GRESHAM COLLEGE

Mathematics Goes to the Movies

Professor Chris Budd OBE

1. Introduction

Everyone likes going to the movies, many of which play a major cultural role in the modern world. This talk is all about the role that maths plays in the movies, and hence indirectly, on the culture of the modern world. At one level it may not appear to play much if a role at all, as the vast majority of movies do not seem to mention mathematics at all. This is not always the case of course. Maths does figure in the plot of such blockbusters. One of these is Jurassic Park, with Jeff Goldblum playing the mathematician concerned with the stability of the park and how this can be explained by using 'chaos theory'. Another is Sherlock Holmes, in which (in the film Sherlock Holmes: a game of shadows) the great mathematician Professor Moriaty can be seen in front of a blackboard (designed by the genuine mathematician Prof Alain Goriely at the University of Oxford), which

gives away the entire plot in mathematical form. Anyone who went to my last lecture will also have seen the role played by the maths of Big Data in the film Bridget Jones's Baby. Some recent films have even featured real or fictional mathematicians as the central character, notable amongst these are The Imitation Game, about Alan Turing, A Beautiful Mind about John Nash, The Man Who Knew Infinity about Ramanujan (and others) and Good Will Hunting. Whilst it is very good (in my opinion) that mathematics is seen in the plot of these films (and I hope that we will see many more such films, for example a film about Galois would have it all in my opinion), it is not that aspect of maths which I will be discussing in this lecture. Instead, I want to talk about the role that maths plays in making modern movies possible at all, even if they never once mention mathematics. In particular I will look at the way that maths is used to create all of the special effects that modern films rely on so much. As an example, the wonderful Lord of the Rings trilogy used 263 special effects artists, all of whom would have been using a huge amount of maths to do their work. Similarly, such giants of the film industry, such as Pixar, employ many mathematicians and computer scientists (such as Ed Catmull and Tony DeRose) in very senior positions. As well as being a really good and challenging use of maths in its own right, it is well worth telling young people about the way that the maths that they learn in school has such exciting and creative applications in such a glamorous and dynamic industry. Hopefully this will give them a strong motivation to study, and enjoy, even more mathematics.

In this talk I will start by showing how maths is used to create the image. We will then see how to use perspective to make big things look small and small things look big. We will then move on to see how some wonderful mathematics, carved into a bridge in Dublin, is now used to animate many of the creations we see in the movies. I will certainly be including some maths, so be warned, but not too much. Readers who want to read more into the maths should have a look at the references at the end of this article. But the real message to take home is that the film industry is one of the places where the true creativity of mathematics (and computer science) can be seen in one of its most undiluted forms.

2. Making the image

Creating   I have some bad news for you if you like Shrek. He is not a real Ogre. I'm sorry if that disappoints you. What 'he' is instead is a collection of points, curves and surfaces, stored inside a computer. These points are moved around to give the impression of a living and breathing ogre, interacting with its environment. However, getting that set of points to look realistic is not easy, and that is the first step of the computer animator. To create the image it is common to build a model of an object that we wish to animate. The object is then scanned with a laser, and the coordinates of the points on its surface are stored in the computer. However, this is just a set of points and does not give any sense of depth or solidity to the character. To do this, we join up the points in a manner which looks like a wire skeletons, which are made up from simple polygons such as triangles which describe the surface between the points and which have a sense of direction, so that we know which side of Shrek is his outside and which is his inside. As the points move, so the triangles move with them, giving the sense of moving a whole body. The more triangles we use, the more realistic the figure appears to be, but also the more computer calculations that are required. Fortunately, modern computers with Graphical Processing Units (GPUs) can do these calculations very quickly, and animations with many millions triangles are not uncommon. The technique of dividing a complex object into many triangles is used in many other applications than the movies. An important example is the finite element method, which breaks down complex problems (such as calculating the air flow over the wing of an aircraft) into many simpler problems, defined on a very large number of triangles. The more triangles we use, the more accurate, but also slower, is the calculation.   On the right we see a picture of Rex from the terrific film *Toy Story 1* (perhaps the first feature length film to be completely computer animated) in which we see the sort of image that you get with just a few triangles.

Rendering and Subdivision Whilst triangles are very useful to define the shape of the object, and (as we shall see) they continue to be useful when lighting it up and moving it, they do not give images which look as realistic as we might like. To produce more realistic images a second process is used to give a smoother picture. There are various ways of doing this. The original methods were based on techniques used in computer aided design (CAD) which were in turn based on ideas used in the field of *Mathematical Approximation*. In these methods a smooth curve or surface is used to recreate a shape defined by its values at a set of points. The workhorse of such methods being curves called non-uniform rational B-Splines (NURBS) which are drawn between the points. Whilst these give excellent surfaces, the methods used are rather too slow to be used easily in film production. Much more widely used in recent film production are subdivision algorithms, which are used iteratively to add more points to an image to give a smoother looking picture. The image below shows a series of applications of a subdivision algorithm.
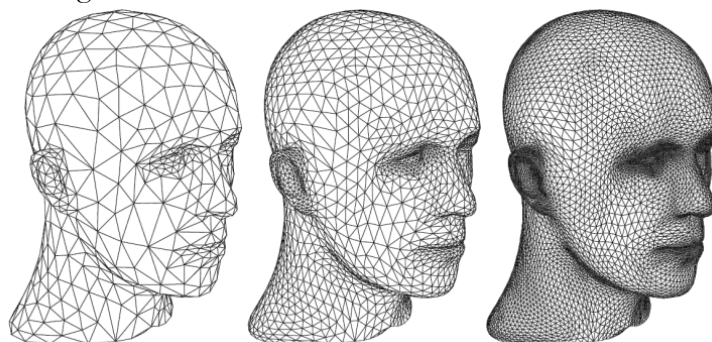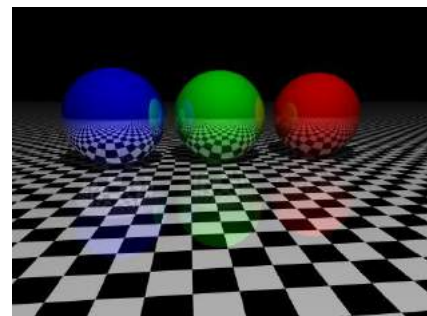
The *Catmull–Clark* algorithm is such a technique. It was devised by Edwin Catmull and Jim Clark in 1978 as a generalization of bi-cubic uniform B-spline surfaces to an the arbitrary shapes met in computer graphics [5] In 2005, Ed Catmull received an Academy Award for Technical Achievement together with Tony DeRose and Jos Stam for their invention and application of subdivision surfaces. Catmull (a computer scientist) is now the current president of Pixar Animation Studios and Walt Disney Animation Studios, where DeRose is also head of research. The lovely short film *Geri's* Game (of an old man playing himself at chess) was one of the first computer animated films to use the subdivision method to produce the realistic surfaces, as well as being a show case for the new realism it gave to skin and cloth surfaces.

Lighting it Up   We see images by the light which is reflected from them. Usually this comes from a natural source such as the sun or an artificial source such as a light bulb. The light can come from a point source such as the sun, or a diffuse source such as the sky. Different surfaces reflect light in different ways, and light from these objects can reach our eyes from many different directions. This complex process must be reproduced accurately if our objects are going to look realistic. The procedure for doing this is called *rendering* and it takes up a great deal of computer time and energy. In the recent production of a James Bond movie, the rendering process took so much computer power that it led to a power cut in the London district of So-Ho where the movie was being produced. The usual method to do this is *ray tracing*, in which a very large number of computer generated light rays are artificially shone onto the object. The light rays are assumed to be straight lines, which reflect off the object in different ways depending upon its orientation and material. The process of ray tracing starts by identifying the viewpoint of the supposed moviegoer. Starting from this point, rays are traced backwards towards the object and reflected off from it with different strengths depending upon the nature of the material (so it will reflect well if the object is metallic and much less well if it is dull cloth for example). The angle at which it is reflected is calculated by using the triangles, which we have used to represent the object. By looking at the point at which the ray meets the triangle, and also the angle between the ray and the (normal vector to the) triangle it is possible to calculate quickly the direction of the reflected ray. If this reflected ray intersects a light source, we shade that triangle in a bright colour so that it appears lit up by the light source. If the reflected ray does not meet the light source, we shade the triangle in a darker colour. We then repeat this for a million or so different rays. This is an example a simple algorithm in practice, which can take a long time to run, so that the time it takes to artificially illuminate a single movie frame can sometimes be measured in hours. In recent years this process has sped up, due to a combination of faster computers, using computers in parallel (to track many rays simultaneously), as well as much better mathematical algorithms. Similar ideas are used to improve the performance of mobile phone and Wi-Fi systems. Computerised ray tracing can produce very realistic scenes but it is very slow. This is usually acceptable for producing movies if you have fast enough computers to make the effects (despite the risk of power cuts – which I will return to in a future Gresham Lecture.) However, the slow speed in doing these calculations becomes a problem when you need the lighting and other effects to change in real-time. This situation arises in applications such as computer games and air flight simulators. In such cases, complex optical phenomena such as shadows, caustics and multiple reflections of complex surfaces, are hard to reproduce quickly, and more sophisticated mathematical methods, for example pre-computed radiance transfer and radiosity, are used to reproduce them. Similar techniques to this are also used to calculate the radiation fields used by engineers to predict the performance of Wi-Fi and mobile phone systems.

Adding Effects Most modern movies make a huge use of special effects. As en example, in one of the final scenes of the Lord of the Rings, Return of the King, we see Gandalf riding on the eagles, rescuing Sam and Frodo from the lava of Mount Doom. Obviously eagles and a volcano were not conveniently available for the film and these effects had to be created inside the computer. Other effects, which need to be created, are the movement of smoke and fire, water, weather or simply the physics of the way things move, interact and deform. In the past, animators would use their best judgement about how to make such effects look realistic (with vary degrees of success) but now we can do better by accurately recreating the underlying physics of the situation that we are watching. To do this we take the laws of physics, which describe the situation, represent these using (partial differential) equations and then solve these (differential) equations. The solution is then added to the

other characters to produce the movie. Solving these equations is far from easy and requires sophisticated mathematical ideas and computer algorithms, drawn mainly from the fields of numerical analysis and scientific computing. Similar techniques are used in industry to simulate many varied processes, such as the flow of air over an aircraft's wing by solving the Navier-Stokes equations of fluid motion (see below),
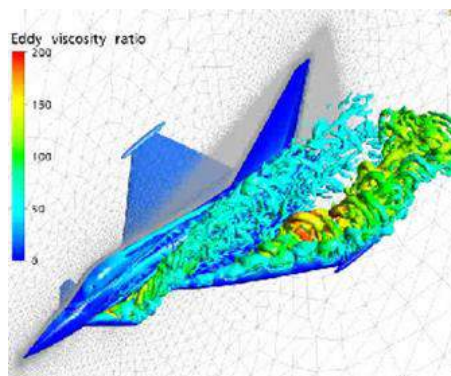
forecasting the weather or climate, and even the application of techniques used to solve cancer. I will return to more of these methods, which relate closely to my own research, in later lectures. However, it is worth saying that the difference between film animation and industrial simulation is that in the former the physics has to look right whilst in the latter it really has to be right. This means that it is possible for films to take short cuts to make the computations faster, creating what Pixar call 'Pixarian physics' in the process.

3. A matter of proportion

All movies have small and large things, which may be people or objects. Some are very large indeed, such as the dinosaurs in the movie Jurassic Park, or they may be small, such as the Hobbits in the Lord of the Rings. In fact I asked myself the question on watching Lord of the Rings for the first time, how did they get such small actors to play the hobbits? The answer was that they didn't. All of the actors were of average height (except the dwarves who tended to be played by taller actors as we shall see presently). This was done, in the main, by using the relatively simple technique of forced perspective. This works from the simple basis that a small object close to the camera will look as large as a bigger object further away from the camera. More precisely, if an object is of height L and it is a distance d from the camera then the angle θ it makes (in radians) is close (if d is large enough) to the ratio L/d. If θ is small then the object will appear smaller to the viewer. This can be achieved either by making L smaller or by making d larger. So to make a 'hobbit' appear to be small the actor (who we assume is of a normal height), has to be placed further away from the camera, than an actor playing a larger character such as a wizard. The same principle applies to houses, bridges and even to dinosaurs. Of course this has to be done very carefully to make sure that the characters all appear to have consistent sizes. If they are moving then the camera has to move with them to keep the same perspective. All of this requires quite careful mathematics, and, I am reliable informed, a team of people armed with rulers and tape measures on the set. This is a really nice use of the geometry and trigonometry taught at schools, and is an excellent way of motivating the teaching of this subject. In fact with the use of cameras on their phones, school students can instantly become film directors and can create their own special effects. Sometimes the use of forced perspective also involves the use of special props. There is a long history of combining forced perspective with small models of objects, such as castles, to recreate the illusion of much larger buildings. Another example was the shot below of Frodo and Gandalf on a cart together. Frodo is made to appear much smaller by using a specially designed cart in which he is actually placed much further away from Gandalf than he appears to the viewer.

It is fair to say that many other, less mathematical, tricks were used in the Lord of the Rings films to give the impression of different heights of the characters. One was to use children or small actors to play hobbits when

they were filmed from behind or from a distance. Another was to use different sizes of props, so that small propos were used when they wanted to make the actors appear to be large, and large props when they were to look small. An example of this is the use of different sizes of hobbit holes in the recreation of Hobbiton. In the images below you will see my son when he recently visited the Hobbiton set in New Zealand. On the left he appears to have the height of a (small) hobbit, and on the right of a (tall) wizard. This is all achieved by having hobbit holes of different sizes. Again maths is used to make sure that the proportionate sizes all look correct.

Finally, we turn to the tricky subject, of the relationship between hobbits and dwarves. In the Lord of the Rings books it is clear that whilst dwarves are smaller than men, they are bigger than hobbits. So, if hobbits and dwarves are played side by side it is logical for the dwarves to be played by larger actors than those playing the hobbits. As the latter are of average size, we reach the paradoxical conclusion that actors playing dwarves have to be taller than average.

4. Making it all move

NOTE This section is mathematically much heavier than the preceding sections. Readers of a more nervous disposition might like to skip ahead to section 5.

The whole essence of movies is that the characters in it move and change. It is important when making a movie that such movement looks realistic. To do this involves both mathematics to create the illusion of movement and physics, to make sure that the movement occurs in the same way that it would in real life. The original forms of animation used a 'stop-go' method, in which the character to be animated was changed slightly between frames. Originally this was done by film studios, such as Disney, by using a very large number of (hand drawn) cartoons. Later on the same technique was used on models, such as the famous skeletons by Ray Harryhausen in the film *Jason and the Argonauts*, or the wonderful 'Claymation' by Nick Park used to create the characters of Wallace and Grommit. However, such techniques (though still wonderful) are time-consuming and labour intensive to create. The modern mathematical methods are not only quicker to use, but can give a much smoother impression of movement

The key idea behind computer generated movement is again to use *coordinates*. Let's suppose that the vertices of the triangles used to make up our figure have the coordinates (x,y,z). If we change these coordinates with time then we can move the triangle. The question is then, how do we change the coordinates to make the movement appear to be realistic.

There are basically three types of motion, *translation*, when a character moves to one side, *enlargement* when the character moves nearer or further away (see the above discussion on proportion and size), and *rotation* where they rotate about some axis. In this motion a whole character can move, or just part, say an arm.

Translation can be achieved by *addition*. If we take the coordinates (x,y,z) and create new coordinates (x+a,y+b,z+c) then we have translated our character in the direction of the vector (a,b,c).

Enlargement can be achieved by *multiplication*. Again, if we take the coordinates (x,y,z) and create new ones (a*x,b*y,c*z) then we have stretched it in three directions. If a,b and c are all equal and greater then one, then this will give the effect of the character moving closer.

Rotation is achieved by a combination of matrix multiplication, trigonometry and complex arithmetic. I'll start by explaining rotation in two-dimensions. If you want to rotate a point with coordinates (x,y) by an angle θ about the origin at the point (0,0), then the new point has the coordinates

$$(x \cos(\theta) - y \sin(\theta), x \sin(\theta) + y \cos(\theta)).$$

Here the functions $\cos(\theta)$ and $\sin(\theta)$ are the lengths of the sides of a right-angled triangle of angle θ with a hypotenuse of length one. The same result can be obtained by multiplying the vector (x,y) by the *rotation matrix* R given by

$$R = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Matrix arithmetic (first developed by Cayley in the 19th Century) is taught to school students towards the end of their time at school. It is often taught in a way that seems like a bunch of tricks with no relation to real life applications. However, as we can see above, it plays a vital role in computer animation.

A different, and arguably more elegant, way of describing rotations is to use complex numbers. We start by taking the *imaginary number* i, which satisfies the equation $i^2 = -1$ and form complex numbers such as $z = x + i\,y$ or $c = a + i\,b$. We can then combine these by complex addition and complex multiplication so that

$$z + c \;=\; x + i\,y + a + i\,b \;=\; (x + a) + i\,(y + b),$$
and
$$z * c \;=\; (x + i\,y)*(a + i\,b) \;=\; (ax - by) + i\,(bx + ay).$$

If you look carefully you will see that addition is again the same as translation, and if we take a = $\cos(\theta)$ and b = $\sin(\theta)$, then complex multiplication is the same as a rotation.

The above describes motion in two-dimensions (which is what we might see in cartoons), however modern animation describes characters which move in three dimensions. This is much more complicated as there are many more ways that a character can move. Again there are essentially two ways that this can be done. A very general and flexible method is to use matrices again. In this case we will use a 3*3 rotation matrix R of the form

$$R = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

multiply a three dimensional vector (x,y,z) by this to rotate it to about any axis by any desired angle. Here the nine numbers a,b,..,h,i have to have certain relationships with each other so that R describes a realistic rotation. In particular the determinant of R needs to be one. It is also not very clear what the axis and the angle of the rotation is. A different, and arguably simpler and more elegant, method, uses the mathematical numbers called *quaternions*. The history behind these numbers is fascinating. Many mathematicians sought to extend the ideas behind complex numbers which I have described above not only for mathematical curiosity (never a bad idea) but also to do exactly what we want, namely to find an elegant way of describing motions in three dimensions. The answer was invented/discovered in 1843 by the great Irish mathematician William Hamilton. The story goes that he was walking in Dublin, and in a moment of inspiration came up with the basic idea for quaternions. Not having a pencil to hand he (as any sensible mathematician would have d0ne at the time) carved his formula into nearby Broome Bridge. Sadly the formula no longer exists in its originally carved form, but a plaque commemorating it does. See below.
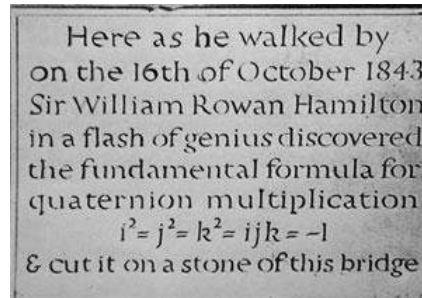
So, what was this marvellous formula? Hamilton's inspired idea was to extend the imaginary numbers, so that there were three different ones i,j and k which he called quaternions. The rules for multiplying them were then

$$ij = k, \ jk = i, \ ki = j, \ ji = -k, \ kj = -i, \ ik = -j$$

together with the wonderful formula

$$i^2 = j^2 = k^2 = ijk = -1.^{[1]}$$

A general quaternion is a combination of these numbers with the usual numbers and it takes the form
$$q = a + b\,i + c\,j + d\,k,$$

where a,b,c, and d are arbitrary Having got used to the quaternions and the way that we can add and multiply them (the laws for which can very easily be programmed into a computer) we are in a position to apply them to computer graphics. To do this we take the point P = (x,y,z) that we wish to move, and form the quaternion number

$$w = x\,i + y\,j + z\,k\,.$$

Now, suppose that we want to rotate P, by an angle θ, about an axis of unit length pointing in the direction (a,b,c). We firstly form the two quaternions

$$q_1 = \cos(\theta/2) + \sin(\theta/2)(a\,i + b\,j + c\,k),\ q_2 = \cos(\theta/2) - \sin(\theta/2)(a\,i + b\,j + c\,k),$$

then from the quaternion w we calculate the new quaternion v given by:

$$v = q_1\,w\,q_2.$$

Remarkably, the quaternion v takes the form

$$v = X\,i + Y\,j + Z\,k$$

where (X,Y,Z) are exactly the coordinates of the point P after it has been rotated about the given axis. Voila! Although this looks complicated, it is in fact very easy to set up and to programme, and it gives a very fast way of moving our computer generated characters around. In fact to do this we need 4 numbers a,b,c and θ , which is a lot less than the nine needed for the matrix method. I think it is wonderful that Hamilton's great discovery of quaternions can be used in this way.

---

[1] The names of the sets of numbers starting with the rational numbers, and increasing in size with the algebraic numbers, the real numbers, the complex numbers and finally the quaternions, spell out the word QARCH. This is the name of the puzzle journal for the Cambridge Undergraduate Maths society the *Archimedeans*.

To animate our character we now apply the quaternion multiplication method to each of the vertices of the triangles that we have used to make up our figure. This can be done quickly and easily to make the character move.

Other applications of quaternions arise in studying the motions of satellites, and even of Olympic gymnasts. Much more detail about this can be found in the lovely article [1] by Joan Lasenby at Cambridge, which also explains how you can use quaternions to deform and interpolate objects to get realistic motion.

5.   Capturing the spirit

The animation effects described above are really only part of the story. Whilst mathematics can create and move characters around, how do we give them a spirit, so that they look, and behave like real people? A lot of modern films (including for example the Lord of the Rings movies, and other blockbusters such as Avatar) do this by using a combination of computer graphics and real life actors using the technique of 'performance/motion capture'. As an example the figure of Gollum in the above movies is an animated version of the actor Andy Serkis shown

above on the right (who has appeared via motion capture in many other films and has even set up his own motion capture workshop, *The Imaginarium Studios*). In motion capture the actor wears a suit covered with many reflecting dots. As the actor moves, several cameras capture the motion of the dots in three dimensions using a fair amount of careful geometry. These points are stored in the computer and then joined in to make a wire frame image, which can move around in exactly the same way as the original actor.
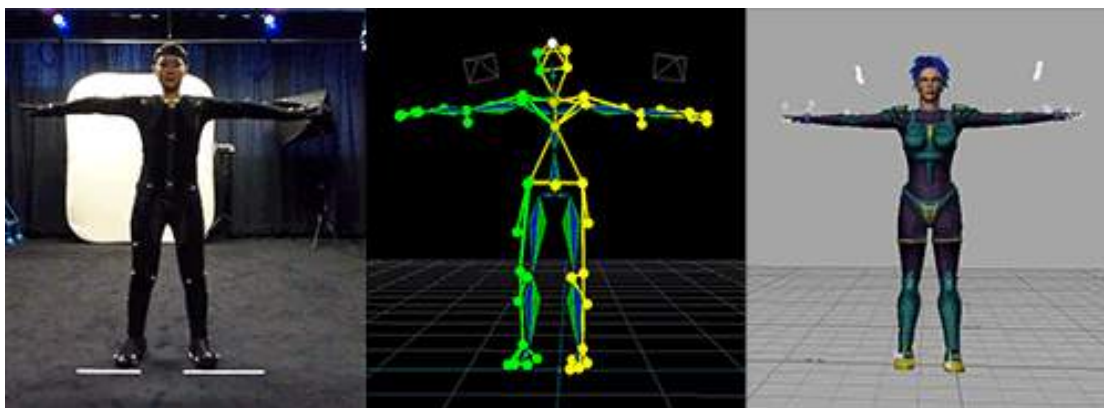
To complete the image, the wire frame model is triangulated and rendered to build up a realistic looking body, using many of the techniques we have already met. A particularly challenging aspect of this is to produce a realistic face which can reproduce the actual expressions of the actor. This is achieved by having a reference face, which is then wrapped around the dots on the actors face using a carefully designed mathematical optimisation algorithm.

I am proud to say that one of the leading centres for this form of animation is the Bath Centre for the Analysis of Motion, Entertainment Research and Applications (Camera) and you can read about its work in [4].

6. What next

As we have already seen, one of the key figures in Pixar is Tony DeRose who started his career as a Professor of Computer Science at Washington University. He moved to Pixar just after its successful release of the film Toy Story and his breakthrough was the film Geri's Game I mentioned above. He was asked in an interview following a Ted Talk [3] *'What advice would you give someone who wanted to use math to make movies?'*

His response was:

> *Learn as much mathematics as you can, particularly applied math. The areas of mathematics we use most heavily today are Euclidean and affine geometry, trigonometry, linear algebra, calculus and numerical analysis. We don't really know what the mathematical tools of tomorrow might be*, **so we're counting on the next generation of employees to tell us.**

Note especially the last few words, which I have emphasized. The mathematics behind the movies is one of the most dynamic and growing applications of modern mathematics, and who knows where it will take us. Mathematics is all about imagining things well beyond our experience (for example who other than a mathematician could conceive of an 11 dimensional universe), and when this is combined with the creative potential of the movies we have a very potent force indeed!

Now, go out and watch some great movies, and maybe think a bit about maths whilst you are doing this!

7. Some References

[1]   J. Lasenby, *Maths goes to the movies*, Plus Maths, March 2007. https://plus.maths.org/content/maths-goes-movies

[2]   T. Whitted, *An improved illumination model for shaded display*, Communications of the ACM, **23**, (1980).

 [3] http://blog.ted.com/the-math-behind-the-movies-an-interview-with-tony-derose-of-pixar

[4] See the Bath Camera website at   http://www.bath.camera/

[5] E. *Catmull, and J. Clark*, *"Recursively generated B-spline surfaces on arbitrary topological meshes"*, *Computer-Aided Design.* **10** *(6): (1978), 350.* *doi*:10.1016/0010-4485(78)90110-0.

Ps. On 23rd May *Dr. Carola-Bibiane Schoenlieb* from Cambridge will be giving the joint LMS-Gresham lecture on *Mathematics can make you Fly*. If you found today's lecture interesting, then I strongly recommend that you attend the lecture by Dr. Schoenlieb. See http://www.gresham.ac.uk/lectures-and-events/mathematics-can-make-you-fly for more details.