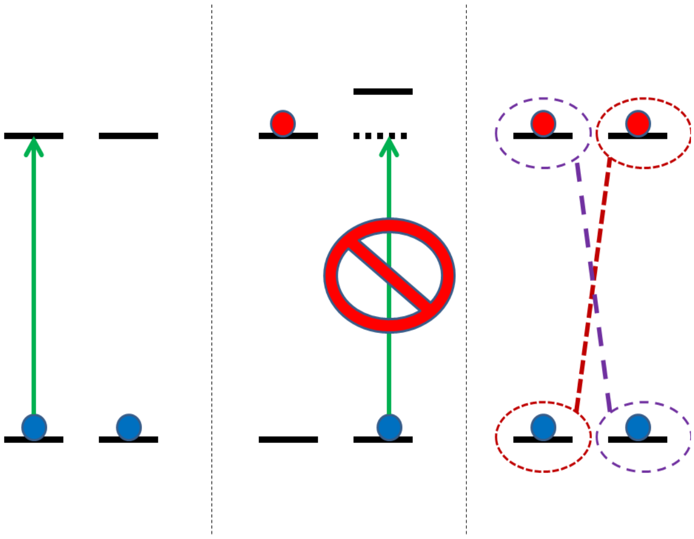


Fig 1. A Rydberg Blockade scheme, with the entangled states on the right side

The second method of entangling photons is to bring them together and let them interact to entangle. This can be done using a 'Rydberg Blockade' where two atoms are arranged in ground state next to each other and then one atom is excited using a laser pulse so that it shifts to a higher energy state. The excited atom should then affect the second, which is not directly affected by the laser, so it ends up in a superposition that is anti-correlated with the first atom. This therefore means that the two atoms are now in an entangled state, as in Fig. 1.

Both these methods have a major drawback to them - they only produce locally entangled particles. This means that the particles are produced from the same source and are in the same place, resulting in limitations for teleportation as it suggests that entangled particles must have a common history. However, teleportation can also allude to basic teleportation within a room itself, and so Quantum Entanglement could still play an important role.

There have been numerous experiments that have sought to test the basis of Quantum Entanglement. The most recreated and worked on experiment and one of the most notable is Bell's Test. John S. Bell developed this in the 1960s to coincide with his famous Theorem, which stated that no 'hidden variables' existed and its physical theory could never compete with the number of properties predicted by Quantum Mechanics. The test was developed from his Theorem and was never actually conducted by Bell himself – he rather set about the framework for someone else to complete the test, as first done in 1972

Bell's test involves observing two particles, often photons, which are entangled through one of the methods mentioned above and then some property of the two photons is measured, such as polarisation or spin. The results of the test can then be compared to what is predicted by Quantum Mechanics, thus determining if Quantum Mechanics is more accurate than the 'hidden variables' theory. In order to further distinguish Quantum Mechanics and hidden variables Bell produced an inequality that, when breached, ruled out the 'hidden variables' as a cause.

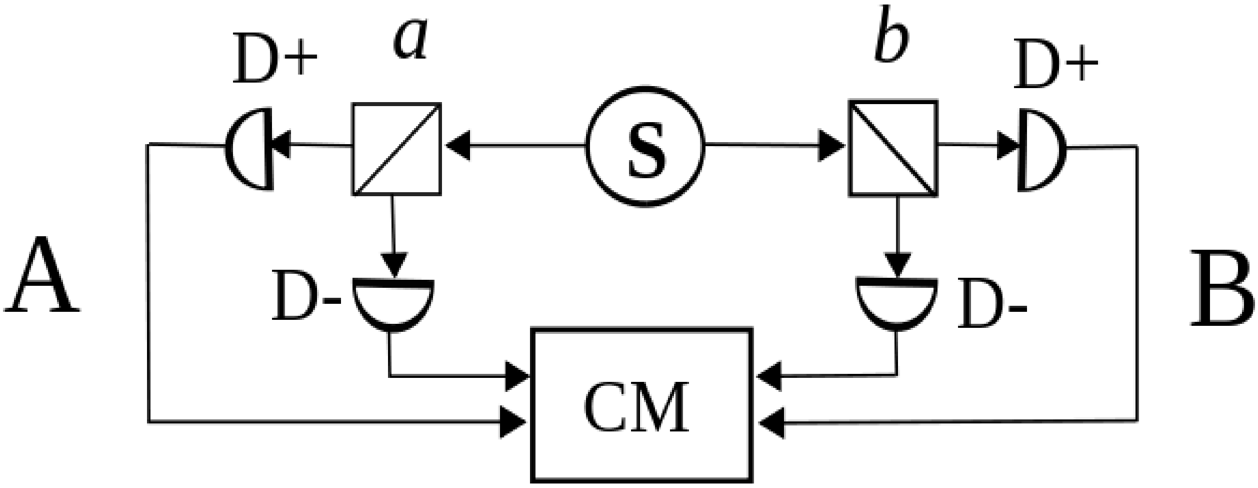


Fig 2. An example set-up of a Bell's Test

Fig. 2 gives an example of the experimental set up of a Bell's Test. The entangled photons are produced at S, normally using the 'Parametric Downconversion' system. These photons are directed towards the dual output polarisers *a* and *b*, whose positions can be easily decided due to the usage of the Downconversion method. Therefore in this specific experimental example the measured property is the polarisation of the photon. The moments where the outputs, which can be seen above as D+ and D-, coincide is measured by the coincidence monitor (CM) and compared to the coincidences measured by Quantum Mechanics.

Numerous Bell's Tests have been performed since the first in 1972 and are still continuing today. The experiments are commonly known to rule out the potential existence of a hidden variable contributing to the Quantum Entanglement paradox. Others' tests such as the 'Cosmic Bell Test' conducted by Johannes Handsteiner et al. in 2017 at MIT and the University of Vienna, have restricted the space-time region in which the hidden variable could be a factor.

Potential uses of Quantum Entanglement:

The use of quantum entanglement has been explored in a wide variety of fields, particularly in communication, especially over extremely long distances. Quantum entanglement has also been explored as a route to making some form of teleportation a reality.

Networks based on quantum processors and computing have been suggested as a method for speeding up networks and sending more information. This has already been turned into a reality by two different teams of Chinese scientists. One team successfully transmitted quantum keys from a satellite to two ground stations in China. They covered a distance between 645 - 1200 km, a complete breakthrough for quantum communication. Yet there is no specific information to suggest that the quantum keys were sent using entangled particles, and so it could have been sent using another quantum method. The other team managed to send two entangled particles from a satellite to a ground station, covering a similar distance. As more such applications are tested, this could provide a breakthrough for faster communication on earth and possibly even for future space travel.

The second and arguably more interesting application of Quantum Entanglement is teleportation. Although this is still in the preliminary stages of research, it is nevertheless a very exciting possibility.

In theory this can be done on a very small scale, 'teleporting' information on photons using a pair of entangled photons. We can take a situation where two people, "A" and "B", have a pair of entangled photons and want to 'teleport' another photon X from one place to other. "A" must measure photon X jointly with their entangled photon, thus changing the information on "B"'s photon to correlate with the combined state of photon X and entangled photon "A". "A" must then communicate the result of the measurement to "B" through conventional means. Depending on the outcome of "A"'s measurement, "B" can perform some measurements to determine the original status on photon X, and therefore 'teleport' its state. Admittedly this is nothing like what we intend teleportation to be! Tests will have to be completed to determine whether this can be safely replicated on atoms in a human body. Mechanical systems will also have to be installed that link households and companies and mass-produced units will need to be available at universally acceptable prices. Furthermore, there is the problem of systems existing only locally, let alone in two different rooms and it will be a long time before systems can exist across country-scale or even space-scale distances. However, by transporting information on photons we have a made good beginning that will be developed with further research.

Conclusion

This essay intended to discuss the phenomenon of Quantum Entanglement and whether it could make teleportation a reality. Although the possibilities have yet to be truly explored, I believe that teleportation may have to emerge from another branch of physics. The constraints of local entangled particles and the current situation where only information, not matter can be transported between particles is a challenge I do not see being overcome. Known physics simply doesn't allow the possibility of teleportation through Quantum Entanglement.

Nevertheless, Quantum Entanglement remains an interesting route for communication in the future. Increasing the speed and information transferred by using quantum systems is a real possibility. So although 'Beam me up, Scotty' teleportation is not a viable possibility from Quantum Entanglement, communications as fast as the speed of light could be.

Thermoacoustics: How To Keep Your Ice Cream Cool

James Tett

In 1999, *Ben and Jerry's* reviewed their refrigeration facilities and determined that they could reduce their greenhouse gas emissions by finding an alternative to the HFC (Hydrofluorocarbon) refrigerants found in nearly all refrigerators. In researching this, they came across the idea of Thermoacoustic refrigeration: using sound waves to chill objects.

What is Thermoacoustics?

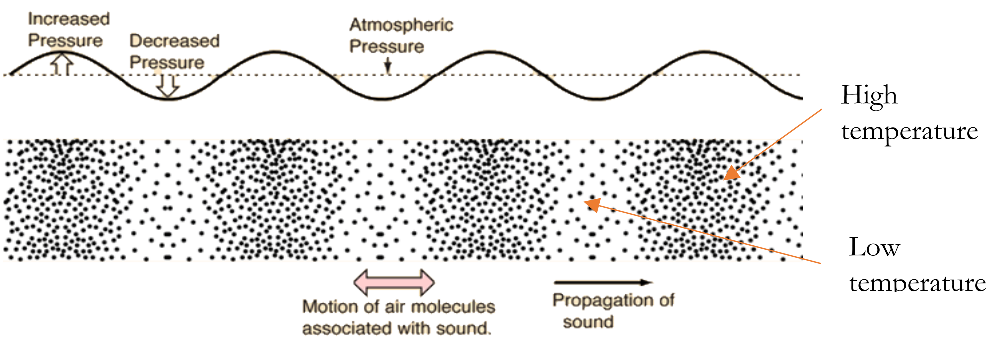
Thermoacoustics is the study of the interaction between acoustic energy and heat transfer. Acoustic (or sound) waves are simple longitudinal waves that propagate through an elastic medium by causing oscillations of pressure, displacement and temperature.

It may seem odd that sound waves cause changes in temperature. However, this temperature change is a natural result of the pressure fluctuations in the medium. Temperature is the average kinetic energy of particles within a system and so increasing the temperature increases the average kinetic energy of the particles in the system. Therefore, the particles can strike the sides of a container with a greater force and so, since $Pressure = \frac{Force}{Area}$, the pressure exerted by the system increases with a temperature increase. It is clear therefore that Pressure and Temperature are related. This is shown in Gay-Lussac's Law, which states that $\frac{Pressure}{Temperature} = k\ (a\ constant)$ (i.e. that temperature and pressure are proportional for ideal gases).

The same is true in reverse; if the pressure of a system (with a fixed no. particles) increases whilst the volume remains constant, the average kinetic energy of the particles <u>must</u> have increased (since collisions with the sides of a container must either be happening faster or with more force, both requiring an increase in the average kinetic energy) and so the temperature <u>must</u> have increased (as the average kinetic energy of the particles is the temperature). Therefore, some energy must have been put into the system. In longitudinal waves, this energy comes from the source of the sound wave (since waves transfer <u>energy</u> with no net transfer of matter) and results in the fluctuations in pressure and temperature within the longitudinal wave (with high temperatures at the high-pressure points and vice versa).

Thus, it is clear that an increase in pressure will cause an increase in temperature and so, since a longitudinal wave has oscillations in pressure, there will also be oscillations of temperature in the wave (see Fig. 1). The waves apparent in a typical conversation, between 50 and 60 decibels, will have temperature oscillations of between 0.0001 degrees centigrade, which is why we do not notice temperature changes when speaking. However, in a thermoacoustic engine with decibel levels of around 180 and 190 decibels, the pressure variations will be on the level of 30 kPa and so the temperature changes can become much more noticeable.

Fig. 1.



What is a Standing Wave?

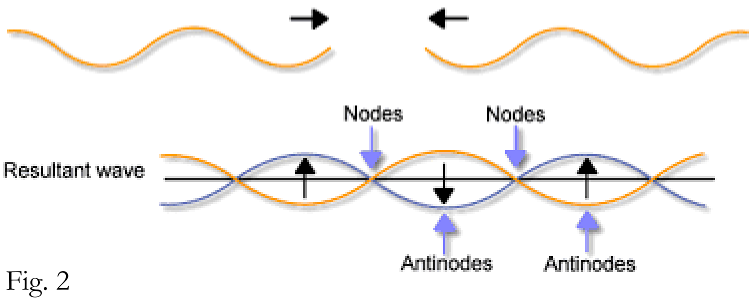


Fig. 2

In order to understand the mechanism of a Thermoacoustic refrigerator, it is imperative to understand what a standing wave is. A *standing wave* is formed when two progressive waves travelling in opposite directions meet and superpose as seen in Fig. 2, so that some particles oscillate with a greater amplitude and others will not oscillate at all. The two main points of interest are the nodes and antinodes. The nodes in a standing wave are points where displacement is 0 *at all times* (since the waves are in antiphase there) whilst the antinodes are points where the displacement varies between +2A and -2A (if A is the amplitude of each of the progressive waves forming it) – this is not shown correctly in Fig. 2, as amplitude should be greater in the resultant wave. In a sound wave, at a motion node, the molecules in the fluid would experience the maximum compressing and expanding so an antinode of pressure would form (and vice versa for the antinodes of motion).

What is wrong with our current refrigerators?

The reason why thermoacoustic refrigeration is so exciting is that there are several problems with current fridges that are not an issue with thermoacoustic fridges. The refrigeration mechanism requires high and low pressures, which in turn require a large amount of energy, so refrigerators consume a lot of energy. Next, there are moving parts in the refrigerators that require lubrication and complex sliding seals (that leave the possibility for leakages). However, the main problem is the refrigerant (the fluid that transfers heat). All fridges use a refrigerant to transfer heat from the fridge to an external 'heat sink' (see Fig. 3).

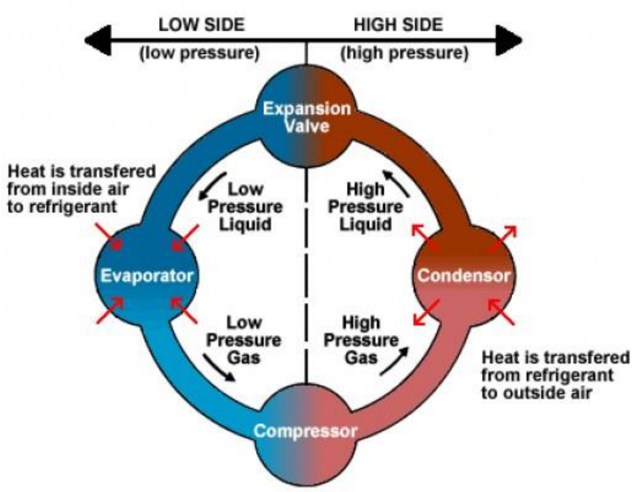


Fig. 3. The refrigerant is throughout the inside of the coil and absorbs heat energy inside the refrigerator as a low-pressure gas (due to the expansion valve). This low-pressure gas, having absorbed heat energy (reduced the temperature) from the refrigerator, is compressed in the compressor to form a high-pressure, high-temperature liquid that radiates heat in the condenser coils out the refrigerator. This cycle then repeats.

However, the refrigerants used each have their own problems. The most commonly used refrigerants are HFCs, Hydrocarbons and Ammonia. Hydrocarbons are flammable and so instantly have risks associated with them since the refrigerant is put under high-pressure and so could combust if the refrigerants are not stored properly. Hydrocarbon gases are also denser than air and so can displace air from the lungs, inhibiting breathing (although this is a very low risk). Ammonia is toxic (it can cause swelling of the brain cells and an energy deficiency to brain cells) and has the same greenhouse effect as carbon dioxide - it is a good refrigerant, it requires extreme care when storing it. Finally, HFCs (or Hydrofluorocarbons) are both greenhouse gases and, although much less than previous alternatives (about 0.035% by 2050), recent research has shown that they contribute to ozone depletion. These refrigerants are also costly, making refrigerators themselves costly (due to the refrigerants and the cost of the energy to run the refrigerator).

Why are some gases greenhouse gases and what is wrong with them?

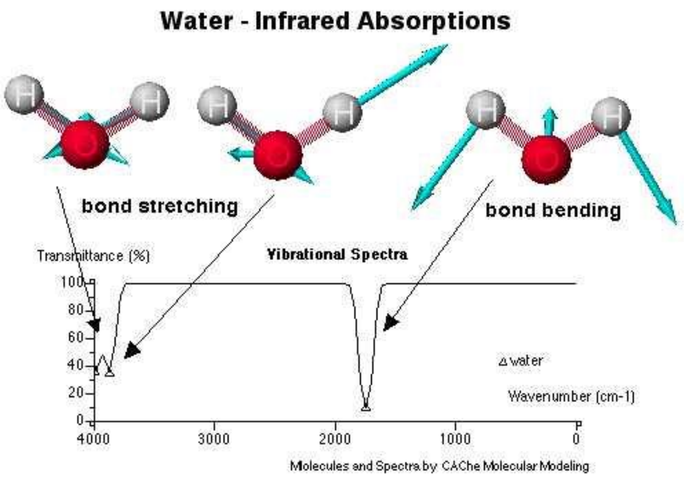


Fig. 4

When EM (electromagnetic radiation) hits a molecule, if the radiation is the correct wavelength (and so energy by $E = \frac{hc}{\lambda}$ where E is energy in joules, h is the Planck constant, c is the speed of light and λ is the wavelength of the light in metres), it can cause the bonds in the molecule to bend and vibrate (form an oscillating dipole) - this is called the absorption of IR energy. The molecule gains kinetic energy by this absorption of IR radiation. This extra kinetic energy may then be transmitted to other molecules such as oxygen and nitrogen and causes a general heating of the atmosphere. Different possible stretching, bending or rotating orientations of the bonds absorb different wavelengths of light. For instance, as is shown in Fig. 4, when a water molecule has its bonds stretched, it absorbs radiation in the region of 2500 nm whilst when its bonds are bent, it absorbs radiation in the region of 5700 nm (a longer wavelength). Since all bonds in all molecules can be stretched or bent in some manner, one might assume all molecules can absorb EM radiation and then potentially reemit it (i.e. go through the greenhouse effect). However, in many molecules such as N_2 and O_2 (the main components of the atmosphere), the molecules are completely non-polar and symmetrical so any possible stretching or bending will only form a different kind of symmetry (there can be no oscillating dipoles so no energy can be absorbed by these gases). Many molecules will oscillate very slightly, but the energy absorbed is so minimal that it can be said to be zero.

Purely absorbing some of the IR radiation being reflected off the earth is not a problem. However, the consequences are the problem. This absorption increases the temperature of the atmosphere around the earth, providing the earth with a 'warm blanket', heating it up. Also, the molecules, having absorbed the radiation will reemit after some time has passed and this will further increase the temperature of the earth. These two consequences are known as 'Greenhouse Effect'. This increase in the earth's temperature is causing melting of the polar ice caps and so increased sea levels amongst other negative climate impacts. These negative consequences are the main reason why many refrigerants are deemed to be bad; fridges tend to leak the refrigerant slightly and so, with refrigerators being so widely spread, the increase in the Greenhouse Effect is very noticeable.

The alternative mechanism to current refrigerators

Therefore, it is clear that refrigerators have a number of problems: the refrigerants, the moving parts, the energy consumption etc. If an ideal refrigerant were found, then this model would indeed be a very good model for refrigeration. However, currently (as seen earlier), all common refrigerants are problematic and it seems unlikely that an ideal refrigerant could be found (that could also keep costs low). Therefore, it would be ideal if a new mechanism could be formulated that contained no need for refrigerant. This is exactly what the thermoacoustic refrigeration model is.

As we can see in Fig. 5, a thermoacoustic refrigerator is a tube (or resonator cavity) with an inert gaseous atmosphere, such as one of argon. A sound wave is sent down the tube from the oscillating piston (or loudspeaker, in many other cases) and because the tube is closed and is made so that it is half the wavelength of the sound wave, the sound wave will be reflected off the end of the tube (the 'rigid termination'). This results in two progressive sound waves travelling in opposite directions. These waves meet and superpose, forming a standing wave. Because the gas is now the medium for a longitudinal standing wave, there will be compressions and expansions of the gas in the tube (as seen in Fig 6). When the gas is compressed, the average kinetic energy of the particles is greater since energy has been imparted to them as the wave propagates through the medium. Therefore, collisions with the walls, since the collisions are not perfectly elastic, will tend to result in the particles losing kinetic energy so at compressions, the gas tends to emit heat whilst at expansions the reverse is true (the gas will absorb heat).

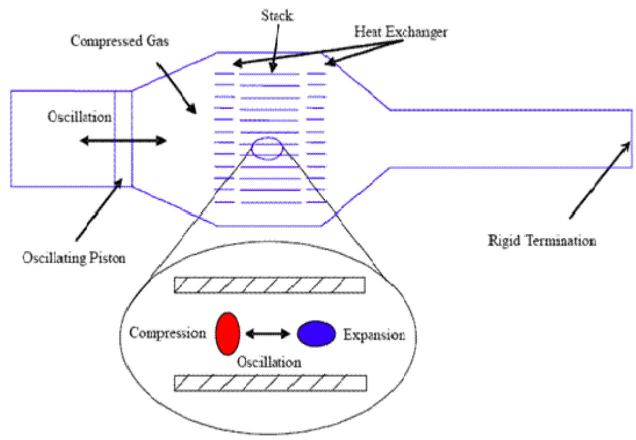


Fig. 5

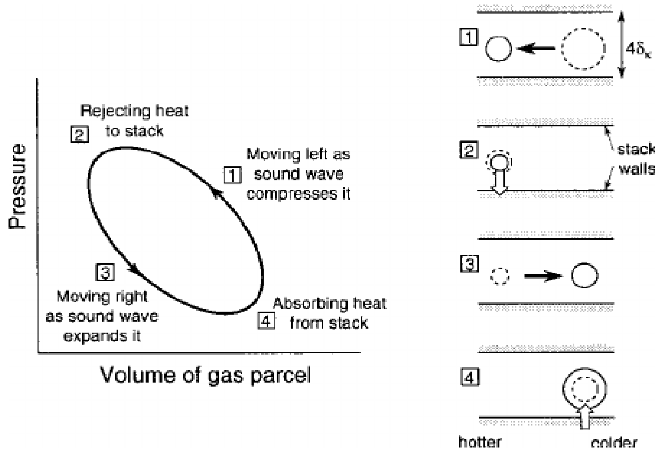
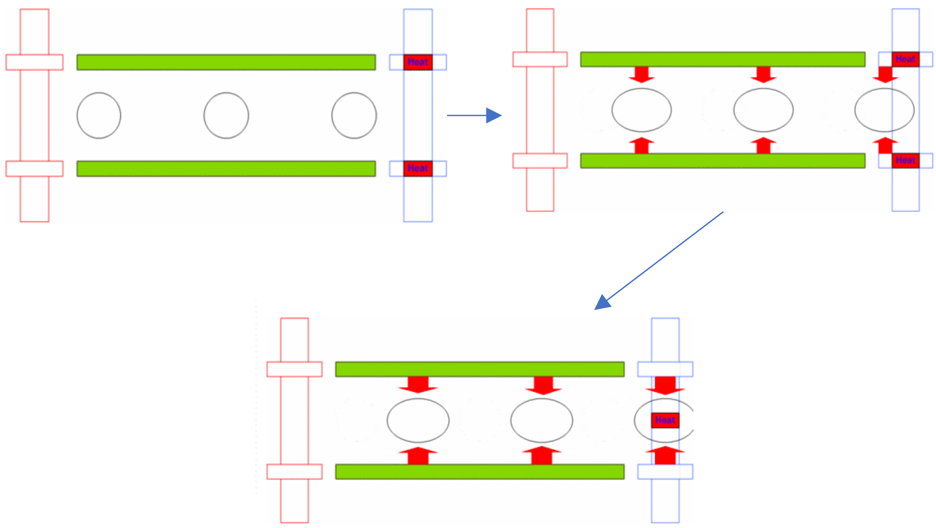


Fig. 6

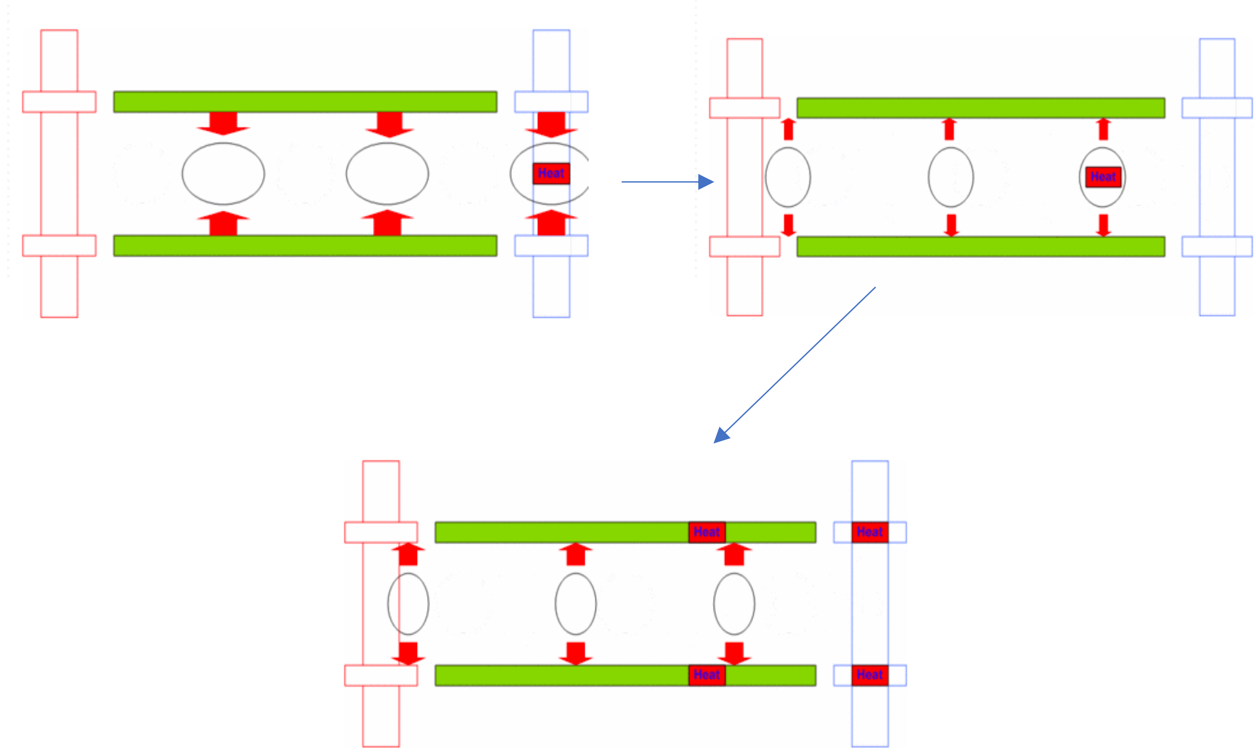
Inside the tube that is the thermoacoustic refrigerator, there is an object called the stack (a porous solid medium that increases the gas-solid interface, enhancing the heat exchange). On one end of the stack is the end of the refrigeration mechanism (i.e. the end that provides the heat that has been absorbed from the refrigerator) and on the other end is the connection to the heat sink (i.e. the place where the heat is removed to the outside world to be used for other purposes). These two are placed in such a way that at the end of the stack connected to the refrigerator, the gas is expanding (i.e. can absorb heat) and at the end connected to the heat sink end of the stack, the gas is compressed (i.e. it can emit heat). Thus, a temperature gradient is formed in the stack (one end is viewed as the hot end and the other as the cool end). Throughout the stack as well, if 'packets' of gas are considered, these packets of gas are constantly being displaced forwards and backwards and at the same time at expanding and contracting and so are absorbing and emitting heat.

As is clear from Fig. 6, as the packet of gas is displaced forward, it compresses and emits heat to the stack. It is then displaced backwards as it oscillates and expands. At its maximum expansion, it will absorb some heat which it can then emit as it again compresses and so the cycle continues until the standing wave is discontinued. This is further explained in the next diagrams.

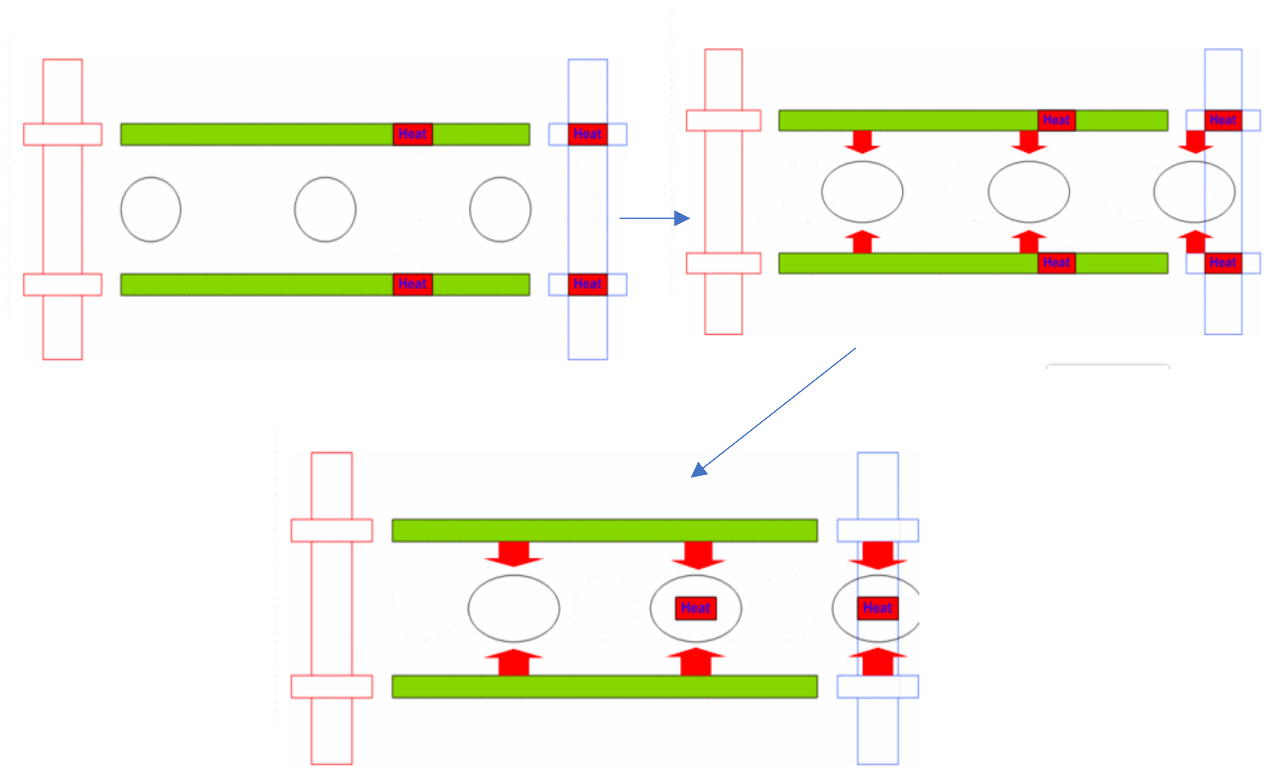
First, the gas packets oscillate to the right and expand and the one on the furthest right absorbs the thermal energy absorbed from the refrigerator:



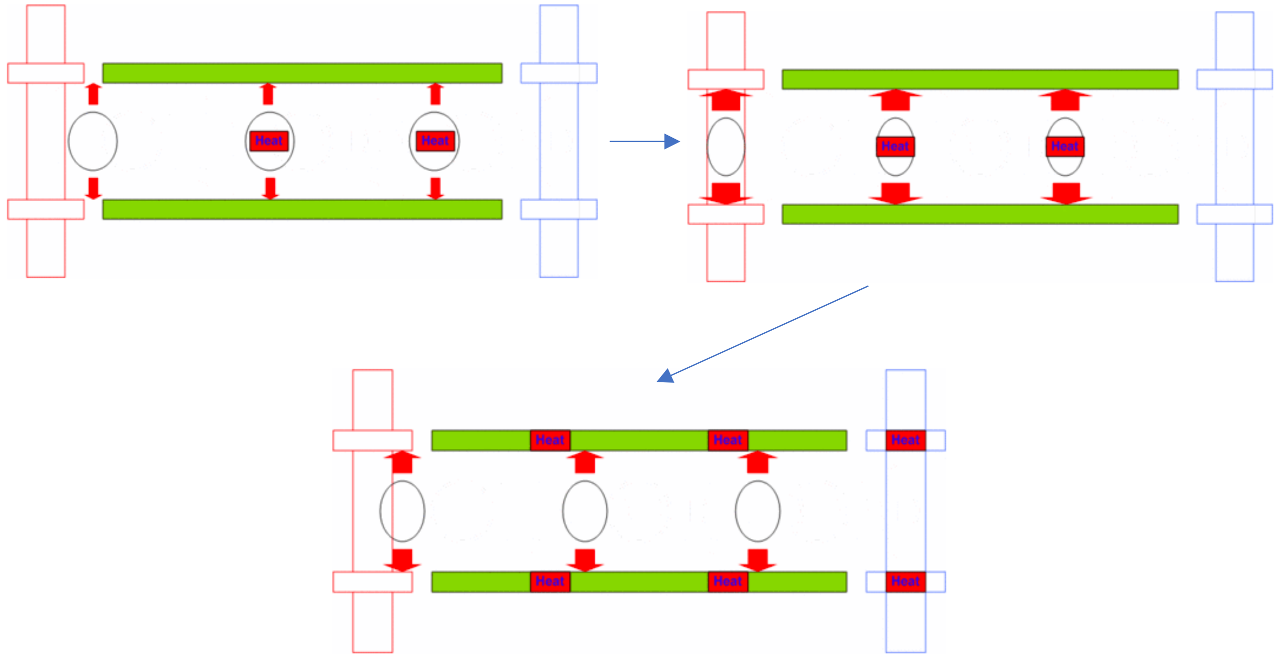
Next, this gas packet is displaced to the left and compresses due to the wave (also increasing its temperature). This gas packet then emits its extra thermal energy gained from the 'cold' (or blue in the diagram) end of the stack:



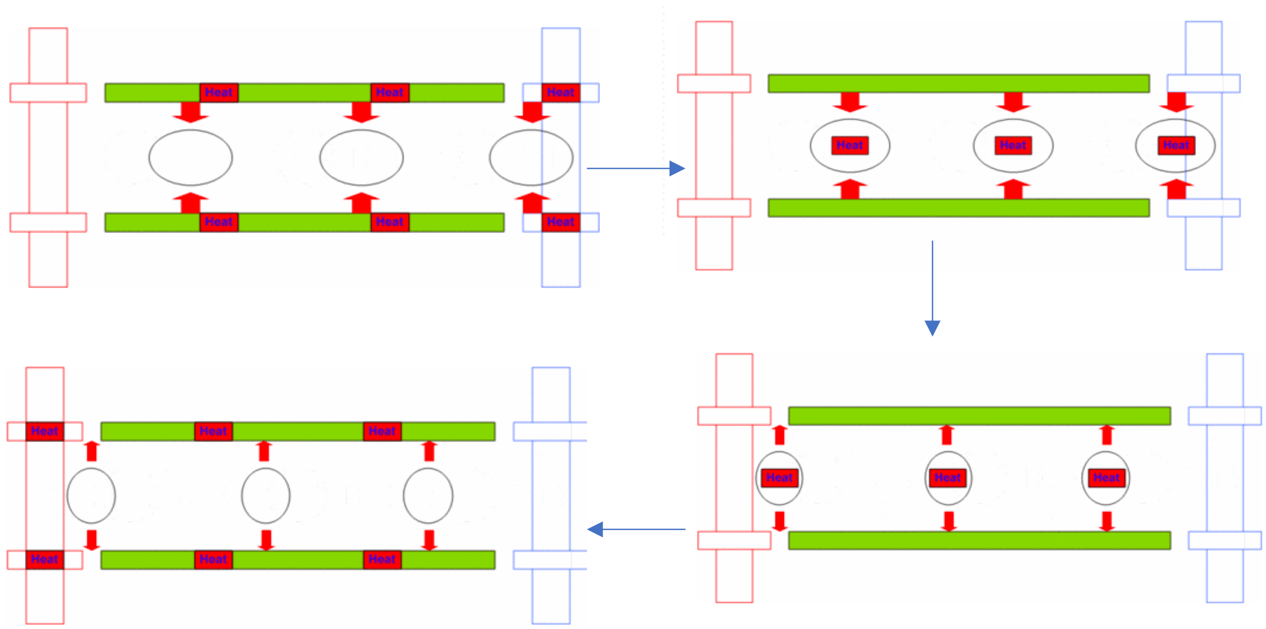
As the three gas packets are displaced to the right and again expand, the second gas packet absorbs the heat emitted from the first gas packet to the stack and the first gas packet absorbs a new 'packet' of thermal energy:



These gas packets then oscillate to the left and are compressed by the wave causing them to emit their excess thermal energy:



Finally, the third gas packet, as it is displaced to the right expands and absorbs the heat from where the second packet emitted it and transfers it to the heat sink (since the gas packet is compressed there so emits its heat).



This cycle repeats indefinitely until source of the standing wave stops; transferring heat and resulting in some cases with temperatures as low as -67 degrees centigrade.

Design of a Thermoacoustic refrigerator

A Thermoacoustic refrigerator consists of five main components: The Speaker, the Resonator, the Stack, the inert gas and the Heat Exchanger.

The Speaker itself is fairly self-explanatory, except that it must be able to produce amplitudes of around 180-200 decibels (a typical Thermoacoustic refrigerator operates at around 190 decibels). A typical speaker used in Thermoacoustic refrigeration can be seen in Fig. 7.

Fig. 7

Next, the Stack is the porous solid that increases the gas-solid interface. Its main purpose is to induce a temperature gradient across itself. Because of this, it must have a high heat capacity (i.e. it must require a large amount of energy to raise the temperature of the substance by one degree) but a low thermal conductivity (i.e. it must be an insulator, so it will not conduct heat well). This is so that the stack can absorb large amounts of heat in specific places whilst not dissipating this heat throughout the solid itself (so that the heat can be 'picked up', absorbed, from where it was 'dropped off', reemitted). A good example of a suitable material, therefore, is a ceramic.

As is clear from the table in Fig. 8, a ceramic has a low thermal Conductivity but a high thermal capacity and so is suitable as the material for the stack (although materials for the stack is still an area of research). The shape of the stack is also under research. However, currently a honeycomb structure (similar to that of the catalysts in car exhausts) is the preferred structure for the stack (see Fig. 9). Finally, the pore size in the Stack is under research - clearly, smaller pores would increase the surface area of the Stack and this would be desirable, but, there will also be some energy lost by the friction between the gas and the solid (this is known as viscous loss). There must be a balancing act between increasing the surface but not making the surface are so large that viscous loss becomes a major factor.

Property	**Unit**	**Value**
Thermal conductivity at 25°C	W/m-K	3
Thermal capacity at 100-200°C	J/g-°C	0.500

Fig. 8

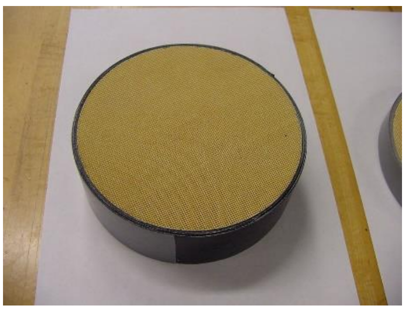


Fig. 9. This is an example of a typical Stack. Although it is not clear in the picture, the stack is actually riddled with tiny holes, significant of the honeycomb structure mentioned above. This increases the solid-gas interface significantly.

Fig. 10. An example resonator for a thermoacoustic refrigerator.

Next, the resonator is the length of tube after the stack that allows the tube to be the correct length to produce resonance in the standing wave (see Fig. 10 for an example). It must be made of a strong material that can cope with high-pressure gases so as to prevent any of the working fluid from leaking. It must also have a low thermal conductivity so as to prevent heat flow from the environment into the 'cold end' of the stack (the end connected to the heat sink). Therefore, a suitable material would be polyvinylchloride (PVC), which has a thermal conductivity of 0.19 W/(mK) (i.e. a very low thermal conductivity) and has a breaking compressive stress of up to 89 MPa and a breaking tensile stress of up to 62 MPa (so, since the pressure fluctuations are around 30 kPa, this is well within PVC's range for stress).

Advantages and limitations of Thermoacoustic refrigeration

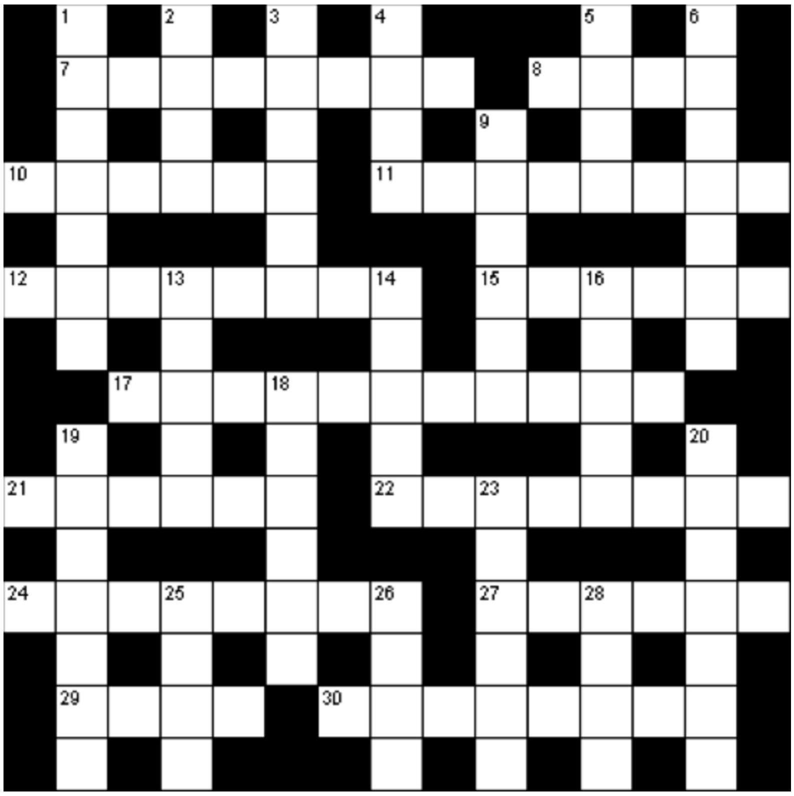
Thermoacoustic refrigeration has a number of significant advantages over conventional refrigeration. Most notably, it has no need for harmful and environmentally damaging refrigerants and so is more beneficial for the environment. It also has no moving parts, so is a much simpler structure that requires no moving seals and it is much easier to manufacture. It needs very simple materials and consumes much less electricity than a conventional refrigerator, once again making it more environmentally friendly (it can be powered on solar or even just waste energy). Finally, there is a far lower need for maintenance in a thermoacoustic refrigerator.

However, there are a number of limitations to thermoacoustic refrigeration that need to be overcome in order to make it a common mechanism for refrigeration. In order to increase the power density of the thermoacoustic refrigerators, very high densities of the operating fluid are required and it is difficult to obtain these high densities of the typical gases. Because thermoacoustic refrigeration is uncommon and requires specially made parts to increase its efficiency, it is difficult to obtain some specific thermoacoustic refrigeration equipment (such as specially made alternators, dynamos generating current, for the loudspeaker). The acoustic standing wave itself is at an unusually high amplitude and this results in many new factors such as turbulence, viscous losses (as mentioned earlier) and harmonic generation of other frequencies. Finally, due to these factors, the thermoacoustic refrigerator is currently less efficient than a current refrigerator.

Conclusion

Therefore, it is clear that thermoacoustic refrigeration is an exciting new area that has already been shown to work well (thermoacoustic refrigeration has been used by Ben and Jerry's, on military ships and even on a space shuttle!). It serves as a viable alternative for conventional refrigeration mechanisms that is simpler and more environmentally friendly. However, it requires more investment in order to become more widespread (so as to make obtaining of special parts easier). There are also a number of different limitations that need to be overcome before these replace conventional refrigerators, in particular improving the efficiency of these refrigerators. However, with more research and investment, this form of refrigeration could soon be available for commercial and even domestic use.

Cryptic Crossword



Across

7. Vote Reagan to be a lepton. (8)

8. Points get a tad backwards. (4)

10. ??? (6)

11. All the clues show an electric field and six foxes' lair next to the church (8)

12. Anguish as first of hussars leaves messy distribution (8)

15. He mixes lithium with uranium and hemp, without power (6)

17. The polisher gets confused above the mantle (11)

21. Reflectance is doable? Possibly (6)

22. Physicist is out of phase with the spread (8)

24. One German beer glass is a genius (7)

27. Biologist has success after return of radiation unit (6)

29. Cut used by young scientist (4)

30 Mix once dead hydrocarbon (8)

Down

1. Iron and many mixtures point to Nobel Laureate (7)

2. Werner gets chopped up and floats off in the sea (4)

3. The strange artist of clouds (6)

4. North's position? (4)

5. Judge reaction speed (4)

6. For metal, mix 1cc aluminium and unlimited minerals initially! (6)

9. Fraction of weight halved? (6)

13. Sticky liquid of entropy and citrus fruit. (5)

14. Cavity for Lewis' lion possibly? (5)

16. Young insect hears hot rock (5)

18. Robert's penny is addicted (6)

19. Orbit otherwise encircles cheek (7)

20. Some gelignite detonated and combusted.(7)

23. Device that's tagged all over. (6)

25. A twirl for fermions? (4)

26. Noble light (4)

28. Numbers for confused Shakespearean king. (4)