Chapter 2

What is Music?

The problem with answering the question "What is music?" is understanding what would constitute a proper answer. Music arises from human behaviour, and the study of human behaviour is part of biology. So any question about music is a question about biology, and every question about biology requires an answer within the framework of Darwin's theory of evolution by natural selection.

2.1 Music is Something We Like

What is music? It's what comes out of the speakers when we play a CD on our stereo. It's what we hear on the radio. Music is singers singing and musicians playing. Music is a sound that we enjoy hearing.

Is this a proper answer to the question "What is music?"?

If I asked "What is a car?", you could answer by pointing at a large object moving up the street and saying "It's one of those." But this may not be a satisfactory answer. A full explanation of what a car is would mention petrol, internal combustion engines, brakes, suspension, transmission and other mechanical things that make a car go. And we don't just want to know what a car *is*; we also want to know what a car is *for*. An explanation of what a car is for would include the facts that there are people and other things (like shopping) inside cars and that the purpose of cars is to move people and things from one place to another.

By analogy, a good answer to the question "What is music?" will say something about the detailed mechanics of music: instruments, notes, scales, rhythm, tempo, chords, harmony, bass and melody. This matches up with the mechanical portion of our car explanation. It's harder to answer the

"What is it for?" part of the question. A simple answer is that music is enjoyable—it makes us "feel good". We could expand on this a bit and say that music creates emotions, or interacts with the emotions we already feel and, sometimes, it makes us want to dance.

2.2 The Biology of Feeling Good

The "feel good" explanation is worth something, but it isn't entirely satisfactory. Or, at least, it's not satisfactory if you're a professional theoretical biologist.

What does music have to do with biology? Music is something that people create and something that people respond to. People are living organisms, and biology is the study of living organisms.

We can compare music to eating. Eating is a well-known activity. People do it. Animals do it. We know what eating is: it is the ingestion of certain substances into our digestive systems. The ingested substances, or **food**, travel through the digestive system, where components of those substances are broken down and extracted by various means for use within the body. Leftover portions of the food get pushed out the other end.

We can explain eating at a psychological level: we eat when we feel hungry because it makes us feel good. Being "hungry" can be defined as a feeling of wanting to eat food. We can determine that we become hungry when we haven't eaten for a while,[1] and that we stay hungry (and slowly get hungrier) until we have eaten.

2.2.1 Having More Grandchildren

A professional biologist would explain the existence of hunger by saying that it is **adaptive** or, equivalently, that it is an **adaptation**.

A biologist calls something an adaptation if it contributes to having *more grandchildren*. Becoming hungry when we need to eat and eating when we are hungry contribute to having more grandchildren in the following ways:

- As children we need to eat food to grow up into adults.
- We need to eat to have the strength and energy to survive, to secure a mate, to do the mating itself, and then do all the work that comes afterwards, i.e. raise the children. In particular, we need to raise our children well enough that they can grow up and have children themselves.
- When a woman is pregnant, and also when she is breast feeding, she needs to "eat for two".

[1] There are other factors that influence hunger, such as whether it's the time of day at which we normally eat.

- We shouldn't eat when we already have enough food in us, because:
  - too much food at once will overload our digestive system,
  - once we have enough food in us, there are other more important things we should be doing instead of eating more food.

I refer to the need to contribute to having more *grandchildren*, rather than just children, to emphasise the importance of the continued cycle of birth, growth, development and reproduction. If something causes us to have more children, but has a negative effect on the ability of our children to raise their own children, to such an extent that it causes us to have fewer grandchildren, then that something is not an adaptation.

Strictly speaking, biologists think in terms of *long-term reproductive success*, i.e. having great-grandchildren and great-great-grandchildren, and so on forever. But, for our purposes, "grandchildren" is a close enough approximation. By the time most people get to having grandchildren, they no longer have the major responsibility to raise them, so whatever enabled their reproductive success to get that far will probably continue indefinitely anyway.

What made biologists think that everything had to be explained in terms of having more grandchildren? Most people would concede that if some species of organism does not have grandchildren, then pretty soon it is not going to exist at all. But does that mean that *every* purposeful behaviour of a living organism has to be explained in terms of long-term reproductive success?

2.2.2 Charles Darwin and His Theory

The most important discovery in the history of biology was Charles Darwin's theory of **evolution by natural selection**.

Even today, when his theory underpins all of modern biology, there are many people who refuse to believe that his theory is correct, or even that it could be correct. More than a hundred and forty years after Charles Darwin published his discovery, there is a whole industry of authors and pseudo-scientists "proving" that evolution does not occur, or that if it does occur then it is not occurring by natural selection.

This book is not aiming to change the minds of people who are skeptical about evolution. This is a science book, and it is based on a scientific point of view that the universe we live in appears to be *comprehensible* in the way that Albert Einstein remarked upon, and that furthermore it is reasonable to proceed on the basis that those bits of the universe that we do not yet comprehend will eventually turn out to be comprehensible.

The specific field of study concerned with understanding human behaviour according to Darwin's theory of evolution by natural selection is **evolutionary psychology**. The basic assumption of evolutionary psychology is that our behaviour is determined in some manner and to some degree by our **genes**.

Genes are the information about how our bodies develop and operate. They are contained in molecules called **DNA**, which can be understood as long strings of text written in a language with a 4-letter molecular "alphabet". If you read molecular biology papers in scientific journals, you will see descriptions of genes written as strings containing the letters A, G, T and C. These are the first letters of the chemical names for the four molecular "letters" in the molecular alphabet: **adenine**, **guanine**, **thymine** and **cytosine**.

AGTTTCTAGGTCGTGAAACTGTTCAGGCTTAAGTTGCGGTA

Figure 2.1. A stretch of (single-stranded) DNA shown as a sequence of A, G, T and C.

For humans the strings of DNA are divided up into 23 pairs of **chromosomes**. Each chromosome is an unbroken stretch of DNA, usually tied up in complex spiral patterns (to keep it safe and out of harm's way when it is not being used). Every cell in your body has these 23 pairs of chromosomes, except for a few types of cell that don't need to reproduce themselves. (Also there are the **gametes** which are the intermediate stage between parent and child, and which have only one of each pair of chromosomes.) The chromosomes in each pair are similar to each other,[2] and we get one of each type of chromosome from each parent (via their gametes). For each pair of chromosomes, each of our parents supplies one chromosome from their own pair of chromosomes, or a mixture of both chromosomes in that pair. Darwin didn't know about DNA, and he didn't understand the mechanics of genetic shuffling and mixing that occurs when we have sex.[3]

When we reproduce, the central thing that reproduces is our DNA. For us, as multi-cellular organisms, this happens when we reproduce to create new organisms (i.e. babies), and also when the cells that make up our own bodies reproduce in order to make our tissues grow. Most of the time the DNA reproduces accurately, but bits of it can get changed or **mutated**. And when these mutations occur, they will on average be preserved, and the next time the DNA reproduces, the parts of the gene that were changed are no

[2]Exception: females have two X chromosomes, but males have one X chromosome and one Y chromosome per cell. Furthermore, one of the female X chromosomes is always rendered inactive within the cell.

[3]Gregor Mendel was the one who first learned about the genetics of sex. The science of genetics as we know it today began when Mendel did his experiments on sweet peas. Darwin's theory of genetics involved a theory of "blending", which didn't work very well. Unfortunately Mendel's work did not become widely known until some time after Darwin's death.

more likely to change the next time than any other part of the gene that was not changed.[4]

What happens to us if our DNA mutates? A lot of the time the answer is nothing, because much of the information in our DNA has little effect on how well our bodies work. In fact the notion of "gene" specifically refers to a portion of DNA which does affect some particular part of how our body develops or operates. Mostly this happens when a gene encodes the makeup of a particular type of molecule called a **protein**. There are many types of proteins that do many different things in our bodies. If DNA in one of your genes changes, then the protein encoded by the gene will change, and this could affect how the protein does whatever it does in your body. Ultimately, the changed protein could change your long-term reproductive success.[5] It might make it better, or it might make it worse (which is actually far more likely). If it makes it better, then you are going to have more grandchildren and great-grandchildren and so on. If it makes it worse, then you are going to have fewer grandchildren and great-grandchildren and so on than everyone else.

An important part of Darwin's theory is the idea that for every species there is some limit as to how many individuals of that species can ever exist at one time. Among other considerations, all life that we know of exists on planet Earth, and the Earth is finite in size. In practice, most species hit some limit long before they get to the point where their members occupy every square and cubic inch of the planet. As the more successful genetic variations form a constantly increasing proportion of the total population, the less successful genetic variations must eventually disappear altogether. When this happens, the species itself has undergone a permanent change. The removal of less successful variations is the **natural selection** and the resulting permanent change is the **evolution**.

Darwin realised that if the process of evolution went on for long enough, species could change into new species that were as different from their ancestors as different species are from each other. And if species sometimes split into separate populations, and those populations happened to evolve in different directions, then one species would turn into two or more species. Taking this idea to its logical conclusion, Darwin supposed that all life on Earth could have evolved from a single ancestral species:

> Therefore I should infer from analogy that probably all the organic beings which have ever lived on this earth have descended from some one primordial form, into which life was first breathed.[6]

[4]This is probably not 100% true, as some locations in the chromosome may be more susceptible to processes that cause mutation. It is more precise to state that the probability of mutation at any given location on the chromosome can be a function of location, but does not depend on whether the location in question has or has not recently suffered a mutation.

[5]A mutation will affect your descendants if it occurs in a **germ cell**, which is a cell from which the gametes (sperms or eggs) are descended.

The modern technical term for this hypothetical "one primordial form" is the **Universal Common Ancestor** (**UCA**).

Evolution by natural selection explains the characteristics of living organisms. Each living organism is the result of a long sequence of individual minor changes, and each minor change became fixed in the population because it resulted in increased reproductive success. There are a few caveats to this reasoning:

- Some changes may have resulted from genetic changes that had only a very marginal effect on reproductive success. There is a certain probability that some changes will become permanent even though they have no effect or even a slightly negative effect on reproductive success. This can happen particularly if a species is occasionally reduced to a very small population, or if a new species evolves from a very small sub-population of its ancestor species.[7]

- In some cases an observable aspect of a species' behaviour will be attributable to the effects of one or more evolved changes that occurred in the past, but this aspect may not currently contribute to reproductive success, even though the corresponding evolutionary changes did contribute to reproductive success at the time they occurred.

2.3 Explaining Purposeful Behaviour

Whether or not a particular aspect of human behaviour requires to be explained within the evolutionary framework is easier to decide if we restrict ourselves to consideration of **purposeful** behaviour.

Purpose can be defined as a type of reverse causality. Causality is something that flows forward in time. What *was* explains what *is*, and what *is* explains what *will be*. With explanations involving purpose it's the other way around: what *is* explains what *was*, and what *will be* explains what *is*.

A normal causal explanation might be applied to a soccer player kicking a ball that goes into goal: the ball with mass m was travelling at velocity v_1, when it made contact with the player's foot (via his boot) at position p_1, which caused it to change velocity to v_2, after which, according to the laws of physics, it travelled in a path that caused it to go into the goal. In the causal explanation, where and how the player kicked the ball determined the ball's path, which in turn determined the ball's final destination inside the goal.

In the purposeful or **teleological** explanation, the ball going into the goal explains the way that the player kicked the ball. That is, the result is treated as the explanation of the events that caused that result. "The player kicked

[6] *The Origin of Species* Charles Darwin 1859

[7] Motoo Kimura developed the **neutral theory of molecular evolution** which emphasises the importance of random (non-selective) processes in evolution.

the ball *so that* it would go into the goal." If the ball had initially been in a different location and travelling in a different direction, the player would have kicked it differently, but *he still would have kicked it in a way that would have caused it to go into the goal.*

Of course players don't always get the ball into goal, even if they try ("try" is a word whose meaning implicitly assumes purpose), but we still accept the explanation that goes backwards in time: the player kicked the ball the way he did because he was trying to get it into goal (and it nearly went in).

This distinction between causal explanations and teleological explanations goes all the way back to Aristotle: he used the term **efficient cause** to describe normal forward causality, and **final cause** to describe reverse teleological causality.[8]

Modern science only admits efficient causes. A very simple way of justifying this is to say that science only allows one explanation for any particular aspect of reality that requires explanation. If we have two explanations of the same phenomenon, either one explanation is not correct, or one of the explanations is redundant and could have been restated in terms of the other.

In the case of the soccer player kicking the ball into goal, we accept the correctness of both explanations: the ball went into the goal because of the way it was kicked, and the ball was kicked the way it was so that it could go into the goal. But these dual explanations only apply to purposeful phenomena. For all other phenomena only the efficient cause type of explanation ever applies. So we may assume that efficient causes are the more basic type of explanation, and we must look for a way to restate the final cause explanation in terms of efficient causes.

At which point we can directly apply Darwin's theory of evolution by natural selection. It is the cycle of reproduction and selection which converts efficient causes into final causes. Various soccer players try to kick the ball into the goal. The ones that get it in are seen as better players. The girls fall in love with the good soccer players, and they have lots of children. The children inherit the genes from their dads who were good soccer players, and some of these genes determine the behaviour that caused their dads to kick the ball into the goal. Maybe the genes give their owners stronger legs, or better coordination, or create a propensity to practice more, or give them a tendency to party less the night before an important match. Whatever the case, in the next generation of soccer players there is a higher proportion of those genes which make the players better at kicking balls into the goal.

This explanation does seem a little trite. The genes that contribute to players being able to kick accurately may be genes that have quite general effects, like being able to focus on achieving a result, or being able to develop coordinated action. The ancestors of a good soccer player may never actually have played soccer (or at least not professionally). They might have been

[8] Aristotle listed two other types of cause: **material** and **formal**, but we would tend to include them as parts of efficient and final causes respectively.

cricket players instead. Or perhaps the skills evolved to help them run away from lions and throw spears at edible prey animals.[9]

But the general idea holds good: natural selection converts a final cause explanation into an efficient cause explanation, protecting and preserving the unity of all scientific explanations.

It also means we can stop feeling guilty about using teleological explanations, *as long as they fit into the theory of evolution by natural selection*.[10]

Final causes can be chained together just like efficient causes. For example, a chain of efficient causes is: I was able to have many grandchildren because the girls liked me because I got rich because I kicked the ball into the goal because I had practiced a lot because I always arrived at practice on time. The corresponding chain of final causes is: I always arrived at soccer practice on time so that I could consistently kick the ball into the goal so that I could get rich from being paid well, so that all the girls would love me and I could choose the best one to marry so that I could have many grandchildren.

We can use Darwin's theory of evolution by natural selection to convert a final cause explanation into an efficient cause explanation, as long as the very last final cause in the chain of final causes is *lots of grandchildren*. If we end up with a final cause of something else, then our teleological explanation is not consistent with our otherwise consistent explanation of reality based on efficient causes.

2.3.1 Incorrect or Apparently Incorrect Sub-Goals

Where does music fit in to this theory of purpose and causality? Certainly we can identify purposeful causality in behaviours relating to music. "I worked at the shop so that I could save up money so that I could buy a fuzz box so that I could plug it into my guitar so that I could play 'Smoke on the Water'." But the chain of final causes seems to stop when we get to the music itself.

Many of the unsolved problems of evolutionary science involve the existence of final causes that appear not to have any explanation in terms of more grandchildren: the chain comes to a stop in a bad place. Any number of human behaviours seem to go directly against what is required for maximising long-term reproductive success, behaviours such as driving too fast,

[9] This is a reference to the **environment of evolutionary adaptedness** (**EEA**): the time when we lived in the jungle in hunter/gatherer tribes. The presumption is that not much evolution has happened between that time and the present day, so any evolutionary explanations must relate to those earlier circumstances as opposed to modern living conditions with cars, roads, supermarkets etc. The EEA (as an explanation for modern human behaviour) is discussed in more detail in Chapter 3.

[10] This is not a complete explanation of the existence of purpose in human (or animal) behaviour: in addition to natural selection, there are selective processes operating *within* the brain, which act to select those behaviours and behavioural strategies that (on average) help us to satisfy our biological goals. The physiological mechanisms that underlie these processes are themselves the result of evolution by natural selection, so there exists a two-level hierarchy of purposeful causality: natural selection has given rise to a purposeful system of internal selection which acts to select purposeful behaviours.

sky-diving, being generous, fighting for your country, eating too much fat (or just eating too much), eating sticky sweets that make your teeth go rotten, and drinking too much alcohol.

How can we explain the existence of these apparently **non-adaptive** purposeful behaviours? Plausible types of explanation include the following:

- The reproductive benefit is there, but just not so obvious to the untrained observer.
- The purposeful behaviour results from some more general purpose which benefits reproductive success on average.
- The behaviour used to benefit reproductive success, but times have changed and now it doesn't.

(The third explanation can be a special case of the second one: the behaviour used to benefit reproductive success, now it doesn't; in the future it may become beneficial again.) Another possible explanation is that the alleged behaviour isn't quite what it seems: for example, maybe generosity isn't quite as common as it appears to be, because people are always doing things to make themselves appear more generous than they really are.

Trying to explain non-adaptive purposes and purposeful behaviours is an ongoing activity in the world of evolutionary psychology, and some of the explanations that have been thought of are more convincing than others.

Here is a sample list of evolutionary explanations for some of the apparently non-adaptive human behaviours given above:

- Wanting to drive too fast used not to be non-adaptive, because there weren't any cars. The instincts that make drivers want to drive too fast had general benefits, encouraging our ancestors to learn how to move quickly and efficiently without crashing into anything.
- There weren't any opportunities to sky-dive in the distant past, on account of the non-existence of parachutes—so a desire to sky-dive would not have been non-adaptive.
- Dying for your tribe or country seems extremely non-adaptive, since dead people can't have children. But if society rewards warriors who risk their lives for the sake of the tribe, then it can be argued that the benefits going to those who risk their lives and survive more than make up for the losses suffered by those who risk their lives and get killed.
- Eating a lot of fat can be beneficial if there is a substantial risk of famine. The extra nutrients stored in the body of a fat person will help them to survive the hard times.

- In the past, most available sweet foods would have been either ripe fruit or honey. These are not quite as bad for your teeth as the boiled sweets and toffees that are available in large quantities in the modern supermarket. A desire to eat anything sweet is of particular advantage to children, as they need the extra energy to play, and play is important because it helps children develop their thinking and general life skills.

- Why people like to drink alcohol requires a different sort of explanation. Alcohol and other recreational drugs, legal or illegal, act directly on those parts of the brain that tell us if we have or have not achieved our goals. The most that evolutionary theory can tell us about drugs is that if a drug was widely available in the distant past, then humans should have evolved some resistance to that drug.

2.4 Proof of our Ignorance About Music

This issue of explaining non-adaptive purposes will come up when we investigate music. With music there is, however, a further complication: *we don't even know what music is.* Music is therefore a double mystery: we don't know the "what" and we don't know the "why". Maybe if we could solve the "what" that would help us answer the "why", or maybe if we could guess what the "why" is we could find out the "what".

There are a number of different ways I have found of demonstrating our ignorance of what music is, and each provides a useful insight into the nature of the problem:

- **Subjective and Objective.** The difference between knowing what something is subjectively and knowing what it is objectively.

- **The Martian Scientist.** Could we explain to a Martian scientist what music is?

- **The Incompleteness of Music Theory.** Here "music theory" refers to the kind of music theory that you learn when you learn to play music, and which will be presented in a basic form in Chapter 4. This music theory tells us something about the structure of music, but beyond a certain point it gives up.

- **Lack of Formula.** Despite common claims that some types of music are "written to a formula", there is no such formula, or if there is one, no one is telling us what it is.

- **The Economics of Music.** Those who compose good music get paid well, because making up good music is a hard problem. The very difficulty of the problem results from our ignorance about what music is.

2.4.1 Subjective and Objective

We know what we know about things in the world around us because information comes into our senses, and we process the information in our nervous systems and brains to create knowledge about those things. Sometimes we can convert this knowledge into symbolic natural language, i.e. by speaking or writing. Sometimes other people can relate our symbolic descriptions of things to their own experiences of the same things (or similar things).

If I see a sparrow, I can describe my observations of that sparrow to you. You can relate that description to memories of sparrows you have seen. If by some chance you have never seen a sparrow, I would first have to explain what a sparrow was, and you would have to relate that to your experience of seeing other types of bird. If you have never even seen a bird, then it becomes more difficult, and I would have to think more carefully about how to describe what a bird is to someone who has never seen one.

If I feel a pain in my leg, I can describe it to you, and you can relate that description to your own experiences of having pain in your legs. But we cannot feel the same pain. I cannot feel the pain in your leg, and you cannot feel the pain in my leg. It is almost impossible for one person to know exactly what pain another person is feeling. In fact we can argue that questions like "Is my pain the same as your pain?" are ultimately meaningless, as there is no meaningful way to make such comparisons.

This problem seems related to questions like "Is my feeling of seeing red the same as your feeling of seeing red?". However, the colour of objects is something that can be specified in terms of physical theories about reflection and absorption of light. We know that human colour perception depends on reception of light by three specific types of colour receptor in the eye. In as much as two people have exactly the same colour receptors (which is mostly the case), there is some sense in which it can be said that they see the same red if they look at the same object under the same lighting conditions. Of course the internal processing of colour perceptions will still be different, because it is very unlikely that two people's brains are wired in exactly the same way.

If we doubt that I am seeing the same red as you are seeing, we can use a spectrograph to measure, for each frequency, the intensity of light falling onto the red surface and the intensity of light reflected off the surface. Then, for each frequency, the ratio between the intensity of light reflected off the surface and the intensity of light falling onto the surface gives us the absolute reflectance of the surface at that frequency. The values of all the ratios for all the frequencies of light define the colour of the surface. We can display these ratios as a function of frequency in a graph, or reduce them to a table of numbers. There is no real possibility of us disagreeing about what the numbers are. We can wonder if my experience of the number 3.567 is different from your experience of the number 3.567, but most of us are prepared to regard the meaning of "3.567" as completely independent of the person who is reading the number.

This independence of observer is what we call **objective**. The opposite of **objective** is **subjective**. The meaning of the number 3.567 is objective. The pain in my leg is subjective.

Somewhere in between objective and subjective is **inter-subjective**. An inter-subjective perception is subjective, but we have some degree of confidence that my experience of it will be the same or at least similar to your experience of it. Most subjective phenomena are inter-subjective to *some* extent, in the sense that there is probably some person somewhere feeling something similar to what you are feeling now, and that person would understand what you were talking about if you described your feelings to them. Even pain is inter-subjective in this sense. Also it could be claimed that the difference between the objectivity of "seeing red" and the subjectivity of feeling pain is not so much that it is impossible to objectively describe what pain means, but just that our current understanding of the human mind and visual perception allows us to be more specific about what "seeing red" means.

2.4.2 The Martian Scientist

In Oliver Sacks' book *An Anthropologist on Mars: Seven Paradoxical Tales* (Vintage, 1996), the "Martian" is Temple Grandin, a well-known autistic, who has difficulty understanding the emotions and intentions of other people, and who has described herself (as quoted on p. 248 in Sacks' book) as feeling like "an anthropologist from Mars".

In general, the concept of the Martian Scientist is a good metaphor for the idea that there are things about ourselves that we are very familiar with, but which might be difficult to explain to an alien from outer space.

There is a presumption in this metaphor that there are at least some things that we could explain to an alien scientist. For example, it is presumed that it would not be too hard to introduce an alien scientist to our mathematical notations, so that we could talk about "3.567", and the alien scientist would know exactly what we were talking about. Similarly we would be confident that we could explain what a spectrograph was, and even explain the characteristics of colour receptors in the human eye, so that our alien friend could understand what we meant when we talked to him about the colour "red".

The concept of the Martian Scientist arises in discussions about consciousness. We all know subjectively what consciousness is, but as yet no one is able to explain what it is in an objective scientific sense. Could we explain consciousness to a Martian scientist? The problem is that a Martian scientist is quite likely to be conscious in exactly the same way that we are. Maybe it is not possible to be intelligent in a way that allows understanding and discussion of scientific concepts, unless one is conscious. So when we talk about consciousness with our friend from Mars, he could indicate that he knows what we are talking about. And yet we cannot say that this proves that either of us (human or Martian) has an objective understanding of what consciousness is, because we may be doing nothing more than sharing our common *subjective* experiences of consciousness with each other.

Music is a bit different in this regard. Our ability to respond to music does not appear to play any essential role in our ability to comprehend the universe. Our perception of music depends in obvious ways on our systems for perceiving and processing sound. But being deaf does not in the least imply a lack of intelligence: quite plausibly our Martian scientist could be deaf. (Maybe the air on Mars is too thin for hearing to be of much use.) A deaf Martian scientist would not have any subjective understanding of what music is. This gives us a straightforward way to ask if we can find an objective description of music: could we explain what music was to a deaf non-musical Martian scientist?

Some people would explain music in terms of what they know about music, saying music is a sequence of sounds according to certain rules, which happens to have certain emotional effects on people. Given this explanation, and given an item of supposed music, the Martian could check if the supposed music satisfied the specified rules, and then check that it also had an effect on human listeners. But what we really want to know is whether the Martian scientist could learn to identify music, and in particular good music, when given only the music itself. In other words, could the Martian scientist *predict* the effect that an item of supposed music would have on human listeners? To use a term that I am going to use a lot throughout this book, would the Martian scientist be able to calculate the **musicality** of music?

2.4.3 The Incompleteness of Music Theory

It seems reasonable to assume that we could discuss mathematics with intelligent aliens. So if we could produce a description of music that was mathematical, then we could easily communicate that description to an alien scientist.

Much of music theory is mathematical. We will see details of this when basic music theory is introduced in Chapter 4. Notes have frequencies. Intervals between notes can be described as vectors and as certain fractional ratios between their frequencies. Notes and percussive sounds occur at certain times according to regular tempos. The relationships between fundamental and harmonic frequencies can be explained in terms of **Fourier analysis**, which is an important and non-trivial area of mathematics.

With all this existing mathematical music theory, we might wonder what the problem is. Can't we just tell our alien audience the mathematics of music theory, and then they will have an objective understanding of what music is? There are two main reasons why this might not be the case:

- Firstly, a mathematical description of music does not necessarily tell the aliens anything about what is going on inside the human brain when we listen to music.

- Secondly, our mathematical theory of music is not *complete*. Although music theory says quite a lot about the mathematical structure of music, it does not say enough to distinguish between really good music and mediocre music. Music theory fails to *predict* the musicality of supposed music.

These two problems are complementary: if we knew exactly what was going on inside the human brain when we listened to music, then this information could be translated into a procedure for calculating the musicality of music. The procedure for calculating musicality would be a simulation of the operation of those parts of the brain that play a role in perceiving music.

On the other hand, it may be possible to develop a complete mathematical description of music *without* developing any understanding about what happens inside the brain when we listen to music. But as you will see when you progress through this book, intelligent guesswork about what is happening inside the brain is the easiest way to make sense of the mathematical structure of music.

The incompleteness of music theory was my major motivation for performing the research which culminated in the development of the theories explained in this book.

Books that discuss music theory tend to skate around the issue of incompleteness. One good question to ask yourself, when reading a book (or paper) that discusses explanations of music, is what, if anything, the book says about why some music is better than other music. If an author ignores or denies the existence of musicality as something that a musical item can have more or less of, this makes it is easier for them to avoid confronting the question of what it is that determines musicality, and they can comfort themselves with discussions of "music", completely ignoring any comparison that can or should be made between "good" music and other music which is still recognisable as music, but not quite so good.

Even when a book does arrive at this issue, the author will admit (sometimes very implicitly), that they do not know what causes the difference between the good and the not so good, or they may just state categorically that this difference cannot be explained by "rules" (generally ignoring the possibility that they are talking about known rules, and that there might be other unknown rules that do explain the difference).

To approach a problem scientifically, we must not be afraid to confront our own ignorance. The more clearly we can state what we think we know, and what it is that we don't know, the more chance we have of finding some way to move forward. A precise statement of our ignorance about something can be an important first step in the development of a new theory, or in the design of an experiment likely to advance our understanding of the problem.

2.4.4 Musical Formulae

When people talk about music "written to a formula", they use this phrase in a derogatory sense, implying that some hack churns out musical items which are all very similar and just good enough to be marketable. The sophisticated listener is bored by this formulaic music, and hungers for musical creativity that comes from an inspired genius whose output could never be captured by anything as mundane as a formula.

No one ever says what the formula is. Or if they do, the formula suffers from the same incompleteness as music theory in general: the formula describes some aspect of the music, but it is not complete enough to generate the same creative output as the output of the person whose output the formula supposedly describes.

Now it is possible that someone somewhere *is* using a formula to generate music, and they are keeping it a secret. If you had a formula to generate music, you might want to keep it a secret too. You could use your formula to compose music which you could sell, but if everyone knew the formula then it would be too easy for anyone to make up good music, and the bottom would drop out of the market.

The type of formula I have just been talking about is a formula to generate music. In the world of mathematical computer science, they would call it an **algorithm** (rather than a "formula"). An algorithm is something that can be written down as a program written in some programming language, and executed on a computer. So we are talking about a computer program that can compose music, and not just any old music, but music that is as good as, or even better than, the best music as currently composed by professional composers and songwriters.

There is another type of algorithm which is relevant to the analysis of music, and that is an algorithm that calculates the quality or musicality of supposed music that is provided as input to the algorithm.

There is some degree of overlap between what these two types of algorithm achieve, but they are not the same thing. The **generative** algorithm produces music which has high musicality. The **predictive** algorithm accepts as input any music, or non-music, and tells what the musicality of that input is, and predicts its effect on the human listener.

If we had a predictive algorithm, then a naïve way to convert this to a generative algorithm would be to attempt an exhaustive search of all possible items of music, apply the predictive algorithm to each candidate, and output each item for which the predicted musicality was found to be high enough. This algorithm would work, but it might not be very efficient, because the set of possible musical items grows large very quickly as we consider items of greater and greater length, and only a very small proportion of all possible tunes might be at all musical.

Similarly, if we had a generative algorithm, there is no guarantee that this could be converted to an efficient predictive algorithm. Firstly, a particular generative algorithm might not generate all possible strong pieces of music. Secondly, even if it did, the only way to use it as a predictive algorithm would be to run the algorithm and generate all possible items until one of them happened to be the same as the input data. If the algorithm terminated, you would know that your input data was musical. If it did not terminate, you would then know that the input data was not musical (but of course it takes an infinitely long time to determine that an algorithm does not terminate, unless you are able to provide a mathematical proof of non-termination).

In practice, we would assume that effective generative algorithms and effective predictive algorithms would both be based on a theoretical understanding of the human response to music, and that given information that could be used to formulate one type of algorithm, we could also formulate the other type of algorithm without undue difficulty.

There are algorithms for which conversion into a related type of algorithm is arbitrarily difficult and suffers from worst-case **complexity**.[11] The standard example is the **cryptographic hash algorithm**. This is an algorithm that produces a fixed length output—the **hash**—typically 128 or 160 bits long, which is derived from arbitrary sized input data, such as a computer data file. The algorithm is irreversible in the sense that it is very difficult to find an input value for a given hash value, unless you happen to already know an input value that generates that hash value. And if you have one input value that generates a hash value, it is equally difficult to discover a second distinct input value that generates the same hash value. In fact a cryptographic hash algorithm is considered broken if anyone ever discovers *any* pair of distinct input values that produce the same hash value.

However, cryptographic hash algorithms have been specially designed to be irreversible. In as much as music does not appear to be part of a biological digital security system, there is no particular reason to suppose that an algorithm for the evaluation of musicality could not be converted into an algorithm for generating music with a high level of musicality. In fact, based on the assumption that the human brain operates according to mathematically specified physical laws, we already have a method which in principle can generate high quality music: simulate the workings of the brains of those people who (at least occasionally) compose good quality music.

2.4.5 The Economics of Musical Composition

I have hinted that finding a musical "formula" would radically change the market for music. But what is the current state of the music composition economy? Who composes the really good music? How do they do it? How hard is it for them?

[11] **Complexity** is a computer science term meaning how much time and memory an algorithm uses when executed in a computer, often specified as a function of the size of the input data.

If existing well-known music theory was complete, then composing good quality music would be relatively easy because the theory would tell us how to do it. I would suggest that the existing economics of music implies that the composition of high quality popular music is far from easy:

- Some composers and songwriters write a lot of music, but others only ever write one or two very good items. This gives rise to the term "one hit wonder" (although this is used more typically of performers, who may or may not also be the composers of the music they perform).
- Some writers write a lot of good songs over a certain period, and then seem to dry up.
- The record industry churns out best-selling albums, many of which contain only one good song, with the rest being "album filler".
- You can get paid a decent amount for making up some good music. Generally nobody ever gets paid a whole lot for doing something that anybody could have done.

We can see that whatever knowledge it is that composers and songwriters have about music that allows them to write music, this knowledge does not exist in a form that enables them to generate arbitrary amounts of new high quality music. It is locked inside their brains as some type of intuitive understanding of music which, when combined with persistence and good luck, enables them to occasionally produce something great.

Trial and error may provide part of the explanation of how music is created: an experienced musician is familiar with many different musical patterns and structures, and combining this knowledge with their own subjective ability to evaluate music, they can generate possible new music, listen to it to see if it is any good, and remember the good stuff. Even when a new piece of music suddenly "comes" to a composer, this may have been the final result of an extended trial and error search that took place within the hidden mechanisms of their brain (a Freudian would say that their **subconscious brain** did all the work).

Although the inner workings of the brains of composers of great music is an interesting topic in its own right, it is not the major purpose of this book to explore the means by which people create new music. My primary focus is on what causes people to respond to the music that they listen to. I cannot rule out the possibility that learning more about musical composition might help us to better understand the listener's response to music, but in practice we will find more direct routes to solving the problem of why and how we respond to music.

The question of creation versus performance versus response cannot be completely ignored when considering the biological purpose of music. Some authors have suggested (and in some cases they just implicitly assume) that the primary biological purpose of music has to do with creation and performance rather than response to music. I do briefly consider these possibilities, but I will show that there are reasons why hypotheses about the biological purpose of creating and performing music are both unnecessary and unconvincing.

Consideration of the economics of music leads to what I call the **luxury yacht test** (**LYT**) for a theory of music. It consists of the following steps:

- Discover a complete theory of music.
- The theory should specify an algorithm for calculating the musicality of music, possibly parameterised for variations in musical taste.
- Reverse this algorithm to create an algorithm for generating new good quality music.
- Sell the new music.
- Use the proceeds to purchase a luxury yacht.

So if you meet someone who claims to know the answer to the question "What is music?", ask them if they own a luxury yacht.

2.5 Universality

In the above discussion of musicality and predictive algorithms, I implicitly assumed that there existed some measure of musicality that was equal for all listeners. In practice there is a lot of commonality in musical taste, but the very fact that the phrase "musical taste" exists in the language tells us that musical preferences do vary from person to person.

It would be over-reacting to conclude that therefore an algorithmic and scientific theory of music cannot be discovered. People vary in how they react to strains of the flu, but that does not mean we cannot come to a scientific understanding of the influenza virus and its effect on people.

What it does mean is that we will have to **parameterise** our algorithms to take account of variations in musical taste. In other words, the algorithms will accept additional input data representing information about the musical taste of the listener. But, having said that, close enough is often good enough, and if a particular algorithm generates high quality music according to *your* tastes, then at least some of that music will also be considered high quality music according to *my* musical tastes. Suppose that I like only 1% of the music that you like, and we have an algorithm that generates new items of music that you like. To generate one item of music that I like, all I have to do is run the algorithm a hundred times. The 1% success rate (of this hypothetical algorithm) is far superior to the (very close to 0%) success rate of any currently known algorithm for generating music that I like.

The major factors likely to cause variations in musical taste are the following:

- Variations in exposure to music over one's lifetime.
- Variations in exposure to other sensory inputs that affect response to music (which could include language, non-verbal utterances, animal sounds and other natural sounds).
- Variations in personality type.
- Genetic variations in whatever it is in our brain that determines our response to music.
- Random/chaotic variations, i.e. points in the development of our bodies and brains where something could just as easily have developed one way as the other.

The most significant variations in musical exposure are where people belong to totally different cultures and each culture has its own distinct type of music. Not only are the tunes different, but the scales that the tunes live on are different (although usually there *are* scales, and those scales usually repeat every octave, but not always). The whole thing becomes relative: we like our music and not their music, and they like their music but not our music.

Cultural relativity spawns political correctness, and political correctness can discourage researchers from following lines of enquiry that they might otherwise follow. It might, for example, be deemed inappropriate to formulate a hypothesis that suggests (or assumes) that the music from one culture is "better" than the music from another culture.

The most politically incorrect candidate for a "best" type of music is probably Western music, as played on Western scales (i.e. the notes on a piano). Western music is coming to dominate over all other types of music, occasionally including ideas and forms from other cultures, but mostly just replacing them.[12] Is this because Western music is better than other music? Is it because Western countries are imperialistic and dominating? Is it all caused by capitalistic marketing machines?

One circumstance which reduces the accessibility of non-Western music to Western musicians is that most readily available musical instruments are tuned to Western scales, i.e. the well-tempered chromatic scale or some subset thereof. There may come a day when electronic keyboards routinely come

[12] The most substantial input into Western music from other cultures happened when American-African slaves and their freed descendants combined aspects of African music and Western music, giving rise to ragtime, jazz and blues. The African influence can probably be held responsible for most of what makes modern popular music different from older Western classical music. Despite this influence, Western popular music remains strongly tied to the diatonic scale and to underlying regular hierarchical tempo.

with options to select alternative tunings, and when that day comes the dominance of Western scales may be reduced somewhat, and alternative musics may be able to reclaim some of their lost ground.

Even ignoring the political questions, there are theoretical issues, like:

- Does a theory have to take account of all known types of music?
- Can I develop a theory that just applies to one musical culture?
- If my theory describes some aspect of music, does that aspect have to appear in all cultures, or in most cultures, or just in the biggest cultures?

There is the idea of **universality** current among those who study music (scientifically or otherwise), which is that theories about music have to apply equally to all known musical cultures. On one level it is a perfectly valid requirement, but if it is applied over-zealously then important sources of information about music can end up being ignored.

The concept of universality is being applied too strongly if it is used to reject any theory or hypothesis that cannot immediately be applied to all forms and genres of music from all musical cultures that have ever existed.

There is a useful analogy with the study of biology and the study of specific biological organisms. In studying biology we expect to find general principles that underlie the workings of all living species. At the same time, the biologist cannot simultaneously study all organisms at one time. He or she must necessarily concentrate their studies on one particular species, and indeed often just on one or a few members of that species. Eventually some of what is learned about particular species will turn out to generalise to theories that apply to many different species, or even to all species, but we cannot expect or require this generalisation to happen immediately every time we develop a new theory about something.

The criterion for accepting a scientific theory as being useful is not whether it unifies all knowledge in a domain, but rather that it unifies at least some set of distinct facts.

For example, it would be entirely possible and legitimate to develop a scientific theory about a single melody. Our observation of the melody could be regarded as a series of observations of individual musical notes. The occurrence of each note in the melody—its time, length and pitch—counts as one fact about the melody. The theory about the melody would be an explanation that described the notes in some way that was simpler and shorter than a full listing of the notes. Having found a theory about this one melody, we would hope that it could be generalised in some way to form a theory about other melodies, or even all melodies. But even if this is not immediately possible, the theory still has value if it can say something significant about just the one melody.

It follows that we should not feel guilty if we happen to develop theories of music that only apply to certain musical cultures, or to certain genres, or to the musical taste of one person (e.g. the person who developed the theory). The eventual aim of a theory of music is to be universal, and the theory I develop in this book certainly claims to be universal. But a theory about some aspect of music is not wrong or irrelevant just because it is not quite as universal as it could be.

2.5.1 Author's Declaration

Having justified the development of non-universal theories of music, it is perhaps now safe for me to declare my own musical tastes and preferences:

- Most of the music I listen to is the sort of thing you will hear on "Top of the Pops".
- Almost all the music I listen to is diatonic music with regular hierarchical tempo.
- I do not listen to, and do not enjoy, atonal music.
- I do not listen to classical music that much.
- I do not think that John Cage's infamous "4 minutes 33 seconds" is music.

The last example gets a mention in the introduction to *The Origins of Music* (see the next chapter for more discussion of the contents of this book and others), as part of the difficulty inherent in defining what music is, and it's not entirely clear if they are joking or not.

2.6 Scientific Theories

2.6.1 Testability and Falsifiability

The relationship between facts and theories is a large part of what science is about.

Consider a simple example: I throw a ball into the air in a certain direction. I take photos of its path with a camera that can take pictures rapidly at regular intervals. From the photos I record a series of positions at different times. The path and the recorded positions will look something like Figure 2.2.

I have a **theory** about the path of my ball. Writing t for time, x for horizontal position and y for height above some baseline, my theory can be written as a pair of equations that specify position as a function of time:

$$x = v_x t$$
$$y = v_y t - \frac{1}{2} g t^2$$

v_x represents initial horizontal velocity, v_y represents initial vertical velocity, and g represents acceleration due to gravity.

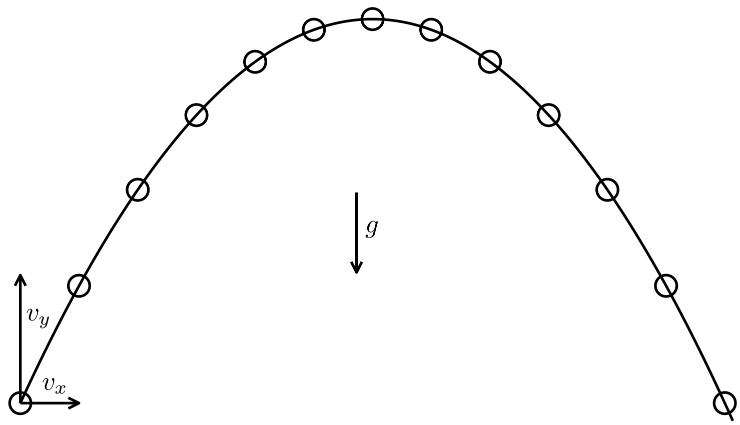

Figure 2.2. A ball thrown into the air with initial horizontal velocity v_x, vertical velocity v_y and downward acceleration g. The camera takes a photo of the ball's position at $t = 0$, $t = 1$, $t = 2$, etc.

The most important thing about the theory in relation to the facts is that the theory is specified using a fixed amount of information (i.e. those two equations), but it can explain a larger number of facts. In this case the number of facts that can be explained by the theory is virtually unlimited, because we can measure a large number of positions each time we throw the ball, and we can throw the ball any number of times, perhaps with different values of v_x and v_y each time.

Sometimes theories explain facts that can only be gleaned by observation, and the supply of facts may be more limited—a good example would be any theory that explains the positions of the planets, as we cannot easily throw new planets into space and observe them (although modern technology does allow one to fire small spaceships out into space). However, as long as the amount of information contained in the observations explained by our theory is larger than the amount of information contained in the specification of the theory, we can be confident that the theory is saying something useful about the world. We can be especially confident if the set of observations explained by the theory keeps on growing, without the theory itself requiring any further improvement or adjustment.

There are a number of things that we can say about the ball example, which reflect on issues that arise generally when doing science:

- The theory can be related to more general theories. For example, the acceleration comes from gravity, and we can form a more general theory about gravity. The theory about gravity will tell us that g depends on height above the Earth, and that it has quite a different value if you happen to do the experiment standing on the moon.

- The theory is only approximately correct, in part because it makes various assumptions that are not quite true. Air resistance is ignored. It is assumed that the gravitational field is constant. (If we threw the ball hard enough to go into orbit, then the equation would turn out to be quite inaccurate.) Any effects due to the ball itself having a finite extent are ignored.

- The measurements of the ball's position will not be made with 100% accuracy. We will have to allow for this when verifying the theory against the data.

- We may not have any independent way of knowing the values of v_x and v_y, and they will have to be estimated from the data itself in each case. One consequence of this is that at least 3 data points have to be taken in order check the theory at all, since for any 2 data points there will be values of v_x and v_y that exactly match the data. If we don't know beforehand what g is, then its value also has to be calculated from the data, and at least 4 data points are required to be able to check anything. (We would, however, expect g to have the same value for different throws of the ball.)

- If we don't have a camera that can take pictures at regular intervals, it will be very difficult to do this experiment at all.

These issues all have to do with the concept of **testability**, or **falsifiability**. If we state a scientific theory, we expect it to make predictions about something; a theory that doesn't make any predictions that can be checked isn't really a theory. We then want to be able to compare the predictions with measurements and observations. If the predictions come out wrong, then the theory is **falsified**, i.e. proven wrong. We can never prove a theory true, but it becomes more convincing if it makes more and more predictions and never gets proven wrong.

This view is somewhat idealised—that a scientific theory is falsifiable by experimental observation and is rejected the moment it is contradicted by just one observation. Sometimes we have to be a bit forgiving of our theories, for various reasons:

- Sometimes a theory cannot be tested by any practical means, at least not when it is formulated, but it is testable *in principle*. Our theory about the thrown ball is difficult to test if we don't have the equipment for measuring its position at known times. Scientists sometimes deal with this difficulty by specifying **thought experiments**, i.e. experiments carried out only in their imaginations. If we don't have a camera that can shoot pictures at regular intervals, we can still imagine the existence of such a camera, and use this possibility to justify the testability of the theory about the position of a ball thrown into the air. Albert Einstein was famous for inventing thought experiments that tested certain aspects of quantum theory.[13]

- Sometimes the "facts" that disprove a theory turn out to be wrong.

- A theory may explain a whole lot of facts, and then fail on just one fact. Even if that one fact is quite reliable, and it disproves the theory, the theory is still telling us something about all the other facts that it does correctly predict. We know that the theory needs to be replaced with a better theory, but we don't throw away the old theory until we have found the new theory. In fact it becomes a requirement that any new theory should explain why the old theory works as well as it does. This sort of thing happened when special relativity "replaced" Newtonian physics,[14] and also when quantum mechanics replaced Newtonian physics (again).

2.6.2 Simplicity and Complexity

Science often progresses in a certain area because someone asks the right questions and does the right experiments. Real life phenomena can be very complicated, and theoretical descriptions of these phenomena must take into account many different factors. It is best if we can separate out the individual factors as much as possible.

In our thrown ball example, we remarked that air resistance was ignored. If we had tried throwing a piece of paper, or a feather, then it would have been impossible to ignore air resistance. We would not have been able to verify the theory contained in our simple equations. Now even an ordinary ball—like a tennis ball—might be affected by air resistance by a noticeable amount. If we had some idea that air resistance was a complicating factor, then we might guess that we could ignore it if the object being thrown was large and dense. Instead of throwing a tennis ball, we might choose to throw a

[13]Einstein was sure that the theory couldn't be correct, and the thought experiments (published in 1935 by Einstein and two other physicists, Boris Podolsky and Nathan Rosen) were intended to prove this—he believed that the results predicted by the theory were too strange to be possible. But when slightly altered versions of the thought experiments were carried out decades later, the results of the experiments confirmed the theory.

[14]But they still teach Newtonian physics in school.

solid iron ball. We would be rewarded by a very close fit to our mathematical equation, because the size and density of the solid iron ball would allow us to ignore air resistance.

By using a heavier ball, we have created a *simpler* phenomenon to study. If we didn't even know what the equation was going to be, we could have made observations on throwing the heavy ball, and looked for simple patterns in the data. For example, using the **method of differences**,[15] it would have been easy to discover the formula for height as a function of time.

In the case of music, we don't necessarily have a clear idea as to what all the complicating factors are, and whether they can be cleanly separated from each other. But there is one easy way we can avoid complexity, and that is to *study the simplest tunes possible.*

This means, given a choice between a symphony and a pop song, where the symphony has hundreds of bars, multiple motifs, several key changes and a whole orchestra of instruments, and the pop song has 12 bars, 3 chords, one melody, no key changes and can be performed by one guy singing while strumming a guitar, study the pop song first.

There is a tendency in musical academia to listen to "difficult" music, such as long complex symphonies, and strange contemporary music that ordinary folk don't listen to. If popular music is studied, this is done so apologetically.

But when we realise that music is a difficult scientific problem, and it has been studied for over 2000 years, and everyone is still clueless as to what music actually is, then no apology should be necessary. We should study the absolute simplest stuff possible. Even when studying pop music, we should simplify it as much as we can without rendering it unmusical. Is it just a melody line? Maybe, maybe not. Can we reduce the accompaniment to a simple chord sequence (like in a "Learn to Play Guitar" book)? Can we reduce the bass to just the root note of the chord? Can we leave out the rhythm accompaniment, or reduce it to a straightforward pattern of regular beats?

Another good example of scientific simplification is found in biology. Biologists have studied many different organisms, both complex and simple. But some of the most important discoveries in genetics and molecular biology have been made using the simplest possible organisms. The relationship between DNA and protein was discovered using viruses, which are usually just a small section of DNA wrapped in some protein. Other problems required self-contained organisms (viruses are always parasites), in which case bacteria were used as the object of study. And to study the mechanisms of development in multi-cellular organisms, a very simple multi-cellular organism was chosen: *Caenorhabditis elegans*, a 1mm soil nematode which not only has a

[15] Given a sequence of values, keep taking the differences of each element in the sequence and the next to get a new sequence, and repeat this procedure. If you arrive at a sequence of all zeros, you can reconstruct a polynomial which describes the original set of values, such that the degree of the polynomial is one less than the number of times the procedure was applied.

relatively small number of cells in its body, it contains an exact number of **somatic** cells as a fully developed adult—959. (**Somatic** cells are non-germ cells, i.e. those cells that are not destined to become ancestors of the cells in the organism's descendants.)

In all these cases, the biologists did not go around apologising for studying organisms that were too easy or too simple.

A more extreme example, where scientists can only solve the easiest version of the problem, is the dynamics of multi-body gravitational systems assuming Newtonian gravity: the interaction of *two* bodies in each other's gravitational fields is soluble with an **analytical** solution,[16] but solving for *three* bodies is too hard, except for certain special cases. Something similar is found when studying the quantum mechanics of the atom: the hydrogen atom with one nucleus and one electron is doable, the helium atom with one nucleus and two electrons is too hard, and scientists must resort to various approximations, or to brute force integration of the relevant equations on big computers.

If the calculations of the consequences of a theory cannot be calculated accurately (because we are not studying the simplest possible system described by the theory), then the predictions of the theory cannot easily be checked against the results of our observations. And if there is no simple equation that describes the behaviour of the system, there is much less chance that we will discover the theory describing the system just by analysing observations of its behaviour. This is demonstrated by the last example: significant discoveries about the quantum nature of the atom were made from observations of spectral lines of the hydrogen atom, which happen to exhibit certain simple regular patterns.[17] Similarly, Newton's discovery of universal gravity was helped by Kepler's discovery of the laws of planetary motion, which take a simple form because for each planet one can (to a first approximation) ignore the gravitational effect of all other bodies besides the Sun.

[16] An **analytical** solution is one that can be written down as a formula that you can work out on a basic scientific calculator, i.e. only containing algebraic operations, trigonometric and exponential functions, and their inverses.

[17] *Hydrogen: The Essential Element* by John S. Rigden (Harvard University Press, 2002) gives a very good account of how the simplicity of the hydrogen atom has contributed to the development of scientific knowledge.