Christian Geometry: the Geometry of Light

Tristram de Piro

"All the darkness in the world cannot extinguish the
light of a single candle", (St. Francis of Assisi)

"No one lights a lamp in order to hide it behind the door;
the purpose of light is to create more light, to open people's
eyes, to reveal the marvels around", (Paolo Coelho)

"I am the light of the world. Whoever follows me will never
walk in darkness, but will have the light of life", (John 8, v12)

List of Chapters

1. Introduction

Seven years ago, I would have described myself as an agnostic. I received a classical education in Mathematics and Philosophy at Oxford and proceeded to do a number of further other degrees in Logic at Manchester and Cambridge. By the time I had finished my qualifying examination for a Ph.D. in Mathematics at M.I.T., I had traversed the full spectrum of pure mathematical thinking and was fully set on the path of scientific academia. The purpose of this book is partly to explain a transformation that I underwent in subsequent years, and how that change came about through accepting and integrating a part of myself which is not exclusively rational. At first reading, that statement sounds quite prosaic. There are many books which explain the importance of leading a "balanced" form of life. This book is different to the extent that it advocates Christianity as one method of finding that balance in a true and meaningful way. To many people, myself included, a few years ago, that statement sounds quite frightening. Christianity seems to be opposed to so many of the things which we take for granted and enjoy in the modern world. It appears to conflict with rationality, our sense of moral freedom and the full expression of the artistic impulse. These are all "pillars" of our age. However, I wish to argue that the opposite is true and, rather, that Christianity not only strengthens and nurtures these forces but, in fact, provides a means of unifying them.

I hope that the title of this book goes some way to negate the sense of hilarity which that last statement inevitably engenders in many people, who would not describe themselves as "spiritual types". I, myself, would have laughed at the possibility of such an argument, and for good reasons. I have a strong rational disposition, describing Mathematics as my "profession". Like anyone, who is not purely ascetic by nature, I value moral freedom. I also have a deep love and enjoyment of art. It is natural to feel suspicious of anything which might remotely challenge these commitments. However, I began to realize, during this period of transition, that this is not the full story. My specific field of study, since I began research for a Ph.D., has been "algebraic curves", a subject which, I believe, and hopefully the name suggests, cannot be easily characterized as purely an art or a science. The subject of geometry, by its very nature, challenges the view that the intellectual and artistic faculties are entirely independent. Now, I believe, that even more is true. The purpose of this book is to argue that geometry is essentially not merely an intellectual and artistic enterprise, but also a spiritual one. More specifically, my central argument is that the rational and artistic sense involved in this study are both guided by what I will refer to as "Christian imagery". I will leave it to the reader to judge, during the course of this book, whether this enterprise is successful. If so, I hope that it goes some way to support the view that Christianity can provide a means of guiding and unifying both our rationality and our artistic sense, in a real and substantial way. It is not my intention here to discuss the moral repercussions. I believe in Christianity because of its recognition of certain fundamental principles of human beings, which elevate both art and science. If you accept these principles on such grounds, then it is hard to deny that these same principles should also guide our moral sense and life in general. I will leave the question, as to whether basic Christian beliefs are the only means of achieving this goal, to the reader.

2. The Geometry of Light

This book was originally entitled “Christian Geometry”. However, after some consideration, I decided to change the title to “Christian Geometry: The Geometry of Light”. The book is essentially structured on the examination of certain Christian imagery, which is reflected in both medieval and Renaissance art, and the role that imagery plays in the understanding of the study of geometry. However, it is also a study of the nature of light, in the association of colours from the spectrum of refracted light, to geometric forms, and the way in which such forms are inter-related. The understanding of such a relationship clarifies our comprehension of light itself.

It would have been possible to take the spectrum of light as the basis for the main chapters. One can divide the visible spectrum into what I would refer to as the “left spectrum”, consisting of the primary colours “violet, indigo, and blue , and a “right spectrum”, consisting of “yellow, orange and red ”. These colours are based on the natural perception of the light of rainbows, or on a more scientific examination of light passing through a prism. One can also associate metallic colours such as silver, gold and copper to light, evidenced in the most natural phenomena of light from the stars, or from the sun and the moon, during an eclipse. Such colours may be associated with the “left spectrum” or “right spectrum” respectively. Ambient or natural light is best described as clear or white, in contrast to the darkness we associate to the absence of light. I will refer to such a quality of light as the “recombined spectrum”, owing to the natural intuition that it is a mixture or combination of primary colours, or of light from different sources in the atmosphere.

The division of the spectrum of light into three components is appealing not only for the natural reasons considered above, but also for the psychological intuition that it corresponds to the structure of the human brain, as divided into hemispheres, which are connected or combined by some physical process. In order to comprehend the structure of light more fully, it is necessary to pass beyond this basic division, and to discern both the nature of its components and their relationships. A naturalist or lover of nature might argue that this can be achieved by the simple observation of the sky, by day and night, the stars, sun and moon, and an intuitive understanding of how these natural sources of light combine. However, such a spiritual, “holistic” philosophy denies the genuine achievements of Western science, in its explanation of many natural optical phenomena such as the diffraction, refraction and reflection of light, through the postulation of a geometric model. The modern, scientific account of light, based mainly on the work of Newton and Huygens, argues that light behaves predominantly either as a particle with mass, by analogy with a snooker ball, or, as a wave, by analogy with water waves. The particle theory, was developed in Newton’s “Opticks”, partially to give a simple mechanical account of the reflection of light in mirrors, and was later espoused by Laplace. The wave theory, developed in “Traite de Lumiere” accounted for the effects of refraction and diffraction of light, which were unsuccessfully explained in Newton’s corpuscular theory. More recently, a version of Newton’s theory, in which light is composed of “massless” photons, has been used to account for phenomena such as the photoelectric effect, in which electrons are released over discrete intervals of time, rather than continuously, in the presence of ultraviolet light. Huygen’s theory is still successful in explaining effects such as the interference patterns of electromagnetic radiation or the propagation of microwaves. An attempt to combine these apparently contradictory viewpoints has been the subject of electrodynamics and the modern quantum theory. Maxwell’s equations describe the evolution of electrical and magnetic potentials in the presence of charge, for example, free electrons on the surface of a metal. If the initial current density is constant, possibly zero, solutions are provided by electromagnetic waves. Electrons in stable symmetric orbits around a nucleus provide an analogy with constant current density on a metallic surface. The corresponding solutions, in the case of a one dimensional surface, provide a basis for the observed frequencies of light. Schrodinger’s equation, which can also be interpreted as describing the variation of electrical potential as a wave, was developed in order to explain the discrete energy levels associated with the emission spectrum of the hydrogen atom. This quantization of energy was developed by Bohr, in order to find the corresponding stable orbits of electrons. Again, the solutions provide intuitions into the component frequencies of light. There are a number of drawbacks to the modern scientific theory. First, it fails to reconcile fully the geometric properties of light being both a wave and a particle, partly because photons, unlike electrons, are weightless. Secondly, the component theories of electrodynamics, quantum physics and thermodynamics still remain to be unified correctly. In this sense, one could argue that the position of the “holistic” spiritualist might be seen as superior.

The purpose of this book is partly to prepare the mathematics and geometry for a unified four-fold model of the theory of light, which addresses the second drawback of the current analytical theory. The reader will find links to current research into these areas in the text, particularly nonstandard methods of finding solutions to Maxwell’s and Schrodinger’s equation, approaches to probability in thermodynamics, and differentials. Each component of the model also attempts to reconcile the wave/particle duality problem, through the notion of projection, reflected in the significance of the Trinity in Christianity as a mediation between the spiritual and physical worlds, and the development of geometrical models. As already mentioned, I have based this theory, primarily, on the study of certain forms of Christian imagery, that fall naturally into four aesthetic and geometric categories, each category grouped into three components. This categorization is strongly supported by four main Christian beliefs in Jesus’s death by crucifixion, his later resurrection, (his transfiguration) and his status as a spiritual ruler. In this sense, one can see the study of a geometric theory of light as a continuation of Christian belief.

The aesthetic understanding of light and its effects is a particular feature of both the art and architecture of the medieval period. Perhaps, the most interesting examples are found in the use of rose windows as filters of ambient light into the interiors of cathedrals. The appearance of such windows depends extensively on the nature of the ambient light, a point of view that is easily demonstrated by photography (see my website http://www.magneticstrix.net) A fuller understanding of Christian imagery and these four main Christian beliefs, may also be achieved through the analogy of Christ, as either a source of light, or as a filter of the light of God. In his description as “the bright and morning star”, reflecting his status as a spiritual ruler, we find the holistic idea of recombination, a source of light between that of the stars and of the Sun. The events of the crucifixion and the resurrection, (transfiguration), may also be usefully understood in terms of the effects of polarizing filters, on the left spectrum and right spectrum of light respectively. The reader should refer to the Christian images, discussed below, of “The Throne of God” , “The Lamp of Heaven”, and (“The Right Hand of the Father”).

The consideration of a four-fold theory of light should take into account a modern alternative, based on the idea that colour is essentially obtained by mixing the three primary colours red, blue and green. Such an idea is supported by the medical discovery of three basic types of receptor in the human eye. However, although successful in explaining some manufactured optical effects, it fails to account for more natural effects, such as the appearance of rainbows or the refraction of light through prisms.

A successful new theory of the structure of light might have interesting implications for both science and Christianity. A central problem in modern science is an understanding of the process of nuclear fusion, the reaction which produces energy inside the Sun. Fusion is obtained by the bonding of hydrogen nuclei. However, modern methods to achieve this form of reaction and produce a steady supply of energy have, so far, been unsuccessful. A practical solution to obtaining a successful fusion reaction, capable of producing a sustainable source of energy, is related to the problem of understanding a three-fold model of light. Such a solution would solve the difficulties of mankind’s future energy supplies. It might also open up the possibility of interstellar space travel, owing to the vast amount of energy released by the reaction and the presence of small amounts of hydrogen in space. Naturally, such statements remain theoretical rather than practical. However, it remains my opinion that a successful approach to such problems requires a radical shift in the current scientific outlook.

The implication for Christianity is the prospect of establishing a new world order, based on the possibility of sustaining a human civilization outside Earth. Not only would this offer the possibility of an enduring harmony for mankind, but would also extend our sense of infinity, that the extent of our surroundings is larger than what now, in the context of globalization, appears to be a fairly small planet Earth.

3. The Images of Christianity

"That if thou shalt confess with thy mouth the Lord Jesus, and shalt believe
in thine heart that God hath raised him from the dead, thou shalt be saved." (Romans 10, vs 10)

For many people, myself included, this statement is a central element of Christian belief. Indeed, Paul, the author of Romans, is telling us that these actions are sufficient for "salvation". We should, therefore, consider carefully, first, what the two distinct parts of this exhortation are asking us to do, and, secondly, why we should do them. I will start with the second part, "and shalt believe in thine heart that God raised him from the dead".

To anyone who has read the testimony of the four Gospels, the meaning of this clause is clear. We should believe that Jesus, after being crucified, died and later ascended to heaven. However, it is extremely important to consider this testimony carefully, in order to understand what it tells us about this sequence of events. The gospels testify that there are four clearly defined stages in this process. I will briefly elucidate these. First, there is Jesus' physical death after the crucifixion. Secondly, there is a period of uncertainty, between his death and his subsequent resurrection. Thirdly, there is the event of the resurrection itself, in which Jesus' spiritual part is reunited with his material body. Finally, there are the events of the transfiguration and ascension, in which both his body and his spirit return to heaven. Let me then give the testimony of the Gospels with regard to each of these four stages. We have the following verses with regard to Jesus' death after the crucifixion;

"Jesus, when he had cried again with a loud voice, yielded up the ghost." (Matthew Ch. 27, vs 50)

"And when Jesus had cried with a loud voice, he said, Father, into thy
hands I commend my spirit: and having said thus, he gave up the ghost." (Luke Ch. 23, vs 46)

This fact is definitely witnessed, for example, we have;

"And when the centurion, which stood over against him, saw that he so
cried out, and gave up the ghost, he said, Truly this man was the son of
God. " (Mark Ch. 15, vs 39)

The period of uncertainty is best documented by the following passage;

"And they found the stone rolled away from the sepulcher. And they entered
in, and found not the body of the Lord Jesus. " (Luke Ch.24, vs 2-3)

"He is not here, but is risen: remember how he spake unto you when he was
yet in Galilee, saying, The Son of man must be delivered into the hands of sinful
men, and be crucified, and the third day rise again." (Luke Ch. 24 vs 6-7)

In other words, there is a period of three days, as Jesus himself predicted, between his death and subsequent resurrection. That some significant period of time does in fact elapse between his death and resurrection is borne out as well by the following passage;

"Saying, Sir, we remember that that deceiver said, while he was yet alive, After
three days I will rise again. Command therefore that the sepulcher be made sure
until the third day, lest his disciples come by night, and steal him away, and say
unto the people, He is risen from the dead: so the last error shall be worst than
the first. Pilate said unto them, Ye have a watch: go your way, make it as sure as

you can." (Mark Ch. 27, vs 63-65)

The passing of three days could be read symbolically. What is important is that some significant time does elapse. There is clearly a period in which Jesus is taken down from the cross and buried;

> "And Pilate marvelled if he were already dead: and calling unto him the centurion, he asked him whether he had been any while dead. And when he knew it of the centurion, he gave the body to Joseph." (Mark Ch.15, vs 44-45)

We now come to the event of the resurrection. Although this is never explicitly described, we can assume that it took place as the resurrected Jesus is witnessed both by Mary Magdalene and his disciples;

> "Sir, if thou have borne him hence, tell me where thou hast laid him, and I will take him away. Jesus saith unto her, Mary. She turned herself, and saith unto him, Rabboni: which is to say, Master" (John, Ch. 20, vs 15-16)

> "Then the eleven disciples went away into Galilee, into a mountain where Jesus had appointed them. And when they saw him, they worshipped him: but some doubted." (Matthew, Ch. 28, vs 16-17)

We should understand that Jesus appears in physical form to his disciples, not as a ghost or as a spirit;

> "And as they thus spake, Jesus himself stood in the midst of them, and saith unto them, Peace be unto you. But they were terrified and affrighted, and supposed that they had seen a spirit. And he said unto them, Why are ye troubled? and why do thoughts arise in your hearts? Behold my hands and my feet, that it is I myself: handle me, and see; for a spirit hath not flesh and bones, as ye see me have." (Luke, Ch. 24, vs 38-39)

Finally, we have the events of the transfiguration, in which the prophets Moses and Elijah appear next to Jesus;

> “After six days Jesus took with him Peter, James and John the brother of James, and led them up a high mountain by themselves. There he was transfigured before them. His face shone like the sun, and his clothes became as white as the light. Just then there appeared before them Moses and Elijah, talking with Jesus.”
>
> (Matthew Ch. 17, vs 1-3)

and the act of the ascension, which was witnessed by his disciples;

> "So then after the Lord had spoken unto them, he was received up into heaven, and sat on the right hand of God." (Mark Ch. 16, vs 19)

These passages elucidate the meaning of the statement "God raised him from the dead". However, it is far from clear why we should believe this, at least on purely rational grounds. As far as I am aware, there is no scientific explanation of the process of "resurrection". Even if there were, one would still have to come to terms with the further miracles of the ascension and transfiguration. On this basis alone, the testimony of a group of disciples, who were so closely attached to Jesus and had much to gain from their testimony, would be highly questionable in a modern court of law. Significantly, though, we are not being asked to accept this statement on rational grounds, rather that we should believe it "in our hearts". Does this mean that the belief should come then from a purely emotional response to this testimony? Many people, myself included, who have experienced the pain of losing a friend or a relative, are able to find resolution and acceptance by a belief in some form of life after death. However, this "feeling" is not strong enough to come to terms with the rather detailed account of the Gospels and the carefully elucidated stages involved in Jesus being "raised from the dead". In my opinion, and part of the subject of this book, the belief should be based on something more fundamental, which informs both a rational and emotional acceptance. This comes about by partly identifying with a series of what I will refer to as "The Basic Images of Christianity". I will consider these images more carefully in the following chapter. For now, let me try to identify some of these images from the four stages involved in the process of Jesus being "raised from the dead".

We first have the following passage describing what occurs before the first stage, Jesus' death;

> "Now from the sixth hour there was darkness all over the land unto the ninth
> hour. And about the ninth hour Jesus cried again with a loud voice, saying, Eli,
> Eli, lama sabachtani? that is to say, My God, my God, why hast thou forsaken
> me?" (Matthew Ch. 27, vs 45-46)

and the following passage, in the same Gospel, describing what occurs immediately after Jesus' death;

> "Jesus, when he had cried again with a loud voice, yielded up the ghost. And
> behold, the veil of the temple was rent in twain from the top to the bottom;
> and the earth did quake, and the rocks rent; And the graves were opened;
> and many bodies of the saints which slept arose." (Matthew Ch. 27, vs 50-52)

This last event is, in fact, the basis for the conversion of the centurion;

> "Now when the centurion, and they that were with him, watching Jesus, saw
> the earthquake, and those things that were done, they feared greatly, saying,
> Truly this was the Son of God." (Matthew Ch. 27, vs 54)

These first two passages from Matthew (vs 45-46 and vs 50-52) clearly convey three images. First, there is a picture here of gathering stress and tension. According to Matthew, it took nine hours for Jesus to die on the cross. In itself, this seems miraculous. Even if you cannot accept that Jesus had divine capabilities, the fact that he was able to remain conscious for these nine hours clearly shows that he had extraordinary strength and endurance. The picture of the prelude to his death is slow, drawn-out and agonizing. Moreover, the sense of foreboding is heightened dramatically by the darkness at the sixth hour and his desperate cry for help in the final moments, the only recorded moment in the Bible that Jesus shows any sign of uncertainty or weakness. This first image of tension is resolved by his death, and accompanied by a second image of "breaking", the temple veil is destroyed, the earth quakes and the rocks are rent. We are left with the sense that something has changed irreversibly. Moreover, this change is a forceful and cataclysmic one, Jesus "does not go gently into that good night". Finally, there is a third image of ascent, which accompanies the initial picture of gathering tension, and is reflected in the splitting of the tomb vertically, from top to bottom, and the bodies of the saints rising from the ground. Taken together, this images have an intuitive and, indeed, geometrical character. In order to support this view, I invite the reader to try the following thought experiment. Imagine a bendable willow stick, with both its ends being slowly moved towards each other. The stick, under increasing tension, will initially curve upwards in the middle, and, finally, break. One half represents the body of Jesus, the other represents his spirit, and, taken together, you have the form of a cross. For later reference, I will refer to these three images taken collectively as the image of crucifixion. The reader should realize that this is a spiritual image and not a physical one, as it is recorded that not a single bone of Jesus's body was actually broken. The picture is an extremely powerful one. The gathering image of tension corresponds to the gradual weakening of Jesus' body but the strengthening of his spirit. This strengthening of spirit is further reinforced by the picture of upward ascent, as Jesus' soul struggles to come closer to God. Finally, at the moment of death, we have the breaking of his spirit from his body, and the image of the body and spirit in a crucified state.

I will examine the image of crucifixion in much further detail later in the book. For now, I will continue with the identification of "Christian imagery" in the further stages leading up to Jesus' ascension.

The second stage of the period of uncertainty, following Jesus' death, begins in John with the following passage;

> "Then took they the body of Jesus, and wound it in linen clothes with
> the spices, as the manner of the Jews is to bury. Now in the place
> where he was crucified there was a garden; and in the garden a new
> sepulcher, wherein was never man yet laid. There laid they Jesus." (John Ch. 19, vs. 40-42)

The image of Jesus' burial in the garden is very different from the image of crucifixion. It has a sense of serenity, which contrasts vividly with the tension of the previous passages. This feeling is heightened by the description of the garden. There is also a sense of grace and respect for the dead, Jesus was not thrown carelessly into a mass grave but entombed in a new sepulcher. Perhaps most importantly, there is an impression of gentle humility here, the physical aspect of Jesus as an ordinary man, rather than as a conquering spiritual leader or potential leader of the Jewish people is emphasized. This impression comes from the sense that, at this moment, Jesus' spiritual part has separated and we are dealing only with his physical body and from the observation that he is buried in the tradition of his people. For later reference, I will refer to this as the image of the Lamb, and, occasionally, as the image of the Branch. I think, for many people, it is one of the most difficult aspects of Christianity to comprehend. In physical terms, Jesus' death is a failure. However, in spiritual terms, it is a potential victory over sin and death, a victory eventually confirmed by his later resurrection. In order to clarify this point, it is important to realize that the image of the Lamb succeeds the image of crucifixion. The images of Christianity that I will explore are essentially spiritual ones, and arise as part of the process of spiritual growth. The ability to identify these images does not just come about from reading the New Testament or this book, but, to some extent, requires the capacity to come to terms with and understand the passage from life into death. As this understanding increases, these images impress themselves more forcefully in a natural sequence. I hope the reader will agree that it is not necessary to die in order to obtain this spiritual awareness. However, I do believe that it is an awareness that can only come about through some understanding of suffering and self-sacrifice, an understanding which is hampered by the easy route of simple materialism. This is not to advocate ascetism. Perhaps, a good analogy is that of a cyclist, who strains to reach the summit of a hill, and then enjoys the combined adrenaline rush of the following descent and the fantastic view. Taken by itself, the image of the Lamb is an image of weakness, however, when viewed in the full context of the images of Christianity, which I will explore, it becomes a strength. As one observes from the above passage, the image of the Lamb also symbolizes grace and refinement. It goes without saying that both of these qualities are essential both to the individual and to the smooth functioning of any civilized society, irrespective of whether that society is a predominantly Christian one. I will discuss the image of the Lamb, from a more abstract point of view, and its connection with the image of the Branch, in greater detail in Chapter 10.

Continuing with the identification of Christian imagery, we further explore the period of uncertainty leading up to Jesus' resurrection. This is the period of the three days after his burial. Although there is no description in the Gospels as to what occurs in these three days, it is clearly significant that some period of time does in fact elapse. Jesus could have been resurrected immediately after his crucifixion. Given that some form of miracle, for example the earthquake, was recorded at the moment of his death, it seems no more improbable that a further miracle, such as his immediate ascension to heaven in a glorious fanfare of trumpets and angels could also have occurred. However, this sequence of events would have excluded the image of the Lamb, and, moreover, the following successive image, which I will now discuss. During the period of uncertainty, as we previously observed in the picture of his burial, Jesus' spirit is divorced from his body. There is a sense that is spirit is searching for something, a way to return to his physical body, a way to reach a higher spiritual plane. As

I will make clearer later on in the book, this image seems to "flow" from the image of the Lamb. This idea is further supported by the following passage from Revelation;

> "And he showed me a pure river of water of life, clear as crystal, proceeding
> out of the throne of God and of the Lamb." (Revelations, Ch. 22, vs 1)

I will, therefore, refer to this as the image of the river of life. It is one of the most beautiful images in Christianity and expresses quite clearly an aspect of Jesus' s artistic nature. The clarity of the water and the reference to a crystalline form complements the masculine image of the crucifixion. Again, I will explore this image in much greater detail, from an artistic and geometric perspective, in Chapter 6.

After the period of uncertainty, we arrive at the third stage of Jesus' resurrection. At this point, Jesus' spirit has reunited with his body, but he is in a far higher, more focused, spiritual state than at the time before his death. This growth in spirit is mirrored by a sense that his physical state is now more delicate and fragile. This view is supported by the following passage, which describes Jesus' meeting with Mary;

> "Jesus saith unto her, Touch me not; for I am not yet ascended to my Father"
> (John, Ch. 20, vs 17)

There is also a sense that Jesus' spirit and body, which were previously joined in a crucified state, are now separated but in close juxtaposition. I will refer to this as the Image of the Resurrection. As with all the images which we will discuss, it informs a fundamental method of geometric thinking. The reader is invited to try another experiment in order, perhaps, to understand this relationship better. Take two pieces of string, representing Jesus' spirit and his body, and attach the four ends to two separate rods. Hold one rod above the other, so that the two strings fall parallel to each other and twist the bottom rod by a full 360 degrees, so that the strings become intertwined or "braided". At this point, if you keep the two rods fixed, the strings cannot be manipulated so that they return to a separated, parallel position, but the strings are also not joined at any point. I have used the analogy of strings, not only because it is impossible to braid two sticks, but to suggest the fragility of Jesus' body and the incorporeal nature of his spirit. This is, of course, not the only interpretation of the Image of the Resurrection, there is a sense of reawakening from sleep, as if by the exposure to sunlight. It is an interesting question as to whether the sepulcher in which Jesus was laid, was closed, and the event happened immediately, or whether, this process occurred gradually, in a period of time after the crucifixion. The analogy of heat propagating through a body, in the form of waves, stimulating a return to mental focus, seems appropriate. Some evidence for this theory can also be found in the debate over the Turin Shroud, in which an almost photographic imprint of Jesus’s face can be found on the garments by which he was covered. This suggests that the unwitnessed event of the resurrection, was accompanied by a surge of energy and light. We discuss the imagery around the resurrection in Chapter 6.

We finally come to the final stages of Jesus’ transfiguration and ascension. As I mentioned before, it is significant that the events are witnessed only by his disciples and do not occur immediately after he is crucified. If they had, we would lose the numerous aesthetic images described after the Image of the Crucifixion. Moreover, these events proceed with a sense of quietness and privacy, as we saw in the passage above from Matthew, (Matthew Ch. 17, vs 1-3), and, in the following;

> "So then the Lord Jesus, after speaking to them, was taken up into heaven
> and sat at the right hand of God" (Mark, Ch.16, vs 19)
>
> "And he led them out as far as to Bethany, and he lifted up his hands, and
> blessed them. And it came to pass, while he blessed them, he was parted
> from them, and carried up into heaven" (Luke Ch. 24, vs 50-51)

There is no suggestion of a grand spectacle of angels here, only the simple description of the events themselves. This reinforces the sense that the Images of the Transfiguration and Ascension have a gentle, feminine quality about them, an idea we will consider again later. There is also the further sense of spiritual ascent, which we noted previously in the Image of Crucifixion, hence, in some sense, one can see that the these images could easily have followed on directly from the Image of Crucifixion. (This is the romantic idea.) This combination of images is completed by the reference to heaven itself. There are many descriptions of heaven in the New Testament, but, perhaps the most striking, which I will refer to as The image of the Throne of God, and The Image of The Right Hand of the Father are sublimely described in the following passages;

> "And out of the throne proceeded lightnings and thunderings and voices;
> and there were seven lamps of fire burning before the throne, which are
> the seven spirits of God. And before the throne there was a sea of glass
> like unto crystal.." (Revelation Ch. 3 vs 5-6)

> "And above the firmament that was over their heads was the likeness of
> a throne, as the appearance of a sapphire stone" (Ezekiel Ch. 10, vs 1)

> "And what is the surpassing greatness of His power toward us, the ones
> believing according to the working of His mighty strength which He worked
> in Christ in raising Him from the dead, and He seated Him at His right hand
> in the heavenlies, far above all principality and authority and power and
> dominion, and every name being named, not only in this world, but also in
> the coming age" (Ephesians Ch. 1, vs 19-21)

The idea of heaven is not exclusive to Christianity, almost all religions that I am aware of, have some notion of a spiritual life after death. Taken on its own, the notion of a heaven and earth represents the basic sense of the distinction between body and spirit, the core element of "spirituality". What distinguishes Christianity from other religions, is its recognition of the complexity of this relationship between our corporeal and spiritual natures, a complexity mirrored by the highly sophisticated sequence of images that recur throughout the New Testament. At its most basic level, simple "spirituality" seems rather weak. It is one half of the romantic sense of wonder that we experience when gazing out at a "sea of glass" on a clear day. As such, it involves no element of suffering, and no recognition of our feminine and artistic qualities. As there is no real feeling of effort required on our part to come closer to "God", we are therefore left with a whimsical sense of the unattainable. Christianity seeks to close this gap through spiritual ascent. Taken on its own, the Images of the Throne of God and The Right Hand of the Father are frightening and unattainable. However, the Ascension shows Christ as a spiritual intermediary between God and man, an image repeated in other forms throughout the New Testament;

> "And John bare record, saying, I saw the Spirit descending from heaven like a
> dove, and it abode upon him...And I saw, and bare record that this is the Son
> of God" (John, Ch. 1, vs 32 and 34)

The dove, another symbol for the image of the Ascension, represents part of the spiritual process by which we are able to traverse this divide between God and man. It also confirms the unique role that Christianity, through the example of its prophet, can play in fostering our ability to grow spiritually.

The spiritual process of the Ascension also represents a return to a harmonious state, described by Jesus's role in the New Jerusalem, as a heavenly ruler;

> "And the city had no need of the sun, neither of the moon, to shine on it; for
> the glory of God did lighten it, and the Lamb is the light thereof" (Revelations, Ch. 21, v 23)

which is translated in the New International Version as;

"The city does not need the sun or the moon to shine on it, for the glory of
God gives it light, and the Lamb is its lamp."

The Lamb is compared to a lamp, which illuminates heaven through the radiation of God's glory, described in the following passage;

"Having the glory of God, and her light was like a stone most precious, even
like a jasper stone, clear as crystal" (Revelations, Ch. 21, v10-11)

The light of the Lamb complements the description of the city as majestic and harmonious;

"And the city has a wall great and high, and had twelve gates,…On the east
three gates, on the north three gates, on the south three gates, and on the
west three gates. " (Revelations, Ch 21, v12)

and of pure gold;

"And the building of the wall of it was of jasper and the city was pure gold,
like unto clear glass" (Revelations, Ch. 21, v18)

The image of harmony is reinforced by the comparison of God to the Sun and the risen Christ as an intermediary between the citizens of heaven and God. The image of the Lamb as a lamp, is evocative of light and heat; reinforcing a notion of harmony and focus, by the intuition of light as a wave. The reawakening of Jesus, through a process of being radiated by light, which I discussed above, reinforces this idea, in the process of resurrection.

A feminine interpretation of Jesus's ascension might find a comparison with the Eastern belief in reincarnation. Reaching heaven, Christ transcends the "wheel of life" and enters into a state of "spiritual nirvana". The process of reincarnation is vividly described in "The Tibetan Book of the Dead". By making a series of unconscious psychological choices, involving the recognition of certain luminosities, the dead person can reach a state of enlightenment (nirvana) or rebirth. Buddhism uses the analogy of a "wheel" to describe this process. One might modify the analogy to that of a ball on a fast spinning roulette wheel or a rotating gyroscope. The final picture of Jesus sitting at the right hand of God, is an imaginative portrayal of God as a creator, intervening through the spirit, with Christ as an intermediary. Considering the verticality of the relationship between God and Jesus, emphasized in the passage from John, concerning his baptism, this passage creates a dynamic sense of asymmetry.

I have tried to identify what I believe are the most important Christian images associated to the sequence of events involved in understanding the meaning of "God hath raised him from the dead". I will leave it to the reader to decide, during the course of this book, whether these images can provide some support for the second part of Paul's original expostulation.

I will now turn to the first part, "That if thou shalt confess with thy mouth the Lord Jesus". In some translations, for example the New International Version, one can find the following reading of this statement, "That if you confess with your mouth, "Jesus is Lord". Again, we should try to understand what this statement really means. It is clearly significant that the word Lord is used. We are not being asked to declare that Jesus was a great man or an important spiritual prophet, but something more, which puts him in a unique position befitting of the title Lord (with a capital L). The normal usage of the word Lord is to denote someone who is a ruler of a group of people, such as a king or queen. However, this has quite an earthly connotation, and ignores the spiritual role that Jesus obviously plays. The previous passages, from Revelations, described him as a ruler of the heavenly city of Jerusalem. As is suggested by these passages though, in which God is compared to the Sun, we are not being asked to declare that "Jesus is God" either. God is clearly represented in the New Testament as the highest point in the spiritual hierarchy, and is a distinct entity from Jesus. We are often reminded that Jesus is the son of God, not God himself;

"For as the Father hath life in himself; so hath he given to the Son to have
life in himself" (John, Ch. 5, vs. 26)

As we observed above, Jesus is represented as a spiritual intermediary between God and man, so the word Lord cannot be construed in a purely spiritual sense either. Given what he have said about the importance of the relationship in Christianity between the spirit and the body, and the central role that Jesus plays in allowing us to understand this relationship, it is clear that we need to understand the word Lord on two different but connected levels. Perhaps the clearest statement of how we should see Jesus as Lord comes in the following passage, which contains the image of the bright and morning star, and what I will later refer to as the image of the Illumination;

"I Jesus have sent mine angel to testify unto you these things in the churches.
I am the root and the offspring of David, and the bright and morning star."
(Revelation Ch. 22, vs 16)

David was the ruler of Israel and Judah for a period of forty years, before the time of Christ, as described in the Old Testament, for example, in the books of Samuel. He is recognized as one of the greatest rulers of Israel and beloved of God;

"And the Lord preserved David whithersoever he went. And David reigned
over all Israel; and David executed judgment and justice unto all his people."
(Samuel 2 Ch. 8, vs 14-15)

In the previous passage, then, Jesus is claiming, in some sense, to be a descendant of David. That he is also the root, is a reference to the destruction of Israel and David'd line, and Jesus' renewal of what is left;

"A shoot will come up from the stump of Jesse; from his roots a Branch will
bear fruit. The Spirit of the Lord will rest on him." (Isaiah, Ch. 11, vs 1)

There is a sense then that we should see Jesus partly as the continuation and head of an earthly kingdom, a source of growth and renewal in this world. This is a striking image which I will refer to as the Image of the Tree of Life. Just as the Image of the River of Life flows from the image of the Lamb, so this image stems from the image of the Branch. It represents a feminine aspect of Jesus' persona and the establishment of a Christian kingdom both on earth and in heaven;

"And he showed me a pure river of water of life, clear as crystal, proceeding
out of the throne of God and of the Lamb. In the midst of the street of it,
and on either side of the river, was there the tree of life, which bare twelve
manner of fruits, and yielded her fruit every month: and the leaves of the tree
were for the healing of the nations." (Revelations 22, vs 1-2)

That this kingdom is both spiritual and corporeal is supported both by this passage and by Jesus' claim to be the root and offspring of David, and the bright and morning star.

Unlike the River of Life, there is no sense of uncertainty in the image of the Tree of Life. It is a triumphant vision of the glory of a Christian kingdom, ruled by Jesus as Lord. There is also a sense in this passage of an interaction between these images, through the depiction of the tree of life standing in the midst of the river. I will also discuss further distinctions and connections, of a more geometric nature, later in this book.

Having discussed the meaning of the "Lord Jesus", we should also consider why Paul emphasizes that we should confess this with "our mouth". He could have just asked us to confess it, or to believe it, in our hearts, as in the previous statement that "God hath raised him from the dead". If Christianity were merely a set of beliefs, which were never communicated, it would have no impact on the world around us. It would become the prerogative of a few isolated individuals, speaking in "tongues" or a purely spiritual language. As we have seen repeatedly in this chapter, Christianity is a religion which seeks to enhance both our spiritual and corporeal natures, to serve as a guide for our whole persona and to help make our lives fulfilled both in this world and the next. We should, therefore, not sit back passively and wait for the afterlife but actively strive to establish a Christian kingdom on this earth! Paul's explicit referral to the word mouth emphasizes the essentially corporeal nature of this task; there is also the sense of contrast between the heart and the mouth, as representing the distinction between spirituality and the more earthly quality of femininity from which the image of "Lord Jesus" flows through the spoken word.

I hope that I have given a fair account of the meaning of Paul's expostulation, presented at the beginning of this chapter. In my opinion, we should accept it on the grounds that the Christian imagery which I have identified can serve as a basis to inform our full persona, spiritually, artistically and rationally. This will be the subject of the rest of this book. Let me end this chapter, then, by reminding the reader of the Images which we have identified, the meaning of the separate columns will be explained shortly;

1. The Crucifixion.	1. The Crucifixion. (left spectrum) (blue, indigo, violet, silver)
2. The Lamb(The Branch).	2. The Ascension.
3. The Resurrection.	3. The Throne of God.
4. Reawakening.	4. The Lamb (The Branch) (recombined spectrum) (white)
5. The Transfiguration.	5. The Illumination.
6,7. The Ascension.	6. The Bright and Morning Star.
8. The Right Hand of the Father.	7. The Resurrection. (right spectrum) (green, yellow)
9. The Throne of God.	8. Reawakening (as reincarnation).
10. The Lamp of Heaven	9. The Lamp of Heaven.
11. The Illumination.	10. The Transfiguration. (orange, red)
12. The Bright and Morning Star.	11. The Ascension (as reincarnation).
	12. The Right Hand of the Father.

I have deliberately used the numbering of the images in the left hand column to suggest a form of spiritual hierarchy. This hierarchy reflects the order of events recounted in the Gospels and Revelations. However, I will also use the ordering of the images in the right hand column. This grouping is very attractive for the reason that it demonstrates an aesthetic balance between masculine imagery (1,2 and 3), feminine imagery (7-9(10-12)), and neutral imagery (4,5 and 6). The images (2,5 and 8(11)) of the second column mediate between the groups of masculine, neutral and feminine images respectively. In the case of the first three images (1-3), this reflects the Trinity of Father, Son and Holy Spirit, which is recounted in the above description of John the Baptist, (John, Ch. 1, vs 32 and 34). The Son is represented by The Image of the Crucifixion, in which Jesus' body and spirit, ascend through the Holy Spirit, represented also by the Image of the Ascension, to the Father, represented by the Image of the Throne of God. In the case of the three images 7-9, we have a similar Trinity of Father, Son and Holy Spirit. The Son is represented by the image of the Resurrection, reawakened through the Holy Spirit, by God, represented by the image of the Lamp of Heaven, (Revelations, Ch. 21, v 23). The mediating act of reawakening should, perhaps, in this case, be given the more feminine interpretation of reincarnation. This occurs also in the three images 10-12, the Son is represented by the image of the Transfiguration, ascending through the holy spirit, and sitting at the right hand of God, (Matthew Ch.17, vs 1-3), (Mark, Ch.16, vs 19). In the case of the three images 4-6, we have a Trinity of the Holy Spirit, represented by the image of The Illumination, and Christ, as both an earthly and spiritual ruler, represented by the image of the Lamb and the Bright and Morning Star, (Revelation Ch. 22, vs 16).

One can find other interesting trinities in the imagery that we have considered, although they are not incorporated directly into the ordering of the second column, and as part of the chapter headings of this book. The images of the Resurrection and the Tree and the River of Life suggest a Trinity of Heaven, Earth and Holy Spirit, reflected in the above passage (Revelations ch22, vs 1-2). The Holy Spirit is represented by The River of Life, in which Jesus' spirit searches, through the process of Resurrection, to reunite with Heaven and Earth, represented by the Tree of Life. One can supply the image of water "weaving" through the tree, as similar to the process of "braiding" that I used to explain the Image of the Resurrection. These images are discussed in Chapter 6.

The images of the Lamb, Lamp and Heaven and Reincarnation suggest a Trinity of God, Christ and Heaven, in which God and the Risen Christ are represented by the image of the Lamp of Heaven, and Heaven and the Risen Christ are represented by the image of the Lamb. Even the people inhabiting paradise and the angels are unable to see the light of God directly. The mediating form of Christ acts as a radiator of the light of God, the analogy of a coloured filter of light seems appropriate. As I explained above, the image of the Lamb has a human, earthly significance to it, though in this case it represents the lower position in the spiritual hierarchy, relative to God, of the people of heaven. One can provide a similar interpretation of the images of the Lamb, the Throne of God and the Ascension. The analogy of a polarizing filter of light can also provide a useful interpretation of the first and last sequence of three images.

The grouping of the second column is the one that I will use for the remainder of this book. The mediating images of The Ascension, The Illumination and Reawakening represent a "process" of understanding the more aesthetic images (1 and 3), (4 and 6), (7 and 9) (10 and 12) respectively. This process is reflected in a form of geometrical thinking that I will describe in the course of this book. The grouping, obtained after removing this process, suggests an interesting physical correspondence between light and matter, which I alluded to in the previous section; namely that of wave-particle duality . The grouping will form an essential component of the overall system of Aesthetics that I will present in the final chapter.

4. The Crucifixion and The Throne of God

In this chapter, we will consider how the images of The Crucifixion and The Throne of God, which we discussed, have been employed in the visual arts. It will be my purpose to show that, on an aesthetic level, they have been of fundamental importance in the development of Western art and architecture. Let me begin, then, by considering some examples of artists who have employed these images.

Gaudenzio Ferrari was a Northern Italian painter and sculptor of the Renaissance period. His major work was carried out in the small town of Varallo, northwest of Milan, in the region of Piedmont. Varallo is mainly famous for the Sacra Monte. As its name might suggest, the Sacra Monte consists of a series of 45 chapels and a Basilica, situated on the slopes of Monte Tre Croci. The chapels depict the life, death and resurrection of Christ, in more than 800 life-size sculptures of terracotta, wood and marble. What is most unusual about these sculptures is their graphic realism, many of the scenes inside the chapels are populated by figures with real hair and beards, enacting events with close attention to actual physical details. The sculptures depicting the Crucifixion of Christ were carried out by Gaudenzio Ferrari, between 1524 and 1529;

1. The Crucifixion of Christ, by Gaudenazio Ferrari, Sacra Monte, Varallo, Italy.

The extreme realism of the figures adds to the sense of suffering in this image, the raw portrayal of the physical death of Christ. There is little feeling of spirituality here. Indeed, the series of chapels are situated on a tortuous ascent of the Monte Tre Croci, no doubt designed to stimulate some fraction of the exhaustion that

Christ must have actually felt as he carried his cross to the top of Calvary. Gaudenzio Ferrari also painted, earlier in 1513, a fresco of the Crucifixion in the Church of Santa Maria delle Grazie;

2. The Crucifixion by Gaudenzio Ferrari, Santa Maria delle Grazie church, Varallo, Italy.

located inside the main town of Varallo rather than the overlooking Sacra Monte;. Perhaps not as straightforwardly realistic as his work in the Sacra Monte, the painting, in its depiction of weeping angels, demons and knights clad in black armour, leaves the viewer with a strong sense of physical suffering and humility in the face of the cruel inevitability of death. Ferrari's work made a deep impression on me when I first visited Varallo, 15 years ago. What struck me then was how different these images were from other, more life affirming works of the Renaissance. Stylistically, this is often explained by the close proximity of Piedmont to Germany, and the influence of the harsh brand of Northern Gothic art in this part of Italy. Perhaps, it is this type of imagery which many people associate with Gothic art.

However, I wish to consider now a very different depiction of the Crucifixion, carried out by a contemporary of Ferrari, Jacopo Robusti, better known as Tintoretto, after his father's trade as a dyer, (Hartt, 2006). Tintoretto is acknowledged as one of the great geniuses of the Late Renaissance Venetian school of painting. In 1564, he won the commission to decorate the Scuola di San Rocco, a benevolent Christian institution situated in the heart of Venice. The paintings that he executed for the school, about 50 in total, are generally recognized as one of the single highest achievements of the Renaissance period. The cycle illustrates episodes from both the Old and New Testaments, and demonstrate not only Tintoretto's virtuosity as a painter but also his deep knowledge of theology and profound religious belief. His painting of The Crucifixion;

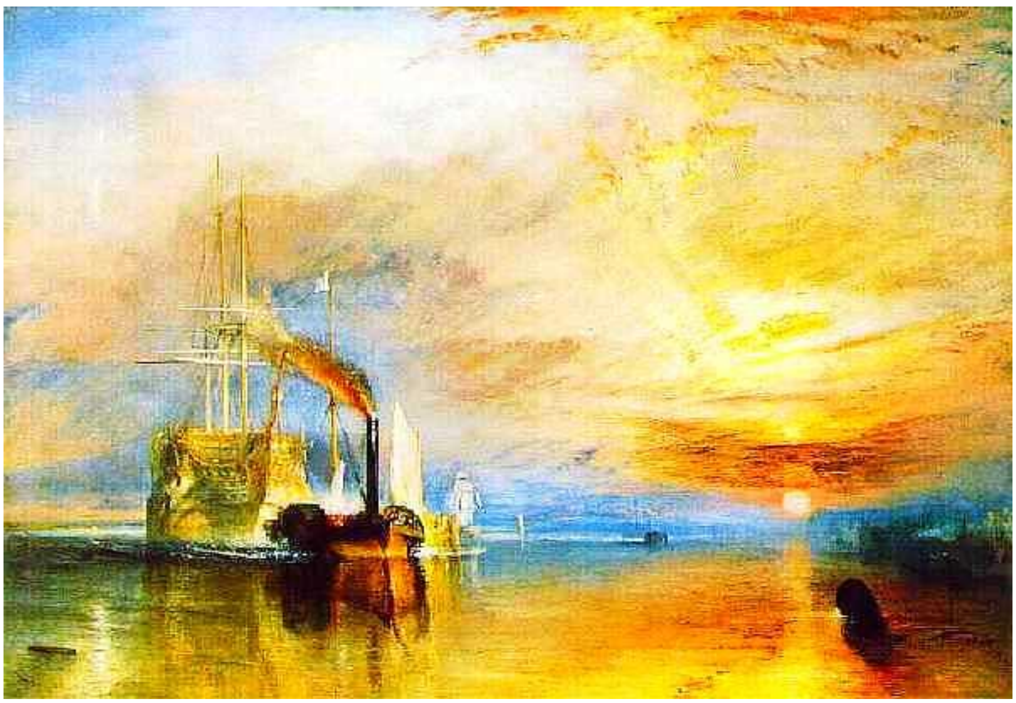
3. The Crucifixion byTintoretto, Sala dell'Albergo, Venice, Italy.

executed in 1565 for The Sala dell'Albergo, is perhaps the most astounding painting of a series of acknowledged masterpieces. In "Stones of Venice", John Ruskin, an outspoken critic of the last century, and usually not someone who withheld his opinions on art, was so awed by the work, that he described it as "beyond all analysis, and above all praise". El Greco, who was heavily influenced by Tintoretto, referred to it as "the greatest painting in the world". It is easy to see why this painting invoked such great admiration. Unlike Ferrari's depictions of the event, Tintoretto emphasizes its spiritual rather than physical significance and he does so on a number of different levels. First, there is a strong dynamism in the painting, compared to the rather static symmetry of Ferrari's fresco. Stylistically, this is achieved, partly by Tintoretto's distinctively energetic brushwork, but, more interestingly, by the geometrical use that he makes of the strong intersecting lines of the crosses. If the reader were to draw the six lines provided by the three crosses, he will be struck not only by their asymmetry but also by the fact that no three of the lines converge to a common apex. A more classical artist would have been careful to ensure that the arrangement of lines was not only symmetrical but also converged to the natural centre of the painting, Christ's head. Indeed, this is the effect achieved by more traditional representations of the Crucifixion, involving three fully raised crosses. I will refer to this aesthetic as fragmentation, we will consider it in much greater detail later in the book. Secondly, we can observe that the painting has a strong sense of verticality. This is achieved partly by placing the centre of Christ's cross at the very top of the painting and extending the vertical line of the cross down the centre of the composition, to the bold mass of onlookers, huddled at the base, but also by the fact that, at the moment depicted, the second cross is in the process of being raised from the ground. I will refer to this aesthetic as ascent, again we will have the opportunity to consider it in greater detail later. Thirdly, there is the intense element of spirituality provided by the sublime depiction of Christ himself. Enveloped in a halo of light, he appears to soar out of the picture over the viewer, an effect which immediately struck me when I first saw it and which no reproduction can fully capture. We are presented with the moment of Christ's death, in which Christ's soul separates from his body in the form, as the light suggests, of either an angel or a dove. This adds to the aesthetic of fragmentation which we noted earlier.

The direction of Christ's flight is deliberately ambiguous, the nimbus of light adding to the feeling of uncertainty that the viewer feels. There is a communion between the physical and something infinite, beyond human comprehension or explanation. This aesthetic, which is more closely connected with the next image that we shall consider, I will refer to as the sublime.

It is much harder to find examples of artists who have directly used the heavenly image of The Throne of God. By its very nature that we discussed in the previous chapter, it is an image which defies representation. However, I would like to consider some who have referred to it indirectly in their work.

Joseph William Turner was an English artist, working mainly in London at the beginning of the 19th century, who is probably most famous for his landscape painting. John Ruskin, who was deeply impressed by Tintoretto's Crucifixion, was also a staunch supporter of Turner. His "Modern Painters", (Ruskin, 1843-60), is to some extent devoted to a defence of Turner's work, in the face of a generally hostile public. Turner's "The Fighting Temeraire" ;



4. The Fighting Temeraire by Joseph Turner, Tate Gallery, London.

was painted in 1839, and is considered to be one of the artist's greatest later works. Now located in the National Gallery, London, it depicts the decommissioning of the gunship "Temeraire", which played a critical role in Nelson's victory at the Battle of Trafalgar. The image of the setting sun over the water is analogous to the vision of the seven lamps seen across a sea of glass, which we discussed in the previous chapter. As we noted before, it is an image of isolation and distance. This reinforces the picture of "The Temeraire" as one of the last of its kind, soon to be replaced by a new generation of steam powered ships, such as the tug which is bringing the battleship to its last berth. It is also an image of power, beyond ordinary comprehension, here this is used to add a sense of pathos to the painting, the setting sun reflecting the faded glory of England's naval power. Finally, it is an image of the infinite, the view of the observer is drawn towards the sun by the path of rays striking the water, but the source of the path is never reached, an effect which Turner emphasizes by deliberately breaking the line as it reaches the horizon. "The Temeraire" moves away from the source of light in the painting, approximately along the path of the rays, as if to reinforce the sense that the source of this path and the ship's past glory is now unattainable. These aesthetics of power, isolation and infinity are features of a more general aesthetic which I have already referred to as the sublime. There is finally a suggested aesthetic of fragmentation, which is more closely related to the previous image. Stylistically, this is achieved by crossing the trajectory of the ship with the ray of light on the water, and also in the theme of the painting, that the Temeraire's final destination will be its destruction.

Another artist who may be said to have attempted a depiction of The Throne of God, at least in the aesthetic from of the sublime, is the romantic painter Kaspar David Friedrichs. His "Wanderer Above The Sea of Fog";

5. Wanderer above the Sea of Fog by Kaspar Friedrichs, Kunstalle, Hamburg, Germany.

completed in 1818, presents us with a lone, isolated figure staring across a seemingly unfathomable landscape. Unlike Turner's painting, the viewer is not drawn to any particular point in the distance, the sense of infinity is here rather portrayed in the sheer mass of fog and mountains. There is again a sense of the awesome power of nature, (Gombrich, 2006).

Having introduced some of the main aesthetic ideas, the rest of the chapter will be concerned with analyzing and developing them at a more abstract level. We will consider first the aesthetics of spirituality and fragmentation, arising from The Crucifixion, and then go on to consider the aesthetics of the sublime. At the end of the chapter, we will consider deeper connections between the two. It is a feature of Tintoretto's work, and to an extent Turner's, that both types of aesthetic are used and related. In the case of Tintoretto's Crucifixion, we are led from the aesthetic of spiritual fragmentation to the sublime. In the case of Turner's "Fighting Temeraire", from the aesthetic of the sublime to fragmentation. This feature is for me what makes these paintings worthy of the title of great art.

Our main basis for the analysis of these aesthetic ideas will be in the writings of John Ruskin and the history of Gothic architecture, mainly in England. These sources are closely related because much of Ruskin's writing is about Gothic architecture. As the figure of Ruskin is so important, I will begin by giving a brief biography of his life and work. John Ruskin was born in London in 1819. After graduating from Oxford, his father's status as a wealthy sherry merchant allowed him to pursue a relatively independent career, writing mainly for architectural magazines. His first major work, "Modern Painters", which we mentioned briefly above, was published in 1843, and argued that modern landscape painting, such as that of Turner, was superior to the classical style of "Old Masters" such as Lorrain and Poussin, as it was "true to nature". His second book, The Seven Lamps of Architecture, (Ruskin, Seven Lamps of Architecture, 1849), was published in 1849, and sets out his main aesthetic ideas. These ideas were further developed in "Stones of Venice", (Ruskin, Stones of Venice, 1851-1853), published in 1851-3, which is essentially a detailed analysis of Gothic architecture in Venice, but also a devastating attack on the Renaissance period, which Ruskin believed to be one of the major causes of the decline of the Venetian empire. By this time, Ruskin had established himself as one of the most influential cultural critics of his day. He was a major figure in the Gothic Revival in England during the 19th century, and one of the main intellectual sources of The Pre-Raphaelite Brotherhood. Ruskin was a committed evangelical Protestant for much of his life, in particular during the period that he wrote his major aesthetic works, although he broke with his original evangelical beliefs in 1858. He spent much of his later life at Brantwood, a beautiful country house situated on the shores of Lake Coniston in Cumbria. He died on January 20th, 1900, at the turn of a century which he had done so much to shape and influence.

I will begin by examining some of Ruskin's views on the emotions associated to the rather harsh aesthetics of Ferrari's work. The Lamps of Sacrifice and Obedience are two of the seven aesthetic criteria that Ruskin examines in "The Seven Lamps of Architecture". In the Lamp of Sacrifice, Ruskin describes the spirit of sacrifice in architecture as the desire to produce small results with cost and thought. This spirit should exhibit itself in either of the following forms, first, the wish to exercise self-denial for self-discipline, and, secondly, the desire to honour or please another by sacrifice. Ruskin is, here, taking a set of moral principles and applying them to the practice of architecture. These moral principles stem directly from his religious beliefs, that self-sacrifice is a way of honouring God, indeed he quotes a passage from Numbers (Ch. 31, v54), on the importance of sacrifice, to support his view. The ability to dedicate art to God, rather than to serve our own narrow self-interest, is, for Ruskin, an important part of what constitutes humanity. The following quote is taken from The Virtues of Architecture (Book 1, Ch. 1 of Stones of Venice);

"Humanity consists in the dedication of them all to Him who will raise them up at the last day"

As someone who believes in the existence of a real spirituality in man, this is a sentiment with which I strongly agree, and I will argue throughout this chapter that spirituality is an important aesthetic criteria in art. As I mentioned in the previous chapter, sacrifice in art and life is one of the conditions required to attain a higher spirituality. This is a rather masculine point of view, but, for this reason alone, I would support Ruskin's view on the importance of sacrifice in art. Ferrari could be said to have fulfilled Ruskin's criteria in both the works described above, through his close attention to physical detail and his choice of the subject matter of the Crucifixion. I can find no record of Ruskin ever visiting Varallo, but, I think he would have admired Ferrari's work to the extent that it demonstrates a form of muscular Christianity which he believed in.

In the Lamp of Obedience, Ruskin adopts a social, rather than spiritual, argument in favour of restraint on the part of the architect. Ruskin argues that architecture should not defy social rules for the sake of singularity, great architecture exists only when it is universal and established as a nation's language. Ruskin, in fact, claims that, only through certain forms of restraint can the artist liberate his expressive potential. His argument, here, is through comparison with language, only by learning the rules first can real expression be formed from it. This exhortation to what Ruskin refers to as "Obedience, Unity, Fellowship and Order" in art is at odds with what we would normally refer to as "Romanticism", that man is essentially a free individual and that art should strive to liberate his full creative energy, even if this might conflict with the needs of society. Ruskin clearly identifies with certain aspects of the Romantic movement, but is skeptical concerning the question of man's liberty. He argues that "liberty does not exist in man, nor in the stars, or the moon, or the sun". He adds, in a footnote to Coleridge's "Ode to Fame", that the verse is noble, but the "sentiment of freedom is wrong".

However, I believe that Ruskin's analogy between artistic creativity and language is wrong. Indeed, the very purpose of this book is partly to argue that original artistic expression derives from the ability to identify certain types of fundamental aesthetic imagery, this has nothing to do with language. We will see later in the book that great art is often produced by individuals who break radically with the social conventions of their day, certainly the major periods of Gothic art, in England, that we will consider in this chapter, could be said to have arisen through such individual efforts. As far as Ruskin's analogy between man and natural phenomenon, with regard to liberty, I also believe this to be flawed, at least to the extent that Ruskin intends to support his essentially social argument. On these grounds, I would disagree with the conclusion that Ruskin draws. Obedience to God serves a spiritual purpose but blind obedience of the individual to society is harmful to both. I believe that only a dynamic, harmonious relationship between man and society is useful, a question I will return to later in the book, when I consider the aesthetics of harmony.

As I mentioned before, Ferrari's work is, to an extent, typical of Gothic art. It is interesting, therefore, to consider what Ruskin writes about "The Nature of Gothic" (Part 2, Chapter 4 of Stones of Venice). He, in fact, divides the "Gothic spirit" into six aesthetic categories;

(i). Savageness.

(ii). Changefulness.

(iii). Naturalism.

(iv). Grotesqueness.

(v). Rigidity.

(vi). Redundance.

The meanings of these terms are fairly self-explanatory, except perhaps for the last two. By rigidity, Ruskin means the independence of control or submission on the part of the artist, an aesthetic which he criticized in The Lamp of Obedience. By redundance, he means the bestowal of the wealth of labour on the part of the artist, an aesthetic which he eulogized in The Lamp of Sacrifice. Certainly, these aesthetics and those of savageness, naturalism and grotesqueness could all be applied to Ferrari's work. This reinforces the idea that Ferrari's art is representative of Gothic. However, in Ferrari's work, none of these aesthetics are used for spiritual effect, and this, for me, is part of what constitutes great English Gothic architecture. In order to understand this distinction more clearly, I wish to consider in more detail the aesthetics of spirituality and fragmentation which we observed in the work of Tintoretto and Turner.

I have already spoken of the importance of spirituality in art. This was far more true of the Medieval age than of today, and it is impossible to assess the significance of art of that period, without taking this into account. In the introduction of (Martingdale, 1985), Andrew Martingdale observes that the Scholastic view of the importance of non-material truth is a defining characteristic of what we should consider to be "Gothic art". He gives the following quotation from the ninth-century scholar John Scotus Erigena;

"We understand a piece of wood or stone only when we see God in it"

In other words, a thing is of true value, only by revealing a divine nature.

We have already touched on the importance of spirituality in Ruskin's criticism of art, the following quote from the Introduction of "Seven Lamps of Architecture" reinforces this view;

"The truth, decision and temperance, which we reverently regard as honourable conditions of the spiritual being, have a representative or derivative influence on the works of the hand, the movements of the frame, and the action of the intellect"

For Ruskin, the aesthetic of spirituality is intimately related to naturalism. The natural world reflects the hand of a divine creator and, through careful study of it, the artist is able to convey spirituality in his work. Perhaps the most detailed account of Ruskin's naturalism can be found in "Modern Painters", where he expounds an aesthetic theory based on the observation of certain natural phenomena. This was the principal basis for his defence of much of Turner's landscape painting. Naturalism is also an important criteria in assessing architecture. In Chapter 3, "The Six Divisions of Architecture" of "Stones of Venice", the wall of a building represents the first division in the assessment of its value. He uses the analogy of a flower with its root, stalk and bell, in order to support his further division of the wall into its base, veil and cornice. In Chapter 5, "The Wall Veil", he cites the Matterhorn as the most important and inspiring example of a natural wall. For Ruskin, then, a wall veil in architecture should reflect the simplicity and mass of such a structure. In his analysis of the arch, Chapter 7, "The Arch", which we will return to later, he judges that the best arches are those which in some way conform to God's arches, those of the rainbow and the rising sun. Indeed, the reader should look at his original diagrams of arches in this chapter, which are all inscribed in these archetypal natural arches. The most detailed account of his naturalistic aesthetics can be found in Chapter 9, "The Material of Ornament", where he discusses the importance of ornamentation in architecture. He claims that;

"All noble ornamentation is the expression of man's delight in God's work"

His assessment of the value of ornament is based on its conformity with his list of certain natural organic and inorganic forms;

1. Abstract Lines.
2. Forms of Earth (Crystals).
3. Forms of Water (Waves).
4. Forms of Fire (Flames and Rays).
5. Forms of Air (Clouds).
6. Organic Forms (Shells).
7. Fish.
8. Reptile and Insects.
9. Vegetation(A) (Stems and Trunks).
10. Vegetation(B) (Foliage).
11. Birds.
12. Mammalian (Animals and Man).

Ruskin goes on to cite a number of examples where these forms are used. Of particular interest here are the comments he makes in relation to the forms 1 and 10, ([1]). With regard to abstract lines, he gives a naturalistic example of what he considers to be a "perfect" line, namely the curve of the Chamounix glacier, the reader should look at the sketch he gives in this chapter of his book. It is an important feature of his philosophy that his appraisal of this abstract form is based on natural rather than other aesthetic or geometric criteria. With regard to the use of foliage, Ruskin makes a number of observations about leaves, "they guide themselves by the sense of each other's remote presence and by a watchful penetration of leafy purpose in the far future",

[1] Ruskin observes, in relation to 12., that the over-reliance on the human form as ornament signals the decline of architecture, which is, of course, an indirect attack on pure classicism.

praising their use in ornament, on naturalistic grounds. We will see later in this chapter the recurrence of foliage forms in English Gothic architecture. An archetypal example of this form is the trefoil;

6. Tomb of Pope Adrian V by Arnolfo di Cambio, St. Francis church, Viterbo, Italy.

It is partly the naturalistic interpretation of this element that motivated Ruskin to write about Gothic in "The Seven Lamps of Architecture";

"If I were asked what was the most distinctive feature of its perfect style, I would say the trefoil. It is the very soul of it, and I think the loveliest Gothic is always formed upon simple and bold tracings of it, taking place between the blank lancet arch, on the one hand, and the overcharged cinquefoiled arch on the other", ([2]).

We should now consider Ruskin's views on the aesthetic of fragmentation. The most obvious example of the use of fragmentation in Gothic architecture is the pointed arch. The question of how the pointed arch form originated in the West is an interesting one, and we will consider it later in the chapter. The pointed arch takes a classical form, namely the rounded arch, and breaks the curve into two distinct pieces , see figure 37, thus creating a node, a simple type of singularity, at the head of the arch. On an aesthetic level, the eye is no longer able to follow a smooth path along the top of the arch, but is, instead, interrupted. This induces a feeling of distortion, and increases the spiritual sense of the observer,([3]). I leave it to the reader's own judgment, to decide whether this description is accurate. Ruskin writes extensively about the evolution and importance of the pointed arch form in "Stones of Venice". In the Introduction, he credits the Arabs, working in Venice, with the first use of the pointed arch form;

"In his intense love of excitement, he points the arch"

This occurs, according to Ruskin, around 800, when the Ducal residence moved to Venice, the arch form, at this stage, being almost identical to one used in Cairo. However, in his chapter on The Nature of Gothic, Ruskin credits the systematic use of the pointed arch form to Gothic architecture;

"Gothic architecture is that which uses, if possible, the pointed arch in the roof proper."

Indeed, in the same chapter, he takes the pointed arch form to be a defining characteristic of such architecture;

"Gothic is nothing more than the development of the group formed by the pointed arch for the bearing line below, and the gable for the protecting line above"

In Chapter 7, The Arch, Ruskin gives a functionalist analysis of the arch form, by which, I mean that he is primarily concerned with a physical justification of its properties. He begins by considering certain forms which fail this test of functionality, such as the reversed arch, which is the same as the pointed arch form, with the modification that the sides of the arch are inverted. Ruskin concludes;

"The natural tendency of such an arch to dissolution by its own weight renders it a feature of detestable ugliness, wherever it occurs on a large scale"

---

[2] For the reader unfamiliar with the terminology lancet, see later in this chapter. Ruskin is also observing a numerical harmony of 1,3 and 5 in his description of the trefoil with the lancet and cinquefoil.

[3] The question of the exact construction of the pointed arch is interesting. According to Jenny Gage, see Bridges Leeuwarden Proceedings, 2008, the sides of the arch are constructed as circle segments, centred at the base vertices of the arch. However, as will become clearer later in the book, this interpretation partly misunderstands at least the Norman construction of such arches, for example at Monreale, and their conception of both the aesthetics of light and linearity.

In Part 2, Chapter 4, The Nature of Gothic, Ruskin uses the analogy with a pair of playing cards, balanced against each other, to further support this argument. Pressure applied to the top of the cards leads to a gradual buckling in the centre, hence, this is where the vertical force is concentrated. The reversed arch form is thus contributing to its own demise. Ruskin pursues this analogy further, by concluding that a pointed arch form, with 'cusps' at the centre points of the sides of the arch, is structurally the most sound. The reason being that the 'cusps' reinforce the arch at the places where it is most liable to fail, ([4]). It is interesting to note that such an arch form is very similar in shape to the trefoil form that we considered earlier and, indeed, Ruskin cites this form as the origin of foliated tracery. Hence, in Ruskin's analysis, both his aesthetic and functionalist analysis arrive at the same conclusions. Although, I do not entirely agree with Ruskin's functional analysis, I will give my own analysis of the simple broken arch form in the next chapter, the way in which his functional and aesthetic argument run in parallel is striking. I believe that this close association between aesthetics and geometry is indeed a feature of all great art and architecture. In particular, we will see in the next chapter how the aesthetic of fragmentation has a striking geometric interpretation.

The Aesthetics of the Sublime and Fragmentation in Medieval Art and Architecture.

Having considered briefly some of the terminology associated to Gothic architecture, it is time to consider the evolution of the aesthetic ideas introduced in this chapter, in more detail, within the history of medieval architecture in England. There is no definitive chronology of such styles. Arguably, Milner's "Treatise on Ecclesiastical Architecture of England during the Middle Ages" (1811), (Milner, 1811), is the first attempt to divide English Gothic into three orders of pointed architecture, later refined in Rickmann's "Attempt to Discriminate the Styles of English Architecture" (1819), (Rickmann A. , 1819) as Early English, Decorated English and Perpendicular English. Incidentally, Milner was also the first Englishman to argue that the pointed arch was the fundamental element of Gothic architecture. In spite of this terminology, the terms Lancet and Curvilinear, to describe the Early English and Decorated English styles respectively, may also be found in the literature, and the origins of these terms is older. As they are also of a more aesthetic nature, describing the geometric shape of a window and a style of tracery, I prefer to use them instead of Rickmann's descriptions, although I will occasionally also use the alternatives. The Norman style of architecture predates what is usually considered to be mainstream Gothic, but, as we shall see, presents the earliest examples of the pointed arch form. The style, Response to Geometric, is my own terminology, though occasionally found in the literature, and denotes the reaction in England to the Geometric style, developed in France. The dates of these respective periods are, approximately, as follows;

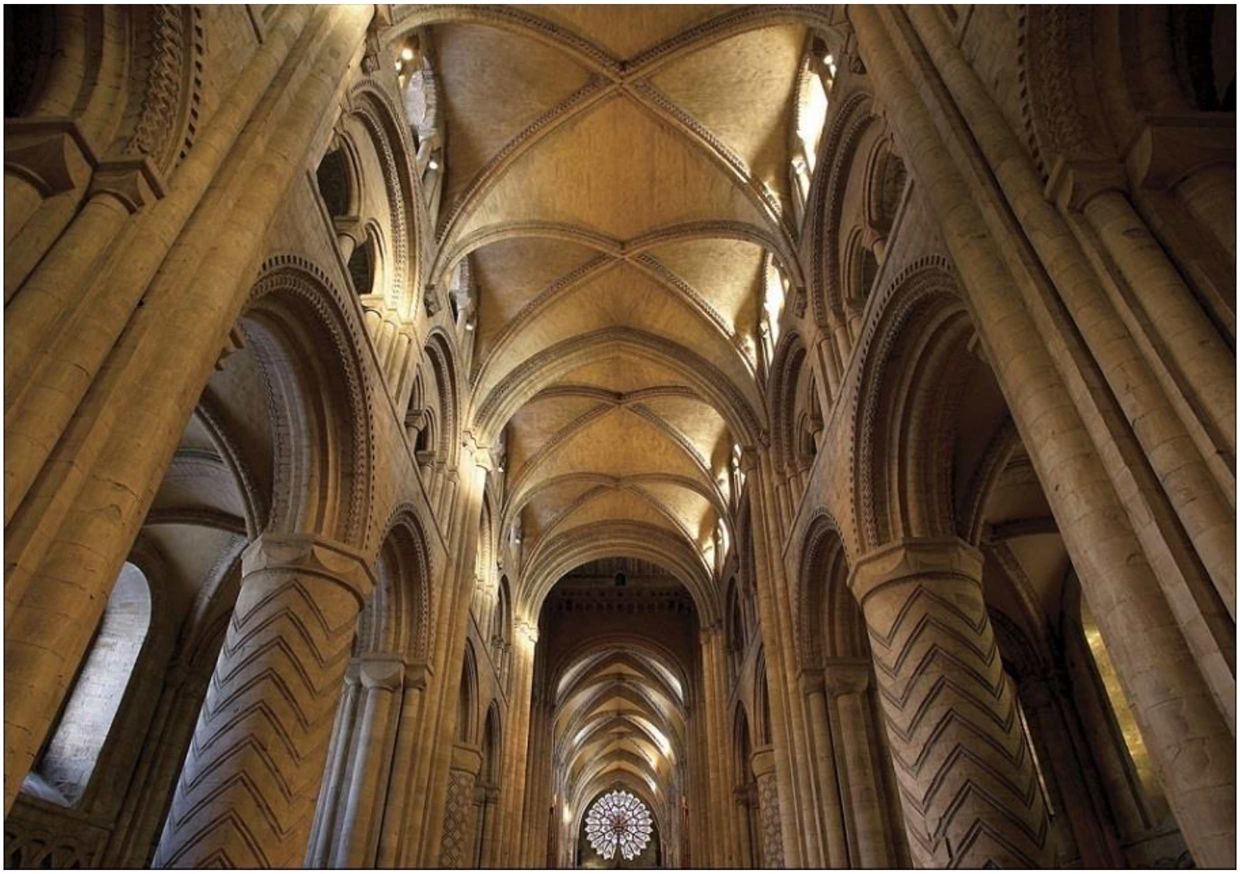
Norman	(c1066-1180)
Lancet (Early English)	(c1180-1275)
Response to Geometric	(c1240-1280)
Curvilinear (Decorated)	(c1275-1380)
Perpendicular	(c1380-1520)

The Norman period of architecture in England begins with the Norman conquest of 1066 and lasts roughly until the end of the Norman dynasty. Within that period, a number of English cathedrals were constructed, although only Durham cathedral survives relatively intact. One of the distinctive features of Norman

[4] Ruskin's use of the word cusp is geometrically misleading, the proper terminology is node, see figure 37.

architecture is the use of large circular columns in the nave of cathedrals, often with a distinctive chevron moulding. This can still be seen in, for example, the naves of Tewkesbury Abbey or Durham Cathedral,

7. Nave of Tewkesbury Abbey, Tewkesbury.

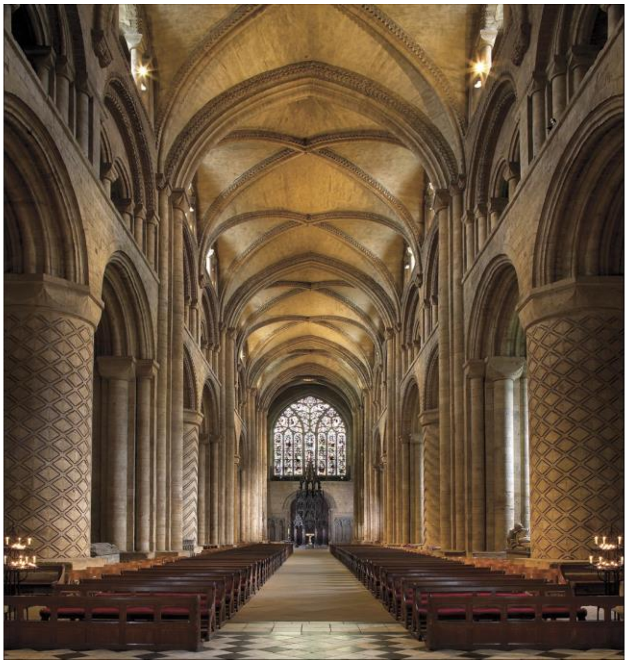


8. Nave of Durham Cathedral, Durham.

At first, the Norman style developed the antique basilica scheme of construction, using circular headed arches in the bays of cathedrals and groin vaults, which are essentially circular arches supporting the entire vault or roof over the nave. However, by 1100, the Normans had made the first innovation in architecture which could be considered to belong to the aesthetic of fragmentation, namely the development of the rib vault, in which crisscrossing beams or stone ribs replaced the previous simpler groin design. Unfortunately, very few examples of such vaults survive. Some notable exceptions are the crossing vault of Worcester Cathedral;

9. Crossing Vault of Worcester Cathedral, Worcester.

c1140, the north choir aisle of Durham Cathedral, built between 1093 and 1096, Durham Cathedral nave vault, built between 1115 and 1130, and the vault of Holy Cross church, Avening, 1105;




10. Nave of Durham Cathedral, Durham, and vault of Holy Cross church, Avening, Gloucestershire.

The linearity of the chevron design and the use of fragmentation in the geometry of the vault are both indications of a more masculine aesthetic style. In some sense, this could be considered as an early indication of the Gothic style, even before the appearance of the pointed arch form. The new aesthetic thinking of the Normans was also accompanied by innovations in the engineering of cathedrals and churches. In "The Gothic Cathedral",((Wilson, 1992), Chris Wilson explains how Norman architects were the first to use large amounts of rubble above the springings of the vault, in order to resolve the thrust lines vertically. This system can still be found in Durham Cathedral and Tewkesbury Abbey. At Durham Cathedral, one finds a further technical innovation, the use of quadrant arches in the nave galleries, in order to reinforce the lateral thrust of the vault, these are precursors to the later use of flying buttresses. As I hope will become clear by the end of the next chapter, this advancement in the aesthetic thinking of the Norman period is accompanied by genuine advances in engineering and functionality, as, for example, in the construction of both the nave and tower at Avening, ([5]).

11. Holy Cross church, Avening, Gloucestershire.

[5] Although, in the author's opinion, the current vault, located directly underneath the tower, has been considerably altered.

The question of whether English Norman architecture is responsible for the first major appearance of the pointed arch form in Western medieval architecture is debatable. The abbey church of Cluny in France, head of the Benedictine order, and built from around 1088, was demolished in the 19th century, but a surviving copy at Paray-le-Monial, built from around 1110, demonstrates the use of sharply pointed arches in the bays. This suggests that the pointed arch form had appeared in France, before its use in Durham cathedral. However, it is certain that the earliest pointed rib vault is the high nave vault of Durham cathedral. As I will explain in the next chapter, the use of a pointed, as opposed to a circular, vault system has a number of structural advantages, apart from aesthetic considerations. For this reason alone, the construction of the Durham vault represents a genuine innovation in architectural thinking. Moreover, the architects of the Durham vault had previously adopted a circular system for the construction of the choir, also built after Cluny, as if coming to their own realization of the advantages of such a system. What is most interesting about the, probably independent, constructions of Durham and Cluny is the way in which a simultaneous advance in both aesthetic and geometric thinking occurs. This close interaction between seemingly distinct disciplines is, in my opinion, essential for such breakthroughs in human development.

It is impossible to examine later periods of English medieval architecture, without, first, briefly considering the development of Early Gothic architecture in France. This period, which dates from approximately 1140 to 1195, is marked, at its beginning by the construction of the ambulatory of Saint Denis, and ends with the construction of the Cathedral of Chartres, ushering in a new period of Gothic architecture in France, which I refer to as Geometric. The ambulatory of Saint Denis is, without doubt, one of the most interesting of medieval constructions;

12. Ambulatory of Saint Denis, Paris, France.

Built under the supervision of the enigmatic Abbot Sugar, between 1140 and 1144, it is unlike any other previous medieval designs. The most striking feature of the work is the use of slender columns, rather than compound piers, to support the surrounding superstructure. Technically, this is made possible by the use of false bearing. The vault is not actually supported by the lower storey columns, but from a middle unseen storey. Aesthetically, this has the effect of creating a sense of lightness and verticality in the observer. As we observed earlier, this aesthetic is intimately connected with that of fragmentation and enhances the spiritual effect on the observer. The aesthetic of fragmentation is also conveyed by the arrangement of the ambulatory chapels, although here we have a fragmentation of space rather than line. Chris Wilson argues convincingly that the use of false bearing suggests that flying buttresses were employed in the construction, even though it was previously accepted that no flying buttresses appeared in France before 1175. His argument relies on the physical requirement of supporting the vault from a middle storey and from the deep buttress projections, which can still be seen, emanating from the outside of the chapel.

It seems clear from consideration of the Norman innovations, particularly at Durham, that the technical breakthroughs of flying buttresses, ribbed and pointed vaults had already been made before the construction of Saint Denis. Abbot Sugar, undoubtedly aware of these developments, seems to have drawn on these achievements, in order to further refine the aesthetics of ascent and fragmentation and light. The remainder of the period of Early Gothic architecture in France may be considered as a time in which the potentiality of these forms was fully explored. In particular, it saw the development of the use of pointed arches, for example at Fonteney, built by the Cistercians between 1139 and 1147. The verticality effect of false bearing was introduced at Laon Cathedral, built between 1165 and 1175, and later developed at Chatres. This aesthetic of ascent is also reflected in the increasing height of the vaults of cathedrals during this period and the fashion for 4-storeyed structures, such as Laon or Notre Dame, as opposed to the traditional 2-storeyed or 3-storeyed cathedrals. Flying buttress systems also became more advanced and widespread, for example at St.Remi in Reims, built between 1170 and 1180, which still exhibits the system that became the standard model for cathedrals all over Europe, namely 2 flyers supporting the second and third storey of the cathedral. A good example can be found in Limoges, begun in 1273.

13. Naves of Reims Cathedral, Reims, and Limoges Cathedral, Limoges.

As Chris Wilson explains in his book, the use of two flyers in buttress systems still remains somewhat of a mystery. It is not clear, even today, how medieval architects perceived the flyers to counteract the main thrust of the vault. However, as I will explain more fully in the following chapter, the two-tiered design of the flying buttress has a geometric explanation, when used in conjunction with a pointed vault. Namely, by breaking the arch of the vault, the thrust lines of the vault are fragmented along the two directions determined by continuation of the sides of the vault arches, see figures 9 and 10. The buttress system is then designed to counteract these two distinct thrust lines. The structural advantages of such a system are clear, the vault thrust is effectively halved and better distributed along the walls of the building. In a circular arch system, the thrust of the vault is concentrated along a single continuous line, making it structurally weaker than the pointed arch system, and more liable to fail, as indeed occurred with many circular Norman vaults, for example, at Tewkesbury Abbey. I am certain that the Norman architects of Durham cathedral, intuitively, understood this geometric picture, and this makes their innovation all the more impressive.

In Chapter 9 of Stones of Venice, The Buttress, Ruskin gives a functional analysis of the buttress system and makes some interesting observations;

"It is, however, very seldom that lateral force in architecture is equally distributed. In most cases, the weight of the roof, or the force of any lateral thrust, is more or less confined to certain points and directions."

"The weight to be borne may be considered as the shock of an electric fluid, which by a hundred different rods and channels, is divided and carried away into the ground"

Ruskin is supporting the idea that the thrust of a vault is concentrated along certain lines of force. This picture makes the problem of supporting vaults a highly geometric one. He is also suggesting that these lines of force may be divided, which enforces the idea that the aesthetic of fragmentation was the guiding principle, leading to the introduction of pointed vaults as a means of solving this geometric problem.

In Otto von Simpson's book, also entitled the "The Gothic Cathedral", (Simpson, 1956), the interesting question of the close relationship between the form and function of Gothic architecture is again discussed. In his opening chapter on Gothic Form, he argues that this relationship between function and form is two way and ambiguous. He supports this conclusion by consideration of the cross rib vault, which, as we observed earlier, makes its first appearances in English Norman architecture, even before the construction of Durham Cathedral. On the one hand, he supports the point I made earlier that the underlying aesthetic of Norman architecture is linear and presents an argument of Bilson's, from "Norman Vaulting in England", (Bilson, 1899);

"The cross rib vault was preceded by the architect's inclination to conduct the ridges, not as the interpretation of curved surfaces, but as the intersection of straight lines"

In other words, as von Simpson explains, the cross rib vault is created by the geometrical "graphism" of Norman design. According to this argument, considerations of form come before considerations of function. On the other hand, von Simpson considers an argument by the writer Bony, writing on the configuration of lines in Gothic architecture;

"Their design transcribes, with some freedom of interpretation, what is going on behind them and expresses what was believed by the architects to be the theoretical framework of the building."

This is the functionalist argument, that considerations of function come before form.

In my opinion, as I have already explained, aesthetic considerations governed the appearance of the cross rib vault. This seems clear from the fact that, even at the time, a cross rib circular vault served no obvious functional purpose, ([6]). The same could be said about the appearance of chevron symbols on Norman piers. However, I believe that Bony is right to the extent that, at a later period, architects had an intuitive geometric understanding of the way in which fragmented vaults could support a building more effectively. This functional advance in thinking seems to have occurred after the construction of the Durham nave vault. I will consider Bony's arguments on vault fragmentation later, in relation to the Curvilinear period in England, where, conversely, geometric, and, therefore, functional considerations refine the aesthetic of fragmentation to even greater levels.

[6] Indeed, the north choir aisle of Durham cathedral, built according to this system, is supposed to have developed cracks, prior to the adoption of the pointed vault.

Von Simpson makes an important observation later about the simultaneous adoption of the pointed arch form at Durham and the arrival of Cistercianism in England, around 1128. This might seem to resolve the perceived ambiguity of form and function which he discovered earlier. The Cistercians, headed at the time by Bernard de Clairvaux, believed that architectural postulates should conform with the ascetic ideals of monasticism. As we found before, the pointed arch induces a spiritual effect on the observer, and, in this sense, conforms with these ideals. However, although, as we shall see, the Cistercian order is responsible for the later dissemination of the pointed arch form through many parts of Europe, it seems highly unlikely that they were solely responsible for the introduction of the form at Durham. The aesthetic of the pointed arch form was certainly known before, although the Cistercians may well have played a part in renewing interest in it. Also, the technical considerations of finding an improved method of supporting the vault suggest that this was also a geometric innovation, as opposed to being a purely aesthetic or spiritual one. Nevertheless, it is impossible to imagine the advances of this period without the aesthetic background which the heightened spirituality of Christianity provided. The image of the Crucifixion, central to the mindset of the period, lies behind every development of the aesthetic of fragmentation. It is this harmonious relationship between geometry, aesthetics and spirituality which makes the period so innovatory and fascinating.

We can now move on, in order to consider the Lancet or Early English period of architecture. As mentioned earlier, the term Lancet originates as a description of the narrow, elongated windows which were a feature of this period;

14. The Five Sisters, York Cathedral, York.

It is impossible to assess the origin of this type of window. As with the pointed arch, it seems likely that the form appeared occasionally before, but its potential as a architectural form was, at this time, fully explored. The aesthetic of the lancet form is that of ascent, the eye is drawn upwards along the vertical lines of the window. When considered in conjunction with the refinement of this aesthetic that had begun earlier in France, it seems clear that such windows, as with the pointed arch, were deliberately used as a means of heightening the spiritual sense of the observer. Unlike constructions such as the pointed arch, however, there is no sense of fragmentation, the eye is encouraged to find a distant but unattainable point at infinity, in this sense, the aesthetic belongs to that of the sublime.

Some of the most spiritual examples of the use of this aesthetic in the form of the Lancet window can be found in the very earliest constructions of this period. The Abbeys of Byland, Rievaulx and Whitby are all within close proximity, in the romantic setting of the Yorkshire dales. Although all three are now in ruins, as a consequence of Henry VIII's brutal suppression of the monasteries in England, many of the original facades survive. Seen in these simple contexts, the geometry of the lancet window is in fact enhanced, I leave it to the reader to decide, whether, even in the form of the reproduced images here, he can experience a sense of serene spirituality.

Rievaulx Abbey was founded in 1131 and was the first Cistercian community in the north of England. The nave of the earliest church, constructed in the 1140's, is now completely ruined, but the presbytery, rebuilt in the 1220's, is one of the finest examples of Lancet architecture in England. It is characterized by an increased 3-storey elevation and numerous lancet windows, deliberately conveying a sense of sublimeness and ascent, and, probably, influenced by the previous innovations of French architecture. The architectural forms of the arcades and windows are enriched by elegant mouldings and adorned with naturalistic details, while the piers are surrounded by highly clustered shafts. These additions, as we will explain below, distinguish the style as English. Lancet windows can also be found in the south wall of the refectory of the Abbey, dating from as early as the 1180's, the very beginning of the Lancet period. Their purely functional use here suggests that the potential of the lancet form as an aesthetic device was realized later, it is unclear whether this realization occurred in England or France, see (Fergusson, 2010).

15. Facade of Rievaulx Abbey, Yorkshire.

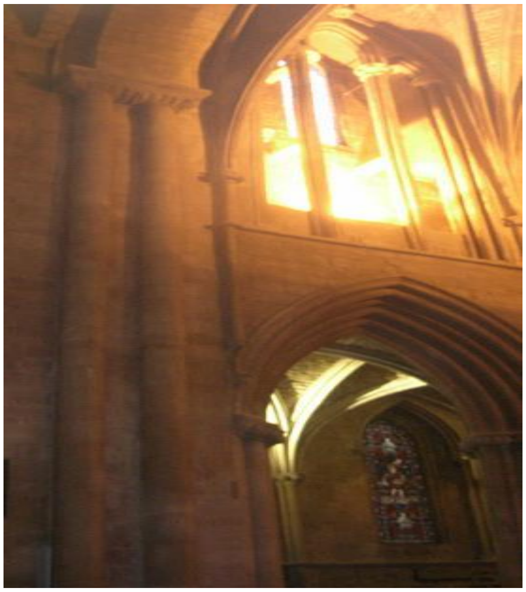
The aesthetics of ascent and linearity, at this watershed period between the Norman and Lancet periods, are combined successfully in the use of highly verticalized shafts, as at Pershore Abbey, precedents of which can be traced to the use of slender corner capitals, combined with cross rib vaulting, as at Avening church, from about 1105, see (N.M Herbert, 1976), Coln St. Denis, Garway and St. German;

16. Holy Cross church, Avening, Gloucestershire.

17. St. James the Great church, Coln St. Denis, Gloucestershire.

18. St. Michael's church, Garway, Herefordshire.

19. St. German's church, St. German, Devon.



20. Shafts at Pershore Abbey, Worcestershire.

Whitby Abbey, situated dramatically on a cliff overlooking the sea, is probably most famous as one of the settings in Dracula. Coming ashore in the form of a dog, during a storm, he attacks his first victim, Lucy Westenra, on a bench by the church. In many ways, Whitby does conform to this stereotyped image of the English Gothic novel. The abbey, in fact, dates from 657, when it was founded by the Northumbrian king Oswy. Although this building was completely destroyed by the Danes in 867, it was rebuilt as a monastery in the late 11th century. By the early 13th century, this building was deemed old fashioned and, in the 1220's, work began on the present choir, characteristic of the Early English style, in its profusion of carved detail and mouldings. The transepts and crossing, begun in the 1250's, are characterized by three levels of lancet windows. The view through the windows on one side, over the North sea, adds to the aesthetic of sublimeness, directing the eye to an unattainable point on the horizon. Perhaps one could feel this more powerfully in the days when the monastery was protected from the harsh, stormy sea winds. For more information about Whitby, see (Goodall, 2003).

21. Facade of Whitby Abbey, Yorkshire.

Byland Abbey was originally founded as a Savagniac house in 1134, but, by 1177, was under the control of the Cistercian order. It is around this time, at the start of the Lancet period, that much of the essential building of the monastery was begun. The completed monastery of the early 13th century, now in ruins, was considered to be one of the finest in Northern England. As at Rievaulx, Byland is anglicized by the use of deeply profiled arches. However, perhaps the most striking feature of the ruined abbey is the large circular window of the facade, a feature which cannot be found at Rievaulx or Whitby. We will discuss this innovation in greater detail in a later chapter.

22. Facade of Byland Abbey, Yorkshire.

The beginning of the Lancet period also saw a great development in cathedrals across England. In 1170, Thomas Becket was murdered at Canterbury Cathedral, and shortly afterwards, in 1174, the eastern part of the Cathedral was consumed by fire. Between 1175 and 1185, the cathedral was rebuilt by William of Sens, and a corona constructed for Becket's burial. At about the same time, between 1176 and 1186, work began on the construction of Wells Cathedral, and by 1215, much of the interior including the nave, transepts and north porch had been completed. According to Clifton Taylor's book (Taylor, 1967), The Cathedrals of England, Wells was the first building in England to use the pointed arch form throughout. Clifton Taylor argues that Wells may be considered as the first truly English Gothic church, possibly with the exception of some of the Cistercian abbeys in Yorkshire, in that, it owes almost nothing to earlier French constructions. On an aesthetic level, a visit to Wells will convince the reader that this is true, ([7]). First, there is no sense of verticality or ascent, a feature which we noted in earlier French Gothic. This is partly due to the size of Wells, small compared to other cathedrals in France, and also the absence of any false bearing. Instead, a sense of spirituality is conveyed by the aesthetic of naturalism, which we encountered in Ruskin's criticism. The piers of the nave in Wells are enriched with distinctive "stiff-leaf capitals". This is an architectural ornament consisting of projecting, three-leaf, foliage at the head of the column. Although three-leaf foliage on capitals may also be found in France, for example at Reims cathedral, at around the same time, the design tends to be flatter, with more emphasis on the geometric outline of the figure, see "A Manual of Historic Ornament" by R. Glazier (Glazier, 1889). In the foliage at Wells, this aesthetic is suppressed by the representation of the leaf in three dimensions. As we observed earlier, foliage appears in Ruskin's list of organic forms and is important in his consideration of desirable aesthetic forms. I think it would be fair to say, that, in comparison with the form at Reims, the stiff-leaf form of foliage conforms more closely to the naturalistic aesthetic on which Ruskin's valuation of ornament depends, mainly because foliage is not found 'flattened' in nature, ([8]). In this sense, the adoption of "stiff-leaf capitals" may be considered a distinctively English innovation. The piers of Wells are also deeply grooved, again conveying a similar aesthetic to the features at Byland, Rievaulx and Whitby, that of an enrichment of the wall surface, which seems to enhance the naturalism of the capitals.

Lincoln Cathedral was also an early creation of the new Lancet style of architecture. The original Norman building was substantially destroyed by an earthquake in 1185 and work began on reconstructing the cathedral in 1192. The renovation introduced several new architectural features, which, at the time of completion, around 1230, became universally accepted in England. One of these is the characteristic wall arcading of the nave aisles. This consists of a series of trefoil arches supported by Purbeck marble colonettes, with a further colonette set within the recess of each archway. The overall effect of the arcading conveys a sense of dissonant rhythm, heightened by the polychromatic tones of the marble. This effect of dissonance is also found in another distinctive feature, the so called "crazy vault", above the nave. Here, the lateral cells of the vault end at different points, some being one third away from the central ridge vault, creating an unusual irregular effect and rhythm. As at Wells, one can also find the same naturalistic effects of stiff-leaf capitals, set at the springing points of the arcade arches, and in the use of the trefoil form in the arcade arches. The aesthetic of dissonance, introduced at Lincoln, is similar to the aesthetic of fragmentation which we saw repeatedly in Norman architecture. In the case of fragmentation, a completely new sense of space and line is being developed, whereas dissonance brings our intuition of time to new psychological levels. In the medieval mindset, demonstrated here, our modern rationalized sense of time as a continuous repeating process is replaced by a naturalistic sense of time as subordinate to other spatial phenomena. Perhaps this underlies the sense of serenity and timelessness that many visitors to Lincoln cathedral feel and led Ruskin to say of Lincoln;

[7] Any visitor to Wells Cathedral should try to ignore the strainer arches, which were added later to the nave and transepts.

[8] As we will see in later chapters, the over-geometrisation of tracery forms, is, for Ruskin, one of the principal causes of the decline in later French Gothic art.

"I have always held and am prepared against all evidence to maintain that the Cathedral of Lincoln is out and out the most precious piece of architecture in the British Isles and roughly speaking worth any two other cathedrals we have."

Worcester Cathedral is probably the leading example of the mature Lancet style of architecture. Although originally a Norman cathedral, the current choir and Eastern transepts were added later, between 1224 and 1240. Here, we see the already familiar architectural devices of lancet windows, stiff leaf carving and mouldings perfected in the arcades of the south-east transept. Moreover, we see two new architectural innovations. The first of these is a two-tiered arcade in the triforium of the eastern transept, with trefoiled arches at the ground level, even more complex than its predecessor at Lincoln. The second is the use of thin black Purbeck marble shafts, which run from the springing level of the pointed arches in the choir, and from floor level in the walls of the south east transept. The effect is, to quote Pevsner's, "The Buildings of England: Worcestershire", one of "unrestrained verticalism". The use of the shafts, in conjunction with lancet windows, is particularly effective, and intentionally emphasizes the aesthetic of the sublime which we observed earlier. The innovation could be said to anticipate the later Perpendicular style of architecture in England, which we will discuss shortly.

Another fine example of the mature Lancet style can be found at York Cathedral. The north transept, which was completed in 1255, contains the tallest medieval lancet windows in England, known as the Five Sisters. The size and multiplicity of the windows adds to their overall sublime aesthetic effect.

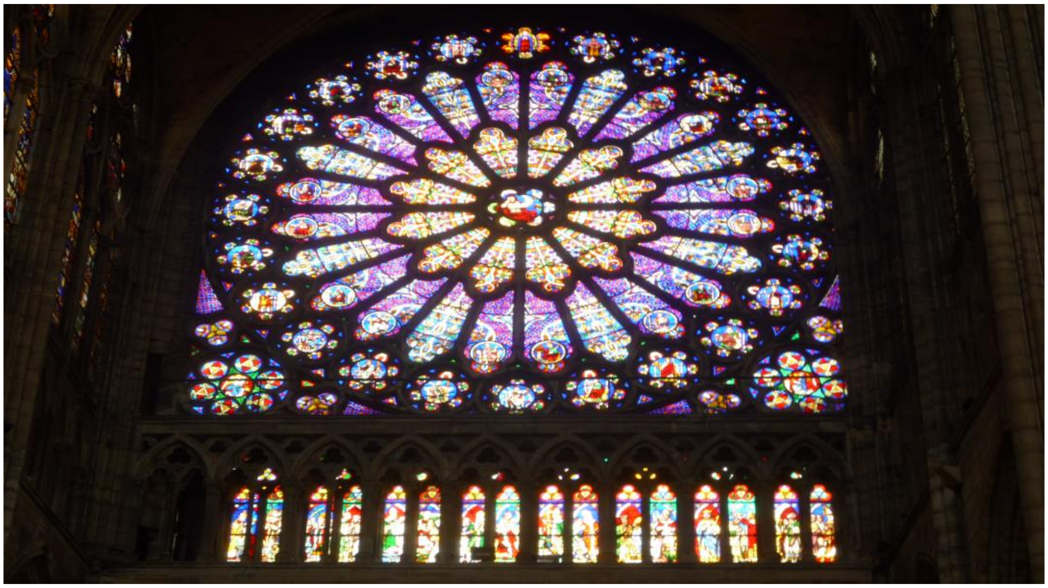
In order to consider the later styles of English medieval architecture, we must again briefly look at developments in style that occurred in France. These developments belong to the Geometric period of French architecture, also known as High Gothic, and are marked by the construction of the cathedrals of Chartres and Reims. We will consider the innovations that occurred here in greater depth later in the book, confining our attention here to the features of interest in England. The reconstruction of Chartres cathedral began in 1194, after an accidental fire destroyed all of the building, leaving only the facade and towers intact, according to (Arthaud, 1962). The nave, completed by 1220, is, undoubtedly, a masterpiece of French Gothic architecture. It demonstrates many of the aesthetic innovations achieved in the earlier period of Saint Denis as well as that of the Lancet period. The most obvious of these are the considerable height of the vault and the tall lancet windows in the clearstorey, both of which enhance the aesthetics of verticality and ascent. This verticality is further emphasized by removing the capitals of the vault responds, an idea which we saw used in a different form, in the marble columns at Worcester. The tremendous high vault thrusts are met by an unprecedented array of three-tiered flying buttresses, developing the previous technical advancements at St. Remi. The breakthrough that distinguishes the style as Geometric involves the use of cusped oculi in the clearstorey above the lancets. These introduce a new aesthetic of light which we will consider in a later chapter, and anticipate the development of the Rayonnant style. The cathedral of Reims was begun at the later date of 1210 and took an extraordinary 80 years to build. Again, we see the incorporation of earlier developments in the French style, but the novel feature, which places it in the category of Geometric, is the use of cusped oculi, as at Chartres, and the development of bar tracery. The clearstorey is treated exactly as the name should suggest, as a void space, into which bar-like mullions are inserted to form the windows. This advance, of course, depended critically on the development of flying buttresses, as a means of supporting the thrust of the high vault, a 2-tiered system, no doubt influenced by the one used at the nearby St. Remi, being employed, see (Demouy, 1995). One should also mention the cathedral of Amiens, begun in 1220, in the development of the Geometric style, which employed the first 4-light tracery in Gothic architecture.

The response to the Geometric style in England can be found in the Angel Choir of Lincoln Cathedral, begun in 1256.

23. Angel Choir of Lincoln Cathedral, Lincoln.

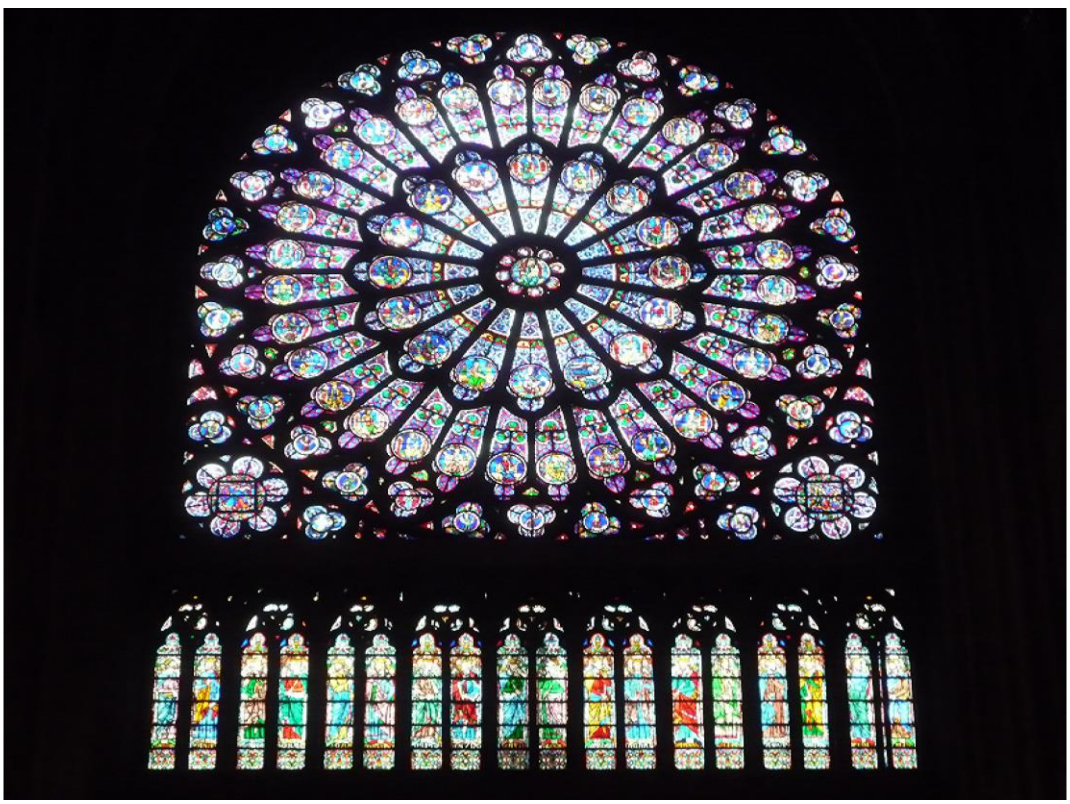
The geometric style of the windows at Chartres and Reims is used, with a profusion of 4-light oculi windows, and also some of the earliest examples of 8-light oculi windows. However, in terms of developing the aesthetic of light, which was done with great effect at Chartres and Reims, the clearstorey composition of the choir at Lincoln is a failure. The reason for this is that the heavy window mouldings, created by the English style, succeed in obstructing the exterior light. The resistance in the English mentality to the development of this aesthetic in France is very characteristic, and, as we shall see, manifests itself in a different form in the reaction to the later French Rayonnant style.

It is the French Rayonnant style that we will now briefly consider, in order to assess the subsequent reaction in England. The term "rayonnant" translates from French as radiating, and was adopted in the 19th century as a means of describing the great rose windows, characteristic of that period, in which a series of radial spokes emanate from the centre of the window. The first of these was created during the reconstruction of the burial church at Saint Denis in 1231



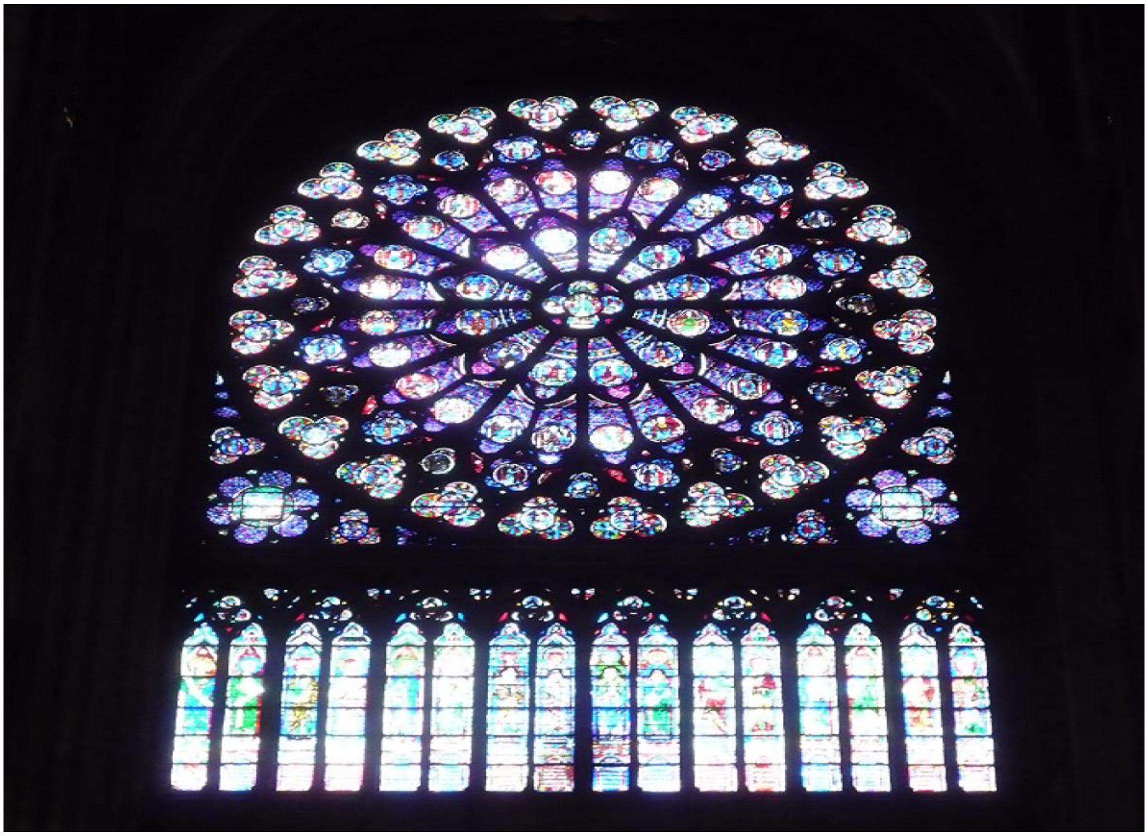
24. Rose Window of Saint Denis, Paris, France.

. A number of further examples were constructed in the area of Ile de France, during the reign of Louis IX, between 1226 and 1270. Most notable of these are the north facade rose window of Notre Dame, constructed in 1250,



25. North Rose Window of Notres Dame Cathedral, Paris, France.

the south facade window of Notre Dame, constructed in 1258,



26. South Rose window of Notres Dame Cathedral, Paris, France.

the enormous rose window of Sainte-Chapelle, built to contain Louis IX's relics, constructed between 1240 and 1248, the rose windows of Carcasonne cathedral, built in 1269, and of St. Urbain, Troyes, built in 1262. As we mentioned before, in relation to the Geometric style, the construction of such enormous windows was made possible by the technical advancements in flying buttresses. However, unlike the purer Geometric style, the rayonnant window suppresses the aesthetic of light, in favour of the radiating geometry of the tracery itself. This geometry has the effect of directing the eye to the central point of the window, rather than considering the aperture as a unified whole. For Ruskin, this was the beginning of the decline of Gothic art, but, we shall consider this question at greater detail later.

The English reaction to the Rayonnant style begins a new chapter in the history of its medieval architecture, namely that of the Curvilinear period. Between 1245 and 1269, Westminster Abbey was rebuilt under the patronage of King Henry III. The large rose windows of the transepts, constructed after 1253, are modelled on the north rose of Saint Denis, which we mentioned earlier. However, apart from this and other isolated examples, ([9]), the rayonnant style of rose window never caught the English imagination, and, indeed, at Westminster, this motif is at variance with many of the other features of the abbey from the same period. The chapter house, begun in 1246, is a compromise between the purer High Gothic and Rayonnant styles, the windows consisting of sexfoil windows above quatrefoils, which are probably the earliest traceried windows to be found in English architecture, ([10]). The tribune windows of Westminster Abbey, constructed after 1250, are further early experiments with bar tracery, in this case the windows are segmented into equilateral triangles

[9] A large rose window, based on Saint Chapelle, could be found at the old cathedral of Saint Paul's, which was rebuilt in 1258 and eventually destroyed in The Great Fire of London, of 1666.

[10] According to Bony's "The English Decorated Style, Gothic Architecture Transformed 1250-1300", (Bony, 1979), there is some debate as to whether the earliest examples can be found at Binham, but the author is skeptical on this point.

with curving sides, filled with three trefoil-cusped circles of bar tracery. Whether these first designs were English inventions, ([11]), the windows at Westminster Abbey demonstrate how the Rayonnant style began a new interest in tracery patterns. The culmination of these English experiments with geometric designs, marking the Curvilinear period, is one of the most interesting stories in medieval architecture.

In order to understand this development more clearly, we need to make some initial distinctions between different types of tracery design. In the course of this discussion, I will distinguish between three basic forms. The first, geometrical tracery, which is often actually referred to in the literature as the Geometrical Decorated style, ([12]), consisting of a basic vocabulary of circles, trefoils, quatrefoils and higher foils, develops the initial experiments with this form at the Westminster Chapterhouse and in the tribune. The second form, curvilinear tracery, a term also found in the literature, uses the dictionary definition of curvilinear, as characterized by free flowing curves in the tracery design. The third form, linear tracery, is my own terminology, and is characterized by a subtle understanding of the central ideas of this chapter, the aesthetics of the sublime and fragmentation. It is also examined in the superb and radical account of this period by Bony, cited above, which will be my main source for the use of this term.

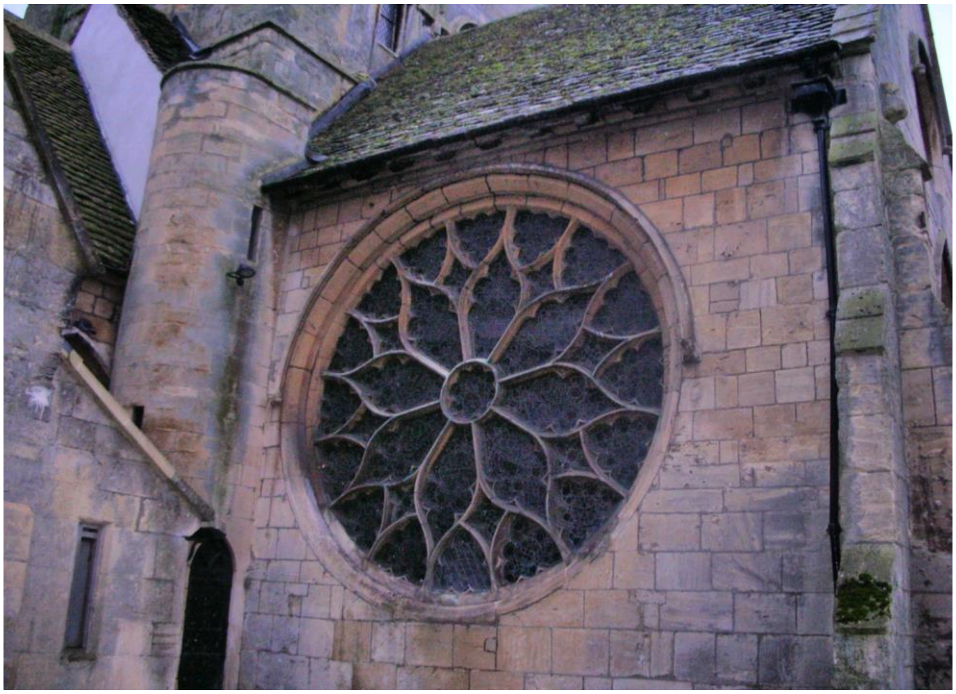
The development of geometric tracery can be found in a number of products of the Early Decorated period, before 1300. The cloisters at Lincoln, often referred to as the finest cloisters in England, were constructed in 1296. Each bay of the cloister consists of two sub-openings, above which are large circles filled with cinquefoils and sexfoils in alternation. At Salisbury chapterhouse, the windows are a simple development of Westminster, namely octofoiled circles placed above quatrefoils. At the chapterhouse of Southwell minster, built in the last decade of the 13th century, one can find an extensive use of trefoil and quatrefoil arches. In all these examples, we see the characteristically English suppression of the aesthetic of light, critical to the original idea of the Geometric style in France. In the cloisters at Lincoln and Salisbury, the main concern is an interest in the geometric design of the apertures themselves, characterized by the large numbers of foils in each opening. In the case of Southwell, we see a return to the naturalistic use of the trefoil form, which appeared at Lincoln and Wells, an aesthetic which is reinforced by the use of extraordinary "crop capitals" over the thirty six chapterhouse stalls, each depicting different types of vegetation such as hops, hawthorn, vines, ivy and buttercups.

The development of curvilinear tracery occurred mainly in the first half of the 14th century. As the name, hopefully, suggests, it derived from the linear style of tracery from the earlier period of before 1300. The idea of a flowing tracery conveys a new feminine aesthetic of fluency, never found before in English architecture. I will examine this aesthetic in much greater detail when we consider the even more feminine geometric forms created by artists in Italy before and during this period, and, it is impossible to deny that the style was partly motivated by such influences. That curvilinear tracery was derived as an exotic form is also supported by its use in conjunction with the ogee, see figure 10 (bottom), also developed at this time. Chris Wilson gives an excellent account of this in his book "The Gothic Cathedral", which I will briefly explain. The lower surviving chapel of St. Stephen's in Westminster Palace, begun in 1292, contains the earliest examples of reticulated windows with ogees, the great east window employing wide ogee cusps. This form was also employed in the stalls of the chapel, which are headed by trefoil-arched canopies that are cusped into an ogee design. These architectural forms were soon repeated, for example in the 4 tall ogee-headed niches of the pulpitum of Lincoln Cathedral, 1300, the nodding ogee of the bishop's throne at Exeter Cathedral, 1313, and the ogee-reticulated tracery of the windows in the East Cloister of Westminster Abbey, 1300. This type of tracery enjoyed its heyday in England, between 1300 and 1340, Wilson describes it as conveying a "sense of swirling, flickering upward movement around a central axis". This aesthetic is partly conveyed by the unique design of the West Window of York Cathedral, 1338, which we will consider shortly. Flowing tracery, typical of the period between 1300 and 1350, is described by Alex Clifton-Taylor, in the appropriately named chapter "The

[11] According to Wilson, the trefoil cusped design appeared at Saint Chapelle five years before.

[12] See Clifton-Taylor.

Flowering of Decorated", as "mullions themselves carried into tracery patterns, swinging to and fro in reverse curves and ogees". In "The Rose Window: Splendour and Symbol", (Cowen, 2005) ,by Painton Cowen, we have a description of curvilinear tracery as "flowing, interweaving tracery with the use of the ogival arch". Whether these descriptions are correct may be judged from the relatively few examples of windows of this period that could be said to explore such aesthetics as fluency or interweaving. These are the rose window of St. Mary's, Cheltenham, early 14th century,



27. Rose Window of St. Mary's church, Cheltenham.

the two traceried windows of the West front of Exeter Cathedral, built between 1346 and 1375, ([13]), the right window of the retro choir in the south aisle of Wells cathedral, ([14]), and probably the most remarkable and famous example, the big central window of the west front of York Cathedral, (1338) .

[13] Although Clifton-Taylor describes these as more geometric than flowing.

[14] Clifton-Taylor describes this as an example of the flowing rather than geometric type of window.

28. West Window of York Cathedral, York.

The west window of York, which embodies a sacred heart, is beautiful and confirms the idea that curvilinear tracery was a luxurious style of English design. However, in spite of these examples, in my opinion, the curvilinear style of tracery was a failure, to the extent that the geometrical ideas underlying aesthetic concepts such as fluency and interweaving were not properly understood in the English mentality. I will continue this discussion at much greater length when we consider the work of Italian artists, but, for now, will make the remark that the use of the cusp or ogival form in tracery patterns actually suppresses the aesthetic of fluency, see also the chapter on duality. In this sense, the conjunction of these aesthetics is unsatisfactory and even slightly unsettling. These remarks do not apply to the linear style of tracery, developed between 1250 and 1300, which was understood at a far more profound level. It is to this subject which I now wish to turn.

I will consider mainly Bony's excellent account cited above. He begins his analysis of this period by looking at what he refers to as "The London Court Style of the 1280's", during which time, the vocabulary of linear tracery was invented. The choir of Old St.Paul's, constructed in the 1270's but now destroyed, can still be examined through surviving drawings. Bony observes that the tracery patterns of the windows now "tend to spread in sheets of a grid like character, instead of a logical law of subdivision", and, moreover, that the "new forms are sharper with tighter tracery patterns", examples of such forms being the pointed trefoil and the linear shapes formed by the lattice tracery. The first of these remarks shows how the English rejected the idea of geometric tracery, in favour of a more linear style. Furthermore, the idea of a grid, seen as a union of intersecting lines, develops the aesthetic of fragmentation to even greater levels. The image of the crucifixion, which underlies the pointed vault, can be seen as a local form of fragmentation, occurring at a single point, whereas the image of a grid globalizes this fragmentation to the whole plane, occurring at the series of points where the lines intersect. The second remark is also interesting, the aesthetic of sharpness conveys a physical rather than spiritual sense of the sublime, and relates to the idea of fragmentation as a physical process, the sharpness of the square grid forms occurring at the points of fragmentation. In this sense, we see new aesthetic relationships between the sublime and fragmentation being explored. We will consider some of the associated geometrical ideas in the following chapter.

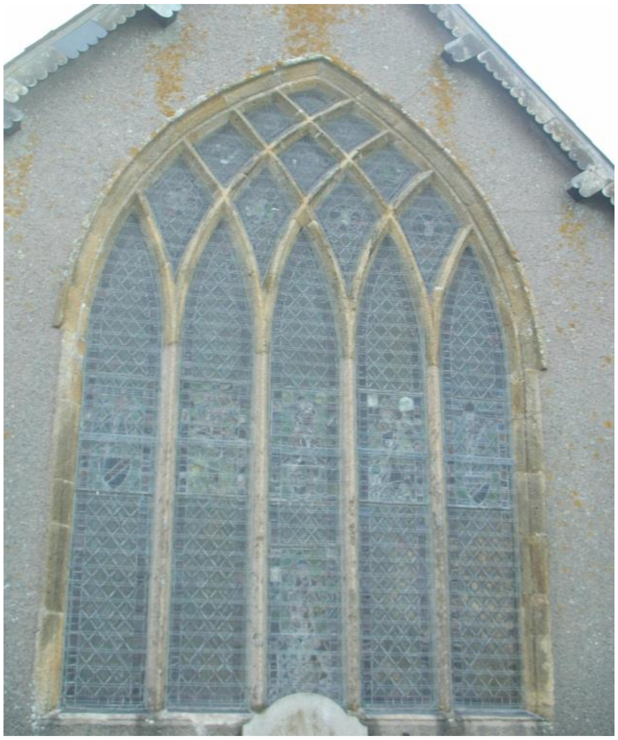
The outcome of the innovations at St. Paul's begins a new cycle of invention, called the London Court Style. The first of these was Y-tracery, so called by the Y-shapes formed by the intersection of lancet-style mullions in window frames. As Bony observes, the origination of this form is Norman, and can be found in Whitby Abbey, from the earlier Lancet period. The second of these is intersecting-tracery, a London invention of this period, which is used on an impressive scale in the window of Nine Alters at Durham Cathedral, 1280's, but developed to its full potential in the East Window of the Chapel of St. Ethelreda, Ely Place, Holborn, where Y-tracery is also used.

29. East window of St. Ethelreda's church, London.

The tracery style found at St. Ethelreda confirms the origination of intersecting tracery as a London design, as, even at Durham, we find a local 1-point fragmentation design, whereas, at Holborn, the design is many-pointed over the whole plane. The idea of many pointed fragmentation raises interesting questions for Christian belief, as it suggests that the image of the crucifixion is not the most fundamental form of physical and spiritual fragmentation, ([15]).

A similar example to the tracery pattern found at St. Ethelreda, occurs in Bere Ferrers, Devon, with, a possibly earlier design based on 1-point fragmentation;

[15] Recalling Tintoretto's image of the crucifixion, many point fragmentation is used, if we consider Jesus as part of a group with the two thieves. Jesus resolves this question in his words "I say to you today, you will be with me in the kingdom of paradise".



30. East windows of St. Andrew's church, Bere Ferrers, Devon.

Within five years, the new linear Court style had spread throughout England. Robert Burnell used Y-tracery in Wells palace, which was finished by 1292, here, appearing in the characteristic double lancet window design. The chapel of Merton College, Oxford, 1289-1294, has windows, which are a very fine example of the linear intersecting style, not only as many-point fragmentation is used, but also because the aesthetic of sharpness is explored in the diamond-shaped forms of the intersecting lines. Another fine example, from around 1285, is the rose window of Boyton church in Wiltshire, which employs the unique design of 3 spherical triangles, inscribed in a circle, intersecting in a cusp, ([16]). Here, we see the idea of degenerating a curve along six fixed lines, tangential to the branches of the triangles, again exploring further aesthetic relationships between the sublime and degeneration. We will consider the geometric technique of degenerating a curve to a configuration of lines in the following chapter.

[16] Although the author found a smaller, virtually identical design at St. Mary's, Cheltenham, from the early 14th century. See also (Coates).

31. Rose window of St. Mary the Virgin church, Boyton, Wiltshire.

32. Rose window of St. Mary's church, Cheltenham.

Bony, himself, makes some interesting aesthetic remarks about the linear form of tracery developed at this time. He observes both the effects of sharpness created by Y-tracery and refers to that created by windows such as at Merton, as quasi-mechanical fragmentation. Bony also makes the crucial observation that the acceptance of the linear style of tracery used by the English, is a result of the previous interest in linearity, developed by the Normans, in rib vaulting, citing the tierceron vault pattern of Exeter Cathedral East End, created around 1290, as an example of this connection. Bony devotes a whole chapter of his book, entitled "Linear Patterns and Networks" to the discussion of this topic, which we will now consider.

Bony begins by making a distinction between the English and French interpretation of linearity. In French medieval architecture, lines had intellectual, scientific value, and underlie the movement and articulation of architecture. In English medieval architecture, linearity is perceived as an invitation to play on patterns, citing the facade of Wells cathedral as "a vision of a cliff face held together by an imprint of woven netting". This sentiment is one with which I partly agree, as we will see from further consideration of Bony's argument, to the extent that English architecture develops the aesthetic of fragmentation, through the configurations of line arrangements in both vaulting and linear tracery, to unprecedented levels. However, it is unreasonable to say that this exploration had little scientific value and, merely, counts as ornamentation. As we saw in the construction of the rib vault of Durham cathedral, the relationship between functionality and aesthetics is a subtle two-way interaction in the English mindset. This point of view will become clearer in the following chapter.

Bony goes on to distinguish between two styles of linearity, both of which were evident in English rib vaulting long before the advent of the linear style of tracery. These are the crazy vault style of Lincoln cathedral and tierceron or star-shaped vault patterns. The tierceron pattern can be traced back to Norman architecture, even though many of the original rib vaults in England have been destroyed, a surviving example is the crossing vault of Worcester cathedral , see previous image. It seems reasonable to suggest that the repertoire of linear styles available to Norman architects extended even beyond this pattern and the conventional rib vault style. We will discuss this question presently, when we consider the influence of Norman art from Sicily at this time. Bony then observes reasonably that these styles explain the later arrival of linear tracery, which becomes "a second element of tight patterning", giving the excellent example of the presbytery and choir of Exeter cathedral, of 1290. Here, we see the full effect of the linearity of vault design and window tracery combined.

Bony continues his analysis with the development of the aesthetic of fragmentation in vault patterns, which occurred after the arrival of linear tracery. One new feature of this development, according to his argument, is the "sense of motif as a self-centred unit for focusing linear power". This powerful observation is supported by his analysis of two new styles of vault pattern, invented in England between 1290 and 1310, the lierne vault and the net vault.

The lierne vault is defined as having little linking ribs which zig-zag from boss to boss, drawing a linear pattern on the vault surface. Perhaps this idea derived from the chevron designs on Norman pillars, but, in any case, Bony attributes the idea to the London court of the time, citing early examples such as St. Stephen's chapel, from between 1295 and 1320, and the choir of Bristol cathedral. Bony also notes the increasing fragmentation of the lierne vault design throughout this period, later examples such as the Lady Chapel vault of Ely Cathedral, from the 1330's, having 48 compartments per bay, compared with the 10 compartments per bay of the St. Stephen's design.

Bony considers two examples of vault patterns which he refers to as transitional between the lierne and net designs. These are the tierceron design of the Lady Chapel of Wells Cathedral, early fourteenth century, and the Prior's Kitchen vault of Durham cathedral, 1366-1371, see also the earlier cited work of Bony;

33. Prior's Kitchen vault, Durham Cathedral, Durham.

Bony notes the Islamic vocabulary of the tierceron design at Wells, and, it seems reasonable to suggest that such designs were heavily influenced by Islamic models. We will consider this connection more closely when we consider the Norman designs found in Sicily. The sense of motif is clearly evident here, and the sharp, angular intersections of the line arrangement have the effect of almost engendering a sense of pain in the observer, due to the associations with a cutting edge. Again, this is a feature of Norman work in Sicily. The design of the Prior's Kitchen Vault, as Bony observes, closely resembles an Islamic pattern found at the Great Mosque of Cordoba, from the 10th century. However, there is no evidence that the architect, John Lewyn, had any knowledge of this design. In any case, the pattern is ingenious to the extent that it solves the problem of constructing a design for a ceiling containing a hexagonal aperture, which was effectively the kitchen's chimney. Here, we see a good example of an aesthetic innovation created by functional considerations. The innovations in linearity, here, are the construction of a pattern of lines in general position, that is no three of the lines intersect in a common point, which create a symmetrical, hexagonal motif. Clearly, this solves the functional problem that the architect was initially confronted with. This type of linear intelligence is another characteristic of the English mentality.

The net vault is characterized by a pattern of squares set on a diagonal. Bony observes that, although the net vault was invented in France, the English designs differ from the French models, in the sense that there is an absence of cross-linkages, that is the crossings of the line arrangements are simplified. We saw this type of intelligence at work in the design of John Lewyn and we will consider it again in the following chapter. An extremely fine example can be found in the vault design for the nave of Ottery St. Mary's in Devon, from the 1340's;

34. Vault of St. Mary's church, Ottery St. Mary, Devon.

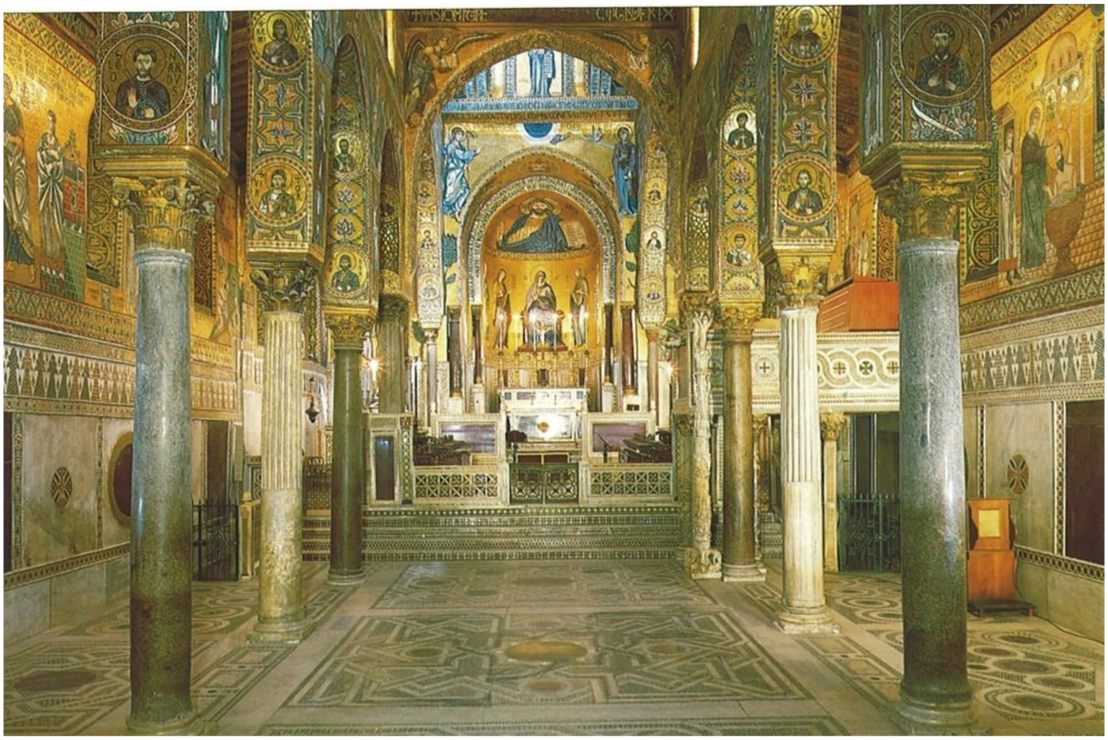
Here, we see a relatively small number of cross linkages, at most six. Again, there seems to be a structural intelligence at work here, the thrust of the vault is spread more evenly across the supporting ribs, hence, if one of the linkages fails, the vault is still likely to be supported. Ottery was the birthplace of the great romantic writer, Samuel Taylor Coleridge, who was known to spend much of his childhood playing around St. Mary's. Perhaps, this design was a motivation for his famous literary image from the Rime of the Ancient Mariner;

"And straight the Sun was flecked with bars,
(Heaven's Mother send us grace!)
As if through a dungeon-grate he peered
With broad and burning face."

Bony concludes his analysis of the net vault, by again noting the increased fragmentation occurring in later designs such as the vaults of Gloucester Cathedral(1337) and Tewkesbury Abbey(1322-6). At Gloucester, Bony observes the use of kaleidoscopic fragmentation, 106 compartments per bay, while at Tewkesbury, a more moderate fragmentation of 32 compartments per bay. He supports his earlier observation by noting the numerous geometric motifs occurring at Tewkesbury, of tierceron, square and octagonal patterns superimposed. These patterns are central to the evolving vocabulary of the linear style, and, I wish now to consider their origination more closely.

One of Bony's main arguments is that the linear style of tracery evolved from the earlier Norman constructions of rib vaulting. This seems to be reasonably supported though, unfortunately, few original Norman vaults survive in England. However, consideration of Norman art in Sicily, from the middle of the 12th century, shows undeniably, that the Normans had, at this time, developed a sophisticated vocabulary of linear motifs and forms. These patterns are mainly found in the designs of pavements, found in the Palace of the Normans and San Cataldo church in Palermo, and, the Cathedral of Monreale, just outside Palermo.

The construction of the Palace of the Normans began in 1130, for King Roger. The private chapel of the King, constructed in 1132, contains examples of a linear eight-pointed star design, centred around five red marble discs.

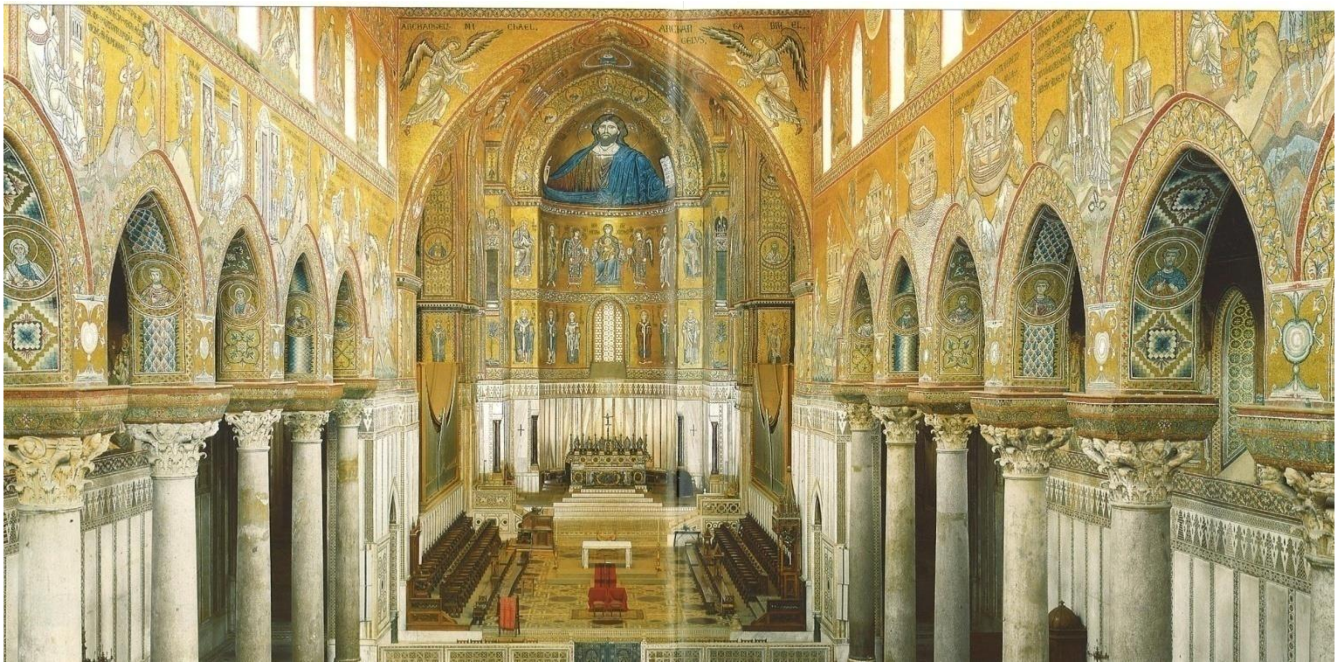


35. Pavement of the Palace of the Normans, Palermo, Sicily.

Here, we see clearly the aesthetic of linearity and the concentration of this aesthetic power in the star form. Moreover, we see the development of the aesthetic of fragmentation in the clever use of five distinct jagged, interweaving lines, corresponding to the number of discs. This development seems to have occurred at the almost simultaneous construction of Durham cathedral. In the royal apartments of King Roger, situated on the floor immediately above the private chapel, we can find a further linear, interweaving design, this time in the form of a ceiling mosaic. Although, there are no centralizing motifs, we again find the same linear aesthetic of jagged, interweaving lines. The Church of San Cataldo was constructed after 1154. It is a fine example of the Norman adaptation of Islamic architecture, with two characteristic hemispherical red domes, reminiscent of Islamic constructions in Spain. The church pavement is typical of the Norman interpretation of traditional Islamic aniconic designs.

36. Pavement of San Cataldo church, Palermo, Sicily.

The finest examples of linear pavement designs can be found in the Cathedral of Monreale, whose construction also begins in the first half of the 12th century, possibly before Durham. In the left section of the choir, one can find variations of the star design of the private chapel. In the right section of the choir, one can find a design, which is even more powerful, in that the linear motifs are an arrangement of diamonds, hexagons and an extremely acute 8-pointed star. Such designs seem to almost convey a sense of pain because of the associations with a cutting edge and the studied regular geometric nature of the linear motifs. Perhaps the most striking design is that of the pavement of the central section of the choir, consisting of a series of 8-pointed stars centred on red porphyry discs. The effect is almost reminiscent of the carnage of a battlefield in which the red discs resemble pools of blood, surrounded by the slashing forms of the star designs. The way in which the stars interlink has a literally castrating effect on the observer, which I leave the reader to consider more carefully. This studied concentration of linear power in the form of particular geometric motifs is unparalleled in any other designs that I have seen. That these designs were transmitted to England seems undeniable, the Cathedral of Monreale ,in fact, contains a mosaic of Thomas Becket, showing that links between the two Norman provinces were well established.



37. Nave of Monreale Cathedral, Palermo, Sicily.

The designs in Sicily show a repeating theme which is summarized by Belfiore;

"Nelle chiese normanne, i pavimenti sono realizzati a mosaico o a tarsea marmorea e svolgono generalmente motivi geometrici aniconici in cui e dominante la large fettuccia che, intrecciandosi, forma figure varie che hanno spesso il loro centro o nella stella ad otto punti o in dischi di porfido"

The eight-pointed star design can be found in earlier Islamic patterns, however, the adaptation of this motif is distinctively Norman. First, it is used as a deliberate method of focusing the linear power of jagged and acutely crossing lines. Secondly, there a suppressed aesthetic of light, which cannot be found in Islamic work, evidenced by the use of a braiding design in which the lines are posited in 3-dimensional rather than 2-dimensional space, a feature which which we will discuss at greater length in a later chapter.

The final medieval style of English architecture, the Perpendicular, is motivated, to a great extent, by the aesthetic of the sublime. In this sense, it has a closer relationship, perhaps, to the earlier Lancet style. Before discussing the work of this period, we should, therefore, consider this aesthetic in more detail.

An important source is the work of Edmund Burke, (Burke, 1757) "Philosophical Enquiry into the Origin of our Ideas of the Sublime and Beautiful". Burke's argument is essentially an early example of psychology, in that his argument relies on the distinction between our sense of pain and pleasure, as a means of differentiating between the sublime and beautiful. Burke, in fact, gives an almost geometrical analysis of forms that engender such senses, arguing that sharp and linear forms are more likely to invoke a sense of pain in the observer. This argument is clearly reasonable to the extent that we associate such forms with a cutting edge. However, as we discussed in the consideration of Norman patterns found in Sicily, the sense of pain engendered in the observer is greatest when concentrated in particular geometrical motifs. The Normans, who were clearly experts in the art of war and swordsmanship, adopted the aesthetic of fragmentation as a means of focusing the less threatening aesthetic of linearity on its own. In this sense, Burke's argument points to a subtle relationship between the aesthetics of the sublime and fragmentation. However, I believe, his argument is weak to the extent that he ignores the relationship between spirituality and the sublime, and, thus, equates the sublime with the sense of pain too closely.

In order, perhaps, to remedy his argument, we should consider Ruskin's extensive writings on the nature of the sublime in "The Seven Lamps of Architecture". Perhaps the most important discussion is contained in the Lamp of Power, entitled, no doubt, to reflect the image of the sublime contained in the vision of The Throne of God. For Ruskin, the Lamp of Power is;

"a sympathy in the form of noble buildings, with what is most sublime in natural things"

What, for Ruskin, constitutes these sublime impressions? Completely unlike Burke, Ruskin claims that the source of the sublime is to be found in weighty masses, not in lines. He cites the natural example of "the surface of a quiet lake" or the architectural example of The Doge's Palace in Venice, in which the principal facade is elongated to the eye by a range of 34 small arches and 35 columns, everything done to convey mass. With regard to what he refers to as "abstract power and awfulness", he writes;

"without breadth of surface, it is vain to seek them"

and talks of the joy that the mind experiences when allowed to "have space enough over which to range...in contemplating the flatness and sweep of great plains and broad seas"

He gives a brief geometrical analysis of the aesthetic of power;

"Thus the square and the circle are pre-eminently the areas of power among those bounded by purely straight or curved lines"

His evaluation of architecture, therefore, depends critically on conformity to this principle of mass;

"I do not believe that ever any building was truly great, unless it had mighty masses, vigorous and deep, of shadow mingled with its surface"

"the majesty of buildings depends on the weight and vigour of their masses, than on any other attribute of their design; mass of everything, of bulk, of light, of darkness, of colour, not the mere sum of any of these, but the breadth of them; not broken light, nor scattered darkness, nor divided weight, but solid stone, broad sunshine, starless shade."

For Ruskin, linearity, was, in fact, a principal cause of the decline of Gothic art;

"the substitution of line for light destroyed Gothic tracery".

Ruskin's evaluation of the aesthetics of light and mass above line, in consideration of the sublime here are clear. However, again, it is an argument with which I disagree. Unlike Burke, Ruskin observes the importance of spirituality in the assessment of the sublime. There is clearly a connection with the power of nature in the images that he considers. The aesthetic of light also invokes a form of spirituality, which we will consider in a later chapter. Ruskin seems to look for some form of connection between the finite and the infinite, in his description of the mind ranging over the ocean. However, he then abandons this connection by considering the bounded images, such as the lake, or the square and the circle. In this sense, there is a feeling of obstructed spirituality about his argument, which can only be resolved by directing the mind to a point at infinity on the ocean horizon. This also resolves the tension between the arguments of Burke and Ruskin on the relative importance of linearity in the assessment of the sublime.

Ruskin also briefly considers linearity in his discussion of the Lamp of Beauty, where he argues that all beautiful forms are composed of curves rather than lines, because "concatenations of straight lines" are not found in nature. I agree with the conclusion of this argument, as we have argued repeatedly, the aesthetic of fragmentation is more likely to induce a sense of pain in the observer. However, for the same reason as his assessment of the sublime in the Lamp of Power, I think Ruskin is wrong in the supporting argument, in discounting the evidence of linearity in nature. I recently purchased a painting of the sun setting over the ocean, as seen from a protruding headland. The three lines formed by the rays of the sun, the horizon and the line of the headland form an almost perfect right-angled triangle, a perfect connection between the aesthetics of the sublime and fragmentation.

However, in his discussion of the Lamp of Memory, Ruskin gives a rather different assessment of the sublime. Here, he makes an interesting distinction between the sublime and the picturesque, defining the picturesque as "sublimity dependent on the accidental characteristics of objects". These accidental characteristics are features of the sublime, which he characterizes aesthetically by "angular and broken lines", "vigorous oppositions of light and shadow" and "grave, deep, totally contrasted colour", most noble when reminiscent of "rocks, mountains or stormy clouds and waves". Ruskin argues that he true sublime though, superior to the picturesque and evident in the work of artists such as Tintoretto, is only obtained when it emerges from the picturesque, independent from the accidental characteristics of the objects themselves. This idea is much closer to connecting the sublime and fragmentation aesthetics and supports the assessment of Tintoretto's painting of "The Crucifixion" which we gave earlier.

We now consider the final period of medieval architecture in England, the perpendicular style. This is usually held to originate in work at Gloucester cathedral in 1331. However, as Bony observes, this theory collapsed in the 1940's when John Harvey pointed out that the style, in fact, dates from the chapterhouse and cloister of St. Paul's, 1330-1331, later destroyed along with the Cathedral by the great fire of London, of which, drawings still remain. Maurice Hastings then claimed that the work at St. Paul's was prepared before by a pre-perpendicular trend, initiated in the 1290's, at St. Stephen's chapel. This argument links the origination of the Perpendicular with that of the Linear style, originating in London, which we have already examined, rather than that of the more regional Curvilinear style. The earliest surviving work of the perpendicular style is that of the north east choir of Gloucester cathedral, in which slender, soaring vaulting shafts are used. Here, we see the aesthetic of flowing curves rejected in favour of the aesthetics of ascent and the sublime. The link with the earlier trend of the linear style and previous uses of such vaulting shafts in England, such as at Worcester, is clear.

Between 1347 and 1350, the most famous example of perpendicular architecture, the East Window of Gloucester, was constructed. About the size of a tennis court, it is the largest stone traceried window in England and was the largest in medieval Europe. The aesthetics of ascent and the sublime are emphasized by the use of highly verticalized tracery, which curves at its highest point. Ruskin criticized this effect, referring to the mullions leading up to the arch head, as "carving knives". However, perhaps this description is insensitive and reflects his misunderstanding of the true nature of the sublime which we discussed earlier. I prefer the description given of the window by Alex Clifton Taylor, in "The Cathedrals of England", as "sparkling like icicles in the sunshine of a crisp winter's day".

38. East Window of Gloucester Cathedral, Gloucester.

The perpendicular style continued to enjoy popularity in England. One of the finest later examples is that of St. George's chapel, Windsor Castle, completed in 1506. Drawing on over a century of experience of Perpendicular design, the windows and columns all follow the pattern of a highly verticalized, parallel aesthetic. The chapel is the seat of the Order of the Garter, in which the 26 piers represent the number of the Garter Knights, upholding the chivalric tradition of knights as pillars of the church.

The English style of medieval architecture was copied in other countries as well, a good example is the Church of Alatri in Italy, of around 1300, whose rose window, in its profusion of trefoils and linear design, acknowledges both the aesthetics of naturalism and fragmentation, perfected during the Curvilinear period.

39. Rose Window of Santa Maria Maggiore church, Alatri, Italy.

Medieval architecture on the island of Gotland, off the coast of Sweden, acknowledges features of the English style, although due to the combining influence of the Geometric , is essentially unique. The octagonal vault, ([17]), and doorway of Helge And, in Visby, from 1200, are distinctive examples of a linear, asymptotic style, which we will refer to at the end of the following chapter.

40. Doorway of Helge And church, Visby, Gotland, Sweden.

[17] The only example of its kind in Sweden, see “Lonely Planet”, Carolyn Bain and Graeme Cornwallis.

Earlier examples of the Norman style of architecture can be found from between 1100 and 1180. The use of corner capitals and pillars, occur at Ribe and Viborg Cathedral in Denmark, Lund Cathedral in Sweden, ([18]), Stavanger Cathedral and St. Paul’s church, Bergen, in Norway;

41. Corner Capitals and Pillars: Ribe and Viborg cathedrals, Denmark, Lund cathedral, Sweden, Stavanger cathedral and St. Paul's church, Bergen, Norway.

The chevron design, on archways can be found, from the same period, at Helsingborg church, Sweden, ([19]), Stavanger Cathedral and the Archbishop’s Palace, Trondheim;

42. Archways of St. Mary's church, Helsingborg, Sweden, Stavanger cathedral and the Archbishop's Palace, Trondheim, Norway.

The diverse medieval style of architecture was largely abandoned in England, until the 19th century, where it regained popularity in the art form known as The Gothic Revival. Michael Lewis gives an excellent account of this period in his book of the same name, (Lewis, 2002).

The 18th century saw a limited revival in interest in Gothic art, reflecting a new literary appetite for what was perceived as the decay and melancholy of the crumbling medieval landscape of England, namely the looted, medieval abbeys of Henry VIII. The advent of the Romantic literary movement during the 19th century, sparked by figures such as Coleridge and Wordsworth, associated ideas such as the sublime and naturalism with the living process of the human imagination, and, partly from this impetus, medieval architecture was elevated to a

[18] Then part of Denmark.

[19] See footnote 17.

higher status in the public consciousness. An important construction, from the early part of the century, is St. Luke's Church, Chelsea, 1820-4, by James Savage, which revived the perpendicular style and raised the aesthetic and technical standards of Gothic architecture. By the 1840's, England was the international leader of the Gothic revival. The development of the doctrines of The Ecclesiological Society identified the perfection of the medieval style, as occurring at the end of the 13th century, and many rural churches were built according to these principles, such as St. Peter, Treverbyn, in Cornwall, by G. E. Street, between 1848 and 1850. Perhaps the most innovative building of the entire period is that of All Saints, Margaret Street, in London, by William Butterfield, 1849, described by Michael Lewis, as "a high keyed chromatic symphony of marble and brick". The second half of the century is inevitably dominated by the aesthetic involvement of Ruskin. This period saw the origin of William Morris's Arts and Crafts Movement, which flourished on the grounds of Ruskin's rejection of The Industrial Revolution, as taking the nobility and spiritual dimension from physical labour. It also saw an increasing interaction with scientific thinking, in which the open forms of Gothic art had a symbolic advantage, suggesting infinity and progress, rather than static knowledge. The Oxford University Museum, designed by Woodward, between 1854 and 1860, is the main product of this connection. The Gothic Revival of this period also flourished to a limited extent in America. Perhaps the greatest figure of this movement was the architect E. H. Richardson. Impressed by Ruskin's writings in "The Lamp of Power", which called for an architecture that "aspires to the state of nature herself, with the fierce dignity of a cliff face, or the swelling sense of geological upheaval", he came the closest to fulfilling these demands. A good example can be found in Trinity Church, Boston, 1872-1876, whose electric colours reflect an interesting interpretation of the sublime as lightning. The most interesting product though of these years is, in my opinion, found in art not architecture, that of "The Pre-Raphaelite Movement", and I will conclude this chapter by considering their work in greater detail.

An excellent account of this may be found in Timothy Hilton's book, "The Pre-Raphaelites", (Hilton, 1970). Although the paintings and artists of this movement are well known, an enigma which seems to trouble many contemporary art critics is identifying a common style and philosophy behind their approach. I hope that some of the aesthetic ideas explored in this chapter will show that they conform to a rediscovery of the medieval mindset of English art and architecture.

Naturalism in English art became increasingly important as a result of Ruskin's defence of Turner's work in "Modern Painters", 1843-1846. Drawing on this new trend, Ford Madox-Brown created the first paintings, which show a renewed interest in medieval ideas, "Chaucer at the Court of King Edward III" and "Wycliffe Reading his Translation of The Bible To John of Gaunt", 1845-1851. Here, we see a return to the flat medieval representation of surfaces and a more naturalistic use of light. The 1830's and 1840's also showed an increased public interest in pre-Renaissance painting, a popular focal point was The Campo Santo in Pisa, now heavily damaged by fire, which exhibited the works of medieval artists such as Orcagna, Buffalmaco and Giotto. Three artists who were deeply affected by this sentiment were Dante Gabriel Rosetti, John Everett Millais and William Holman Hunt. Millais and Hunt agreed that the classical teaching of Reynolds, a seminal figure of the time, had led to harmful tendencies in English art. Their attack on classical art was focused on Raphael's "Transfiguration", condemned by Millais for its "grandiose disregard of the simplicity of truth, the pompous posturing of the Apostles, and the unspiritual attitudinizing of the Saviour". In the Summer of 1848, Rosetti, Millais and Hunt formed The Pre-Raphaelite Brotherhood.

The first meeting of The PRB, as it was often abbreviated, took place in Millais' studio, in September 1848. Here, the philosophical basis of the brotherhood was formed, Hunt, in all probability, being the intellectual focus of the group. The brotherhood adhered to a renewed spirituality, by forming a pyramid of "immortals", with Jesus Christ at the apex, followed by Shakespeare and The Author of The Book of Job. Although the group might be said to show a certain intellectual uncertainty at this stage, in their first paintings, a deep understanding of the aesthetics behind English medieval art is clearly evident. Millais' first PRB painting, Isabella, 1849, is an illustration of Keats' poem of the same name. In it we see the same naturalism and flatness of spatial construction, evident in the earlier work of Ford Madox-Brown. The distinctive feature of the out thrust leg, kicking the dog, draws the attention of the viewer towards the two lovers Isabella and Lorenzo.

Distinctly non-classical, it reflects the sharp, physical manifestation of the sublime that we discussed earlier, central to the aesthetics of English medieval art. Rosetti's "The Childhood of Mary Virgin", 1848-49, is another fine example. Again, we see a disregard of perspective, and a demonstration of the aesthetic of fragmentation in the strong horizontal and vertical trellise lines, which form crosses in the background. There is also a peculiarity in the organization of the framework, as in Tintoretto's work, demonstrating the idea of fragmentation being an essentially unfocused and disorganized representation of lines on a surface.

43. Isabella by Millais, Walker Art Gallery, Liverpool.

44. The Childhood of Mary Virgin by Rosetti, Tate Gallery, London.

A number of paintings by the brotherhood followed, often called the "Gothic" phase. Unfortunately, these drew considerable criticism from the British Establishment and a second style followed, which is more concerned with the aesthetic of light. We have said very little about this subject up to now, as it is the principal theme of a later chapter. However, it was a question that concerned the PRB, and they took their main stylistic leads from the writer of The Moonstone, Wilkie Collins, who argued that paintings should be "clean, fresh and genuine", and their own invented system, propose by Ruskin, that of "the sun", which "illuminates everything". A painting which could be said to conform to the first system of Collins is Rosetti's "Ecce Ancilla Domini", 1849-50, in which a cool light suffuses the scene. This aesthetic is, as we shall see, distinctively Italian and, no doubt, was also an influence on Rosetti's style of this time. A painting which conforms to the second system is that of Holman Hunt's, "Our English Coasts", 1852, of which Ruskin commented;

"It showed us for the first time in the History of Art, the absolutely faithful balance of colour and shade by which sunshine might be transported into a key in which the harmonies possible with material pigments should yet produce the same impressions on the mind which were caused by the light itself"

and

"An accomplished artist, Holman Hunt sees light and shade as bluish green barred with gold"

Ruskin's preoccupation with light, at this time, was a result of his work "The Elements of Drawing", in which he argued that the technical power of painting depended on the recovery of "the innocence of the eye". It could be argued that Hunt intended the painting to be a rendering of the sublime, however, it differs strikingly from Turner's construction, in that the light is focused, rather than drawn as lines across the sea.

Hunt's next painting "The Light of The World", 1853, is one of England's most famous religious paintings. Although not well received, it was praised by Ruskin for its use of symbology, Christ knocks at the door of the human soul, which is barred with ivy and leaves. In some ways, it is a depressing picture as there is a sense of obstructed spirituality about it. However, Hunt's characteristic use of gently focused light is highly original and conforms to an advanced aesthetic of light, which we will consider in a later chapter.

45. The Light of the World by Holman Hunt, Keble College chapel, Oxford.

In 1851, the PRB held its last meeting, in which there was a complete division about the functioning of the brotherhood. Hunt moved to Palestine, where, in 1854, he painted "The Scapegoat", quite probably a symbolic rendering of Christ. The ghastliness of the depiction of the goat is matched by the location and story behind the painting, apparently the first goat that Hunt bought died while the lake of Oosdoom is a place devoid of all vegetation and life. Although, there is a hard nastiness about the painting, it is an interesting interpretation of the sublime as isolated and strange, beyond human comprehension or appetite.

The relationship between Ruskin and Hunt is worth considering in greater detail, particularly in relation to the aesthetics of light. Hunt, in fact, wrote a pamphlet entitled "Your Good Influence on Me", in which he explains this relationship from his point of view. In my opinion, both Ruskin and Hunt aimed to describe the sublime, but failed to capture its true nature, however, in doing so, they found different but interesting new aesthetics of light and form.

The later paintings of the PRB are more governed by growing social concerns than aesthetic or spiritual motivations. An important contributor was Ford Madox Brown, whose "Work" considers the importance of labour in society, and "The Last of England" reflects on the problems of emigration. The very last years are governed by the new figure of Edward Burne-Jones, whose style is, in my opinion, too different to fit into the overall spiritual aesthetics behind the PRB, and the later paintings of Rosetti. Rosetti's work continued to have an intense spiritual dimension, one of his last paintings, "The Blessed Damozel", 1871-79, was also the subject of his poetry. It depicts the damozel, inspired by the of love of his final years Lizzy Siddal, at the golden bower of heaven, looking reverently down to earth.

Perhaps, the 19th century was the last period to reflect again on the mindset of the medieval age, in a way that the last century has failed to do. In some respects, the aesthetic expression of that age has been exhausted and we now live in a society that requires us to push forward in new, more scientific fields of endeavor. However, it is my purpose in the next chapter, to show that the aesthetic considerations which we have discussed are essential in the field of geometry and, therefore, of science, if it is to be regarded as more than a merely analytic enterprise. In a substantial way, this was understood in the medieval period, and, I hope that this chapter has gone some way to make this point of view clearer.

5. NEWTON AND THE SUBLIME GEOMETRY OF LINES

TRISTRAM DE PIRO

This chapter will be concerned with the way in which geometric thinking is able to draw on the aesthetic ideas outlined above. In the second chapter, I referred to the intelligence which mediates between the images of the Crucifixion and the Throne of God as the Ascension. In some sense, the geometric ideas that we will consider in this chapter conform to this general notion. However, in order to illustrate the specific geometric ideas, I will prefer to rely on the more detailed aesthetic terminology which I introduced above. The geometric ideas which I will consider here are all concerned with the theory of algebraic curves. One does not need to be an expert mathematician, in order to understand such objects at a rudimentary level. Anyone who has studied mathematics to secondary school level in England, will have come across the idea of using functions to draw the graphs of curves or lines in a plane. A simple example is given by a line, with the equation;

$y = ax + b$

which we represent using cartesian coordinates as shown in the diagram (insert image). The simplest example of a curve (a parabola) is given by the equation;

$y = ax^2 + bx + c = 0$

We are all taught the formula for solving such an equation at school;

$x = -b \pm \frac{\sqrt{b^2-4ac}}{2a}$

Geometrically, this is interpreted as finding the two points where the curve meets the x-axis or the line $y = 0$, ([6] is a good reference for secondary level mathematics.). One may generalise the theory of functions, by considering equations in which the y -variable may appear as a higher power. For example, the equation;
$y^2 = x^3 - x$

may be represented as follows (insert image). In modern terminology, this is referred to as an elliptic curve. Generalising further, one may consider arbitrary polynomials $p(x, y) = 0$ in the variables x and y;

$p(x, y) = \sum_{i+j \leq d} a_{ij}x^i y^j = 0$ $(*)$

The solutions of such polynomial equations are referred to as algebraic curves. The important point to remember in the definition of an algebraic curve is that the polynomial equation defining it has a finite number of terms. The index d appearing in the polynomial equation above is referred to as the degree of the curve and measures its complexity.

In order to make the theory more precise, one has to clarify the domain in which the coefficients of the equation (*), and the variables $\{x, y\}$ are allowed to vary. At secondary school, we confine our attention to the real numbers, intuitively any number that can be represented by a possibly infinite decimal expansion, and denoted in modern mathematics by the symbol $\mathcal{R}$. Such numbers are convenient, as they can be programmed easily on a computer or a calculator. However, they also suffer from the following drawback, that there are many polynomial equations which have coefficients in the real numbers, that admit no real solutions. The simplest example is given by the equation;

$$y^2 + 1 = 0$$

This has no real solutions, and, in order to solve it, we are forced to introduce the imaginary number, denoted by i. The domain obtained, by extending real numbers, to numbers of the form;

$$a + bi \text{ with } \{a, b\} \subset \mathcal{R}$$

is referred to as the complex numbers, denoted in modern mathematics by the symbol $\mathcal{C}$.

For the purposes of algebraic curves, it is more convenient to work with the domain $\mathcal{C}$, rather than $\mathcal{R}$. For example, let us consider the algebraic curve, defined by the polynomial equation $y^2 + 1 = 0$. Over the real numbers, the solution set is empty, but over the complex numbers, it consists of 2 lines defined by $y = i$ and $y = -i$, (see diagram). Intuitively, we would like any polynomial equation to define a "curve" rather than the empty set, hence, the complex numbers seem to be a better choice of domain. It might be objected, however, that there could still be polynomial equations, with coefficients in $\mathcal{C}$, that admit no complex solutions. However, remarkably, this objection turns out to be false, as a consequence of the following property of $\mathcal{C}$, often called "The Fundamental Theorem of Algebra";

Any polynomial equation of the form $y^n + a_1 y^{n-1} + \ldots + a_n = 0$, with coefficients in $\mathcal{C}$, admits n complex solutions (possibly with multiplicity). In particular, any such equation admits at least one complex solution. (**)

Here, one should clarify what is meant by a solution with multiplicity, the equation $y^2 - 2y + 1 = 0$ factors as $(y - 1)^2 = 0$, hence, 1 is counted as a solution *twice*. As a consequence of this theorem, any polynomial $p(x, y)$ with coefficients in $\mathcal{C}$ admits an infinite number of solutions in the plane $\mathcal{C}^2$. Moreover, we have the following further property of the solution set of $p(x, y)$, which I will denote by C;

Either the solution set C consists of finitely many lines of the form $x = a$, for $a \in \mathcal{C}$, *or*, for all but finitely many elements $\{a_1, \ldots, a_n\} \subset \mathcal{C}$, if $x \in \mathcal{C} \setminus \{a_1, \ldots, a_n\}$, there exist finitely many solutions of the form (x, y) in C.

If we represent the complex numbers $\mathcal{C}$ as a line, corresponding to the horizontal x-axis, then this theorem tells us that any algebraic curve C consists of either a finite number of vertical lines, or is essentially a finite cover of the x-axis, see diagram. Intuitively, this is exactly the property that we would expect a "curve" drawn in the plane to have. In modern mathematics, this property is referred to, by saying that the dimension of the curve C is 1, abbreviated as $dim(C) = 1$. In contrast, we say that a finite number of points in the plane has dimension 0, while, the complement of an algebraic curve or the complement of finitely many points, has dimension 2. In particular, the whole plane $\mathcal{C}^2$ has dimension 2.

There is still, however, another consideration to be taken into account, for the representation of algebraic curves over the complex numbers. It is more natural to represent the real numbers $\mathcal{R}$ as a line, in which case, the domain of the complex numbers corresponds to a real plane, given by $\mathcal{R}^2$. In this interpretation, any algebraic curve may be viewed as a surface, in the sense that it looks 2-dimensional, when considered in relation to the real line $\mathcal{R}$. In the aesthetic representations of algebraic curves that we will consider, both interpretations of an algebraic curve as a "curve" and a "surface" will prove to be useful.

For the reader interested in a rigorous account of the mathematical construction of the real numbers, see, for example, [5]. The construction of the complex numbers and more advanced properties of functions on them, is dealt with in [28]. The history behind the construction of the real numbers is also interesting, and usually accredited to Dedekind in the nineteenth century. The mathematical development of the complex numbers is also an important field of study, and is associated with the names of a number of great French mathematicans from the nineteenth century, such as Cauchy and Liouville. If the reader is interested in the basic construction of algebraic curves, excellent modern accounts can be found in [35] or [7]. For a series of pictures of algebraic curves, the reader is strongly recommended to consult the St. Andrew's directory of algebraic curves, a link can be found on my website at http://www.curveline.net

Although the rigorous construction of algebraic curves using algebra is a fairly modern development, ([1]), the basic theory of polynomials has been known for a long time, probably with the invention of algebra by the Arabs. The intuitive representation of algebraic curves, in terms of a planar coordinate system, is, accredited to the French philosopher Rene Descartes. However, it seems feasible that this representation was known, in some form, even before. In order to illustrate the aesthetic ideas of the previous chapter, however, I wish to confine my attention to, without doubt, the greatest of all English geometers, Isaac Newton.

Isaac Newton was born on 4 January 1643, in the tiny village of Woolsthorpe, Lincolnshire, ([2]). An amusing anecdote records that he was so small at birth, that he could be fitted into a pint pot, a clear demonstration, if one was needed, that a towering intellect has nothing to do with physical size. In 1661, he attended Trinity College, Cambridge, graduating in 1665 without honours or distinction. In the middle of 1665, due to plague which had broken out in London, Newton returned to his family home, Woolsthorpe Manor. In the course of the next two years, Newton made a series of extraordinary mathematical breakthroughs.

The first of these, in 1665, was his discovery of the generalised Binomial Theorem. The

[1] Part of the modern subjects of algebraic geometry and commutative algebra, good references are [3], [20], [24] and [33].

[2] There are a number of good biographies of Newton, see, for example, [2] and [37]

binomial theorem gives a procedure for finding the expansion of a *positive* power of a polynomial expression. It is normally stated in the following form;

$(x+y)^n = \sum_{k=0}^{n} C_k^n x^{n-k} y^k$ where $C_k^n = \frac{n!}{(n-k)!k!}$, $n \geq 0$

This result was known in some form, possibly as early as the 3rd century BC. What is remarkable about Newton's generalised version of this theorem, is that it includes negative and even rational powers and introduces the idea of an infinite power series, an innovation also due to Newton. It is now formally stated in the following way;

$(x+y)^n = \sum_{k=0}^{\infty} C_k^n x^{n-k} y^k$ where $C_k^n = \frac{n(n-1)\ldots(n-(k-1))}{k!}$, n rational

In order to give simple demonstrations of this theorem, in the style that Newton originally gave, consider the problem of finding an inverse to the polynomial $(1+x)$, that is an expression $q(x)$ involving x, such that $q(x)(1+x) = 1$. A simple calculation shows that any finite expression of the form $a_0 + a_1 x + \ldots + a_n x^n$, for some $n \geq 0$ is insufficient. The correct answer, provided by the generalised binomial theorem, is;

$1 - x + x^2 - \ldots + (-1)^n x^n + \ldots$

where the summation is taken over an infinite number of terms. Consider the problem of finding a square root to the polynomial $(1+y^2)$, that is an expression $q(y)$ such that $q(y)^2 = 1+y^2$. Again, a simple calculation shows that a finite expression is inadequate. This time, the correct answer is provided by;

$1 + \frac{1}{2}y^2 - \frac{1}{8}y^4 + \ldots + \frac{1}{2}\frac{-1}{2}\frac{-3}{2}\ldots\frac{1-2(n-1)}{2}y^{2n} + \ldots$

where, again, an infinite summation is used. Both these problems are considered and solved successfully by Newton in his first major work "Analysis of Equations of an Infinite Number of Terms", ([3]).

The second achievement of the years between 1665 and 1667, when Newton returned to Cambridge, was his development of the infinitesimal calculus, probably the single greatest mathematical breakthrough in human history. Newton's mathematical writings on calculus of this time, which were never published, still survive and can be found in the prodigious "Mathematical Papers of Isaac Newton", by D.T Whiteside ([4]). Together, these papers are an invaluable source for understanding the evolution of his work in this field. The most important papers are collated under the headings;

(i). "Normals, Curvature and the Resolution of the General Problem of Tangents", (1664-1665)

---

[3] He, in fact, considers the more general cases of inverting the expression $b+x$ and finding the square root of a^2+y^2

[4] Sadly, D.T Whiteside died on 22 April, 2008.

(ii). "The Calculus Becomes an Algorithm", (1665), (Whiteside's terminology)

(iii). "The General Problems of Tangents, Curvature and Limit Motion Analysed by the Method of Fluxions", (1665-1666)

(iv). "The October 1666 Tract on Fluxions", (1666)

We will take the opportunity to consider them, in greater detail, later in the chapter. In 1669, Newton wrote;

(v). "Analysis of Equations of an Infinite Number of Terms".

which we mentioned above. Not only does this deal with the innovation of infinite series, but also develops a general method for finding infinite series solutions to polynomial equations, now known as Newton's Theorem, and work on the theory of integration. However, for some reason, Newton decided against publication of his manuscript and it was only circulated amongst a close circle of friends. The complete manuscript did not appear in print for decades later, ([5]). Shortly afterwards, in 1671, Newton wrote

(vi). "On the Method of Fluxions and Infinite Series".

This work not only incorporated his results on infinite series from his work of 1669, but also developed his ideas on the infinitesimal calculus, in particular his work on the theory of differentiation. Again, Newton decided to withold the manuscript from publication, a full version being printed much later.([6]) In 1676, Newton completed his last major work on calculus;

(vii). "The Quadrature of Curves".

([7]), which gave his most rigorous exposition of the theory of fluxions and tangency.

In the period between 1670 and 1700, Newton made a number of pioneering breakthroughs in the field of physics. The first of these was in the subject of Optics, the theory of light. Newton's first writings on this subject date back to his years as an undergraduate and his return to Woolsthorpe Manor, collated in Whiteside's "Mathematical Papers of Isaac Newton";

(viii). "Early Notes on Reflection and Refraction", (1664)

(ix). "The Essay 'Of Refractions'", (1665-1666)

[5] A translated version was printed in 1745 along with another publication, "The Quadrature of Curves", edited by Stewart, together referred to as "Geometria Analitica"

[6] The latin manuscript "Analysis per quantitem series, fluxiones ac differentias; cum enumeratione lineares tertius ordinis" appeared in 1711, the translated version, edited by Colson, "The method of fluxions and infinite series, with its application to the geometry of curve lines" appearing in 1736.

[7] A published version appeared in 1704, appended to his "Opticks"

(x). "Refraction of Light at a Spherical Surface", (1666?)

Between 1670 and 1671, Newton collated the results of this research in "The Lucasian Lectures on Optics", (1670-1671), delivered in 1672, after his appointment in 1669, as successor to Isaac Barrow, to The Lucasian Professorship of Mathematics. In 1672, he published his first paper on Optics, in the journal "Philosophical Transactions", entitled "A New Theory About Light and Colours". Two other important unpublished optical papers survive from this period;

(xi). "Miscellaneous Researches into Refraction at a Curved Interface", (1671)

(xii). "Miscellaneous Optical Calculations", (1670)

A final version of these ideas appeared in 1704, in;

(xiii). "Opticks: A Treatise of the Reflections, Refractions, Inflexions and Colours of Light", ([8]).

Here, Newton formulated his corpuscular theory of light, as consisting of particles propogated along straight line paths, and developed his theories on refraction and reflection. Newton conducted experiments with prisms to show that light could be refracted into a continuous spectrum of coloured light, and knives, observing the hyperbolic nature of the fringes of their shadows, an optical phenomenon now known as diffraction, ([9].) The work also contained the design of the first reflecting telescope and numerous results on the theory and construction of lenses for focusing light. Appended to his treatise on Optics was a new geometric work;

(xiv). "Enumeration of Lines of the Third Order"

dealing with the classification of cubic curves, that is algebraic curves of degree 3. This involved a number of new geometrical ideas, such as the use of asymptotes to understand such curves in terms of hyperbolas, a connection that I will argue was intimately connected to his research into optics.

The second important development occurred in the theory of motion and gravitation, which Newton published in;

(xv). "Philosophiae Naturalis Principia Mathematica", (1687), (See [26]).

Newton formulated three laws of motion, summarised as follows;

[8] See [25]

[9] Newton was unable to give a clear explanation of diffraction. This led to his corpuscular theory being largely rejected outside England, in favour of the circular wave model, proposed by his contemporary Huygens, in [22]

(i). A body's velocity will remain constant if the forces acting on it are balanced.

"Every body perseveres in its state of rest, or of uniform motion in a right line, unless it is compelled to change that state by forces impressed thereon", (Book I, Axioms, or Laws of Motion).

(ii). A body's acceleration is proportional to the net force acting on it, and described by the formula $F = ma$, where F denotes force, m denotes mass and a denotes acceleration.

"The alteration of motion is ever proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed", (Book I, Axioms, or Laws of Motion).

(iii). Every action has an equal and opposite reaction. That is any applied force meets with an equal force in the opposite direction.

"To every action there is always opposed an equal reaction: or, the mutual actions of two bodies upon each other are always equal, and directed to contrary parts", (Book I, Axioms, or Laws of Motion).

Newton formulated a Universal Principle of Gravitation in (Book III, Proposition VII, Theorem VII);

"That there is a power of gravity pertaining to all bodies, proportional to the several quantities of matter which they contain"

and the inverse square law of gravitation in (Book III, Proposition VIII, Theorem VIII) (see also (Book I, Section XII);

"In two spheres mutually gravitating each towards the other, if the matter in places on all sides round about and equi-distant from the centres is similar, the weight of either sphere towards the other will be reciprocally as the square of the distance between their centres"

These laws are still accepted as fundamental to modern physics. Newton used the inverse square law of gravitation to explain the elliptical orbits of planets around the sun, confirming in a striking way earlier observations made by Johannes Kepler in 1660. He was one of the first people to consider the mathematics behind three body gravitational problems in (Book I, Proposition LXVI, Theorem XXVI). The discovery of the law of gravitation is usually associated with the famous story of an apple falling on Newton's head. It is not known whether this event is apocryphal or not.

As well as scientific studies, Newton was deeply interested in theological issues. After 1690, he wrote a number of religious tracts, concerned with the literal interpretation of the Bible. Perhaps the most important of these is his "Observations on the Prophecies of Daniel, and the Apocalypse of St. John", published in 1733. Here, he makes the interesting observation, which, is not contained in Revelations, that the sign of the Antichrist is "+++", (three crosses), rather than the more traditionally held "666". Newton conducted his research into theology with the same precision that characterised his scientific writings, his Observations are full of historical details and numerology. These clearly show that Newton believed the Bible to be the unchangeable word of God, although subject to rational interpretation. Newton was able to reconcile his rational, scientific thinking with his theological views by taking God to be a supremely rational creator, designing the universe according to prescribed, mechanical laws. Many later writers and artists, such as Blake, took this to mean that Newton enjoyed a unique status of being priviliged to the working of God's mind. His portrait of Newton (1795) is comparable to his depiction of Urizen in "Ancient of Days", measuring the world with a compass.

The rationalist philosophy to which Newton subscribed, in many ways, shaped the thinking of much of the eighteenth century. However, part of the subject of this chapter is to argue that Newton's work in geometry is underpinned by a number of aesthetic ideas, which we explored previously. In this sense, there is a latent spirituality inherent in Newton's work, which, perhaps, he was not consciously aware of, but, nevertheless, informed his highly creative scientific work. The picture of Newton drawing on aesthetic and spiritual ideas to inform his rational thinking, as opposed to his pure rationality accessing the mind of God, is, in my opinion, a truer and more attractive image of him. Given his unquestionably devout nature, Newton himself admitted that he spent more time reading the Bible than any other book, perhaps, it is the way in which he, himself, would have wanted posterity to remember him.

The later years of Newton's life were mainly involved with what has become known as the Newton-Liebniz controversy. Liebniz, a German mathematician and philosopher, began working on a variant of the infinitesimal calculus, usually known as the differential calculus in 1674, and published his first paper on the subject in 1684. As we observed earlier, Newton didn't give a full printed account of his version of the calculus until 1704, but Newton's circle of friends claimed that he had obtained the results earlier, around 1667, and before Liebniz. In 1711, members of the Royal Society accused Liebniz of plagiarism, based on previous allegations of Facio and Kiel, and the fact that he had obtained a copy of Newton's manuscript (v) from a colleague of Newton, Tschirnhaus, in 1675. It is not known whether Liebniz made use of this manuscript in his work on calculus, or, whether he had, in fact, already invented the calculus previously. However, those who question Liebniz's good faith allege that, to a man of his ability, the manuscript sufficed to give him a clue to the methods of calculus. We will consider the question of the priority dispute in greater detail shortly.

Newton moved away from Cambridge in 1696, after obtaining a position as Warden and Master of the Mint in London. It is during this time that he is alleged to have carried out much of his alchemical research, some of which can be found in the final sections of his "Optics". He was elected as President of the Royal Society in 1703, an office that he retained for the rest of his life, and was knighted in 1705. He died in London on March 31st, 1727, at the age of 84, and is buried in Westminster Abbey. His tomb, the work of the artists William Kent and Michael Rysbrack, was completed in 1731. A good description of the tomb can be found in The Gentlemen's Magazine of the same year;

"On a Pedestal is placed a Sarcophagus (or Stone Coffin) upon the front of which are Boys in Basso-relievo with instruments in their hands, denoting his several discoveries, viz.

one with a Prism on which principally his admirable Book of Light and Colours is founded; another with a reflecting Telescope, whose great Advantages are so well known; another Boy is weighting the Sun and Planets with a Stillard, the Sun being near the Centre on one side, and the Planets on the other, alluding to a celebrated Proposition in his Principia; another is busy about a Furnace, and two others (near him) are loaded with money as newly coined, intimating his Office in the Mint.

Behind the Sarcophagus is a Pyramid; from the middle of it a Globe arises in Mezzo Relievo, on which several of the Constellations are drawn, in order to shew the path of the Comet in 1681, whose period he has with the greatest Sagacity determin'd. And also the Position of the solstitical Colure mention'd by Hypparchus, by which (in his Chronology) he has fixed the time of the Argonautic Expedition - On the Globe sits the Figure of Astronomy weeping, with a Sceptre in her Hand, (as Queen of the Sciences) and a Star over the Head of the Pyramid."

The frontispiece of the tomb displays a trefoil design, characteristic of the English innovations in Gothic design. As I hope to demonstrate later in this chapter, Newton's work draws heavily on the aesthetic ideas inherent to this period, hence, his tomb seems to be an entirely fitting memorial to his life and work.

Newton's work in calculus, and the priority dispute over its invention with Liebniz, provide an excellent introduction to the deep geometric and aesthetic ideas behind his thinking. A good survey of this subject can be found in my paper [10], but I will repeat a number of the essential ideas here.

The science of calculus is well known to any college student of mathematics. Its modern rigorous formulation is primarily due to the work of the 19th century mathematicians Cauchy, Riemann and Weierstrass. However, the geometrical ideas underlying the theory are, essentially, due to Newton. I hope to make this point of view clearer in the course of this chapter. The foundations of the calculus are the ideas of differentiation and integration. In order to express the modern formulations of these ideas, we need to first briefly introduce the mathematical idea of a *limit*, the modern formulation of which is essentially due to Weierstrass. Roughly speaking, we define the limit of a function $f(x)$, at a given point x_0, to be the value obtained by the function as the variable x approaches infinitely close to x_0. In the diagram,(insert diagram "limit"), we are given two elementary functions, $y = x^2$ and $y = sign(x)$, defined by;

$$sign(x) = \ \ 1 \text{ if } x > 0$$
$$sign(x) = -1 \text{ if } x < 0$$

The former function has a well defined limit at $x = 0$, namely $y = 0$. We express this by saying that $lim_{x \to 0} x^2 = 0$, ([10]). The latter function, however, has no well defined limit at $x = 0$, as when x approaches 0 from values strictly greater than 0, the function $sign(x)$

[10] It is possible for a function $f(x)$ to have a well defined limit at x_0, without this actually being the value of the function at x_0. In the given example of the function $y = x^2$, we could declare the function to take the value 1 at 0, without effecting the limit calculation. If $f(x_0) = lim_{x \to x_0} f(x)$, we say that the function is continuous at x_0

approaches 1, whereas, when x approaches 0 from values strictly less than 0, the function $sign(x)$ approaches -1. The concept of differentiation originated in the problem of finding the tangent line to a curve. In the diagram, insert diagram "tangents", we are given an ellipse and want to find the equation of the tangent line to the ellipse at the marked point O. The critical geometrical observation, we will discuss its origination in greater detail later, is that the tangent line is obtained as an approximation of lines of the form OP, where P approaches infinitely close to O. The modern formulation of this idea, due to the German mathematician Karl Wierstrass, is the following;

Definition 0.1. *Differentiation*

Let $f(x)$ be a real-valued continuous function on the open interval (a,b), then we say that $f(x)$ is differentiable on (a,b) if;

$lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$ $(*)$

exists for every $x \in (a,b)$. In this case, we define the derivative of f to be $\frac{df}{dx}$.

It is possible that this limit does not exist, a simple example is given by the function $y=|x|$, insert diagram "modulus". It is not possible to draw a tangent line to this function at the marked point 0, equivalently, the limit defined in $(*)$ is undefined. The idea of integration originated in computing the area under the arc of a curve, as a sequence of approximations of areas under a series of rectangles. The modern formulation is mainly due to the mathematicians Bernhard Riemann, and Henri Lebesgue, (see [4] and [36]);

Definition 0.2. *Integration*

Let $f(x)$ be a real-valued continuous function on the closed interval $[a,b]$, then, if $\epsilon_n = \frac{b-a}{n}$, and;

$s_n = \epsilon_n \sum_{j=0}^{n-1} f(a+j\epsilon_n)$

we define the integral;

$\int_a^b f(x)dx = lim_{n\to\infty} s_n$

The fundamental theorem of calculus relates the notions of differentiation and integration, showing that they are inverse procedures. Roughly speaking, if I begin with a given function $f(x)$, and then proceed to integrate and differentiate it, I return to the original one. The modern formulation and proof is credited to the French mathematician Augustin Louis Cauchy;

Theorem 0.3. *Fundamental Theorem of Calculus*

Let $f(x)$ be a real-valued continuous function on the closed interval $[a,b]$, then, if;

$F(x) = \int_a^x f(y)dy$

$F(x)$ is differentiable on the open interval (a, b) and continuous on the closed interval $[a, b]$. Moreover;

$\frac{dF}{dx}(x) = f(x)$ *for* $x \in (a, b)$

Proof:
The proof is a simple consequence of the definitions, we refer the reader to the paper [10] for more details.

As a result of the theorem, one obtains a simple method of computing integrals;

Theorem 0.4. *Suppose that $f(x)$ is a continuous function on a closed interval $[a, b]$, and $G(x)$ is an antiderivative of $f(x)$, that is a continuous function on $[a, b]$, with the property that $\frac{dG}{dx}(x) = f(x)$, on (a, b). Then;*

$\int_a^b f(x)dx = G(b) - G(a)$

Proof:
If F is the function given by the previous theorem, then the function $F - G$ is continuous and;

$\frac{d(F-G)}{dx} = \frac{dF}{dx} - \frac{dG}{dx} = f(x) - f(x) = 0$ on (a, b)

It follows easily that $F - G = c$, where c is a constant, and;

$\int_a^b f(x)dx = F(b) - F(a) = (G(b) - c) - (G(a) - c) = G(b) - G(a)$

Newton's approach to calculus is, in many ways, radically different from the modern approach. His method can be unravelled from the unpublished papers (ii), (iii) and (iv) that he wrote between 1665 and 1666, referred to above, collected in [38], and his published papers, (v),(vi),(vii),(xiv) and (xv). As we explained above, Newton's major work in the field of calculus occurred in close proximity to that of his contemporary, the German mathematician, Gottfried Liebniz, who published his results in 1684.

The first point of departure, in the methods of both Newton and Liebniz, with the modern approach is the replacement of the notion of a limit with that of an infinitesimal quantity. Roughly speaking, an infinitesimal quantity, which Newton usually denoted by o or δ, is a quantity which is non-vanishing, and yet smaller than any other finite quantity. Newton essentially justified such quantities on a geometric level, by using them to find the tangent line to a curve at a given point. In the paper (vii), sections 5 and 6 of the Introduction, we find his explanation of this method, see figure 1;

"5. Let the Ordinate BC advance from its Place into any new Place bc. Complete the Parallelogram BCEb, and draw the right Line VTH touching the curve in C, and meeting the two lines bc and BA produced in T and V: and Bb, Ec and Cc will be the Augments now generated of the Abciss AB, the Ordinate BC and the Curve Line ACc; and the Sides of the Triangle CET are in the *first Ratio* of these Augments considered as nascent, therefore the fluxions of AB, BC and AC are as the Sides CE, ET and CT of the triangle CET, and may be expounded by these same Sides, or, which is the same thing, by the sides of the Triangle VBC, which is similar to the triangle CET.

6. It comes to the same purpose to take the Fluxions in the *ultimate Ratio* of the evanescent Parts. Draw the right line Cc, and produce it to K. Let the Ordinate bc return into it's former place BC, and when the points C and c coalesce, the right line CK will coincide with the tangent CH, and the evanescent triangle CEc in its ultimate Form will become similar to the Triangle CET, and its evanescent Sides CE, Ec and Cc will be *ultimately* among themselves as the Sides CE, ET and CT of the other triangle CET, are, and therefore the Fluxions of the lines AB, BC and AC are in the same Ratio. If the points C and c are distant from one another by any small Distance, the right line CK will likewise be distant from the Tangent CH by a small Distance. That the right Line CK may coincide with the Tangent CH, and the ultimate Ratios of the lines CE, Ec and Cc may be found, the Points C and c ought to coalesce and exactly coincide. The very smallest Errors in mathematical Matters are not to be neglected".

Newton argues that, in order to find the slope of the tangent line to the curve at C, given by the ratio $\frac{ET}{EC}$, it is necessary to find the "ultimate" ratio $\frac{Ec}{EC}$, as c "coalesces and exactly coincides" with C. Clearly, if c were identified with C, there would be no "evanescent triangle" CEc for which to compute such a ratio. Whereas, if c and C are "distant from one another by any small Distance", one still only obtains a line CK, "distant from the tangent CH by a small Distance". Newton is suggesting at an infinitesimal quantity to solve the problem, and, indeed, in Section 11 of the same Introduction, he shows how to compute the gradient (or the derivative of Definition 0.1) of the function $f(x) = x^n$, which he refers to as its Fluxion;

"11. *Let the Quantity x flow uniformly, and let it be proposed to find the Fluxion of x^n.*

In the same Time that the Quantity x, by flowing, becomes $x + o$, the Quantity x^n will become $(x + o)^n$, that is, by the Method of infinite Series's, $x^n + nox^{n-1} + \frac{n^2-n}{2}oox^{n-2}+$ etc. And the Augments o and $nox^{n-1} + \frac{n^2-n}{2}oox^{n-2}+$ etc are to one another as 1 and $nx^{n-1} + \frac{n^2-2}{2}ox^{n-2} + etc$. Now let these Augments vanish, and their ultimate Ratio will be 1 to nx^{n-1}."

In other words, if the curve in the previous figure is given by the graph of the function $y = x^n$, Newton computes the gradient of the line Cc as the ratio;

$$\frac{y(x+o)-y(x)}{o} = \frac{(x+o)^n - x^n}{o}$$

He then expands the expression in the numerator, and cancels the quantities involving o. In order for this calculation to make sense, it is clearly necessary that o should represent a non-zero quantity. After arriving at the final expression;

$nx^{n-1} + \frac{n^2-n}{2}ox^{n-2} + \dots$

Newton then supposes that o may be taken to be so small, that it can be set to 0 in the above expression, leaving the final fluxion to be(correctly) nx^{n-1}.

On a purely logical level, there is a problem with Newton's argument in 11. Newton concerns himself with a single infinitesimal quantity o, which he needs to be both zero and non-zero at different stages of the calculation in 11. This logical paradox was heavily criticized by the philosopher George Berkeley, in his tract, "The Analyst; Or, A Discourse Addressed to an Infidel Mathematician" (1734);

"XIV...Hitherto I have supposed that x flows, that x hath a real increment, that o is something. And I have proceeded all along on that Supposition, without which I should not have been able to have made so much as one single Step. From that Supposition it is that I get at the increment of x^n, that I am able to compare it with the Increment of x, and that I find the Proportion between the two Increments. I now beg leave to make a new Supposition contrary to the first, i.e I will suppose that there is no Increment of x, or that o is nothing; which second Supposition destroys my first, and is inconsistent with it, and therefore with every thing that supposeth it. I do nevertheless beg leave to retain nx^{n-1}, which is an Expression obtained in virtue of my first Supposition, which necessarily presupposeth such Supposition, and which could not be obtained without it: All which seems a most inconsistent way of arguing, and such as would not be allowed of in Divinity"

There is also a similar latent logical paradox in Newton's argument $(5, 6)$ on tangent lines. On the one hand, Newton needs to find a non-vanishing triangle CEc, in order to compute the ratio $\frac{Ec}{EC}$, while, on the other hand, he needs this triangle to collapse to the point C, in order for this ratio to coincide with the slope of the tangent line given by $\frac{ET}{EC}$. In the same tract, Berkeley again observes this paradox;

"XXXIV...It is supposed that the Ordinate bc moves into the place BC, so that the Point c is coincident with the Point C; and the right Line CK, and consequently the Curve Cc, is coincident with the Tangent CH. In which case the mixtilinear evanescent Triangle CEc will, in its last form, be similar to the triangle CET: And its evanescent Sides CE,Ec and Cc will be proportional to CE, ET and CT the Sides of the Triangle CET. And therefore it is concluded, that the Fluxions of the lines AB,BC, and AC, being in the last Ratio of their evanescent Increments, are proportional to the Sides of the triangle CET, or, which is all one, of the triangle VBC similar thereunto. [NOTE: Introd. ad Quad. Curv.] It is particularly remarked and insisted on by the great Author, that the points C and c must not be distant from one another, by any the least interval whatsoever: But that, in order to find the ultimate Proportions of the Lines CE, Ec, and Cc (i.e. the Proportions of the Fluxions or Velocities) expressed by the finite sides of the triangle VBC, the points C and c must be accurately coincident, i.e one and the same. A Point therefore is considered as a

Triangle, or a Triangle is supposed to be formed in a Point. Which to conceive seems quite impossible. Yet some there are, who, though they shrink at all other Mysteries, make no difficulty of their own, who strain at a Gnat and swallow a Camel."

Leaving aside the additional vitriol that Berkeley pours into his argument, ([11]), the logical inconsistencies that he observes are valid. However, there is clearly a sense that Newton's geometrical intuition is correct. The rigorous formulation of the calculus in the 19th century was able to resolve the logical problems by replacing the notion of an infinitesimal quantity with that of a limit, arriving at Definition 0.1 to replace Newton's arguments on tangents that we have considered ([12]). However, the notion of a limit draws on certain topological properties of real numbers, that they are in a sense infinitely divisible, which, perhaps, Newton was keen to avoid. Indeed, he gives the following description of Time in (vi);

" 3. Fluxions are very nearly as the Augments of the Fluents generated in equal but very small Particles of Time,..."

In the last fifty years, the theory of infinitesimals has enjoyed a modern renaissance, in the field of what is now referred to as non-standard analysis. The first major pioneer in this area was Abraham Robinson, whose book "Non-Standard Analysis" is still a definitive account of the subject. More recently, the mathematicians Zilber and Hrushovski have further developed the theory of infinitesimals in the context of Zariski structures, finding new approaches to current problems in algebraic geometry, [39] is an excellent reference.

Robinson was able to resolve the paradox, observed by Berkeley, by finding a logically consistent structure $\mathcal{R}^*$, extending the real numbers $\mathcal{R}$, which contains infinitesimal elements. ([13]). In such a structure, every bounded element r^* has a uniquely defined standard part,

[11] Berkeley also attacks the logical contradictions involving succession of infinities, which are introduced by the use of infinitesimals. This problem is resolved for the case of complex algebraic curves in [15]. With some effort, one could provide similar arguments in the case of real algebraic curves.

[12] There is no possibility that Newton could have derived his arguments on tangents from Liebniz's work, even though his paper (vii) was published later than Liebniz's work on calculus. The reader can find one of Newton's first use of the infinitesimal o notation in his unpublished paper (iii), where he uses the method outlined above, to compute the tangent line of the curve defined by $rx + x^2 - y^2 = 0$;

"Now if the Equation expressing the relation of the lines x and y be $rx + xx - yy = 0$. I may substitute $x + o$ and $y + \frac{qo}{p}$ into the place of x and y because (by the lemma) they as well as x and y doe signify the lines described by the bodys A and B. By doeing so there results $rx + ro + xx + 2ox + oo - yy - \frac{2qoy}{p} - \frac{qqoo}{pp} = 0$. But $rx + xx - yy = 0$ by supposition: there remaines therefore $ro + 2ox + oo - \frac{2qoy}{p} - \frac{qqoo}{pp} = 0$. Or divideing it by o tis $r + 2x + o - \frac{2qy}{p} - \frac{oqq}{pp} = 0$. Also those terms in which o is are infinitely less than those in which o is not therefore blotting them out there rests $r + 2x - 2\frac{qy}{p} = 0$. Or $pr + 2px = 2qy$."

Here, the quantities p and q denote the fluxions $\frac{dx}{dt}$ and $\frac{dy}{dt}$. Hence, Newton derives the correct formula for the derivative $\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{q}{p} = \frac{r+2x}{2y}$.

[13] The reader should look at the paper [11], for the construction of $\mathcal{R}^*$. For technical reasons, it is convenient to work with what I refer to as analytic non-standard extensions, see [11], this assumption was in force throughout the paper [10]. The construction is carried out more rigorously in [16], where an important logical step is the use of compactness, see [21], and saturation, see [27], to guarantee the existence of infinitesimals and other technical logical properties such as "overflow" and "underflow".

which he denoted $st(r^*)$, with the property that the difference $r - st(r^*)$ is an infinitesimal. Real analytic functions $f(x)$, defined on $\mathcal{R}$, extend to well-defined functions on the non-standard structure $\mathcal{R}^*$. Each element $r \in \mathcal{R}$ has a set of elements which are "infinitely close" to it, in $\mathcal{R}^*$, which is now called an infinitesimal neighborhood $\mathcal{V}_r$;

Definition 0.5. *Infinitesimal Neighborhood*

If $r \in \mathcal{R}$, we define its infinitesimal neighborhood $\mathcal{V}_r$, to be;

$\{r^* \in \mathcal{R}^* : st(r^*) = r\}$

One may then give a consistent definition of differentiation, using infinitesimals, in the following way;

Definition 0.6. *Non-Standard Differentiation*

Let $f(x)$ be a real-valued analytic function on the open interval (a, b), then we say that $f(x)$ is non-standard differentiable on (a, b), if, for every $x \in (a, b)$, there exists $c_x \in \mathcal{R}$, such that;

$\frac{f(x+o)-f(x)}{o} \in \mathcal{V}_{c_x}$, *for every infinitesimal $o \in \mathcal{V}_0$.*

We then define the non-standard derivative of f at x to be c_x.

It is not hard to show that the non-standard definition of differentiation is equivalent to the modern Definition 0.1. The interested reader should look at Robinson's book [29] for a detailed account of his construction or my notes [11]. With this definition, it is easy to see that Newton's calculation, in (5,6) of (vii), is no longer paradoxical. If the ordinate b is taken an infinitesimal distance away from B, the gradient of the corresponding line Cc lies infinitely close to the gradient of the tangent line CT. By Robinson's construction, the gradient of the tangent CT is then determined uniquely from this information.

Although the logical problems with Newton's original method are now resolved, there is still something geometrically unsatisfactory about the resulting use of infinitesimals. This originates in Newton's comment from (6) of (vii), that "the points C and c ought to coalesce and exactly coincide". In the Definition 0.1 of differentiation using limits, given any prescribed neighborhood of B, it is possible to take the ordinate b to any point within this neighborhood. In this sense the line Cc genuinely approaches and converges to the tangent line CT. This is not the case with infinitesimals. If I was to choose two distinct infinitesimal quantities, say $\{o, o'\}$, consider the ordinates $B + o$ and $B + o'$, and the corresponding values $\{c, c'\}$, then, although the gradients of the lines Cc and Cc' both lie infinitely close to the gradient of the line CT, it is not important that one gradient lies closer to the gradient of CT than the other. In this picture, there is no *motion* of the line Cc towards the tangent line CT.

That Newton intended the geometrical picture of a sequence of lines converging towards the true tangent line, is supported by his description of mathematical quantities in 1. of (vii);

"I consider mathematical Quantities in this Place not as consisting of very small Parts; but as decrib'd by a continued Motion. Lines are describ'd, and thereby generated not by the Apposition of Parts, but by the continued Motion of Points..."

his description of fluxions as velocities in 2. of (vii);

"Therefore considering that Quantities, which increase in equal Times, and by increasing are generated, become greater or less according to the greater or less Velocity with which they increase and are generated; I sought a Method of determining Quantities from the Velocities of the Motions or Increments, with which they are generated; and calling the Velocities of the Motions or Increments, with which they are generated; and calling these Velocities of the Motions or Increments Fluxions, and the generated Quantities Fluents, I fell by degrees upon the Method of Fluxions..."

and his description of the Method of Fluxions in (vi);

"In Finite Quantities so to frame a calculus, and thus to investigate the Prime and Ultimate Ratios of Nascent or Evanescent Finite Quantities, is agreeable to the Ancients; and I was willing to shew, that in the Method of Fluxions there's no need of introducing Figures infinitely small into Geometry. For this Analysis may be performed in any Figures whatsoever, whether finite or infinitely small, so that they are imagined to be similar to the Evanescent Figures..."

Professor Goldblatt makes the important observation in [19], that this last passage supports the view that Newton was prepared to dispense with the use of infinitesimals, if he had a coherent notion of a limit available. In his later paper (xv), Newton developed what he referred to as "The method of first and last ratios of quantities", in which he comes close to formulating a reasonable definition of a limit, and, therefore, avoiding the logical paradoxes of infinitesimals;

From (viii), Lemma 1 of Section 1, Book 1; "Quantities, and the ratios of quantities, which in any finite time converge continually to equality, and before the end of time approach nearer to each other than by any given difference, become ultimately equal.

If you deny it, suppose them to be ultimately unequal, and let D be their ultimate difference. Therefore they cannot approach nearer to equality than by that given difference D; which is against the supposition."

In the case of ratios of quantities, if one takes the definition of "ultimately equal" to be the one provided by his definition of fluxions in Section 11 of (vi), that we considered above, and the definition of a limit as that provided by the remaining statement of the lemma, then Newton's proof attempts to show that these definitions coincide. Again, both the definition and proof are unrigorous by modern standards, but further support the view that Newton favoured the use of the geometrical model using limits, introducing infinitesimals as a practical technical solution.

In my opinion, the geometrical limit model that Newton uses draws on the aesthetic idea of the sublime that we considered in the previous chapter. Namely, Newton relies on the notion of an infinity, "an ultimate ratio", which is, in itself, unattainable, but may be approached through a series of finite calculations. As we will see later in the chapter, Newton refines this aesthetic idea, in his use of asymptotes as a method of analysing algebraic curves.

There is an alternative, more geometrically satisfying picture of tangency, which preserves the use of infinitesimals, see figure 2. The aesthetic idea behind this geometric model belongs to the following chapter, and we will consider its geometric implications in greater detail, when we consider Severi's work. However, as will become clearer below, Newton also considered this model, ([14]) and, hence, it is natural to consider it now, to the extent that it features in Newton's work. In this sense, the remark that he made on Time from (vi), which we considered above, reflects a more profound understanding of the geometric significance of the use of infinitesimals. In this example, I consider two complex ([15]) plane algebraic curves C and D, that is curves defined by the polynomials;

$p(x,y) = \sum_{(i+j)\leq n} a_{ij}x^i y^j$, with $a_{ij} \in \mathcal{C}$

$q(x,y) = \sum_{(i+j)\leq m} b_{ij}x^i y^j$, with $b_{ij} \in \mathcal{C}$

in the complex plane $\mathcal{C}^2$, intersecting in a point O, and sharing a common tangent line l. ([16]). I now consider what happens if I vary the coefficients of the polynomials by infinitesimal amounts([17]), that is I choose infinitesimal quantities $\{\epsilon_{ij} : i+j \leq n\}$, $\{\delta_{ij} : i+j \leq m\}$ and consider the curves C_ϵ and D_δ, defined by the polynomials;

$p_\epsilon(x,y) = \sum_{(i+j)\leq n}(a_{ij} + \epsilon_{ij})x^i y^j$

$q_\delta(x,y) = \sum_{(i+j)\leq m}(b_{ij} + \delta_{ij})x^i y^j$

In general, ([18]), one would expect to obtain 2 points of intersection, marked by the points $\{u, v\}$ in the diagram, which are at an infinitesimal distance from the point O. This is, in fact, an intuition that Newton observes in his discussion of curvature from (vi);

[14] Newton, therefore, had two geometric approaches to the theory of tangency

[15] This is a different situation to real plane curves, which Newton studied extensively in the cited papers (ii)-(vii). We will consider this case later in the chapter

[16] If a plane algebraic curve C, (real or complex), is defined by a polynomial $p(x,y)$, with $p(x_0,y_0) = 0$, then we say that C is singular at (x_0,y_0), if $\frac{\partial p}{\partial x}(x_0,y_0) = \frac{\partial p}{\partial y}(x_0,y_0) = 0$, otherwise we say that C is non-singular. If C is non-singular at (x_0,y_0), we define its tangent line l, to be the line defined by the equation $\frac{\partial p}{\partial x}(x_0,y_0)x + \frac{\partial p}{\partial y}(x_0,y_0)y = 0$. In the particular case of a real function $y = f(x)$, which we have considered, this rule gives the tangent line l to be $\frac{df}{dx}(x_0)x - y = 0$, as we saw in the explanation of differentiation above.

[17] The construction of non-standard extensions of the complex numbers $\mathcal{C}$ may be done algebraically. The reader should look at my paper [12] for more details.

[18] By which I mean, the infinitesimal quantities are chosen generically, that is there are no algebraic relations between the elements of the tuple $\{\bar{a}+\bar{\epsilon}, \bar{b}+\bar{\delta}\}$, and the tuple $\{\bar{a}, \bar{b}\}$ defining the original curves C and D is generic in the space of curves, tangent to the line l at O

"54. And now I have finish'd the Problem; but having made use of a Method which is pretty different from the common ways of operation, and as the Problem itself is of the number of those which are not very frequent among Geometricians: For the illustration and confirmation of the Solution here given, I shall not think much to give a hint of another, which is more obvious, and has a nearer relation to the usual Methods of drawing Tangents. Thus if from any Center, and with any Radius, a Circle be conceived to be describd, which may cut any Curve in several points; if that Circle be suppos'd to be contracted, or enlarged, till two of the Points of intersection coincide, it will there touch the Curve. And besides, if its Center be suppos'd to approach towards, or recede from, the Point of Contact, till the third Point of intersection shall meet with the former in the Point of Contact; then will that Circle be aequicurved with the Curve in that Point of Contact..." ([19])

A more precise formulation of this geometric intuition is the following;

Theorem 0.7. *Let C and D be plane complex algebraic curves, intersecting at a point O, which is non-singular for both curves. Then C and D share a common tangent line at O iff;*

For a generic choice of infinitesimal quantities $\{\bar{\epsilon},\bar{\delta}\}$;

$$Card(C_{\bar{\epsilon}}\cap D_{\bar{\delta}}\cap\mathcal{V}_O)\geq 2$$

The proof of this result may be found in my paper ([13]), ([20]) One can reformulate this result in the particular case of a curve intersecting a line;

Theorem 0.8. *New Geometric Formulation of Tangency*

If C is an irreducible complex algebraic curve, in particular, if C is the graph of a polynomial function $f(x)$, that is an algebraic curve of the form $y-f(x)=0$, passing through the point $O=(0,0)$, then a line of the form $y=c_O x$ is tangent([21]) to the curve at O, iff, for any non-zero choice of infinitesimal $\epsilon\in\mathcal{V}_l$, there exists a non-empty collection of points $\{O_1(\epsilon),\ldots,O_n(\epsilon)\}$, distinct from O, such that;

$$C\cap(y-(c_0+\epsilon)x=0)\cap\mathcal{V}_O=\{O_1(\epsilon),\ldots,O_n(\epsilon)\}$$

The proof of this result may be found in the paper [14]. The reader should consider figure 3, in which, by moving the tangent line l to the curve C at O, to the new position l_ϵ, one

[19] Newton defines the centre of curvature in (i) and (iii), as the limit meet of normals to a curve C. The proof that this is well defined is intrinsically connected to his proof of "The Fundamental Theorem of Calculus", that we will consider below.

[20] In this paper, I show an even more general result; that the notion of algebraic intersection multiplicity between curves C and D, coincides with a non-standard notion of intersection multiplicity, defined, as in Theorem 0.7.

[21] One may still formulate a coherent definition of tangency for singular points on an irreducible curve C, the reader should look at the paper [14] for such a definition.

obtains two new intersection points $\{A(\epsilon), B(\epsilon)\}$ infinitely close to the original point O. In this geometric picture, it is not important that these points converge to the fixed point O, but only that they are *released* from the original position O. The distinction is a subtle one, but important for more advanced geometrical constructions. Moreover, it is the most natural geometric picture to use, in conjunction with infinitesimal quantities, and one which Newton, at least superficially, seems to have considered. The method is explored, in greater detail, in the paper [14].

In order to extend these considerations to the case of real plane curves, which Newton considered, that is curves defined by a polynomial $p(x, y)$ in the plane $\mathcal{R}^2$, one needs to overcome certain technical difficulties resulting from the existence of real solutions to such polynomials. A simple example is given by the polynomial $x^2+1=0$, which has *no* solutions in $\mathcal{R}^2$. One would, therefore, hesitate to call this a curve. Another problematic example is given by a polynomial such as $x^2+y^8+x^4+y^2=0$, which only has one real solution at $(0,0)$, due to the fact that this is a singular point for the polynomial. In such cases, it is impossible to formulate a coherent notion of tangency. The technical solution to this problem may be found in my paper [10], which the reader is encouraged to read, as it uses Newton's method of constructing power series solutions to polynomial equations, first developed in (v), which we will consider later in this chapter. The final result obtained in [10] is the following;

Theorem 0.9. *Let C be a real plane curve, such that $O=(0,0)$ is a non-singular point for C, then a line of the form $y=c_O x$ is tangent to the curve at O, iff, in a non-standard extension of $\mathcal{R}$;*

Either, *for any positive infinitesimal ϵ, there exists at least one solution $O(\epsilon)$, distinct from O in;*

$$C \cap (y-(c_O+\epsilon)x=0) \cap \mathcal{V}_O$$

Or, *for any positive infinitesimal ϵ, there exists at least one solution $O(-\epsilon)$, distinct from O in;*

$$C \cap (y-(c_O-\epsilon)x=0) \cap \mathcal{V}_O$$

In this theorem, one should think of a line passing through the curve at O as rotating, either "clockwise" or "anti-clockwise" about the fixed point O. Tangency is then characterised by an intersection $O(\epsilon)$, moving away continuously from the fixed point O, along the curve, in one of these cases. This form of definition seems to improve on the slightly clumsy use of infinitesimals appearing in Newton's original definition of tangency, and, from previous remarks, is one which we he considered.

The methods of Newton and Liebniz also differ considerably from the modern approach to theory of integration. Both formulated a non-standard definition of integration, based on the idea of finding the area under a curve by summation over a series of rectangles of infinitely small width, see figure 4. The formal non-standard definition is the following;

Definition 0.10. *Non-Standard Integration*

Let $f(x)$ be a continuous function on the closed interval $[a, b]$, and let R_f be the Riemann sum, defined for a real number with $0 < c < (b - a)$, by;

$R_f(c) = \sum_{j=0}^{N(c)} f(a + jc)c$

where $N(c)$ is the greatest positive integer n such that $(a + nc) < b$. Then, we define;

$\int_a^b f(x)dx = st(R_f(\epsilon))$

where ϵ is a positive infinitesimal.

The proof that this is a good definition, that is doesn't depend on the choice of infinitesimal ϵ, and gives the same value as Definition 0.2, can be found in my notes [11]. Arguably, Liebniz was the first to formulate this definition in what would be called a rigorous way. However, that the geometric idea was known to Newton, before Liebniz's publications on calculus, is clear from his proof of the Fundamental Theorem of Calculus, given in his unpublished documents (i), see p304 of [38], which we will consider shortly.

Newton's first published explanation of integration can be found in (vii). Newton uses the term quadrature to mean integration, the term quadrature referring obliquely to approximating the area under a curve by a series of quadrangles. However, the paper is essentially a tabulation of integrals for various curves, and there is no recognisable proof of the Fundamental Theorem of Calculus (Theorem 0.3) and its corollary (Theorem 0.4), which allows one to compute explicit integrals. Newton gives a much clearer explanation of quadrature in his earlier paper (v), see figure 5;

" The Demonstration of the Quadrature of Simple Curves belonging to Rule the first.

Preparation for demonstrating the first Rule.

54. Let then $AD\delta$ be any curve whose Base $AB = x$, the perpendicular Ordinate $BD = y$, and the area $ABD = z$, as at the Beginning. Likewise put $B\beta = o$, $BK = v$; and the Rectangle $B\beta HK(ov)$ equal to the Space $B\beta\delta D$.

Therefore it is $A\beta = x + o$, and $A\delta\beta = z + ov$: Which Things being premised, assume any Relation betwixt x and z that you please, and seek for y in the following Manner.

Take at Pleasure $\frac{2}{3}x^{\frac{3}{2}} = z$; or $\frac{4}{9}x^3 = z^2$. Then $x + o$ $(A\beta)$ being substituted for x, and $z + ov$ $(A\delta\beta)$ for z, there arises $\frac{4}{9}$ into $x^3 + 3xo^2 + 3xo^2 + o^3 =$ (from the Nature of the Curve) $z^2 + 2zov + o^2v^2$. And taking away Equals ($\frac{4}{9}x^3$ *and* z^2) and dividing the Remainders by o, there arises $\frac{4}{9}$ into $3x^2 + 3xo + oo = 2zv + ovv$. Now if we suppose $B\beta$ to be diminished infinitely and to vanish, or o to be nothing, v and y, in that Case will be equal, and the Terms which are multiplied by o will vanish: So that there will remain $\frac{4}{9} \times 3x^2 = 2zv$, or $\frac{2}{3}x^2 (= zy) = \frac{2}{3}x^{\frac{3}{2}}y$; or $x^{\frac{1}{2}} = \frac{x^2}{x^{\frac{3}{2}}} = y$. Wherefore conversely if it be $x^{\frac{1}{2}} = y$, it shall be $\frac{2}{3}x^{\frac{3}{2}} = z$.

55. Or universally, if $\frac{n}{m+n} \times ax^{\frac{m+n}{n}} = z$; or, putting $\frac{na}{m+n} = c$, and $m + n = p$, if $cx^{\frac{p}{n}} = z$; or $c^n x^p = z^n$: Then, by substituting $x + o$ for x, and $z + ov$ (or which is the same $z + oy$) for

z, there arises c^n into $x^p + pox^{p-1}$, etc $= z^n + noyz^{n-1}etc$, the other Terms, which would at length vanish being neglected. Now taking away $c^n x^p$ and z^n which are equal, and dividing the Remainders by o, there remains $c^n px^{p-1} = nyz^{n-1}(= \frac{nyz^n}{z}) = \frac{nyc^n x^p}{cx^{\frac{p}{n}}}$, or, by dividing by $c^n x^p$, it shall be $px^{-1} = \frac{ny}{cx^{\frac{p}{n}}}$; or $pcx^{\frac{p-n}{n}} = ny$; or by restoring $\frac{na}{m+n}$ for c, and $m+n$ for p, that is m for $p-n$, and na for pc, it becomes $ax^{\frac{m}{n}} = y$. Wherefore conversely, if $ax^{\frac{m}{n}} = y$, it shall be $\frac{n}{m+n}ax^{\frac{m+n}{n}} = z$. Q.E.D"

Here, Newton demonstrates how to find the quadrature of simple curves. If a curve is given by the equation $y = x^{\frac{m}{n}}$, in Article 55, he deduces correctly the formula for the integral of the curve as $\frac{n}{m+n}x^{\frac{m+n}{n}}$. The argument he gives for this formula, however, would not be called rigorous by modern standards. In the diagram pertaining to the problem which Newton gives, figure 5, Newton sets the area $B\beta\delta D$ to be $o.v$, where $o = length(B\beta)$ and $v = length(BK) = y$ (†). Assuming the quadrature(area) of the curve (area(ABD)) can be expressed as a polynomial expression $z(x) = x^q$, q rational, of the base $x = length(AB)$, Newton then forms the equation;

$$z + o.v = z(x+o) = (x+o)^q$$

By expanding this expression and setting the term o to be nothing, (††), he then derives the expression $y(x) = qx^{q-1}$. Finally, he argues that these steps may be reversed to give the formula for quadrature, given only the equation of the curve $y(x)$, (†††). The first argument (†) relies on the intuition that, as $length(B\beta)$ becomes sufficiently small, the quadrangle $B\beta HK$ is a good approximation to the area under the curve $BD\delta\beta$. In order to make the idea rigorous, one needs to formulate a definition of integration, using infinitesimals, similar to that given in Definition 0.10. However, the intuition is still one of two main geometric components behind such a definition, and, as we will find the other geometric component in an earlier unpublished paper, it is reasonable to say that Newton had, at this stage, formulated a clear geometric idea of integration, although, he fails, here, to formulate the idea in its entirety. The second argument (††) contains the germ of the idea of fluxions, but the exposition of it, here, is unclear. Finally, the third argument (†††), suffers from the same deficiencies as the first (†), that of a clear definition of integration, which is required, here, to make the argument rigorous.

The contents of this paper, in particular the argument given here, was the centre of the later dispute, as to whether Liebniz had obtained the idea of the differential calculus from Newton, previous to his publication of 1684. However, it is interesting to note that in Liebniz's "Excerpta from Newton's De Analisi",([22]), October 1676, he makes no comment whatsoever on this fundamental passage. Given the lack of justification in certain steps of Newton's argument, it seems reasonable that Liebniz's formulation of a rigorous definition of integration and differentiation, using infinitesimals, was a genuine independent achievement. This last conclusion is similar to that which Charles Bossut gives in "A General History of Mathematics from the Earliest Times to the Middle of the Eighteenth Century" (1802);

[22] De Analisi was the latin title of Newton's paper (iv)

"All these considerations appear to me to evince that, if the piece De Analysi per Aequationes and the letter of 1672 contain the method of fluxions, it was at least enveloped in great darkness." ([23])

I wish to now consider an argument that Newton gives in his paper (ii), see figure 6, which I will, first, quote in full;

Prop:

Haveing an equation of 2 dimensions to find w^t crooke line it is whose area it dothe expresse. suppose y^e equation is $\frac{x^3}{a}$. naming y^e quantitys, $a = dh = kl$. $bg = y$. $db = mk = x = gp$. y^e superficies $dbg = \frac{x^3}{a}$. supose y^e square $dhkl$ is equall to y^e superficies gbd; y^n $dk = z = bm = lh = \frac{x^3}{aa}$, and $aaz = x^3$. w^{ch} is an equation expressing y^e nature of y^e line fmd.
Next makeing $nm = s$ a line w^{ch} cutteth dmf at right angles. $nd = v$.

$ss - vv + 2vx - xx = \frac{x^6}{a^4} = mb$ squared.

0. 0. 1. 2. 6.

(w^{ch} is an equation haveing 2 equall rootes and therefore multiplyed according to Huddenius his method, produceth another.)

$2vx = 2xx + \frac{6x^6}{a^4}$.

$v = x + \frac{3x^5}{a^4}$. and $nb = v - x = \frac{3x^5}{a^4}$.

Now supposeing $mb : bn :: dh : bg$. that is, $\frac{3x^5}{a^4} : \frac{x^3}{aa} :: y : a$. $3xx = ay$ and $3xxa = a^2y$. Which is y^e nature of y^e line dgw. and y^e area $dbg = dklh = \frac{x^3}{aa}$, makeing $db = x$. $dh = a$. Or. $diw = deoh = \frac{x^3}{a}$, determining (di) to be (x), etc.

The Demonstration whereof is as followeth.

Suppose $w\Pi\Omega$, Ωmz, zfv etc are tangents of y^e line dmf. and from theire intersections z, Ω, v, w draw $va, zq, \Omega s, wx$ and from theire touch points draw $fw, mg, \Pi\xi$ all parallell to kp. also from y^e same points[s] of intersection draw $v\sigma, z\lambda, \Omega v, \omega\zeta$. And $mb : nb :: bt : bm :: \Omega\beta : \beta m :: kl : bg$. wherefore $\Omega\beta \times bg = \beta m \times kl$. that is y^e rectangle $kl\nu\mu = b\rho sg$. And $\pi\rho s\breve{o} = \theta\lambda\nu\mu$. in like manner it may be demonstrated y^t $aq\pi n = \theta\lambda\sigma\rho$, and $\rho\omega xy = \mu d\nu h$ etc so y^t y^e rectangle $\rho\sigma hd$ is equall to any number of such like squares inscribed twixt y^e line $n\psi$ and y^e point d, w^{ch} squares if they bee infinite in number, they will bee equall to y^e

[23] The letter referred to is one that Newton wrote to Collins, claiming to have found a general method of finding the tangent to a curve, but without giving any demonstration. As we have also observed, the method of quadrature(integration), is also obscure. Given the argument which we will establish below, that Newton had, by this stage, a full understanding of the methods of calculus, it seems likely that this obscurity was deliberate.

superficies $dn\psi\omega g\xi$, $(*)$. ([24])

The argument is essentially a proof of The Fundamental Theorem of Calculus. However, before considering it in detail, we will make a preliminary observation. This is the argument $(*)$, at the end, that the area $dn\psi\omega g\xi$ (that is the area above the lower curve between d and n) is equal to the sum of an infinite number of rectangles inscribed between the lower curve and the axis rdh. This observation, together with the argument we considered above, from Newton's paper (v), suggest that Newton had, at least, an intuitive idea of a formal definition of integration, using infinitesimals, by the time he wrote (v), ([25]). Given his use of infinitesimal arguments in the context of differentiation, and his proofs, here and in (v), of The Fundamental Theorem of Calculus, it seems clear that Newton had formulated his own version of the calculus, by 1669, independently from Liebniz, and, even at the same level of precision. (The logical paradoxes observed by Berkeley are problematic for both the work of Liebniz and Newton).

Now, considering the argument in more detail, I will show that not only does Newton give a reasonable proof of The Fundamental Theorem of Calculus, but also one which is geometrically superior to the modern proofs of the result. Newton begins his argument, by setting the fixed length dh to be a, the area dbg to be $\frac{x^3}{a}$, the length of the ordinate db to be x, the length of the coordinate bg to be y and the length of the coordinate bm to be $\frac{x^3}{a^2}$. In modern terminology, Newton takes the bottom curve to be described by the equation $y(x)$ and the function;

$z(x) = \frac{1}{a}\int_0^x y(x)dx$ $(\dagger)$

which describes the top curve. He assumes that $z(x) = \frac{x^3}{a^2}$ $(**)$. Newton now makes the unusual step of drawing the *normal* to the top curve at m. Setting $length(nm) = s$ and $length(nd) = v$, he calculates;

$(v-x)^2 + length(mb)^2 = s^2$ (Pythagoras' Theorem)

deriving the formula;

$length(mb)^2 = s^2 - (v-x)^2 = s^2 - v^2 + 2vx - x^2 = (\frac{x^3}{a^2})^2 = \frac{x^6}{a^4}$ $(\dagger\dagger)$

He then differentiates the expression $(\dagger\dagger)$, to obtain;

---

[24]Newton uses the shorthand notation w^t for which, y^e for the, y^t for that, and y^n for then. I have made the occasional modifications to Newton's original text, as done in footnotes 48-50 of Whiteside's commentary on the paper. Finally, one should observe that the letter ζ in Newton's argument coincides with the letter h of the attached diagram. In an original, cancelled version of the figure, Newton drew the line ωd distinct from rdh, and $\omega\zeta$ a little above the horizontal axis rdh, see footnote 47 of Whiteside. Also the letter n denotes both the intersection of the vertical line $va\psi$ with the horizontal axis rdh, *and* the intersction of the normal to the curve at m with the base rdh, see Whiteside's footnote 45 in [38]

[25]In his paper (xv), Lemmas 2,3 of Section 1, Book 1, Newton gives a much clearer account of a formal definition of integration. Although he previously introduces the notion of a limit in this paper, his definition is closer to the previous Definition 0.10 using infinitesimals, than Definition 0.2, using limits

$(v-x)=\frac{3x^5}{a^4}$[26]

In modern terminology, $v-x$ gives the length of the subnormal to the top curve defined by $z(x)$. The general formula for the length of the subnormal is $z\frac{dz}{dx}$, (***),[27] as, indeed, Newton deduces correctly in this particular case.

Newton continues by assuming that $\frac{length(bn)}{length(bm)}=\frac{length(bg)}{length(dh)}$. By $(***)$, this is equivalent to;

$\frac{dz}{dx}=\frac{z\frac{dz}{dx}}{z}=\frac{y}{a}$ $(****)$

By Newton's previous argument, which we reformulated in $(\dagger)$, the assumption is exactly a statement of The Fundamental Theorem of Calculus, Theorem 0.3. From this assumption, and the explicit equation of the function $z(x)$, given in $(**)$, Newton is then able to derive the formula for the original "crooked line" $y(x)$, namely $y(x)=\frac{3x^2}{a}$, which was the purpose of the original proposition.

The crux of Newton's argument is, then, contained in his proof of $(****)$, "The Demonstration whereof is as followeth". Newton begins by drawing a series of tangent lines to the top curve, centred at the points $\{f, m, \Pi\}$. From the intersections of these tangent lines and the points of tangency themselves, he draws a series of lines, parallel to the lines defined by kp and dh. He then claims that the ratio $\Omega\beta : \beta m$ is equal to the ratio $kl : bg$, from which he deduces that $\Omega\beta \times bg = \beta m \times kl$, $(\dagger\dagger\dagger)$. [28] The rectangles formed by the series of lines form a partition, which approximates the area under the bottom curve. Using the equality $(\dagger\dagger\dagger)$, he shows that the area defined by this partition is equal to the area defined by the rectangle $\rho\sigma hd$. By refining the partition, that is taking an infinite number of tangent lines and intersections, he deduces the formula $(\dagger)$.

The circularity in Newton's argument, see footnote 25, is easily remedied by, instead, defining the function $z(x)$ by the formula $(****)$ and deducing the formula $(\dagger)$. A more precise version of Newton's geometric idea is formulated and proved in [10], which the reader is encouraged to read. For convenience, we state the result, although the interest is in the mechanism of the proof.

Theorem 0.11. *Newton's Version of The Fundamental Theorem of Calculus*

[26] The reference to Huddenius' method is unclear. The sequence of numbers 0.0.1.2.6 refer to the weights of x in the expression $(\dagger\dagger)$

[27] The proof is an elementary exercise in differentiation and trigonometry, which we give for the convenience of the reader. Suppose that $z(x)$ defines a differentiable function and let (x_0, z_0) be a fixed coordinate on the curve C defined by $z-z(x)=0$. The vector defining the tangent to C at (x_0, z_0) is given by $(1, \frac{dz}{dx}|_{(x_0,z_0)})$. Hence, if (α, β) is the vector defining the normal to C at (x_0, z_0), we obtain, by the use of the dot product, $\alpha+\beta\frac{dz}{dx}|_{(x_0,z_0)}=0$. Hence, $\frac{\beta}{\alpha}=\frac{-1}{\frac{dz}{dx}|_{(x_0,z_0)}}$. The equation of the normal through (x_0, z_0) is given by $z-z_0=-\frac{(x-x_0)}{\frac{dz}{dx}|_{(x_0,z_0)}}$. The point of intersection x_1 of this line, with the axis $z=0$, is, therefore, $x_1=x_0+z_0\frac{dz}{dx}|_{(x_0,z_0)}$, hence, the length of the subnormal is $x_1-x_0=z_0\frac{dz}{dx}|_{(x_0,z_0)}$, as required.

[28] Unfortunately, this claim is equivalent to the assumption $(****)$ that he is trying to demonstrate. However, we will show presently how to remedy Newton's argument.

Let $y(x)$ be an analytic, ([29]), function, defined on the closed interval $[a,b]$, and let $z(x)$ be an analytic function, defined on $[a,b]$, with the additional property that $\frac{dz}{dx}=y$. Then;

$\int_a^b y(x)dx = z(b)-z(a)$

Newton's Version of the Fundamental Theorem of Calculus, Theorem 0.11, although slightly different in structure to Theorem 0.3, is equivalent for analytic functions. We give the proof, here, for the convenience of the reader;

Theorem 0.12. *Newton's Version of The Fundamental Theorem of Calculus and Theorem 0.3 are equivalent.*

Proof:
Suppose that $f(x)$ is an analytic function on the closed interval $[a,b]$, and we have shown Theorem 0.11. Let $F(x)$ be the function defined in Theorem 0.3, and let $G(x)$ be an antiderivative of $f(x)$, that is $G'(x)=f(x)$ on $[a,b]$, ([30]). By Theorem 0.11, we have that $F(x)=G(x)-G(a)$. In particular, we obtain $F'(x)=G'(x)=f(x)$, for $x\in(a,b)$. Hence, Theorem 0.3 holds for such an analytic function f. Conversely, suppose that Theorem 0.3 is shown for f, let $G(x)$ and $F(x)$ be as defined above, then we obtain that both $G'(x)=F'(x)=f(x)$ on (a,b). Applying elementary results, not depending on the theory of integration, we have that $G(x)=F(x)+c$, where c is a constant. Then $G(b)-G(a)=F(b)-F(a)=\int_a^b f(x)dx$, by definition of F. Therefore, Theorem 0.11 holds for the analytic function f as well.

The attentive reader may, at this stage, be wondering why Newton makes the step of introducing the normal to the curve defined by $z(x)$ in his Proposition which we considered above. Although unnecessary for his calculation to go through, Newton makes a connection between his definition of curvature, see footnote 16, and his proof of The Fundamental Theorem of Calculus. The relationship is the following;

(i). His calculus proof depends on the fact that;

For a given analytic function $f(x)$, if l_{x_0} denotes the tangent line to the curve C at $O=(x_0,f(x_0))$, defined by f, then, for an infinitesimal ϵ, if $l_{x_0+\epsilon}$ denotes the tangent line to the curve at $(x_0+\epsilon,f(x_0+\epsilon))$, the intersection $O_\epsilon = l_{x_0}\cap l_{x_0+\epsilon}$ lies in the infinitesimal neighborhood $\mathcal{V}_O$.

(ii). In his definition of curvature;

With the same conditions on f, if n_{x_0} denotes the normal to the curve at $O=(x_0,f(x_0))$, defined by f, then, for an infinitesimal ϵ, if $n_{x_0+\epsilon}$ denotes the normal to the curve at

[29] The formal definition of analytic is given in [10], but, roughly speaking, it is a function defined by power series, which Newton introduced in (v)

[30] Such a function is easily constructed for an analytic function f, by integrating each term of its power series expansion. This technique was also developed by Newton in (vi).

$(x_0+\epsilon, f(x_0+\epsilon))$, the intersection $N_\epsilon = n_{x_0} \cap n_{x_0+\epsilon}$ lies in the infinitesimal neighborhood $\mathcal{V}_K$, where K is the centre of curvature of the curve C at O.

This connection supports the idea that Newton's work in geometry is deeply related to his research into Optics, a connection which will develop further later in the chapter. The interested reader can find out more about Newton's work on curvature in, for example, the mathematical papers (i),([31]), and (iii) , and his more scientific papers $(xiii)$ and (xv).

Although the mechanism of Newton's proof of The Fundamental Theorem of Calculus is longer than the modern approach, Theorem 0.3, the geometric idea behind it is more elegant, for the following reasons. First, his argument relies on a "global" geometric relationship between the function $y(x)$ and its antiderivative $z(x)$, thus the global definition of integration is incorporated directly into the proof. In Theorem 0.3, a local property of the integrated function $y(x)$ is used, which is, on a geometric level, slightly unsatisfactory. Secondly, as we have just noted, Newton explores, in his argument, a connection between curvature and integration. The notion of locus of curvature may be defined for any algebraic curve C, using Newton's method, ([32]). Understanding the geometry of the locus of curvature was, for Newton, see footnote 29, intrinsically related to understanding the geometry of the curve C itself, and Newton could be said to establish some geometric link with his theory of integration, here. This is not only interesting in itself, but, further confirms the primary role that aesthetic and geometric connections play in Newton's work.

Although Newton's work in calculus is fundamentally important in the history of scientific ideas, in my opinion, his deepest geometric work can be found elsewhere. This work draws heavily on the aesthetic terminology that we considered in the previous chapter, in particular the images of ascent, crucifixion, fragmentation and the sublime.

The most important motivating idea towards this connection is Newton's introduction of power series, modern terminology for infinite series, in his papers (v) and (vi). We briefly encountered the idea of power series in the earlier discussion of the Generalised Binomial Theorem. Newton's discovery of power series methods for algebraic curves, which he refers to as "species", partially resulted from his consideration of "affected equations", that is polynomial equations $p(y)$ in a single variable. In Sections 19 and 20 of (vi), we find an explanation of his method of solving such equations;

---

[31]In this paper, one can find Newton's first definition and computations involving the centre of curvature. He first computes the subnormal to a given curve, as in footnote 23, which is done on p278 of [38], "The Perpendiculars to crooked lines and also y^e Theorems for finding them may otherwis more conveniently be found thus." Then, using an argument, involving similar triangles, and an infinitesimal o-argument, see footnote 8, he correctly deduces the coordinates (x_1, y_1) of the centre of curvature K, at $(x, f(x))$ for the graph of $y = f(x)$ as;

$$y_1 = f(x) + \frac{1+[f'(x)]^2}{f''(x)}, \; x_1 = x + \frac{f'(x)(1+[f'(x)]^2)}{f''(x)} \; (*)$$

in particular cases, see, for example, p245-248 of [38]. In Newton's diagram for the problem, see figure 7, $y_1 = length(gh)$ and $x_1 - x = length(hf)$ after setting $length(hi) = o$ to be zero in the limit calculation. The formula $(*)$ is given in full generality, using Newton's rather cumbersome differential notation, on p290 of [38]. Newton also considers the method that we discussed above, section 54 of (vi), on p262 of [38]

[32]Newton's interest in curvature is related to his idea that, if the curve C could be considered as a perfect refractive surface, light shining on the curve would be focused along the locus of curvature

" Of the Reduction of affected Equations.

19. As to affected Equations, we must be something more particular in explaining how their Roots are to be reduced to such Series as these; because their Doctrine in Numbers, as hitherto deliver'd by Mathematicians, is very perplexed, and incumber'd with superfluous Operations, so as not to afford proper Specimens for performing the Work in Species. I shall therefore first shew how the Resolution of affected equations may be compendiously perform'd in Numbers, and then I shall apply the same to Species.

20. Let this Equation $y^3 - 2y - 5 = 0$ be proposed to be resolved, and let 2 be a Number (any how found) which differs from the true Root less than by a tenth part of itself. Then I make $2 + p = y$, and substitute $2 + p$ for y in the given Equation, by which is produced a new Equation $p^3 + 6p^2 + 10p - 1 = 0$, whose Root is to be sought for, that it may be added to the Quote. Thus rejecting $p^3 + 6p^2$ because of its smallness, the remaining Equation $10p - 1 = 0$; or $p = 0.1$, will approach very near to the truth. Therefore I write this in the Quote, and suppose $0.1 + q = p$, and substitute this fictitious Value of p as before, which produces $q^3 + 6.3q^2 + 11.23q + 0.061 = 0$. And since $11.23q + 0.061 = 0$ is near the truth, or $q = -0.0054$ nearly, (that is, dividing 0.061 by 11.23, till so many figures arise as there are places between the first Figure of this, and of the principal Quote exlusively, as here there are two places between 2 and 0.005) I write -0.0054 in the lower part of the Quote, as being negative, and supposing $-0.0054 + r = q$, I substitute this as before. And thus I continue the Operation as far as I please, in the manner of the following Diagram:" ([33])

In other words, Newton makes a series of approximations to solutions of the original equation $y^3 - 2y - 5 = 0$. At the second and third stages of the approximation, the remainder terms, 0.1 and -0.0054 become smaller, hence, one is guaranteed that the infinite sum obtained by adding the remainders to the original quote 2, will converge to a genuine solution of the original equation. The reader can think of such a solution as an infinite decimal expansion on a calculator.

Newton extends this method to find power series solutions to polynomial equations $G(x, y)$ in two variables, in Section 36 of (vi), "The Praxis of Resolution";

"36. These things being premised, it remains now to exhibit the Praxis of Resolution. Therefore, let the Equation $y^3 + a^2y + axy - 2a^3 - x^3 = 0$ be proposed to be resolved. And from its Terms $y^3 + a^2y - 2a^3 = 0$, being a fictitious Equation, by the third of the foregoing Premises, ([34]), I obtain $y - a = 0$, and therefore I write $+a$ in the Quote. Then because $+a$ is not the complete Value of y, I put $a + p = y$, and instead of y, in the Terms of the Equation written in the Margin, I substitute $a + p$, and the Terms resulting ($p^3 + 3ap^2 + axp$, etc) I again write in the Margin; from which again, according to the third of the Premises, I select the Terms $+4a^2p + a^2x = 0$ for a fictitious Equation, which giving $p = \frac{-1}{4}x$, I write $\frac{-1}{4}x$ in the Quote. Then because $\frac{-1}{4}x$ is not the accurate Value of p, I put $\frac{-1}{4}x + q = p$, and in the marginal Terms for p I substitute $\frac{-1}{4}x + q$, and the resulting Terms ($q^3 - \frac{3}{4}xq^2 + 3aq^2$, etc) I again write in the Margin, out of which, according to the foregoing Rule, I again select the

[33] The diagram is omitted, as the written text is a clear enough explanation of Newton's method.

[34] We explain this reference below.

Terms $4a^2q - \frac{1}{16}ax^2 = 0$ for a fictitious Equation, which giving $q = \frac{x^2}{64a}$, I write $\frac{x^2}{64a}$ in the Quote. Again, since $\frac{x^2}{64a}$ is not the accurate Value of q, I make $\frac{x^2}{64a} + r = q$, and instead of q I substitute $\frac{x^2}{64a} + r$ in the marginal Terms. And thus I continue the Process at pleasure, as the following Diagram exhibits to view." ([35])

Here, Newton finds a power series solution to the equation defined by $y^3 + a^2y + axy - 2a^3 - x^3 = 0$. After finding a as an initial approximation, the remainder terms, $\frac{-1}{4}x$ and $\frac{x^2}{64a}$, at the second and third stages, involve successively higher powers of x. Although Newton doesn't give a completely rigorous proof of the fact, he argues that the process will continue to yield successively higher powers of x as remainder terms. Hence, one eventually obtains a power series solution of the form $\sum_{i=0}^{\infty} y_i(x)$, where $y_i(x)$, the i'th remainder term, is of the form $a_i x^{j(i)}$ with $a_i \in \mathcal{R}$, $j(i) \geq i$ is an integer, and $j(i') > j(i)$, for $i' > i$. There is an obvious parallel with the previous example, the higher powers of x correspond to the successively higher decimal places of the numerical approximation, ([36]). A completely rigorous proof of Newton's construction is given in the following lemma;

Lemma 0.13. *Let C be a real (or complex) plane curve, defined by a polynomial $G(x,y)$, passing through the origin O, such that O is a non-singular point for C. Then, possibly after a linear change of variables, one can find a power series of the form;*

$s(x) = \sum_{j=1}^{\infty} a_j x^j$, with $a_j \in \mathcal{R}$ (or $a_j \in \mathcal{C}$ respectively)

such that $G(x, s(x)) \equiv 0$.

Proof:

After making a linear change of variables, one can assume that the tangent line to C at $(0,0)$, defined algebraically as in Footnote 13, does not coincide with the axis $x = 0$. Equivalently, one may assume that $G(0,0) = 0$ and $\frac{\partial G}{\partial y}(0,0) \neq 0$, (†). We follow Newton's method in Section 36 above. The assumption (†) allows us to write;

$G(x,y) = p_m(x)y^m + \ldots + p_1(x)y + p_0(x)$ $(*)$

with $p_0(0) = 0$ and $p_1(0) \neq 0$. For (x,y) "small", Newton observes that it, therefore, makes sense to take as a first approximate solution to this equation;

$y_0 = \frac{-\lambda_0 x^{i_0}}{p_1(0)}$

where $\lambda_0 x^{i_0}$, $(i_0 \geq 1)$, is the first term in the expression for $p_0(x)$, this is the content of his "third of the foregoing Premises", ([37]).

[35] Again, I have omitted the accompanying diagram.

[36] The interested reader can find similar calculations on "affected equations" and "species" in Sections 21-32 of (v)

[37] The content of the third premise in (*vi*) is the following;

"27. Thirdly, when the Equation is thus prepared, the work begins by finding the first Term of the Quote; concerning which, as also for finding the following Terms, we have this general Rule, when the indefinite

Now, Newton makes the substitution $y = (y' + y_0)$ in $(*)$, this results in a further polynomial equation of the same form;

$q_m(x)y'^m + \ldots + q_1(x)y' + q_0(x) = 0 \ (**)$

By a straightforward algebraic calculation, using the fact that $i_0 \geq 1$, one checks that $q_1(0) \neq 0$ and $ord(q_0(x)) > ord(p_0(x))$. Hence, one can take as the second quote;

$y_1 = \frac{-\lambda_1 x^{i_1}}{q_1(0)}$

where $\lambda_1 x^{i_1}$, $(i_1 > i_0)$, is the first term in the expression for $q_0(x)$. Continuing in this way, one obtains a sequence of approximate solutions;

$s_n(x) = y_0(x) + y_1(x) + \ldots + y_n(x)$, for $n \geq 0$

Either this process terminates after a finite number of approximations, giving a polynomial solution to $(*)$, or one obtains an infinite power series;

$s(x) = \sum_{i \geq 0} y_i(x)$

A straightforward algebraic calculation shows that;

$ord(G(x, s_{n+1}(x))) \geq ord(G(x, s_n(x))) + 1$, ([38])

Species (x or z) is supposed to be small; to which Case the other two Cases are reducible.

28. Of all the Terms, in which the Radical Species (y,p,q, or r, etc.) is not found, chuse the lowest in respect of the Dimensions of the indefinite Species (x or z, etc) then chuse another Term in which that Radical Species is found, such as that the Progression of the Dimensions of each of the fore-mentioned Species, being continued from the Term first assumed to this Term, may descend as much as may be, or ascend as little as may be. And if there are any other Terms, whose Dimensions may fall in with this Progression continued at pleasure, they must be taken in likewise. Lastly, from these Terms thus selected, and made equal to nothing, find the Value of the said Radical Species, and write it in the Quote."

(In the proof of the lemma, x is the indefinite species and y is the radical species.)

[38] ord is the modern abbreviation for *order*, that is the value of the highest power of x appearing in a polynomial $p(x)$

Hence, by elementary arguments, we are guaranteed that $G(x,s(x))\equiv 0$, ([39]).

The assumption of "non-singularity" in Lemma 0.13 is necessary, in order to ensure that the obtained power series does not involve fractional powers of x, that is powers of the form $x^{\frac{m}{n}}$, for $n\geq 2$. If we consider the example of a curve C, defined by $y^2-x^3=0$, then the only possible power series solution of this equation is given by $y=x^{\frac{3}{2}}$, the curve in this case has a "cusp" singularity at the origin O. However, Newton also gives a general method for finding power series solutions to polynomial equations of the form $G(x,y)=0$, without the simplifying assumption that $\frac{\partial G}{\partial y}(0,0)\neq 0$. This is done by the introduction of "Newton's parallelogram", in Paragraph 29 of his paper (vi), and fractional exponents. However, although Newton's method is ingenious, as I will demonstrate below, the introduction of what are now referred to as Puiseux series, is not particularly useful, in the sense of understanding the global geometry of algebraic curves. It is to this subject that I now wish to turn.

In my opinion, Newton's finest contribution to the study of algebraic curves and also his most important geometric work is the paper (xiv), which will be my major source for the appraisal of his work in this field. Before considering this paper in greater detail, I wish to give a geometric interpretation of Newton's construction of power series in relation to algebraic curves. This interpretation begins with what is today known as "Newton's Theorem".

The modern formal statement of Newton's theorem, given in [1], is the following;

Theorem 0.14. *Let $\mathcal{C}$ be an algebraically closed field, and let;*

$$G(x,y)=y^n+a_1(x)y^{n-1}+\ldots+a_n(x)\ \textit{in}\ \mathcal{C}(x)[y]$$

[39]Newton's method of constructing a solution to the polynomial equation $G(x,y)=0$, given in the Lemma, is closely related to what is now known as the Newton-Raphson method. Namely, one considers the function;

$$G:\mathcal{R}[x]\rightarrow\mathcal{R}[x],\ G(y)=p_m(x)y^m+\ldots+p_1(x)y+p_0(x)\ (\dagger)$$

Having obtained a first approximation $y_0=s_0$ to the equation $G(y)=0$, the Newton-Raphson method gives the further approximation;

$$s_1=y_0-\frac{G(y_0)}{G'(y_0)}=y_0(x)-\frac{q_0(x)}{q_1(x)}$$

where $q_0(x)$ and $q_1(x)$ are obtained from the transformed polynomial $(**)$ in the Lemma. By a similar argument to the above, replacing the successive approximations $-\frac{\lambda_1 x^{t_1}}{q_1(0)}$ by $-\frac{q_0(x)}{q_1(x)}$ and noting that;

$$\frac{q_0(x)}{q_1(x)}=\frac{\lambda_1 x^{t_1}}{q_1(0)}u(x)\text{ for a unit }u(x)\in\mathcal{R}[[x]]$$

one is similarly guaranteed that this method also yields the same power series solution $s(x)$ to the equation $G(y)=0$ in $(\dagger)$. The Newton-Raphson method is usually applied to functions of a real variable, rather than the rings $\mathcal{R}[x]$ or $\mathcal{R}[[x]]$. However, that Newton intended his method of finding power series solutions to polynomial equations of the form $G(x,y)=0$, ("species"), to be a partial generalisation of this method is borne out by his consideration of the case of "affected equations", that we considered above. Newton's work in (v), and similar calculations in (iv), should, therefore, be considered as a principal origination of this idea.

be a monic polynomial of degree $n > 0$ in y, with coefficients $\{a_1(x), \ldots, a_n(x)\}$ in $\mathcal{C}(x)$. Then, there exists a positive integer m, such that;

$G(t^m, y) = \prod_{i=1}^{n}[y - \eta_i(t)]$, *with* $\eta_i(t)$ *in* $\mathcal{C}((t))$

On an algebraic level, one may view this theorem as a generalisation of "The Fundamental Theorem of Algebra" to polynomials in 2 variables, see also Footnote 36 and the above consideration of "affected equations". Just as any polynomial in 1 variable, of degree n, admits n complex number solutions, any polynomial in 2 variables of degree n, as defined above, admits n infinite series solutions, if we allow for fractional exponents and a finite number of inverse powers. Let us give some simple examples to illustrate the theorem. Consider the polynomial $y^2 - x^3$. This factorises as $(y - x^{\frac{3}{2}})(y + x^{\frac{3}{2}})$, demonstrating the necessity for fractional powers. The polynomial $y^2 - (1+x)$, on the other hand, factorises as $(y - \eta(x))(y + \eta(x))$, where $\eta(x)$ is given by;

$\eta(x) = 1 + \frac{x}{2} - \frac{x^2}{8} + \ldots + \frac{(1.(-1).\ldots.(1-2(n-1)))}{2^n n!}x^n$

demonstrating, here, the necessity for infinite series. Finally, consider the polynomial $x^2y^2 - 1$, which factors as $(y - \frac{1}{x})(y + \frac{1}{x})$, showing that inverse exponents are needed.

The proof of the theorem is a straightforward generalisation of Newton's method, given in Section 36 of (vi), for finding power series solutions to curves, see also Lemma 0.13, his invention of the parallelogram method for fractional exponents, and his analogy with "affected equations", in Sections 19 and 20 of (vi). However, I wish to argue that the real justification for the theorem is reflected not in this numerical analogy, ([40]), but, in the way it is intuitively used by Newton, at a deeper, more profound geometric level, in the study of plane cubic curves.

In order to do this, it is necessary to find geometric interpretations of some of the terminology that Newton uses in (vi) and (xiv). In doing so, we will find deep connections with the aesthetic ideas that we explored in the previous chapter.

The interpretation of Newton's power series solutions, that we discussed in Lemma 0.13, can be found in [13]. Here, the notion of an etale cover of the plane is introduced. Such a cover can be thought of as a surface S, which locally imitates the geometry of the plane P near a given point O, but, on a global level, is more complicated, ([41]). An infinite series of the form $y - s(x)$, as in Lemma 0.13, corresponds to a curve $C' \subset S$, passing through the point O^{lift}, which projects onto the curve $C \subset P$, defined by the polynomial $G(x, y)$, see figure 8. By the nature of the ambient surface S, the curve C' reflects the geometry of C in a local neighborhood of O^{lift}, though, again, the global geometry of C' is, in general, more complex, ([42]).

The intuition of lifting a curve C, lying in the plane, to a surface in 3-dimensional space, recalls the aesthetic idea of ascent that we considered in the previous chapter. The curve C' is a 3-dimensional manifestation of the planar curve C. In Christian terminology, see

[40] Indeed, over non-algebraically closed fields, such as $\mathcal{R}$, the theorem fails. As we explain below, Newton uses a geometric version of the result, in his study of real plane cubic curves.

[41] A formal definition is given in [13]

[42] An excellent discussion of the etale topology can be found in [23]

my paper [9], it can be seen as a spiritual 3-dimensional interpretation of a 2-dimensional physical reality. It is, perhaps, no coincidence that the modern invention of etale covers occurred in Paris, one of the major medieval centres in the development of the aesthetic of ascent.

In the context of Newton's Theorem, the power series solution guaranteed by Lemma 0.13, corresponds to a factorisation of the polynomial $G(x,y)$ into the form;

$$G(x,y) = (y - s(x))u(x,y) \ (\dagger)$$

where $u(x,y)$ is an infinite series in the variables $\{x,y\}$ such that $u(0,0) \neq 0$. As before, we can interpret the infinite series $u(x,y)$ as a curve $C'' \subset S$, which projects onto C but does not pass through the point O^{lift}, see figure 9.

In this example, we see the aesthetic idea of fragmentation, the curve C breaks into 2 components on the surface S. More specifically, the configuration of the curves C' and C'' is reminiscent of the image of the crucifixion. This idea becomes even more geometrically intuitive when we consider S as part of a continuous family of surfaces S_t, for $t \in [0,1]$, moving between the plane $P = S_0$ and the surface $S = S_1$. For $t \in [0,1)$, the curve C lifts to a single curve C'''. At the critical point $t = 1$, we observe the breaking effect into the component curves C' and C''.

In this interpretation, we can see the geometric subtlety of Newton's Theorem in the effective combination of the aesthetics of ascent and fragmentation. It is this form of geometric intuition, paralleled by aesthetic considerations, which I suggested, in the previous chapter, motivated such creative and innovative constructions as the rib vault of Durham Cathedral.

One can refine this intuition further, by introducing an algebraic result that generalises the decomposition, given in ($\dagger$), known as Weierstrass preparation, ([43]). This is formally stated in [1] as follows;

Theorem 0.15. *Weierstrass Preparation*

Let $\mathcal{C}$ be an algebraically closed field, and let $G(x,y)$ be a polynomial with $G(0,y) \neq 0$ and $d = ord_y G(0,y)$. Then there exist unique series $U(x,y)$ and $N(x,y)$, with $U(0,0) \neq 0$, such that;

$$G(x,y) = N(x,y)U(x,y)$$

and

$$N(x,y) = y^d + c_1(x)y^{d-1} + \ldots + c_d(x)$$

with $c_i(x) \in \mathcal{C}[[x]]$ and $c_i(0) = 0$, for $1 \leq i \leq d$.

In the case considered by Lemma 0.13, the infinite series $N(x,y)$ is given by $y - s(x)$. However, in general, the curve C may posess a more complicated singularity at the origin

[43]Weierstrass was a German mathematician of the 19th century, best known for his work in the field of analysis, which we briefly alluded to at the beginning of the chapter

O, for which Newton's construction, summarised in Lemma 0.13, no longer applies. The relevant case to our discussion occurs when the singularity is a node(ii), a particular example of a double point(i), both definitions are given rigorously in the paper [17]. Intuitively, a double point occurs when any infinitesimal variation of a line l, passing through O, intersects the curve C in two points, (see footnotes 14 and 15), while a node occurs when the curve C intersects itself in a cruciform shape, see figure 10. Newton could be said to consider both definitions in (xiv);

(i). Section V; Of Double Points in Curves.

"We have remarked that curves of the second genus can be cut by a straight line in three points. Sometimes two of these three points coincide, as in the case when the straight line passes through an infinitely small oval, or through the intersection of two parts of the curve cutting each other, or meeting in a cusp. And whenever all the straight lines, extending in the direction of any infinite branch, cut the curve in only one point (as occurs in...(text omitted)...) we must conceive that those straight lines pass through two points in the curve at an infinite distance. Two intersections of this sort when they coincide, whether at a finite or an infinite distance, we shall call a double point."

(ii). Section III; The Names of the Curves.

"that which intersects and returns in a loop upon itself, the nodate hyperbola."

Section IV; The Enumeration of Curves.

Newton's description of the 2nd species as a "nodate figure", see figure 11.

Newton's description of the 7th and 8th species, "the figure will be cruciform", see figure 12.

Newton's description of the 34th species, using the term "node", see figure 13.

When a node of $C \subset P$ occurs at the origin O, being a double point, the Weierstrass Theorem guarantees that $d = 2$ for the infinite series $N(x, y)$. Using the elementary method for solving quadratic equations and geometric properties of a node, one can obtain the factorisation;

$$N(x, y) = (y - s_1(x))(y - s_2(x))$$

where $s_1(x)$ and $s_2(x)$ are distinct infinite series, with the properties of $s(x)$, given in Lemma 0.13. Using the method above, we then obtain the following geometric interpretation of this result. The infinite series $s_1(x)$ and $s_2(x)$ correspond to curves C_1 and C_2 on an etale cover S, passing through O^{lift}. As the geometry of S imitates the geometry of the plane P near O, the curves C_1 and C_2 also intersect in a node, see figure 14. In this example, we see that the process of fragmentation occurring on the surface S, imitates the geometry of C itself, the curve fragments over the nodal point at O. This phenomenon can be used to understand the geometry of a planar curve C in a neighborhood of such points, by reducing a potentially difficult problem concerning singularities to that of a simpler problem concerning the intersection of distinct curves.

This method is used in the paper [17] to deduce local properties of a family of nodal curves, namely that one obtains 2 intersections between "infinitely close" curves, in a neighborhood of a node, see figure 15. We will see later how generalisations of this result are important in the study of certain types of degenerations of plane curves. The proof of this result introduces a new idea, that of conic projections, to relate the geometry of the 3-dimensional "cruciform figure", determined by C_1 and C_2, with the plane "nodate figure", determined by C, see figure 14. The method of conic projections is a subject that we will consider further, in a later chapter. Newton was also aware of this method, which he refers to in (xiv) as "The Generation of Curves by Shadows";

Section V. The Generation of Curves by Shadows.

" If the shadows of curves caused by a luminous point, be projected on an infinite plane, the shadows of conic sections will always be conic sections; those of curves of the second genus will always be curves of the second genus; those of the third genus will always be curves of the third genus; and so on ad infinitum.

And in the same manner as the circle, projecting its shadow, generates all the conic sections, so the five divergent parabolas, by their shadows, generate all other curves of the second genus. And thus some of the more simple curves of other genera might be found, which would form all curves of the same genus by the projection of their shadows on a plane."

The use of conic projections introduces a new aesthetic, that of light, which we will discuss at greater depth in the following chapter. The reader is referred to the paper [17], where some aesthetic remarks on light and fragmentation are made, in relation to the use of circular windows with cruciform arches. This combination of designs occurs frequently in medieval abbeys and cathedrals, particularly in France and Italy.

The considerations we have made on the geometrical interpretation of Newton's Theorem, have, so far, taken into account only local properties of plane curves. However, there is a sense that the theorem provides information on a curve's global geometry. In order to make this idea precise, it is necessary to introduce the geometric concept of asymptotes.

A formal definition of asymptotes can be found in the paper [18](Section 2), along with a method of computing the asymptotes of any given plane algebraic curve. Newton was the first writer to make extensive geometric use of asymptotes in his paper (xiv). Intuitively, an asymptote is the limiting direction of the tangent to a point on a plane curve C, as the point moves towards infinity. In the following example, when the curve C is defined by the equation $y = \frac{1}{x}$, the asymptotes are given by the lines $x = 0$ and $y = 0$. It is clear from figure 16 how these correspond to the limiting directions of tangent lines to the curve. Newton gives such a definition in (xiv), when he discusses hyperbolic and parabolic branches;

Section II. 5. Of Hyperbolic and Parabolic Branches, and their Directions.

"All infinite branches of curves of the second and higher genus, like those of the first, are either of the hyperbolic or parabolic sort. I define a hyperbolic branch, as one which constantly approaches some asymptote, a parabolic branch to be that which, although infinite, has no asymptote. These branches are easily distinguished by their tangents, for supposing the point of contact to be infinitely distant, the tangent of the hyperbolic branch will coincide with the asymptote, but, the tangent of the parabolic branch, being at an infinite distance, vanishes, and, is not to be found"

In this passage, Newton makes a distinction between two types of behaviour of a curve C at infinity. In the hyperbolic case, the curve C approaches a limiting direction, defined by a line in the (x, y) plane, as in the previous example, the curve C posesses 2 hyperbolic branches at infinity. In the parabolic case, there is no such limiting direction, a simple example is given by the curve $y = x^2$. This distinction is developed in the paper [18](Section 3), when we consider covers of the line, having certain geometrical properties, which we consider later. Such covers have the feature that the curve intersects itself numerous times at a given point O_∞ at infinity. Using Newton's interpretation of asymptotes and his distinction between hyperbolic and parabolic branches at infinity, one can adopt a more convenient representation of such curves, see figure 17. A famous example of such a representation is "Newton's Trident", fig 74. of (xiv), see figure 18;

Section IV. 12. Of the Trident.

"In the second cited case of the equations, we had the equation $xy = ax^3 + bx^2 + cx + d$. In this case, the curve will have four infinite infinite branches, of which two are hyperbolic about the asymptote AG, extending on contrary sides, and two are parabolic, converging and making with the other two a sort of trident-shaped figure."

In the paper [18], the representation is used to study the behaviour of such curves C, when projected to the x-axis. Formally speaking, such projections are not defined at the point O_∞. Using vertical asymptotic lines, one can still, however, analyse the behaviour of the projection at O_∞, as a limit of points on the curve.

The notion of an asymptote recalls the aesthetic of the sublime that we considered in the previous chapter, the eye is drawn towards a fixed point at infinity along the curve. A particularly geometric use of this device can be found in the numerous examples of lancet windows, belonging to English medieval buildings. There is also the connection with the spiritual image of the Throne of God, represented by an inaccessible point at infinity.

Perhaps, Newton's greatest achievement in (xiv), is his use of the asymptotic method, to understand the global geometry of cubic curves. This begins with the following striking observation;

Section III. The Reduction of all Curves of the Second Genus to Four Cases of Equations.

"All line of the first, third, fifth, seventh, or odd orders, have at least two infinite branches extending in opposite directions; and all lines of the third order have two branches of the same kind, proceeding in opposite directions, towards which no other of their infinite branches proceed, (except only the Cartesian parabola).

Case I (fig 1.)

If the branches be hyperbolic, let GAS be their asymptote, and let CBc be any line drawn parallel to it, meeting the curve on each side(if possible); let this line be bisected in X, which will be the locus of a hyperbola, say $X\Phi$, one of whose asymptotes is AG. Let its other asymptote be AB and the equation defining the relation between abciss AB and ordinate BC, if $AB = x, BC = y$, will be of the form always;

$xy^2 + ey = ax^3 + bx^2 + cx + d$", (i), $(\sharp)$, see figure 19, labels (a) and (b) are not part of Newton's original sketch, see Footnote 41 below) ([44])

In this passage, Newton argues that any cubic curve C must have at least 2 infinite(opposite) branches. This is due to the observation that any real cubic polynomial in 1 variable has a real root. Hence, the curve C must have a limit as $x \to \infty$ or $x \to -\infty$. In the case that these branches are hyperbolic ,by taking the asymptote to these branches as a coordinate axis, we obtain the following picture, see figure 20. At $x = 0$, the curve C has a solution with multiplicity 2 at ∞, due to tangency of the axis, hence, at most 1 real(complex) solution in finite position. This reduces the equation of C to the following form;

$xy^2 - (ex^2 + fx + g)y = ax^3 + bx^2 + cx + d$ $(\dagger)$

as if there were any isolated terms of the form y^2 or y^3 in the equation, then, setting $x = 0$, one would obtain a polynomial of degree at least 2 in y, implying the existence of at least 2 finite solutions along the axis $x = 0$. The remainder of Newton's argument implicitly uses the method of completing the square, namely, we obtain;

$y = \frac{ex^2+fx+g}{2x} \pm \sqrt{(\frac{ex^2+fx+g}{2x})^2 + \frac{ax^3+bx^2+cx+d}{x}}$ $(\dagger\dagger)$

The equation of the hyperbola $X\Phi$ that Newton refers to, is then given by $y = \frac{ex^2+fx+g}{2x}$. Newton then simplifies the equation of the hyperbola $X\Phi$ to the form $y' = \frac{e}{x'}$, by taking the asymptotes of the hyperbola as coordinate axes. This reduces the equation $(\dagger)$ to the form (i), under generic assumptions on distinct real roots appearing in the equation $(\dagger\dagger)$. Equation (ii) follows from consideration of the degenerate case, when the expression under the square root is set identically to zero. Equations (iii) and (iv) occur from consideration of the case when there are two infinite(opposite) parabolic branches.

The major part of Newton's work in (xiv) is concerned with analysing the different cases

[44] Newton makes a further analysis, reducing to the further equations;

$xy = ax^3 + bx^2 + cx + d$ (ii)
$y^2 = ax^3 + bx^2 + cx + d$ (iii)
$y = ax^3 + bx^2 + cx + d$ (iv)

The four cases of equations account for Species 1-65, 66, 67-71 and 72 in Newton's classification.

that occur for the equation (i). Curves satisfying the Case I equation all have the property that they possess at least 1 asymptote. Newton is able to make some straightforward distinctions between such curves, by considering, first, the number of asymptotes, either one or three, and, in the case of three asymptotes, their arrangement in the plane. Examples of Case I curves with a single asymptote are given in Species 33-56 (fig. 43-64) of Newton's classification, described in Section IV, 5-8. A description of some of these, parabolic hyperbolas, can be found in the paper [18], the following diagram figure 21 (fig. 64), is an example of such a curve. For Case I curves with 3 distinct asymptotes, Newton's analysis proceeds by considering their possible configurations. One would usually expect that 3 such asymptotes intersect in a triangular configuration, which I will refer to as lines in general position. Indeed, these account for Species 1-23 (fig. 5-34) of Newton's classification, described in Section IV, 1-3. A description of some of these, triple hyperbolas, can be found in [18], the previous diagram, figure 19, appearing as (fig. 1 and 5) in Newton's sketches, is an example "Of the nine redundant Hyperbolas, having no Diameter, and three Asymptotes, making a Triangle". Occasionally, one obtains the exceptional configuration of three asymptotes intersecting in a point, these account for Species 24-32 (fig. 35-42) of Newton's classification, described in Section IV, 4, "Of the Nine Redundant Hyperbolas, with three Asymptotes, converging to a common point", the following diagram, figure 22, appears as (fig 42) in Newton's sketches. Another exceptional configuration is that of 3 parallel lines in the plane, accounting for Species 57-60 (fig. 65-68), described in Section IV, 9, "Of the four Hyperbolisms of the Hyperbola", the following diagram, figure 23, appears as (fig 68) in Newton's sketches. Curves satisfying the Cases II-IV equations are more straightforward to analyse, accounting for Species 66-71 (fig. 74-81), described in Section IV, 12-14. Newton's Trident, (fig. 74), which we considered above, is the only example of a Case II curve, the following diagram, figure 24, (fig. 81), is an example "Of the Cubic Parabola", the only Case IV curve. The use of asymptotes, however, is not the only method by which Newton distinguishes between cubic curves. Newton is able to use a more geometrically subtle method to distinguish, for example, between the Species 1-23, having 3 asymptotes in general position. The basis for this method is the use of hyperbolas to decompose the global geometry of the curve. In the previous passage, ($\sharp$), where Newton derives the Case I equation, we find a description of the hyperbola $X\Phi$, given by the equation $y = \frac{ex^2+fx+g}{2x}$ in ($\dagger$). Further hyperbolas, corresponding to the actual locus of the curve C, are obtained, by using Newton's method to expand the root term as an infinite series. ([45])

Newton makes use of this interpretation intuitively. It seems likely, from previous considerations and the confident geometric manner in which the text is written, that Newton, in fact, carried out a number of algebraic calculations with such series, which were never

[45] Rigorous calculations can be found in Talbot's analysis of Newton's text, [34]. From the Case I equation $xy^2 - ey = ax^3 + bx^2 + cx + d$, he derives;

$$y = \frac{e}{2x} \pm \sqrt{ax^2 + bx + c + \frac{d}{x} + \frac{e^2}{4x^2}}$$

$$y = \frac{e}{x} + \frac{d}{e} + Ax + Bx^3 + \dots \text{ (a)}$$

$$y = \frac{-d}{e} - Ax - Bx^2 - \dots \text{ (b)}$$

describing the hyperbolas, labelled (a) and (b) respectively, in figure 19.

included. In the following passage, he gives a geometric description of the hyperbolas which are used;

"Section III The Names of the Curves.

In the enumeration of these cases of curves, we shall call that which is included within the angle of the asymptotes in like manner as the hyperbola of the cone, the *inscribed* hyperbola; that which cuts the asymptotes, and includes within its branches the parts of the asymptotes so cut off, the *circumscribed* hyperbola; that which, as to one branch, is inscribed, and, as to the other, circumscribed, we shall call the *ambigenous* hyperbola; that which has branches concave to each other, and proceeding towards the same direction, the *converging* hyperbola; that which has branches convex to each other, and proceeding towards contrary directions, the *diverging* hyperbola; that which has branches convex to contrary parts and infinite towards contrary sides, the *contrary branched* hyperbola; that which, with reference to its asymptote, is concave at the vertex, and has diverging branches, the *conchoidal* hyperbola; that which cuts the asymptote in contrary flexures, having on both sides contrary branches, the *serpentine* hyperbola; that which intersects its conjugate, the *cruciform* hyperbola; that which intersects and returns in a loop upon itself, the *nodate* hyperbola; that which has two branches meeting at an angle of contact, and then stopping, the *cusped* hyperbola; that which has an infinitely small conjugate oval, i.e a conjugate point, the *punctate* hyperbola; that which from the impossibility of two roots, has neither oval, node, cusp or conjugate point, the *pure* hyperbola."

Newton distinguishes between the behaviour of several types of hyperbola. This relies on his analysis of the behaviour of the hyperbola towards its asymptotes, in the description of the terms *inscribed, circumscribed, ambigenous, diverging, contrary branched* and *conchoidal*, on local properties of the hyperbola, at finite position, away from infinity, in the description of the terms *serpentine, cruciform, nodate, cusped, punctate* and *pure*. In the following of Newton's sketches, we see examples of different types;

(i). 3 pure hyperbolas. (fig 14), see figure 28.

(ii).1 pure hyperbola and 2 punctate hyperbolas, see figure 25.

(iii). 2 pure hyperbolas and 1 nodate hyperbola, (fig 8), see figure 11.

(iv). 2 pure hyperbolas and 1 cusped hyperbola, (fig 10), see figure 26.

(v).1 pure hyperbolas and 2 cruciform hyperbolas, (fig 17), see figure 12.

(vi). 2 pure hyperbolas and 1 serpentine hyperbola, (fig 20), see figure 27.

As we have seen in Newton's original text, and Talbot's commentary, the analysis of the behaviour of the hyperbola, at infinity, depends on finding series that describe the behaviour of the curve, towards the relevant asymptote. The local properties of the hyperbola, away from infinity, depend on an analysis of the relationship between two distinct series, obtained from an asymptotic calculation, (†). This idea is suggested by the following passages from Section IV, The Enumeration of Curves;

"If two roots are equal, and the other two, either impossible, (figs 16,18) or real, and with different signs from the equal roots, (figs 17,19), the figure will be cruciform, two of the hyperbolas *intersecting* one another, either at the vertex of the asymptotic triangle, or at its base, (figs 16, 17)." (figure 12 corresponds to Fig 17)

"If the two greatest roots $A\pi$, Ap (Fig 8), or the two least (AP,$A\omega$, fig 9), are equal to each other, and all of the same sign, the oval and circumscribed hyperbola will coalesce, their points of contact η and t, or T and t, and the *branches* of the hyperbola, *intersecting* one another, run on into the oval, making the figure nodate." (see figure 11 corresponding to fig 8)

"If the three greatest roots Ap, $A\pi$, $A\omega$ (fig 10), or the three least roots $A\pi$, $A\omega$, AP (fig 11), are equal to each other, the node becomes a sharp cusp, because the two *branches* of the circumscribed hyperbola *meet* at an angle of contact, and extend no further." (see figure 26 for fig 10) (††)([46])

Newton is continuously aware of the relationship between the global geometry of the curve and its three asymptotes. Although never formally stated in the text, he intuitively uses the idea that a continuous family of degenerations of the curve, will reduce it to these asymptotes. In this degeneration, the local singularities that we have considered, and the vertices of the hyperbolas, (points on the hyperbolas, furthest from infinity), reduce to the intersections of the lines themselves. This is evidenced by his sketches, perhaps, in his description of cubic curves as "Lines of the Third Order", and his reference to the relation of hyperbolas with the angles formed by the asymptotes, both in (††), and the following passage from Section IV;

"If two roots are impossible and the other two unequal, and of the same sign, three pure hyperbolas result, without oval, node or conjugate point, and these hyperbolas will either lie at the sides of the triangle made by the asymptotes, or at its angles, thus forming either the 5th (figs 12,13) or the 6th species (figs 14,15)." (see figure 28 for fig 14)

There is also a sense of linear symmetry in the way that Newton relates a curve to certain configurations of lines.

In order to understand the, perhaps, four distinct geometric ideas that Newton employs in "the method of hyperbolas", I refer the reader to [18], where it is used in the study of plane nodal curves,([47]), without restriction on the degree of the curve.

Suppose, then, that we are given such a curve C of degree m. We begin, by observing that there exists a line l, which cuts C transversely in m distinct non-singular points of the curve. We generalise Newton's construction in (†) and (††) of "The Reduction of all Curves of the Second Genus to Four Cases of Equations", as follows. We take the line l as a coordinate

[46] The italics are not in the original text, but illustrate the argument (†). Newton's use of the term branch in the final two passages, together with his previous definition in Section II, 5, further support this argument, and the choice of labelling (a), (b) of the hyperbolas in figure 19. In the paper [18], we interpret Newton's branches as "flashes", this interpretation will become clearer, later in the chapter.

[47] The assumption that the curve has at most nodes as singularities is probably unnecessary

axis $x = 0$, ensuring that all of the intersections between C and the line are in finite position, not at infinity, ($\dagger$), see figure 29. The condition of transversality to the axis $x = 0$, allows us to apply Newton's Theorem in the following form;

Theorem 0.16. *Let $\mathcal{C}$ be an algebraically closed field, and let $G(x, y)$ define the plane curve C in the coordinate system defined by ($\dagger$). Then we can find infinite series $\{\eta_1(x), \ldots, \eta_m(x)\}$ in $\mathcal{C}[[x]]$, with $\{\eta_1(0), \ldots, \eta_m(0)\}$ distinct, such that;*

$$G(x, y) = (y - \eta_1(x)) \cdot \ldots \cdot (y - \eta_m(x))$$

The proof may be found in the paper [18]. By ensuring the line $x = 0$ avoids any singular points or infinite points of the curve C, and is not tangent to it, we avoid the technical problems of fractional series (Puiseux series), and inverse terms (Cauchy series), ([48]). We now draw the tangent lines to the curve C, at each intersection with the axis $x = 0$, see figure 29. These have the form;

$$y = n_j x + c_j, \text{ for } 1 \leq j \leq m$$

corresponding to the first two terms of the infinite series defined by $\eta_j(x)$ in the previous theorem, ($\star$). In line with Newton's construction, we can visualise the line $x = 0$ as the line at infinity in the projective plane. The tangent lines that we have drawn now correspond to the asymptotes of C. In this representation, the observation ($\star$) supports the idea that the infinite series $\eta_j(x)$ imitate the geometry of C towards the corresponding asymptote. Indeed, the fact that the infinite series $\eta_j(x)$ define functions of x, suggest that they describe the locus of C between infinity and the marked points ($*$) in the following diagram, figure 30. This supports the first of Newton's geometric methods on the asymptotic behavior of hyperbolas.

The problem of understanding the behaviour of the curve C, at the vertices of the hyperbolas, which we previously suggested was due to the intersection of loci, described by infinite series, requires a more sophisticated geometric interpretation, which is achieved in the paper [18]. Again, using the method of etale covers, we construct a projective surface $S \subset P^3$, for which the infinite series $\eta_j(x)$ correspond to curves $C_j \subset S$, projecting onto C, see figure 31. We then consider the intersections of the family of curves $\{C_1, \ldots, C_m\}$ on the surface S. In the case that C is a nodal curve, at these intersections, the corresponding projected point is indeed either a vertex or a node of the original plane curve C. Conversely, for any vertex or node of C, there exists a corresponding intersection in the family of curves on S. By the nature of the curve C, it follows that no three of the curves $\{C_1, \ldots, C_m\}$ can intersect in a point, giving a geometrically appealing representation of the family as a net, ([49]),($\sharp$). This, perhaps, clarifies the second of Newton's geometric intuitions in his discussion of hyperbolas. In the case of a general plane curve C, the intersecting geometry of the corresponding

[48] The modern terminology for an infinite series without fractional powers or inverse terms is a Taylor series, denoted in the previous theorem by $\mathcal{C}[[x]]$

[49] The intersecting geometry of the curves $\{C_1, \ldots, C_m\}$ partly motivated the terminology "flash" in the paper [18].

curves $\{C_1, \ldots, C_m\}$ imitates the curve's singularities. More precisely, we have the following;

Theorem 0.17. *If C is a plane curve of degree m and $p \in C$ is a singular point with coordinates (a,b), then;*

$I_p(C, x=a) = r < m$, ([50])

iff there exists a corresponding intersection between exactly r of the curves $\{C_1, \ldots, C_m\}$.

The proof can be found in the paper [18]. The result is useful in the subsequent discussion, see also figure 33.

In order to understand Newton's third intuition, we introduce the notion of an *asymptotic degeneration* of a plane curve C. In the case that C is a nodal curve, we mean by this a continuous family of curves $\{C_t\}_{t \in P^1}$, such that;

(i). $C_0 = C$.

(ii). $deg(C_t) = m \ \forall t \in P^1$.

(iii). The asymptotes of each curve C_t are fixed.

(iv). For $t \neq \infty$, C_t has at most nodes as singularities. ([51]) (♯♯)

The condition (iii) ensures that the family of curves $\{C_t\}_{t \in P^1}$ all bear the same relation to the fixed set of asymptotes, see figure 32. Moreover, we obtain the following uniformity in Newton's Theorem;

Theorem 0.18. *Let $G(x,y,t)$ define the plane curves C_t in the definition of an asymptotic degeneration, and let $\{\eta_1(x,t), \ldots, \eta_m(x,t)\}$ be the power series given by Theorem 0.16, then there exists a corresponding continuous family of curves $\{C_{1,t}, \ldots, C_{m,t}\}_{t \in P^1}$, with the properties outlined in (♯) above and in* [18].

The proof can be found in [18]. With the results of the previous two theorems, the problem of understanding the limit C_∞ of the degeneration, is equivalent to understanding the intersection geometry of the limit series of flashes. In the limit of such a series, the intersections of flashes may become more complex, for example, we may obtain that 3 of the flashes in the set of curves $\{C_{1,\infty}, \ldots, C_{m,\infty}\}$ intersect in a point. In this case, the corresponding projected point of the limit curve C_∞ defines a more complex singularity, see figure 33. However, we

[50] The reader should look at [13] for a definition of I_p using infinitesimals, generalising the description above of tangency between curves, see Theorem 0.7. In this case, it is clear how such a definition measures the complexity of the singularity p.

[51] More precisely, we require that the d nodes of C are preserved throughout the degeneration. Formally, this requires the parametrisation to take place inside the space $V_{d,m}$, consisting of the closure of irreducible curves of degree m, having d nodes as singularities. The mathematically more advanced reader should look at [32] for the original construction of this space, or [31] for a modern approach.

maintain a geometric understanding of the problem by keeping track of the intersections of flashes. This is made possible by imposing condition (iv), any point of intersection between two flashes is preserved throughout an asymptotic degeneration.

By repeating the process of asymptotic degenerations, ([52]), we can force the limit curve C_∞ to have a higher contact with any given asymptote l_j, ([53]). By a geometrical result, known as Bezout's Theorem, we are eventually guaranteed that the limit curve, C_∞, splits into $(l \cup C'_\infty)$, where l is the corresponding asymptote and C'_∞ has lower degree $m-1$, see figure 34, with corresponding flashes $\{C_l, C'_{1,\infty}, \ldots, C'_{m-1,\infty}\}$.

We then repeat the process for the plane curve C'_∞. In the flash interpretation, the $m-1$ flashes $\{C'_{1,\infty,t}, \ldots, C'_{m-1,\infty,t}\}$, corresponding to degenerations of C'_∞, may now separate from the flash C_l, corresponding to l. However, again we solve the problem by tracking the original intersection throughout the degeneration, see figure 35. Eventually, we obtain a degeneration to the original set of asymptotes.

The method of asymptotic degenerations is connected to a conjecture of Severi that any plane nodal curve $C \subset P$ of degree m can be degenerated to a series of m lines in general position, that is there exists a degeneration of C satisfying conditions $(i), (ii)$ and (iv) of $(\sharp\sharp)$ above. Severi's argument proceeds in the following stages;

(a). Construct a cone $Cone(C)$ over C.

(b). Slice the cone through a continuous family of plane sections;

$(Plane_{(t \in P^1)} \cap Cone(C))$

with $Plane_0 = P$ and $Plane_\infty$ passing generically through the vertex of the cone.

(c). Observe that the intersection $(Plane_\infty \cap Cone(C))$ consists of a series of m distinct lines passing through a point, the vertex of the cone.

(d). Move the m distinct lines passing through a point to m lines in general position.

Unfortunately, Severi's argument fails as the degeneration in the arguments $(a)-(c)$ is incompatible with the degeneration in (d), the reason being that the nodes of the original curve C may not converge to nodes formed by the lines in (c). This problem was observed by Zariski. The difficulty is resolved in the case of asymptotic degenerations, by Theorem 0.18 , and the observation that the curves in the family $\{C_{1,t}, \ldots, C_{m,t}\}_{t \in P^1}$ remain irreducible and non-coincident, $(\sharp\sharp\sharp)$, hence the nodes converge to intersections of lines, a precise statement of this result can be found in [?]. For degenerations which are not asymptotic, such as the example given by Zariski, it may still be possible to construct a corresponding continuous family, but the property $(\sharp\sharp\sharp)$ fails, that is the limit curves $C_{i,\infty}$ may become reducible and

[52]For a general curve C, one needs to modify condition (iv), to allow for more complex singularities. The existence of such degenerations follows from dimension calculations on the space of certain adjoint curves to C. The reader should look at the paper [17], to generalise the remarks above on "infinitely close" intersections between curves, in relation to (newton8.jpg)

[53]This implies that the corresponding infinite series, $\eta_{j,\infty}(x)$, has the form $a_j + b_j x + O(x^n)$, for large n.

components of the limit curves $C_{i,\infty}$ and $C_{j,\infty}$, for $i \neq j$, may coincide. In such cases, the intersection geometry of the limit series may change, making it impossible to track the position of the nodes throughout the degeneration, and use Theorem 0.17. However, Severi's observation (d), his argument in [32], and the property of asymptotic degenerations, that the positions of the original nodes of C vary continuously, ([54]), allows us to potentially solve his conjecture, by either altering the final configuration of the asymptotes, or choosing a set of asymptotes for C that are in general position. Clearly, a solution to the conjecture supports Newton's original geometrical thinking in [34].

Newton's fourth intuitive use of symmetry in his work on plane cubics, is also an important feature of the geometrical method used in the construction of flashes and asymptotic degenerations. More specifically, we have the following result from [18];

Theorem 0.19. *Let C be a plane(nodal) curve with $\{C_1, \ldots, C_m\}$ constructed as in the discussion $(\sharp)$ above. Then there exists a finite group G and a definable, transitive action of G on this set of curves.*

The finite group G is obtained as the Galois ([55]) group of the polynomial defined by the original curve C. Such an action is useful in understanding the intersection geometry of the set of curves $\{C_1, \ldots, C_m\}$, as, for any point p belonging to $C_i \cap C_j$, and $g \in G$, we have that $g.p \in (C_{g.i} \cap C_{g.j})$. In certain cases, this allows us to find bounds on the number of intersections between two distinct flashes C_i and C_j. The reader should consider the following diagram, see figure 36, of typical intersections between three flashes, in which the Galois group fixes the curve C_2 and permutes the curves C_1 and C_3, the corresponding intersections p_{12} and p_{23} are permuted, while p_{13} is fixed. Such considerations and Theorem 0.17 are useful in understanding the symmetry of the original plane curve C, in terms of the arrangement of its singularities or vertices, in the asymptotic interpretation, (see newton30.jpg). The use of Galois symmetry also simplifies the study of the structure of individual curves appearing in the set $\{C_1, \ldots, C_m\}$, namely that one can assume such curves have no singularities outside their set of intersections, the reader should look at [18] for more details.

The aesthetic of fragmentation in a plane is also reflected in medieval art. Particularly good examples are given by the work of the Cosmati artists in Sicily, which we considered briefly in the previous chapter. The pavements of the Palazzo Normanni and Monreale Cathedral in Palermo are excellent examples of configurations of lines in general position, that is no three of which intersect in a point. There is a highly developed sense of linear symmetry in such patterns, reflecting the advanced geometric intuition of Norman design. There is also an inherent 3-dimensionality in the arrangements, as if the lines are projected from an ambient space. This partly motivated the author's use of the flash construction in the problem of curve degenerations. Interesting examples of the geometry behind asymptotic degenerations can be found in the medieval art of Sweden, particularly around the island of Gotland. One can find extensive use of symmetric linear designs, such as the hexagonal

[54] This last property allows one to vary the families of curves within an irreducible component of $V_{d,m}$, see footnote 47 and [32]

[55] The theory of such groups is based on two papers by the French mathematician Evariste Galois, "Memoire sur les conditions de resolubilite des equations par radicaux" and "Des equations primitives qui sont solubles par radicaux", published in 1846, 14 years after his death. The interested reader can find out more about Galois theory in [8]

vault, and cusped windows, both highly suggestive of the aesthetic of sublime. The reader is referred to the previous chapter.

References

[1] Algebraic Geometry for Scientists and Engineers, S.S Abhyankar, AMS Mathematical Surveys and Monographs, Volume 35, (1990)

[2] Newton, P. Ackroyd, Knopf Doubleday Publishing Group, (2008)

[3] Introduction to Commutative Algebra, M. Atiyah and I. MacDonald, Addison-Wesley, (1994).

[4] Introduction to Real Analysis, R. Bartle and D. Schubert, Wiley, (2011).

[5] Calculus, K.G.Binmore, CUP, (2001).

[6] Pure Mathematics, Volumes 1 and 2, L. Bostock, S. Chandler, Nelson Thornes, (1978).

[7] A Treatise on Algebraic Plane Curves, J.L.Coolidge, Dover, (2004).

[8] A Course in Galois Theory, D.J.H Garling, CUP, (1987).

[9] Cosmati Pavements: The Art of Geometry, Tristram de Piro, Bridges Leeuwarden Proceedings, (2008).

[10] An Interpretation of Newton's Work in Calculus, Tristram de Piro, available at http://www.magneticstrix.net, (2008).

[11] Some Notes on Non-Standard Differentiation and Integration, Tristram de Piro, available at http://www.magneticstrix.net, (2009).

[12] Infinitesimals in a Recursively Enumerable Prime Model, Tristram de Piro, available at http://www.magneticstrix.net, (2005).

[13] A Non-Standard Bezout Theorem for Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2004).

[14] A Theory of Branches for Algebraic Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2006).

[15] A Non Standard Approach to the Theory of Algebraic Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2012).

[16] Applications of Nonstandard Analysis to Probability Theory, Tristram de Piro, available at http://www.magneticstrix.net, (2013).

[17] Some Geometry of Nodal Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2007).

[18] Flash Geometry of Algebraic Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2010).

[19] Lectures on the Hyperreals: An Introduction to Nonstandard Analysis, Rob Goldblatt, Graduate Texts in Mathematics 188, Springer-Verlag, New York, (1998).

[20] Algebraic Geometry, R. Hartshorne, Springer, (1997)

[21] Logic for Mathematicians, A.G. Hamilton, CUP, (1988)

[22] Traite de la Lumiere, C. Huygens, (1690).

[23] Lectures on Etale Cohomology, J.S. Milne, available at http://www.jmilne.org/math/CourseNotes/lec.html, (2013)

[24] The Red Book of Varieties and Schemes, D. Mumford, Springer, (1999).

[25] Opticks, Isaac Newton, Prometheus Books, (1993).

[26] The Principia, Isaac Newton, tr. A. Motte, Prometheus Books, (1995).

[27] An Introduction to Stability Theory, A. Pillay, Dover, (2008).

[28] Introduction to Complex Analysis, H. Priestley, OUP, (2003).

[29] Non-Standard Analysis, Abraham Robinson, Princeton University Press, (1996)

[30] Real and Complex Analysis, Walter Rudin, McGraw-Hill, (1987)

[31] Deformations of Algebraic Schemes, E. Sernesi, Springer, (2006)

[32] Vorlesungen uber Algebraische Geometrie, (Anhang F), F. Severi, Liepzig: Teubner (1921)

[33] Basic Algebraic Geometry, Volumes 1 and 2, I. Shaferevich, Springer, (1984).

[34] Notes and Examples on Sir Isaac Newton's Enumeration of Lines of the Third Order, C.R.M Talbot, H.G Bohn, Covent Garden, London, printed by G. Barclay, Castle St., Leicester Square, (1861)

[35] Algebraic Curves, R. Walker, Springer, (1978).

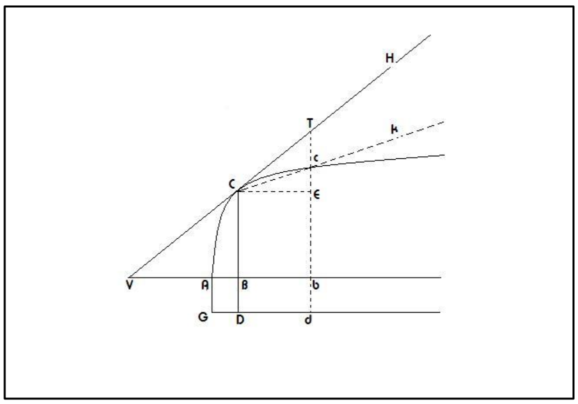
[36] Lebesgue Integration and Measure, A.J.Weir, OUP, (1973).

[37] Isaac Newton, the Last Sorceror, M.White, Perseus Books, (1999) .

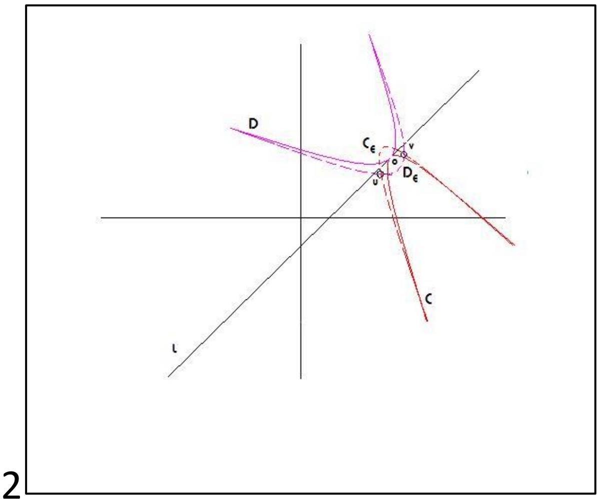
[38] The Mathematical Paper of Isaac Newton, Volume 1, D.T Whiteside, CUP, (2008).

[39] Zariski Geometries, Geometry from the Logician's Point of View, Boris Zilber, CUP, (2010).

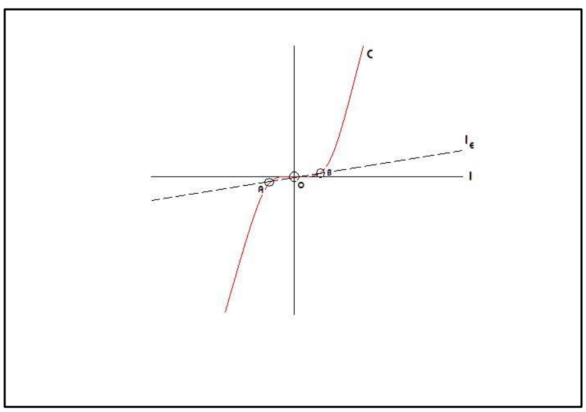
Figures



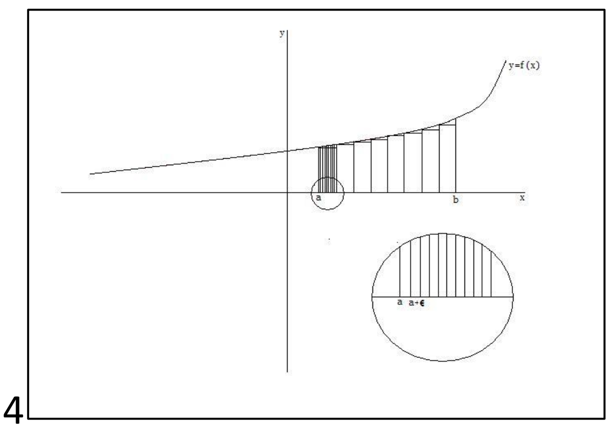
1



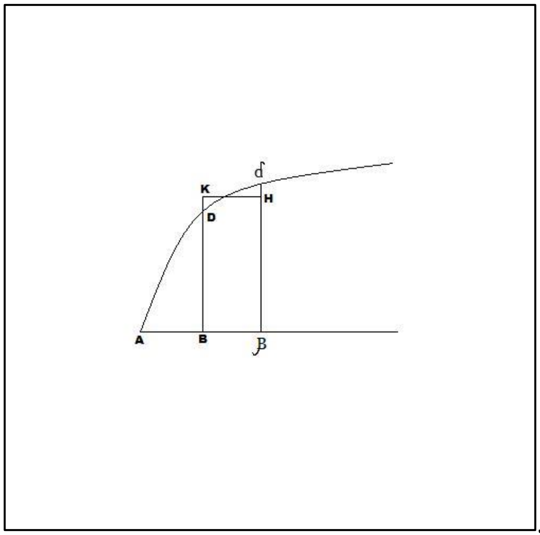
2



3

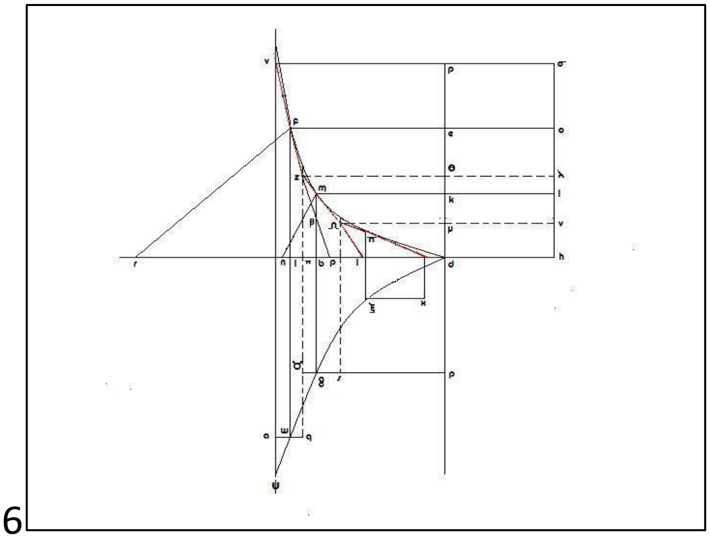


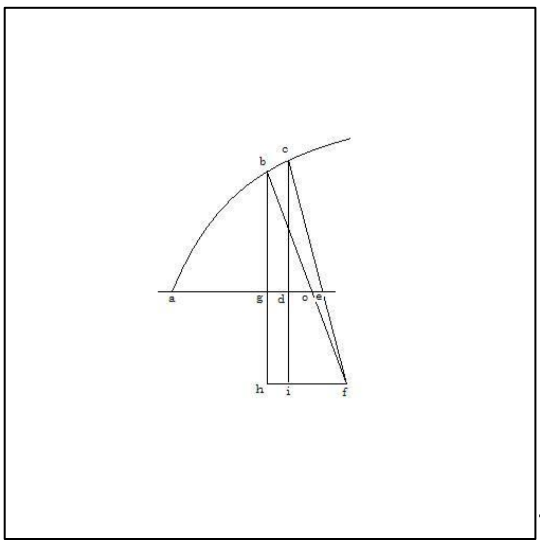
4



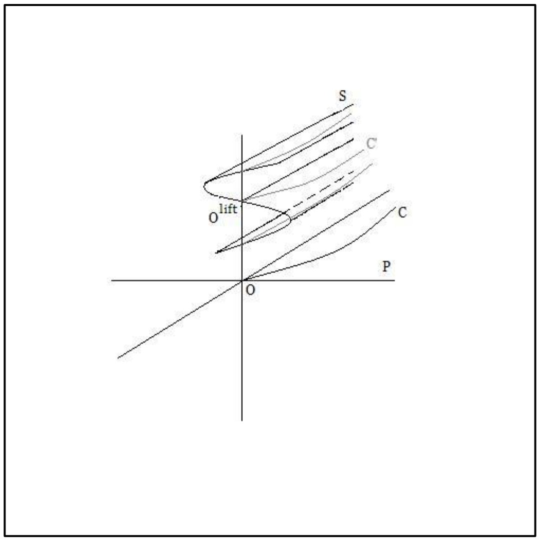
5

6



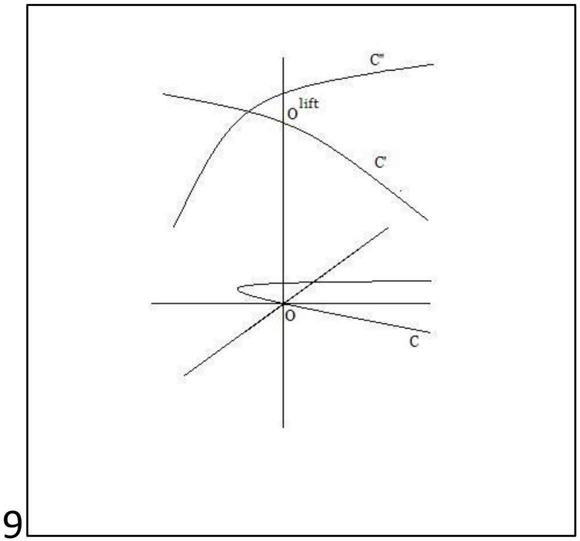


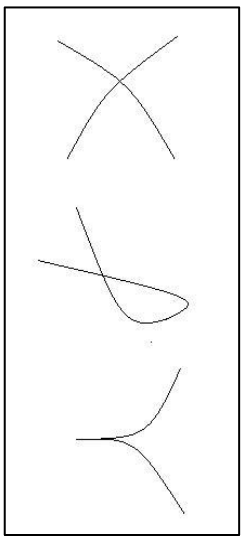
7



8

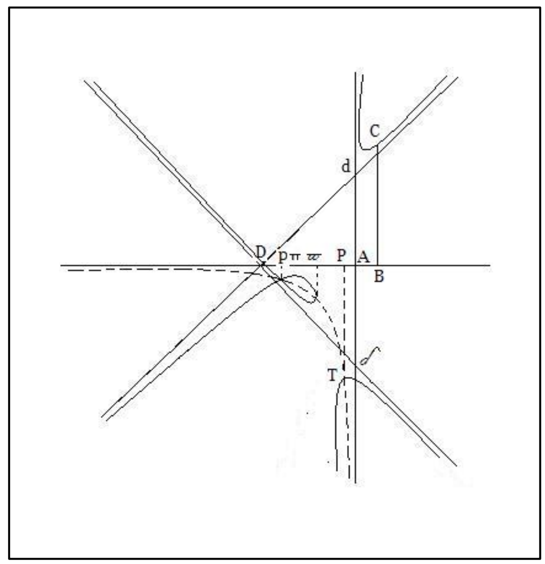
9

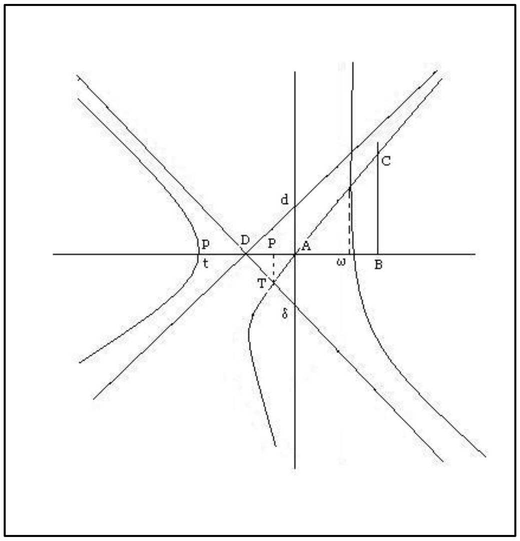




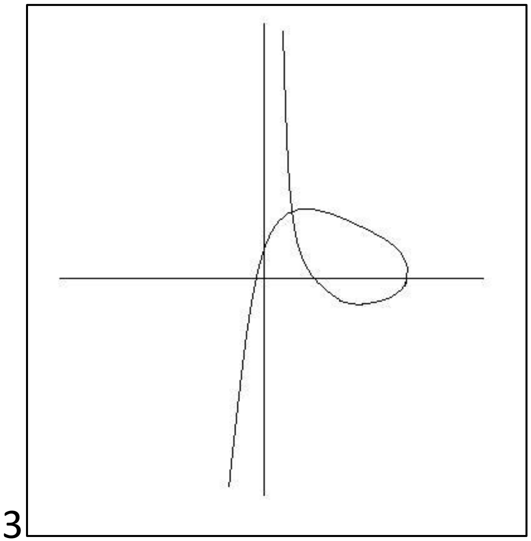
10

11

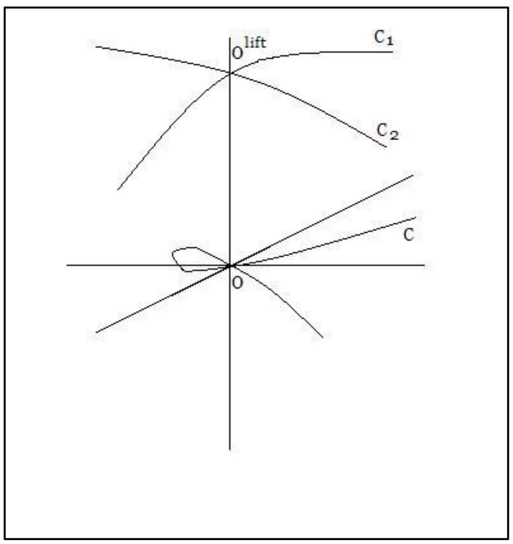




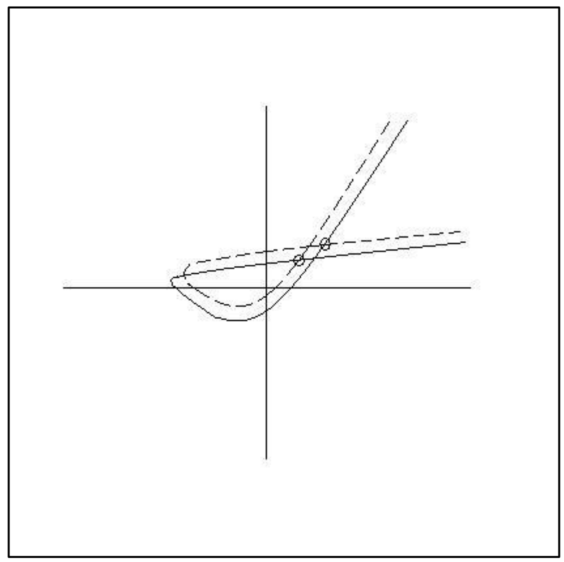
12



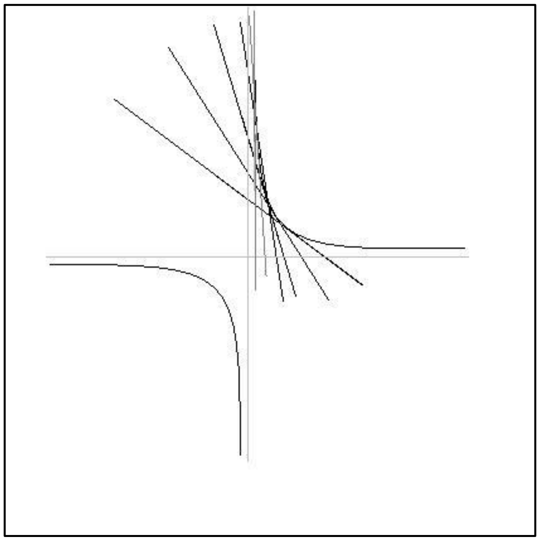
13



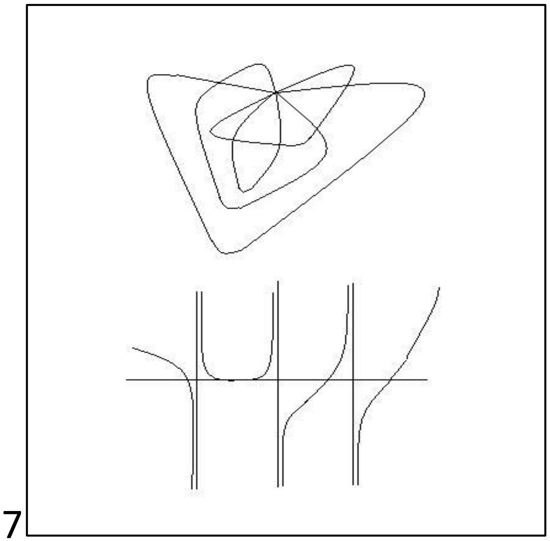
14



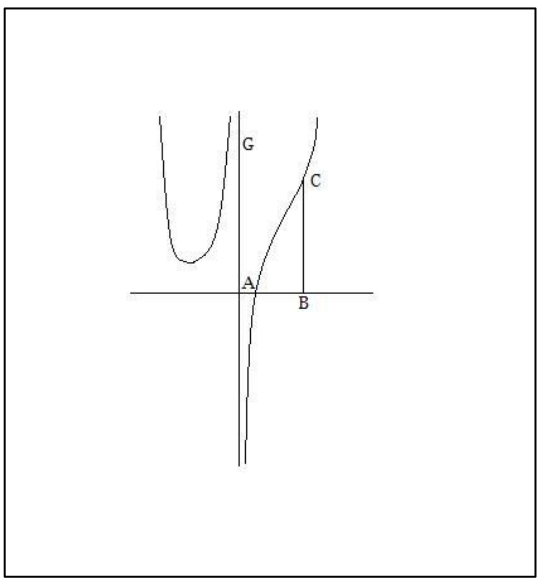
15



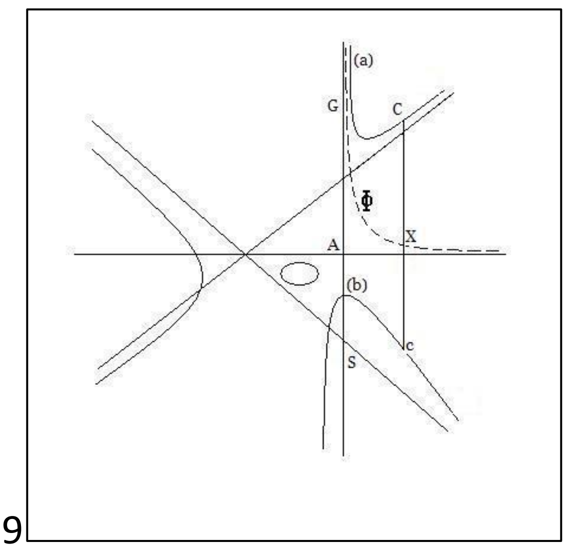
16



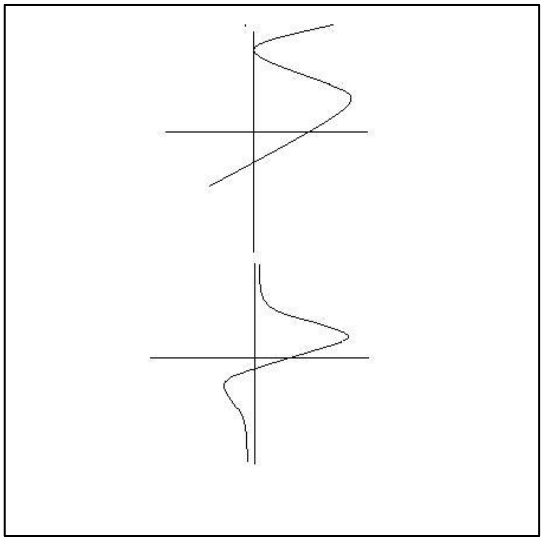
17



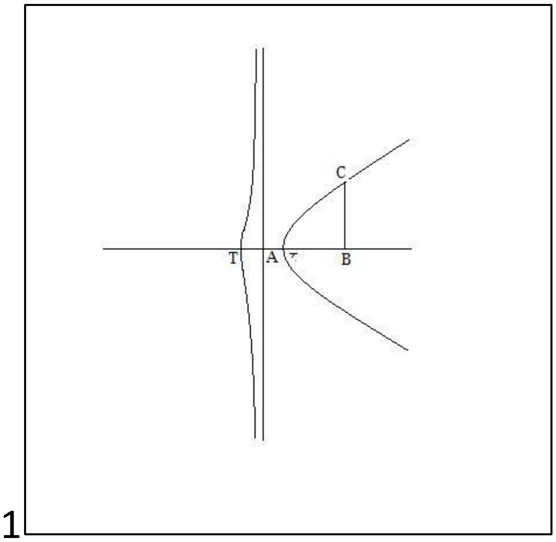
18



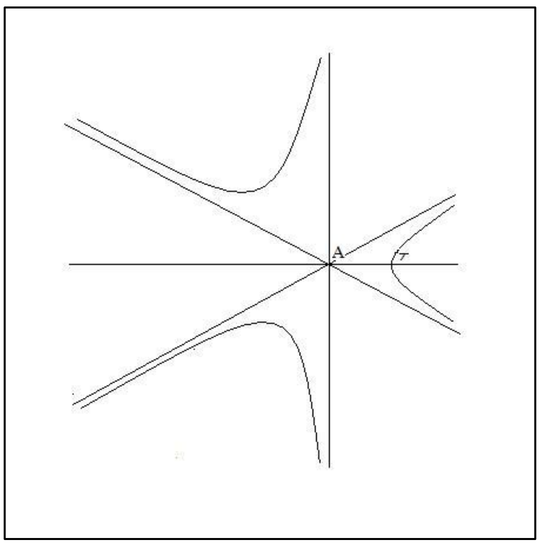
19



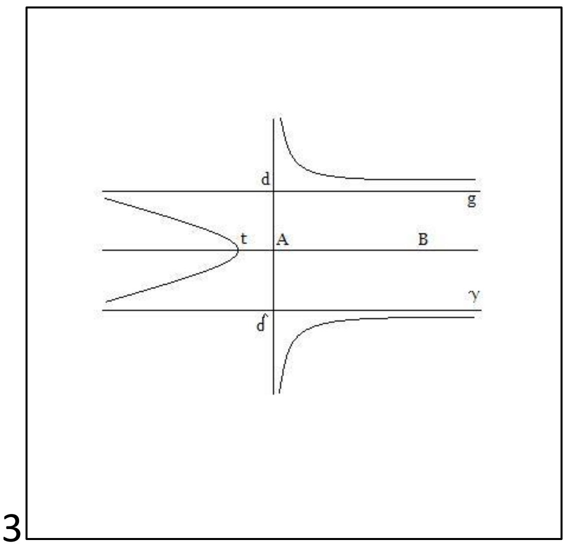
20



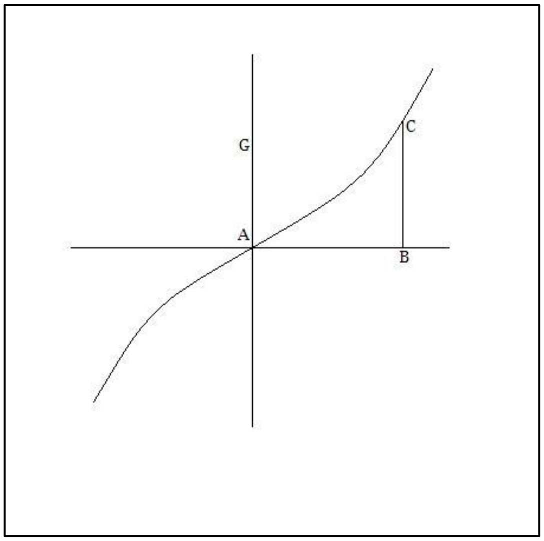
21



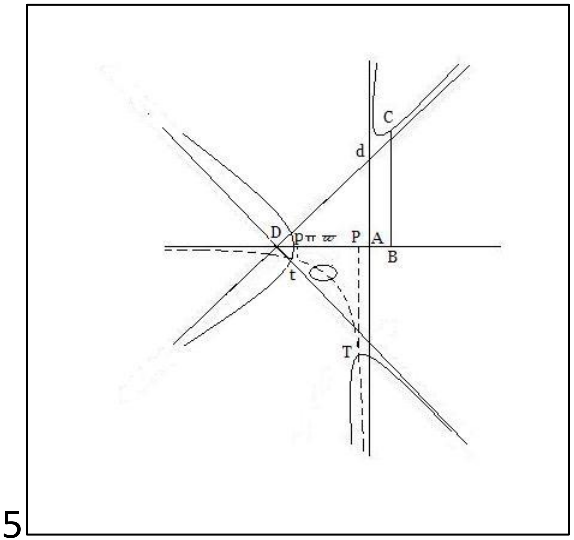
22



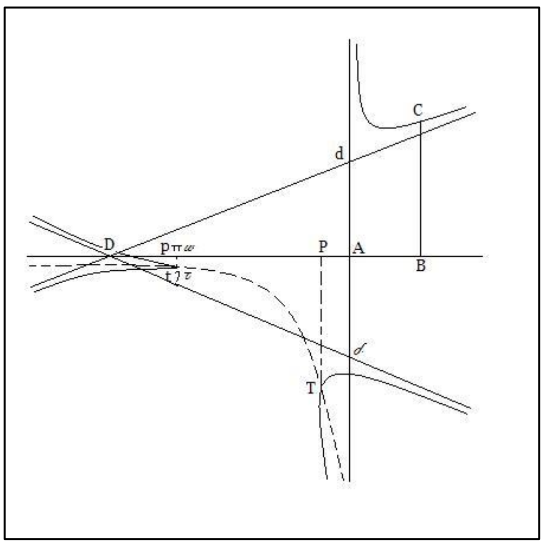
23



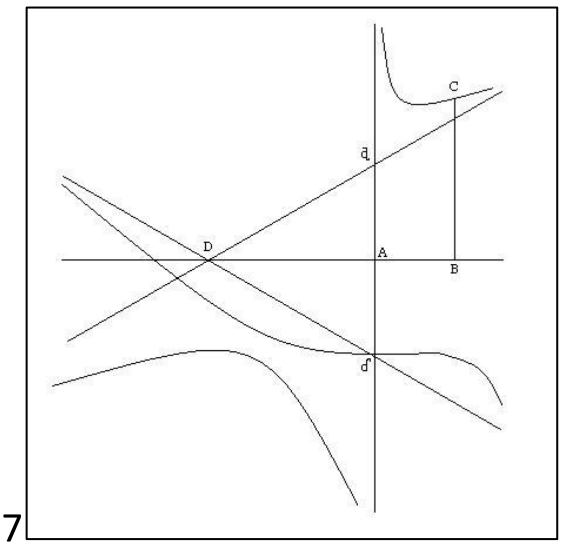
24



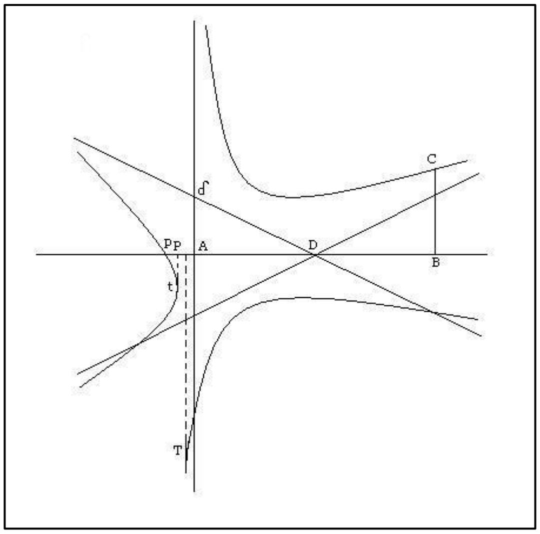
25



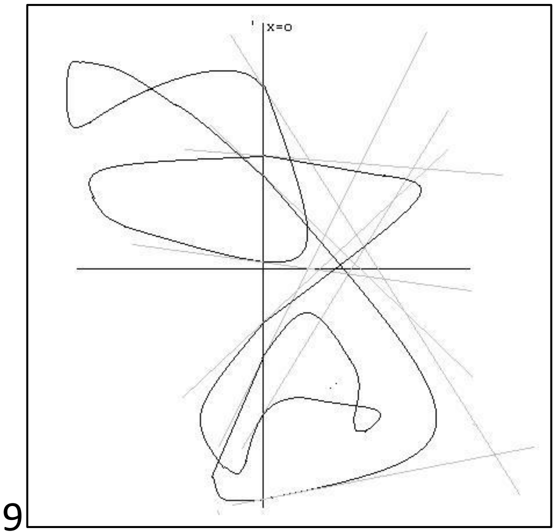
26



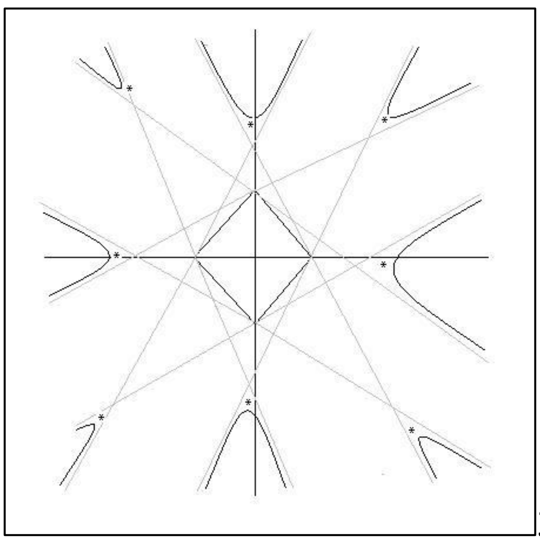
27



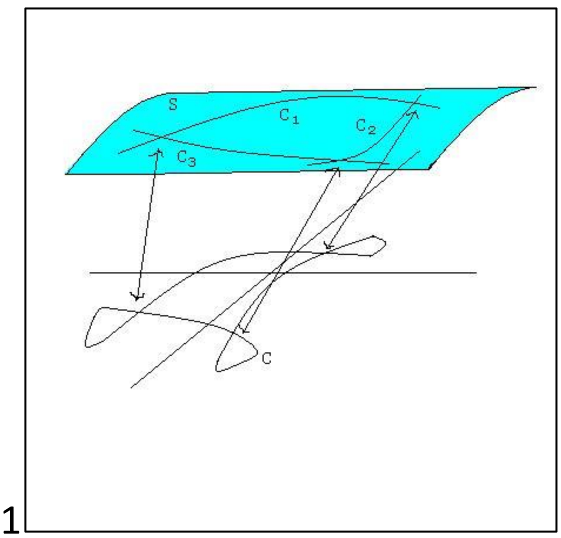
28



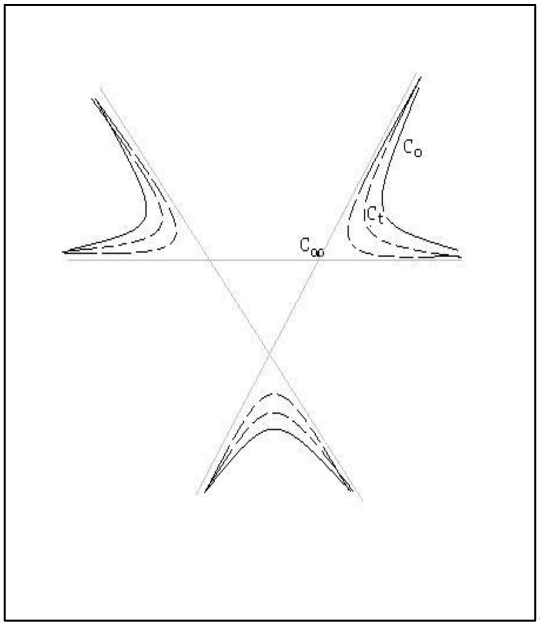
29



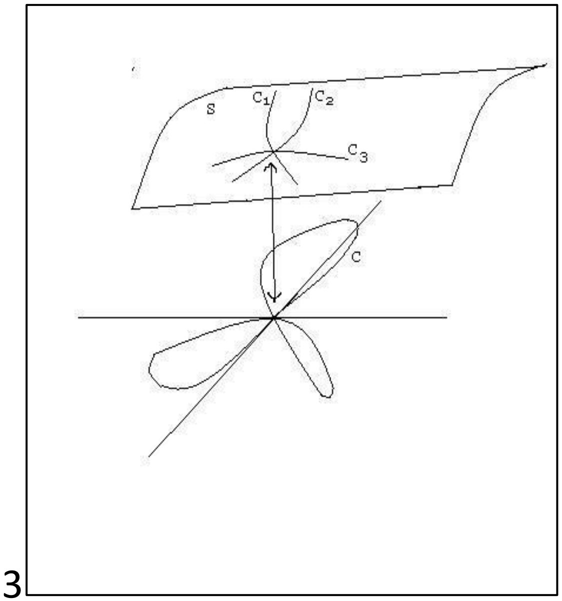
30



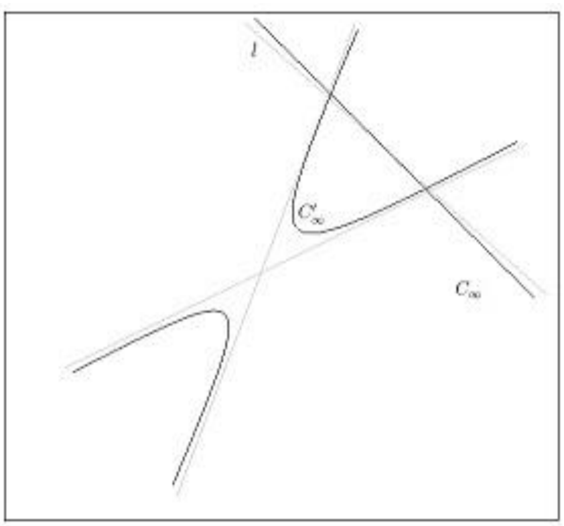
31



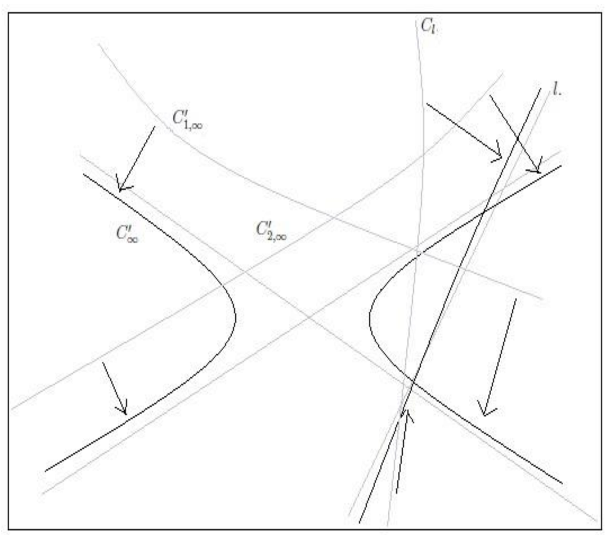
32



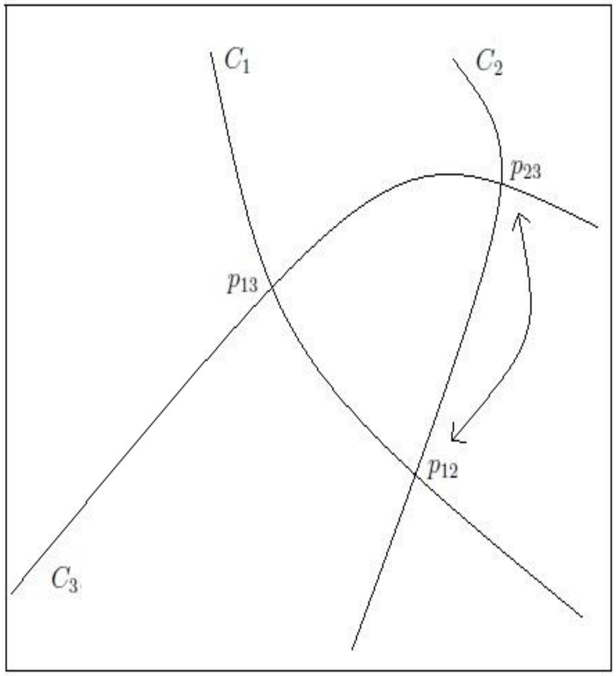
33



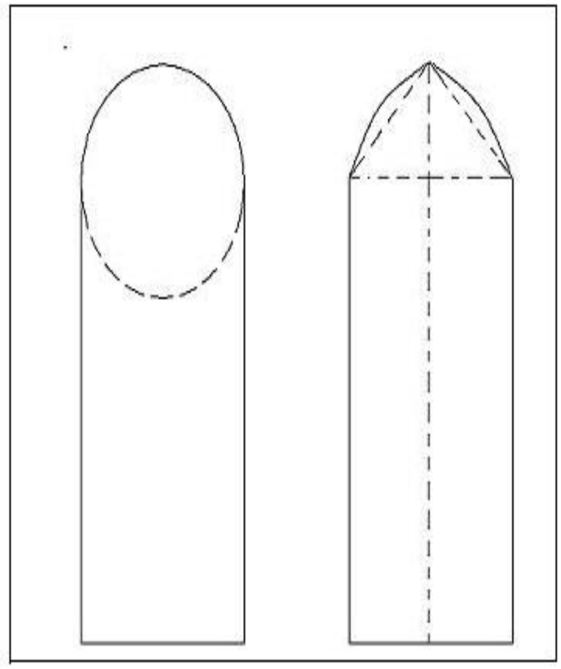
34



35



36



37

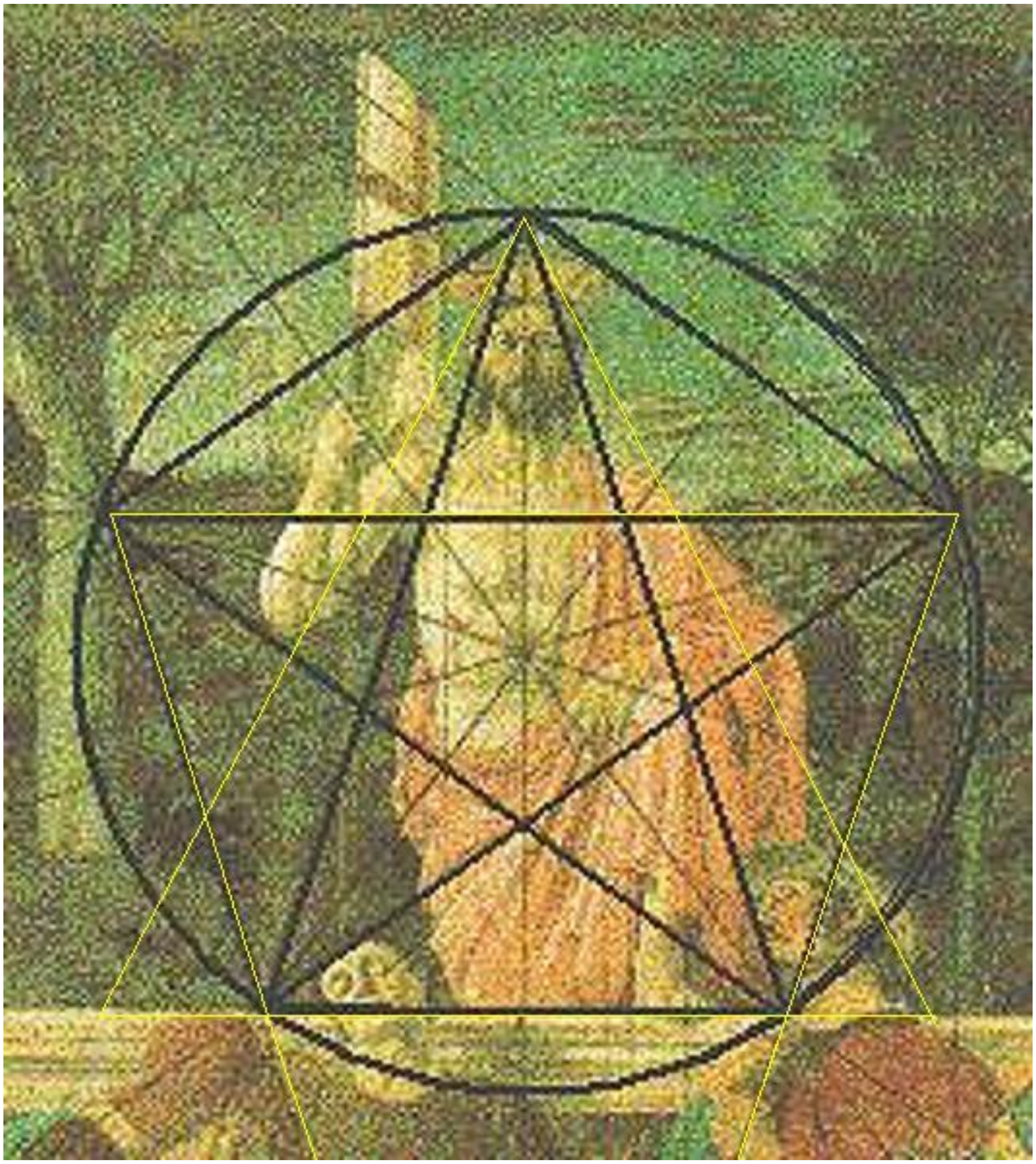
6. The Resurrection and the Lamp of Heaven

The purpose of this chapter is to consider the images of the Resurrection and the Lamp of Heaven, and to understand how they motivate a number of new aesthetic ideas, which we will develop further in the course of the book. Let me begin, then, by considering how some artists have employed these or closely related images.

Perhaps one of the most famous depictions of the Resurrection is by the Renaissance artist Piero della Francesca.

46. The Resurrection by Piero della Francesca, Museo Civico, Sansepolcro, Umbria, Italy.

In many ways, this painting is atypical of much of his work, which is concerned with light and perspective, subjects that we will consider, in depth, later in the book. The painting, commissioned for the town hall of Sansepolcro in the early 1450's, depicts Christ rising from the tomb, above a group of sleeping soldiers. There is a sense of majesty about the scene, emphasized by the solidity of the figures and the dignity of the facial expression of Christ. There is also a feeling of change, accompanying the event of the resurrection. The artist expresses this in his depiction of the landscape as in a process of rebirth, and a carefully constructed dynamism of colour and form in the figures. Piero uses the colour red for Christ's cloak, the banner that he holds in his right hand, and in the faces of some of the figures. There is an upward movement in the figure of Christ himself, reinforced by the artist's use of two point perspective, and the diagonal sweep formed by his cloak and the group of soldiers, seated at the left of the tomb. However, the artist is careful to subordinate this dynamic, by his harmonious placement of the figure of Christ, in relation to the soldiers beneath. This is achieved by the use of triangular constructions in the figure of Christ and the horizontal line of the tomb, and the two groups of soldiers with the central line of the landscape. The two triangular compositions, together, form a hexagonal arrangement, a geometric device, which, Piero della Francesca employs in other earlier works, ([20]). An alternative interpretation places the triangles in a pentagonal configuration, ([21]). The following diagram illustrates both possibilities.



47. Diagram of the Resurrection by Piero della Francesca, Museo Civico, Sansepolcro, Italy.

Another Italian artist who employs the image of the Resurrection, at least indirectly, is Raphael, in his painting "The Transfiguration". This event, recorded in the gospels of Matthew, Mark and Luke, occurs before the disciples Peter, James and John, who witness Jesus's appearance change "so that his face shone like the sun and his clothing became dazzling white" (Matthew Ch. 17, v2). The similarity with the event of the resurrection is reinforced by Jesus's commandment , "Don't tell anyone what you have seen, until I, the son of Man, have been raised from the dead".

[20] The interested reader should look at "Piero della Francesca" by Marco Bussagli.

[21] The interested reader should look at the article "Pentagonal Geometry in Piero della Francesca's Resurrection", by Chris Miles.

48. The Transfiguration by Raphael, Pinacoteca, Vatican, Rome.

Raphael's painting, created at the height of the Renaissance in Rome, in 1517, was commissioned by Cardinal Giulio for the Cathedral of Narbonne in France. However, such was the quality of the painting, that it remained in Rome. Vasari described it in his "Lives of the Artists", (Vasari, 2005), as "his most beautiful and most divine work". Compared to Piero's interpretation, there is an even greater dynamic energy in the representation. Raphael achieves this through his characteristic use of spiral movement, running through the prostrate figures of James, Peter and John. Like Piero, he is careful to subordinate this energy to a harmonious placement of the individuals in the scene. In this case, a large figure of eight unites the composition, extending from Christ's outstretched arms, down to the crowd of people below, a reference to the later event of Jesus healing a demon possessed boy. The mystical awareness that Raphael displays in this interpretation was possibly due to his attendance of meetings of "The Oratory of Divine Love", a group of Catholic laymen and priests in Rome, anticipating the reforms of Martin Luther. In this, his last work, painted shortly before his death in 1520, at the age of 37, and placed above his tomb in the Pantheon, Raphael expresses a profound, feminine understanding of the idea of resurrection. Although many critics would suggest that he was essentially motivated by Classical ideals, I will argue, later in the chapter, that his style and interpretation, even of classical themes, was extremely original and heavily influenced by medieval art.

In order to partly support this idea, and to continue with artistic interpretations of the resurrection, it is interesting to consider the work of the Cosmati, that we briefly alluded to in a previous chapter. The Cosmati were a group of artists, active in Rome roughly between 1200 and 1300, ([22]). Although primarily designers of pavimenti, they were also architects, responsible for many of the campaniles attached to Roman churches and some of the most beautiful cloisters in Italy. Their interest in the image of the Resurrection lies in the use of a certain frequently recurring motif within their designs, the guilloche. This design may be seen, for example, in Santa Maria de Trastevere, Rome.

49. Pavement of Santa Maria de Trastevere church, Rome, Italy.

It consists of two interweaving threads that run along the central section of the nave. The braiding image is repeated in the architecture of twisting columns, which can be found in cloisters such as Subiaco, San Paolo Fuori de Le Mura, San Giovanni in Laterano and Sassovivo

[22] The interested reader can find out more about their work in “Cosmati Pavements: The Art of Geometry” , Tristram de Piro, Bridges Leeuwarden Proceedings, 2008.

49. Cloister, Santa Scholastica Abbey, Subiaco, Lazio, Italy.

50. Cloister, San Paolo fuori de Le Mura church, Rome, Italy.

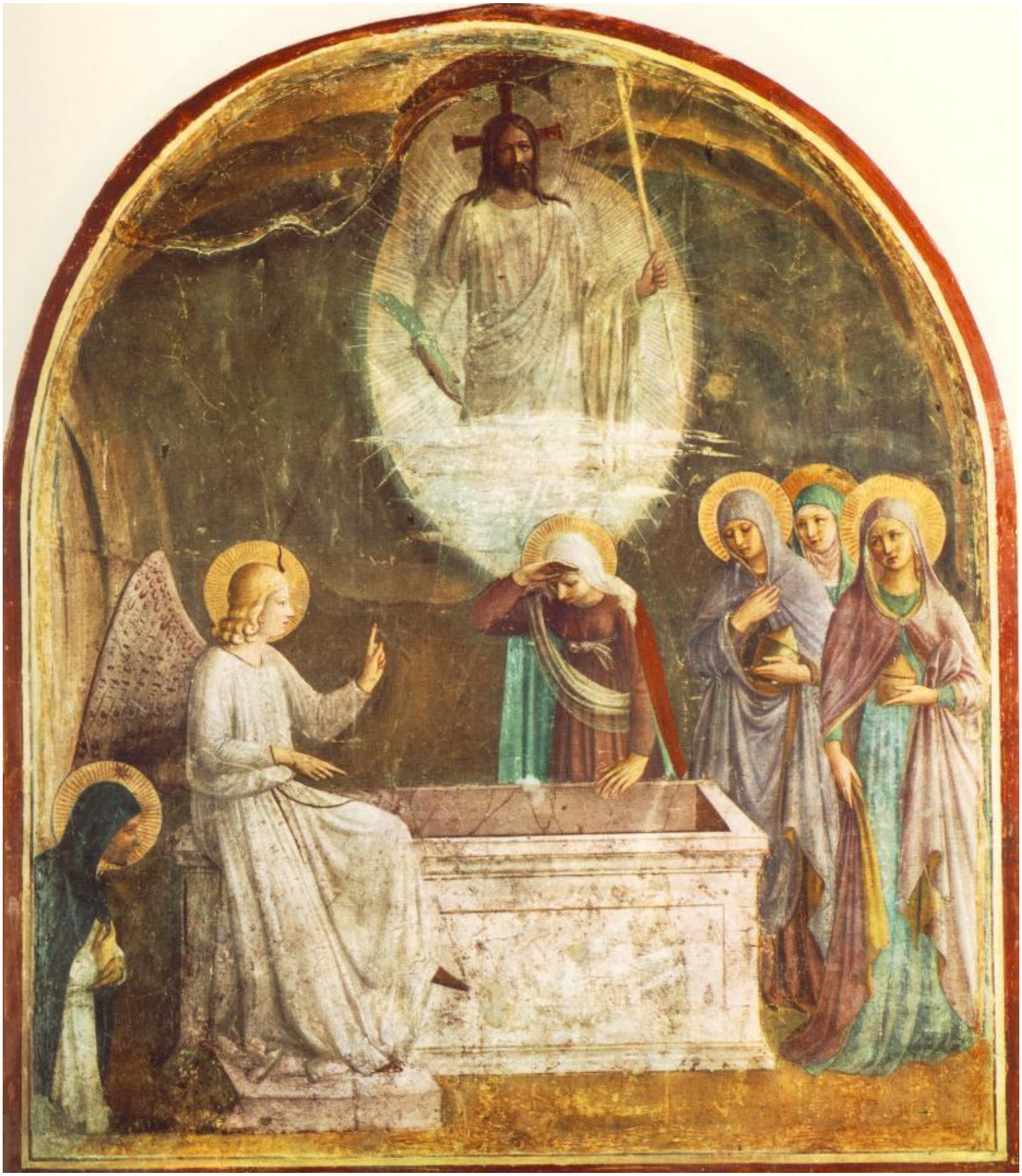
51. Cloister, San Giovanni de Laterano church, Rome, Italy.

52. Cloister, Sassovivo abbey, Umbria, Italy.

and also in a number of Paschal Candelabra, for example, at Ferentino. One can find other interpretations of the guilloche form, in particular, as being symbolic of the River of Life, ([23]). The flowing movement, that is a feature of Cosmati art, is undoubtedly an influence on Raphael's later work, after he arrived in Rome, around 1509. The idea of braiding, peculiar to the guilloche design, is, perhaps imitated in Raphael's use of the "figure of eight" composition from the Transfiguration, although, we will find it repeated in other of his works that we consider later.

Three other images of the Resurrection, which are interesting to compare, are by the artists Fra Angelico and Grunewald. The painting of the resurrection by Fra Angelico, from the 1400's,

[23] This is suggested by Kim Williams, in her book "Cosmati Pavements:Patterns in Space", based on the flowing nature and length of the design, which runs along the entire length of the nave in Santa Maria de Trastevere. However, given the aesthetic of fluency, which may be found in nearly all Cosmati patterns, and the functional requirements of pavements, this interpretation of the guilloche form is perhaps over simplistic.



53. The Resurrection by Fra Angelico, San Marco Museum, Florence, Italy.

shows the figure of the risen Christ, hovering above an empty tomb, surrounded by a group of angelic onlookers. Unlike previous interpretations that we have considered, the artist makes no attempt to convey a sense of harmony or dynamism to the image, through a geometric device. However, the colour scheme of the artist, a combination of gold, in the haloes of the onlookers, and red, in the crossed halo of Christ, to some extent, transmits these aesthetic ideas in a simpler, more visual form. The painting of the transfiguration by the same artist,

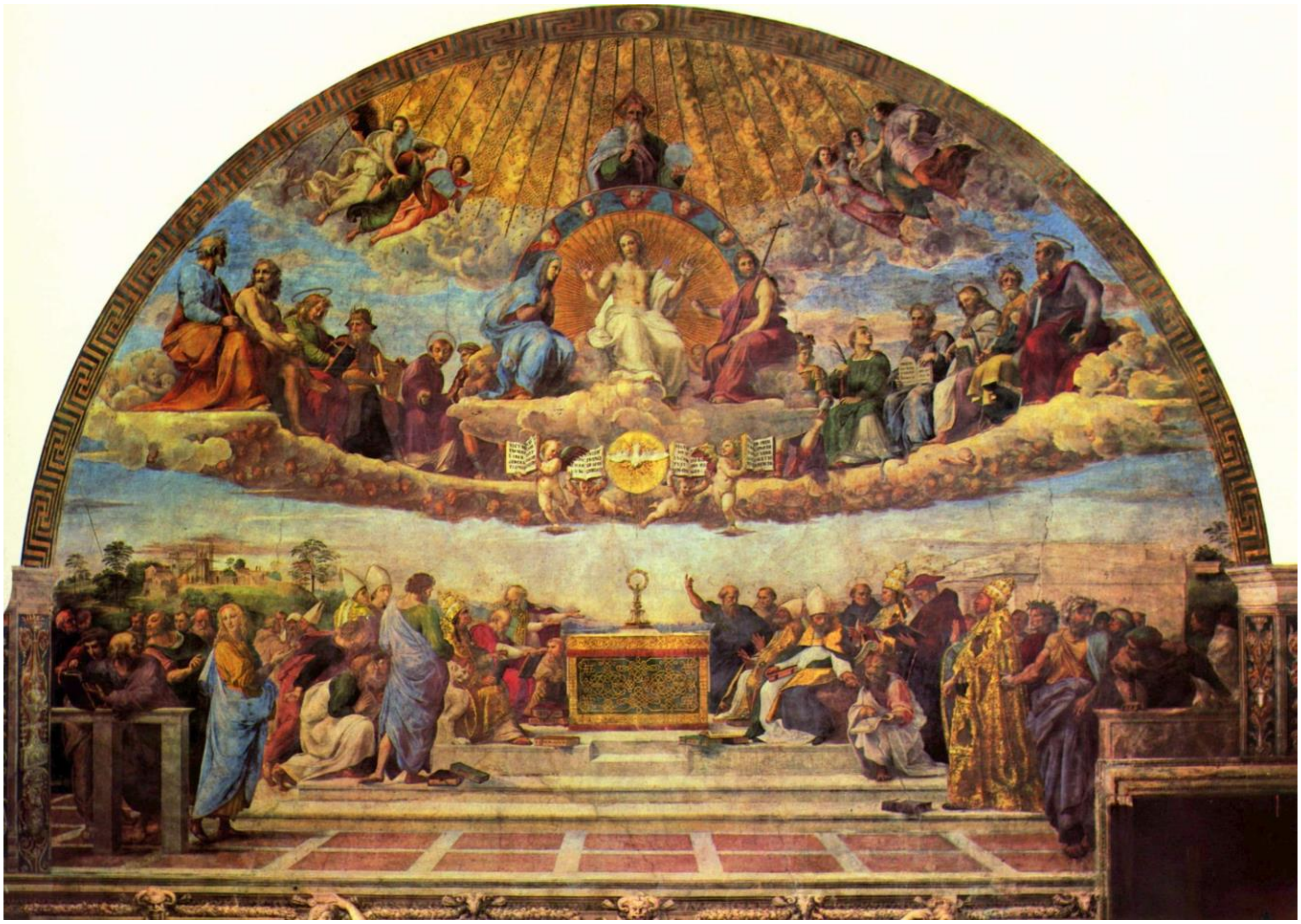
54. The Transfiguration by Fra Angelico, San Marco Museum, Florence, Italy.

from the later period of the 1430’s, uses a similar colour scheme. A clear golden light radiates outwards from the figure of Christ, again wearing a red crossed halo. In both paintings, there is a gradation of colour, from gold to red, as one moves away from the centre, indeed his first painting is surrounded by concentric bands of these colours. The painting of the resurrection by Grunewald,

55. The Resurrection by Grunewald, Unterlinden Museum, Alsace, France.

from 1515, as part of the Isenheim altarpiece, is similar, in form, to the compositions of Fra Angelico, with Christ hovering above an empty tomb. However, the colour scheme that he uses reverses the one used by Fra Angelico. Christ is clothed in a red cloak, surrounded by a circle of yellow light, following the traditional intuition of the structure of the sun. The reason for this reversal is an interesting question, which I will consider more fully later.

The description of light in these last three paintings demonstrates the close association of the event of the resurrection with that of the image of the Lamp of Heaven, which we will now consider. The artist Raphael provides two interesting examples of depictions of either heaven or visions associated with heaven. His “Disputa”, painted for the “Vatican stanze” between 1509 and 1510, traces the origin of the sacrament.



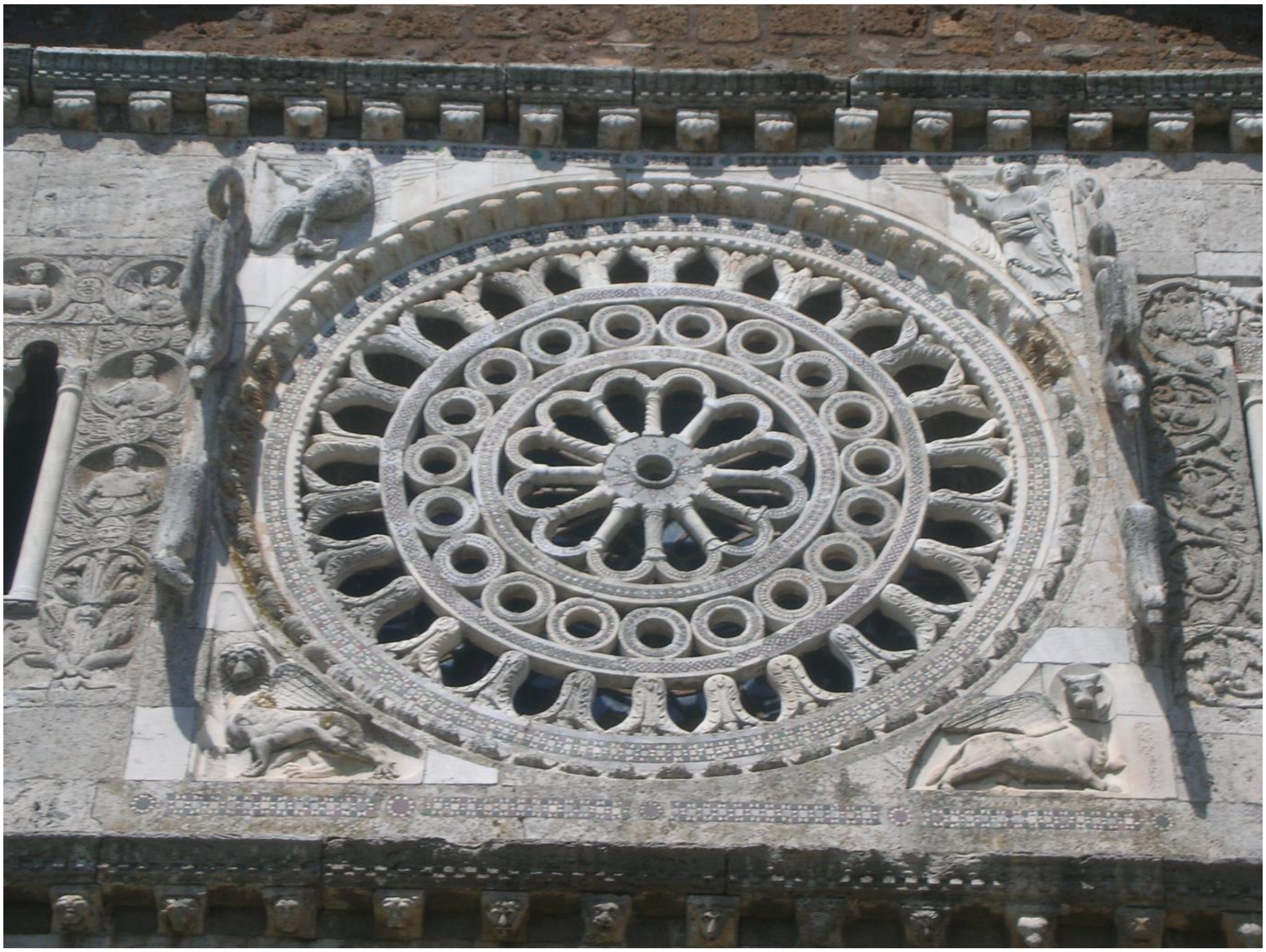
56. The Disputa by Raphael, Vatican Stanze, Rome, Italy.

God, Christ, the Virgin and St. John the Baptist are enthroned, surrounded by a majestic semicircle of figures, with angels spiraling above them in a golden sky. There is a stronger sense of harmony than in his depiction of the transfiguration. “Ezekiel’s Vision”,

57. Ezekiel's Vision by Raphael, Pitti Palace, Florence, Italy.

painted in 1518, now located in the Pitti Palace, Florence, is a Blakean vision of God, surrounded by a piercing golden light, recalling the event from the Old Testament,([24]) . The description of the light of God, as being of jasper, which we observed in the identification of Christian imagery, is again suggestive of the image of the Sun, although the artist prefers to convey this impression through the dramatic portrayal of the central figure.

The same symbolism underlies the design of many Romanesque rose windows in Italy, a particularly good example being that at San Pietro in Tuscania.



58. Rose window of San Pietro church, Tuscania, Tuscany, Italy.

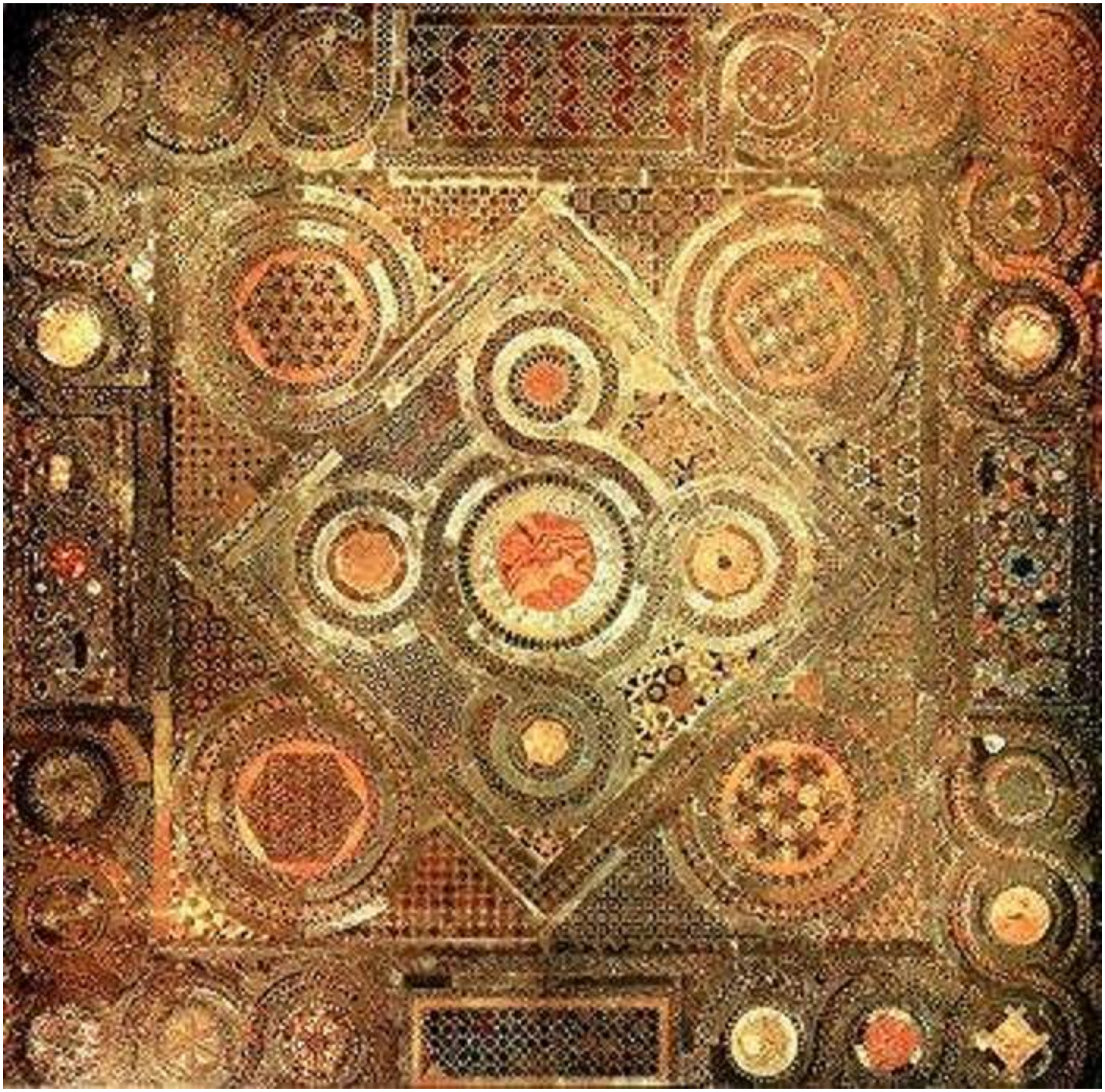
This window is set as a wheel, within a square, containing the symbols of the four evangelists as hybrid beasts at each corner, both features of Ezekiel's vision from the Old Testament, ([25]).

The geometry of the window at Tuscania is used repeatedly in another motif of the Cosmati artists, the quincunx, which can be seen at Santa Maria in Cosmedin, Rome, see (Massimi, 1953), and also at Westminster Abbey.

[24] Two of the surrounding figures are hybrid beasts, representing the apostles.

[25] On either side of the window are carvings of dragons, placed perpendicular to the façade of the church. This creates a dynamic effect, typical of the Italian style of Romanesque architecture. It is worth noting Ruskin's observation on the recurrence of the "dragon form" in this style, "The most picturesque and powerful of all animal forms and of peculiar interest to the Christian mind."

59. Pavement of Santa Maria in Cosmedin church, Rome, Italy.



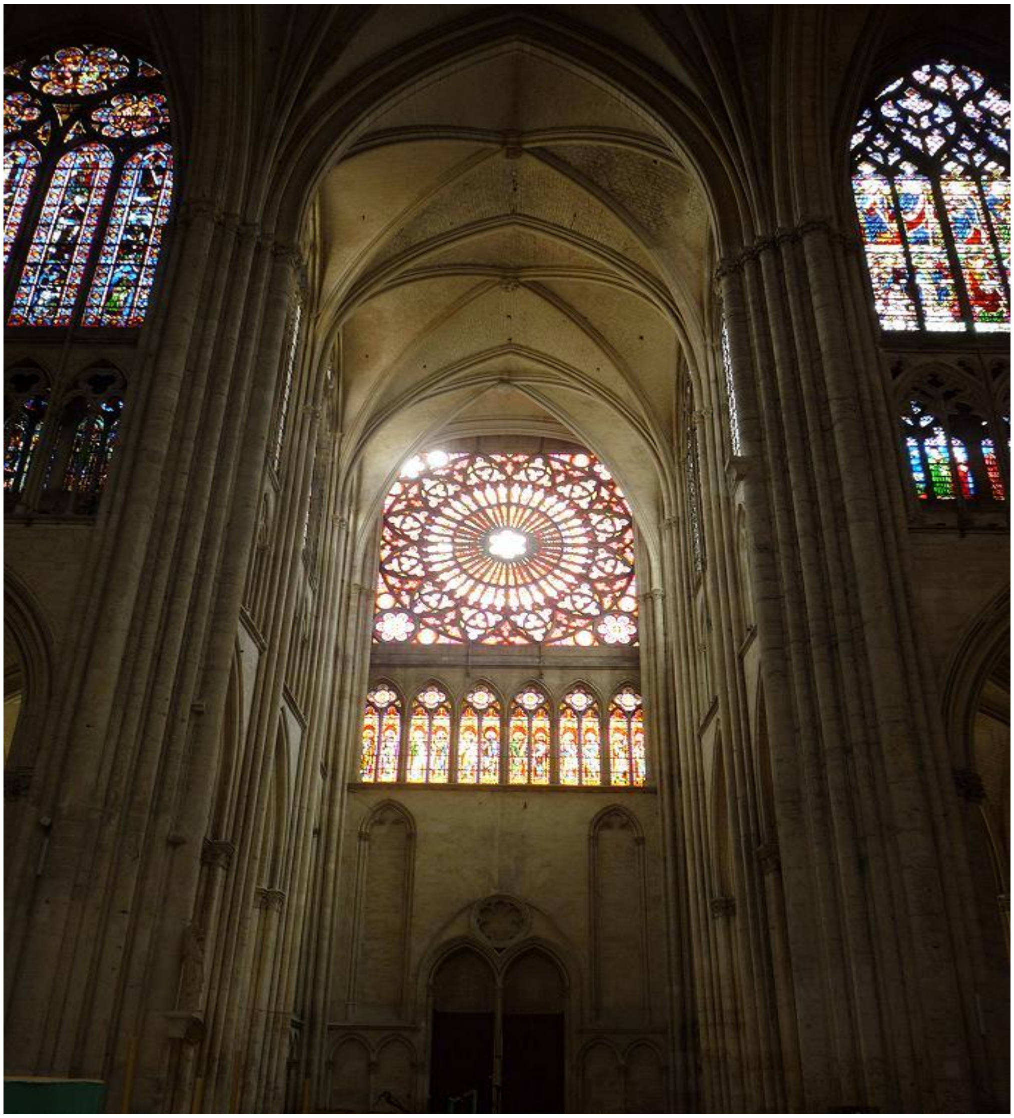
60. Pavement of Westminster Abbey, London.

It consists of a central roundel and four smaller surrounding eyelets, around which a single thread is intertwined. In Ayuela's "Cosmatesque Ornament", (Pajares-Ayuela, 2002), we are given the following description of the symbology of the quincunx;

"A 2-dimensional abstract representation, that is, the monogram or coded representation that signifies the 3-dimensional reality associated with the medieval Christian cult".

As we have seen, this 3-dimensional reality is the tetramorph, consisting of the evangelists, surrounding Heaven or the symbolic representation of Christ as a circle. As an image of the Lamp of Heaven, one could also view it as an abstract representation of light, radiating from a central source.

The strong association of light with the image of the Lamp of Heaven can be seen in other designs of medieval windows. The following picture is of the rose window of Troyes in France.



61. Window of Troyes Cathedral, Troyes, France.

Again, one sees the characteristic combination of red and yellow in the stained glass. There is an ambient sense of focus towards the interior of the window, and a stronger focus at the centre and the exterior. This is partly the optical effect of the photograph but, also, the colour of the glass, which radiates in red bars, on a yellow background, from the centre to a surrounding circle. The effect of light through the window is almost tangible, perhaps due to the analogy of extremely focused light with heat or flame.

The images that we have considered lead to a number of new aesthetic ideas. The rest of the chapter will be mainly concerned with examining these ideas through examples in art and architecture. My major sources will be medieval art and architecture, in particular the development of the Romanesque style and the work of the Cosmati, and Renaissance art, with particular reference to the work of Raphael.

The Aesthetics of Focus and Harmony in Medieval Art and Architecture

The term "Romanesque" is mainly used to refer to a style of architecture that became popular throughout Europe, between 1000 and 1200. Its origins can be traced back to the architecture of the Roman period, and , it is, perhaps, for this reason, that many people discount the originality of its designs. Its most characteristic feature is the use of the rounded arch, instead of the pointed form, adopted by the later Gothic style. The rounded form was undoubtedly a Roman invention, with antecedents in the triumphal arches, found near the Forum of Rome, ([26]). However, the Romanesque style is responsible for a number of aesthetic innovations, independently of Roman predecessors, which make it an important field of study in its own right, not least because of the geometric intelligence that underlies many of them.

The first Romanesque buildings were the Christian basilicas. Admittedly many of these were plagiarisms of previous Roman constructions. Santa Sabina, AD 422-432, was built on an ancient Roman temple, creating the interesting aesthetic effect of free standing columns, used to support the weight of the arches. This was also a Roman invention, found in the Diocletian palace of Split, AD 300-330. Old St. Peters, on the site of the present Vatican, from the 4th century, displayed a classical dignity, enhanced by the use of horizontal architraves, an architectural feature pioneered by the ancient Greeks, in the 5th century BC. A visitor to San Paolo fuori le Mura, partly destroyed by a fire in the 19th century, is impressed by the cathedral's classical gravitas, intensified by its size and numerous arches, each decorated by an image of one of the popes of Rome. Santa Prudenziana, AD 384-399, was built on the site of a Roman thermae. All of these buildings inherit the classical sense of form and mass, evident in Roman and Greek architecture.

Following the collapse of Rome, the style lost much of its original impetus. Some developments of the basic arch form, implementing a horseshoe design, occurred in Visigothic Spain, at San Juan de Banos de Cerrato, in 661. In England, the basilica form was adapted, by the Anglo Saxons, to the construction of simple, aisle less stone churches, for example, at Escomb in Hexham, 672-678. However, it was during the second half of the 8th century that the style became transformed, in the period known as the Carolingian renaissance.

Under the influence of the emperor Charlemagne, the Carolingian empire became the dominant force in central Europe. It saw a redefinition of Christian architecture, mainly in France and Germany, but also as far afield as Italy, where the remains of a Carolingian basilica can be found in the Benedictine abbey of San Vincenzo al Volturno, 60 miles south of Rome. For the first time, we see a genuine departure from the purely classical influence of Greece and Rome, in a number of new architectural themes.

The first of these is the invention of the "westwork", literally a castellum or turris, at the west end of a church. Possibly used, initially, for defensive purpose, the westwork emerged as a "harmonic facade", consisting of a pair of towers, at the entrance. The first example can be found at St. Riquier, in France, built between 790 and 799, but later destroyed in 881. A surviving example exists at Corvey in Westphalia, Germany, 873-885. The westwork creates an aesthetic effect of balance and harmony, but also has a spiritual significance, in that the overall form is not enclosed. It serves to focus the eye towards the building, a theme which we will find repeated in later Romanesque constructions. For these reasons, we begin to see the true spirit of the Romanesque, as both developing and incorporating the classical principles of harmony, inherited from its initial Roman impetus,([27]).

The second important innovation occurs at the palace chapel of Aachen.

[26] Some good examples are that of Septimus Severus, AD 203, the Arch of Titus, AD 83, Trajan's Arch, AD 114, and the Arch of Constantine, AD 326.

[27] These principles can be seen in the T-transept designs of St. Peters and San Paolo fuori le Mura, an idea copied by the Carolingians at the monastery of Fulda, Germany, 790-819, and, perhaps, motivating the westwork.

62. Interior of Charlemagne's Palace chapel, Aachen, Germany.

Built for Charlemagne, who was the ruler of the Carolingian empire from 768-814, it consists of a double shell design with an octagonal core. It was intended as a model of the heavenly Jerusalem, as can be seen by noting the numerical coincidence that the inner octagon has a circumference of 144=8x18 Carolingian feet, while the walls of the heavenly Jerusalem measured 144 cubits. This correlation is further supported by an inscription, in Latin, on the interior cornice;

“When the living stones are assembled harmoniously, and the numbers coincide in an equal manner, then rises resplendently the work of the Lord who has constructed the entire hall.”

63. Detail of interior of Charlemagne's Palace chapel, Aachen, Germany.

The interest in harmony is here demonstrated on several levels. First, there is a numerical relationship between the lengths of the interior walls, as referred to in the inscription. Secondly, there is a geometrical spatial harmony, evidenced by the use of an octagon, as the basic form. Thirdly, there is an association with the heavenly Jerusalem, which, we argued, was a symbol of harmony in the Christian ideology. These first two associations are further confirmed by the following construction; in order to draw a regular polygon with an even number of sides, it is sufficient to construct a polygon with half that number inside a circle, draw a concentric exterior circle, and mark off the intersections of the circle with the initial lines. This is, in fact, exactly the construction used in the double shell design of Aachen.

64. Exterior of Charlemagne's Palace chapel, Aachen, Germany.

It is supposed to have inspired a number of architectural copies, including The Bishop's Chapel of Hereford Cathedral. Hans Boker, in his article on the subject, (Boker, 1998), rightly observes the difference in the design to Hereford, but misses Aachen's geometrical innovation, and, hence, perhaps misleadingly, attributes its source to Byzantine origins. One may also compare the chapel at Aachen to previous classical constructions, such as San Vitale in Ravenna, which also employs a double shell structure with an octagonal core. However, the designs are different in so far as a 16-sided figure is used for Aachen's outer wall, although San Vitale may have inspired the central section.

The development in the aesthetic of harmony is also found in a further Carolingian innovation at San Germigny des Pres. Built in 805, it is one of the oldest churches in France. The original plan of the oratory is a square lantern shaped bell tower, with four radiating apses. The design, importantly, prefigures the quatrefoil form, that came to be used, extensively in the geometric period of Gothic art and castle construction. Examples, in the form of windows, were mentioned in an earlier chapter, the reader interested in the connection with castles, should consider the ground plans of French buildings such as Etampes, and, to a lesser extent, Ambleney and Provins, ([28]).

65. Oratory of Saint Germigny des Pres, Loire Valley, France.

[28] It is the author's opinion that the diagonal embrasures of the lantern tower were added later to the original building. This view is supported by the extensive use of embrasures in Norman buildings, and a careful inspection of the colour of the stonework.

These innovations were, to a lesser extent, continued during the subsequent period of the Ottonian dynasty, 936-973, which became a ruling force, after the Carolingian empire disintegrated. Some new architectural features may be found, at Gernrode, established in 961, including the use of alternating piers and columns to support the main arcades, ([29]).

The next important developments occurred during the period known as The First Romanesque style, of the 11th century, mainly in the South East of France. The style was spread, primarily, through the work of Lombard masons and its new architectural features include the use of bell towers and Lombard bands, thin pilaster strips which run up a wall, between a string of small blind arches. Two of the finest examples can be found at Anzy-le-Duc, in Burgundy, and St. Michel-de-Cuxa, in Roussillon, both from the early 11th century. The use of Lombard bands subtly increases the height of a building, hence, the aesthetic effectiveness in conjunction with features such as bell towers. At Anzy-le-Duc, the effect is carefully harmonised by the polygonal profile of the bell tower, possibly inspired by Carolingian work,

66. Priory church of Anzy-le-Duc, Burgundy, France.

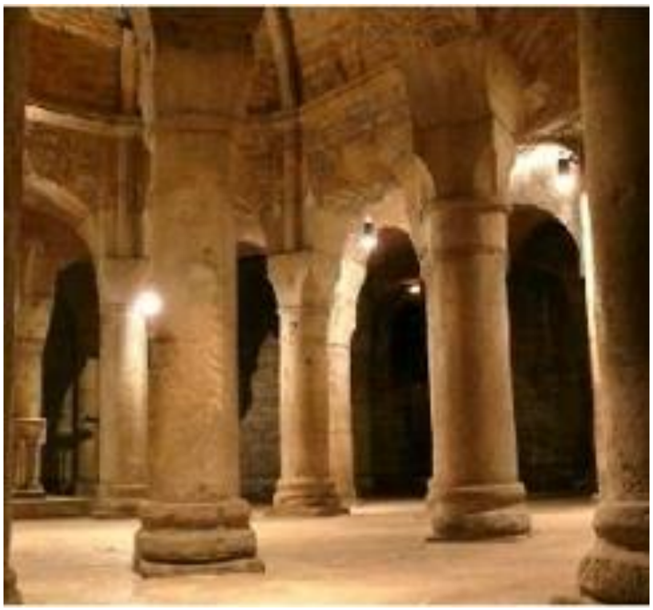
while at St. Michel-de-Cuxa, it is counterbalanced by the broad sweep of the nave.

[29] This became a prevalent feature in later German Romanesque art, after 1000, see (Carl, 2008). The period also also inherited the legacy of Carolingian architecture, in, for example, the two-sided arrangement of masses at Hildesheim, 1013-1033.

67. Saint Michel-de-Cuxa Abbey, Roussillon, France.

In both buildings, we find a continuity in the aesthetic of focused light, enhanced by the soft colour of the stonework.

The Carolingian influence may be found in buildings of this period, particularly in the architectural features of crypts and semi-circular apses. The crypt of Chartres cathedral, dating from the 9th century, is possibly the origin of this former invention. A fine example may be found at Dijon-St-Benigne, from the early 11th century, in which the columns form three concentric rings. It originally formed the base of a rotunda, which was later destroyed, during the French revolution.



68. Crypt of Saint Benigne-de-Dijon church, Burgundy, France.

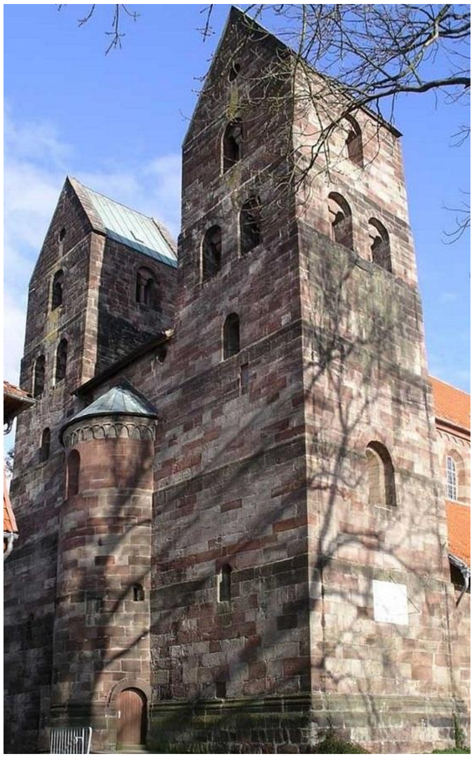
The semi circular apse is possibly derived from San Germigny du Pres, and is present at both Anzy-le-Duc and St. Michel-de-Cuxa, as well as other churches we will mention later, such as Toulouse St. Sernin, from the end of the 11^{th} century. The theme of the “westwork” is repeated in 11^{th} century constructions such as Morienval, in Picardy;

69. Abbey church of Morienval, Picardy, France.

which also has fine early examples of Romanesque capital carving;

70. Corner capital of the Abbey church of Notre Dame, Morienval, Picardy, France.

The westwork also reappears in Germany, in the Ottonian period (c920-1024), at Gernrode and Hidesheim, and, at the abbey churches of Maria Laach and Fredesloh , (1025-1200);



71. Fredesloh church, Saxony, Germany.

An enigma of the French Romanesque is the introduction of the chevet, consisting of a series of radiating chapels or apses in echelon, that circle an outside sanctuary, the interior forming an ambulatory or walkway at the east end. It is the author's opinion that this development coincided with the knowledge of concentric castle design, first practically implemented about the time of the Second Crusade, in the 1140's. Several authors, such as (Strafford, 2005), and (Hubert, 1938), trace the development to the earlier First Romanesque style, citing, for example, the monastery of St. Martin du Canigou from the early 11th century. There is, perhaps, some reason to be suspicious of this claim, owing, first, to the fairly remote location of the building in the Pyrenees, and, secondly, to the construction of the chevet on a flat wall surface, different in conception to the use of the ambulatory, based on concentric circles. However, the idea is latent in the combination of concentric polygonal designs, developed by Carolingian architects, and the use of apses in echelon, derived from Lombardian art of the early 11th century, as at Lomello, and developed later extensively in Spain, see the next chapter.

72. Santa Maria Maggiore church, Lomello, Lombardy, Italy.

The introduction of the chevet is truly innovative in that it develops the aesthetic of harmony to new levels. There is a sense of concentric radiation, analogous to the natural growth of petals from a flower, or light emerging from the sun. However, it is, at the same time, deeply geometrical, a question we will explore in the next chapter. Certainly, two of the finest mature examples may be found in the Auvergne, Saint Austremoine, in Issoire, from the middle of the 12th century, and Brioude, built between the middle of the 11th and the end of the 12th centuries. At Brioude, the chevet is on two levels, with 5 chapels radiating from the apse. The construction is beautifully continued inside, with a semicircle of 4 columns, and a ring of windows, forming the ambulatory.

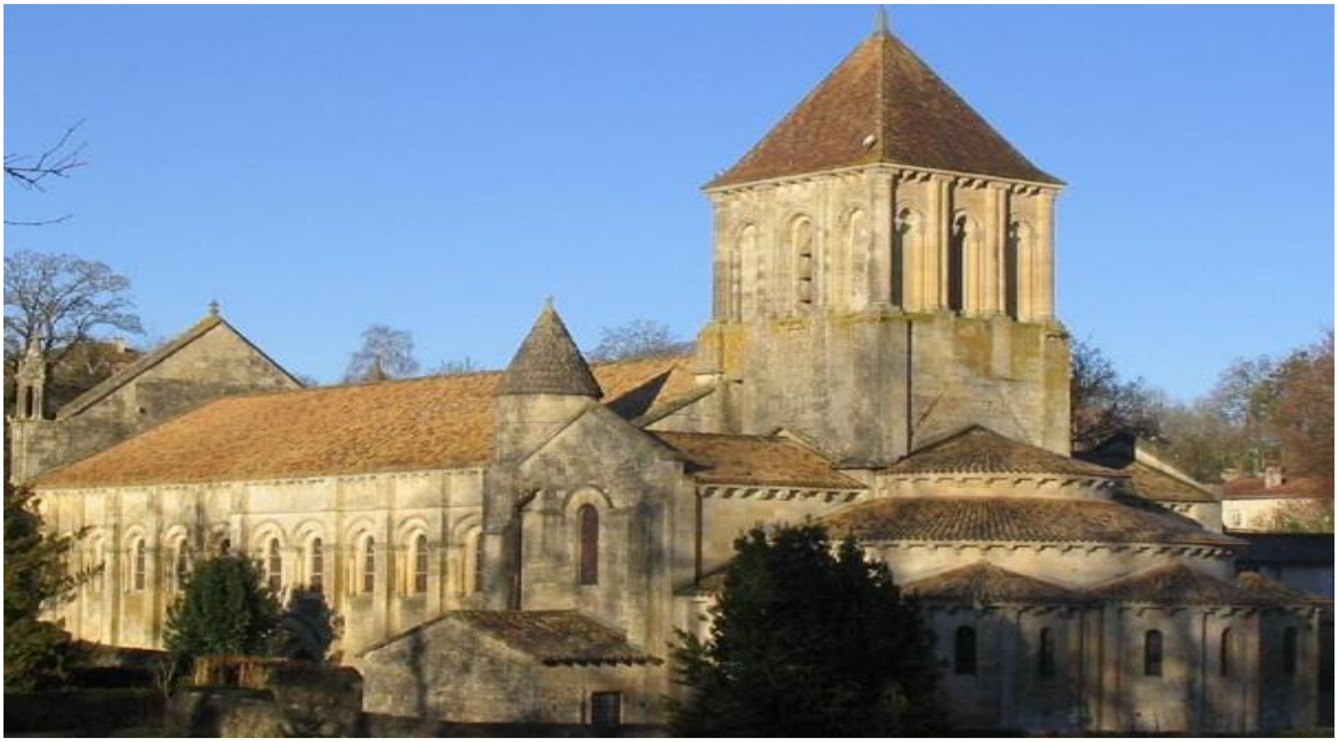
73. Basilica of St. Julien, Brioude, Auvergne, France.

At Issoire, the chevet is combined with a lantern transept and a balanced arrangement of masses at the north and south ends, recalling the idea of the westwork. The result is an elegant and compact three dimensional building, which, as a model of harmonious construction, seems difficult to surpass.

74. Saint Austremoine church, Issoire, France.

Further examples can be found at Notres Dame du Port, Clermont Ferrand, from the middle of the 12th century, and St. Hilaire, Melle and St. Foy, Conques, from the end of the 11th century;

75. Notres Dame du Port church, Clermont-Ferrand, France.



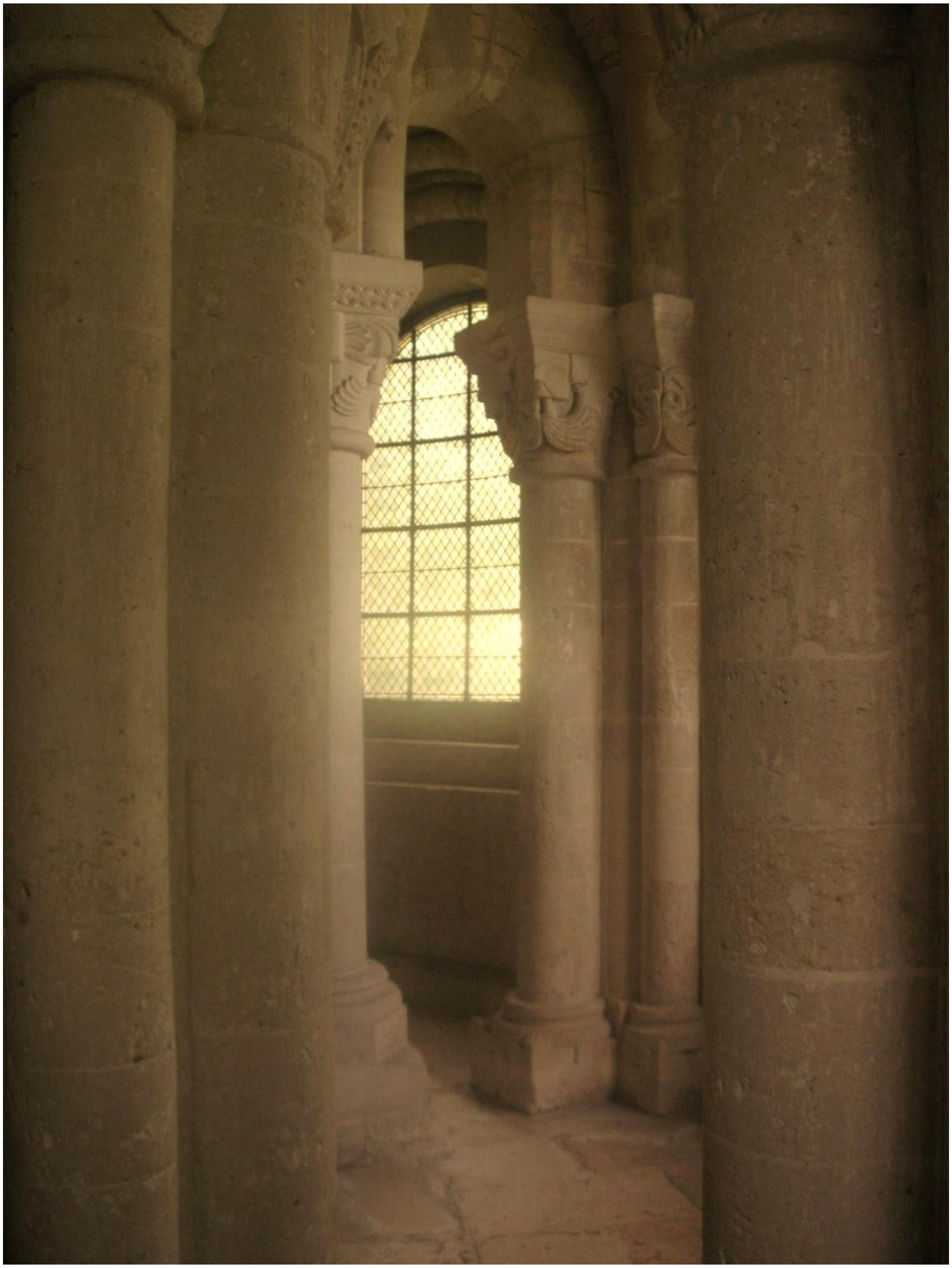
76. St. Hilaire church, Melle, Poitou-Charentes, France.

77. Abbey church of St. Foy, Conques, Aveyron, France.

A further harmonic device, introduced at Conques, and later at Paray-le-Monial, is the employment of a central tower, in conjunction with the use of a westwork, which was previously introduced in Germany, serving to focus the eye around the centre and periphery of the building. It is interesting to note that the abbey church of St. Foy is an immediate predecessor of Durham Cathedral, where a similar configuration of towers can be found.

78. Abbey church of St-Foy, Conques, Aveyron, France.

At Morienval, an ambulatory from the middle of the 12th century, which, in its simplicity, might even be a precursor to that of Saint Denis, echoes the form of the chevet, and demonstrates motifs of another important period of French Romanesque art, the style usually referred to as the Second Romanesque.



79. Corner capitals, Abbey church of Notre Dame, Morienval, Picardy, France.

These designs, with the use of wings and foliage, echo forms which date from as early as 1075, as found at the now ruined Abbey of Saint Martial in Limoges;

80. Capitals from the Abbey of Saint Martial, Limoges, Musee des Beaux Arts, Limoges, France.

A similar design in the capitals can be found at Buckfast Abbey, in Devon;

81. Corner Capitals, Buckfast Abbey, Devon.

This saw the development of barrel vaulting and groin vaulting, occurring when two barrel vaults intersect transversely. Both of these techniques were enhanced, in order to overcome the problem of vault thrusts, a completely satisfactory solution to which, only the Gothic style, which we have discussed, was able to find. The cathedral at Tournus, built mainly at the end of the 11th century and beginning of the 12th century, uses both transverse barrel vaulting and groin vaulting, one of the few remaining buildings of its kind.

82. Vault of Tournus Cathedral, Tournus, France.

A contemporary structure is that of the cathedral of St. Sernin in Toulouse, built between 1080 and 1120, featuring a majestic nave with barrel vaulting. St. Sernin was one of a series of five great pilgrimage churches on the route to Santiago de Compostela, itself a fine Romanesque cathedral of the second style.

83. Nave of St. Sernin church, Toulouse, France.

A fine example can also be found at the abbey church of St. Benoit sur Loire, built mainly in the 12th century, which features the use of blind arcades to counteract vault thrusts. In all of these French churches, we can find the characteristic feature of the chevet.

84. Abbey church of St. Benoit, St. Benoit-sur-Loire, Picardy, France.

There is more of an interest in the aesthetic of light, as opposed to harmony, in these examples, the chevet serving to focus the eye on the periphery of the east end, rather than comprising a significant part of the whole building. One should compare them with another contemporary structure, the abbey built at Cluny, in 1088, which later collapsed in 1128, also known as Cluny III. This had a unique example of a double ambulatory with five radial chapels.

These examples point to the date of the early 12th century, for the introduction of the rounded chevet form in French architecture. This date coincides with that of the First Crusade, suggesting that, at least a latent knowledge of concentric castle building existed at this time, even though it wasn't practically employed until, later, in the 1140's. According to (Nicolle, 2004), projecting towers on a curtain wall were used in the Byzantine province of Cilicia at the beginning of the 12th century, knowledge that the crusaders would have obtained during the siege of Antioch, 1097-1098.

An interesting variant of French Romanesque is the abbey church of Montmajour in Provence, built in the second half of the 12th century. In some sense, this develops the chevet form, with the introduction of a polygonal apse, surrounded by radiating chapels. We will discuss the geometrical implications of this arrangement in the following chapter.

85. Abbey church of Montmajour, Arles, Provence, France.

The interest in the aesthetic of focused light, through the use of the chevet, is also enhanced by the introduction of rounded windows, on a smaller scale to the earlier use of Romanesque arches. It seems possible that such forms were previously thought to be impractical, but that the new interest in exploring the possibilities of illuminating the inner spaces of cathedrals and churches, led to their introduction on a wider scale. An early example of the use of such forms can, again, be found at St. Foy, Conques;

86. Ambulatory of St. Foy, Conques, Aveyron, France.

87. Windows at St. Foy, Conques, Aveyron, France.

where, such windows are employed in a three-tier array, above the ambulatory. The height of the abbey church, the confidence with which the array is employed, and the earlier noted innovation of the turrets, might suggest that a sophisticated system of vaulting could have been used, based on the idea of using the chevet as a buttress, with the central tower as an apex, in some sense focusing, rather than pinpointing the vault thrust. This idea is reinforced by the configuration of windows in the north transept and in the western facade, mirroring the construction of the towers;

88. North Transept of St. Foy, Conques, Aveyron, France.

89. Façade of St. Foy, Conques, Aveyron, France.

A similar façade (from the 19th century) can be found at Buckfast Abbey, though there is some debate as to whether it resembles that of the original building.

This interest might be said to reach a culmination in the early Gothic masterpiece of the ambulatory of St. Denis, in a sense, continuing a Romanesque tradition. The aesthetic of light and its relation to harmony, is a feature of the metaphysics of this period. Light was seen as the source and essence of all visual beauty, to the extent that even something's objective value was determined by the degree to which it partakes of light.

This is the language that Abbot Sugar used to describe the choir at St. Denis;

"The entire sanctuary is thus pervaded by a wonderful and continuous light, entering through the most sacred windows"

The Neo-Platonism that underlies this philosophy saw the stability of the cosmos as grounded in proportion and geometry as the art by which the world soul, a metaphor for light, could access a spiritual harmony, symbolized by the image of the heavenly Jerusalem. At St. Denis, the axes of the sanctuary radiate from the centre of the choir to the windows between the surrounding chapels, enabling light to penetrate to the interior, forming the zone of "continuous light" that Sugar refers to. It is this additional interest in the effects of focusing light that distinguishes the ambulatory from its Romanesque predecessors, where the interiors are often relatively dark.

The philosophy of harmony in French Romanesque art is continued in the great Gothic edifice of the geometric period, Chartres cathedral, built between 1194 and 1220, after a fire destroyed almost all of the previous Romanesque church, ([30]). The interest in proportion and light is nowhere more evident than in the design of Chartres, indeed Chartres is the first architectural system in which these principles are completely realized. Von Simpson, in "The Gothic Cathedral", argues convincingly that the golden mean, the ratio of the length of a side of a pentagon to the radius of the circle in which it is inscribed, ([31]), was employed in the design of the crossing, specifically the ratio of the width to the length, as well as the elevations. This point of view is developed by Gordon Strachan, who observes that by locating two concentric hexagons at the crossing, one can locate virtually all the columns of the building by extending their diagonals. The latent use of the Carolingian device, employed at Aachen, seems to be a plausible Romanesque influence on Chartres' design.

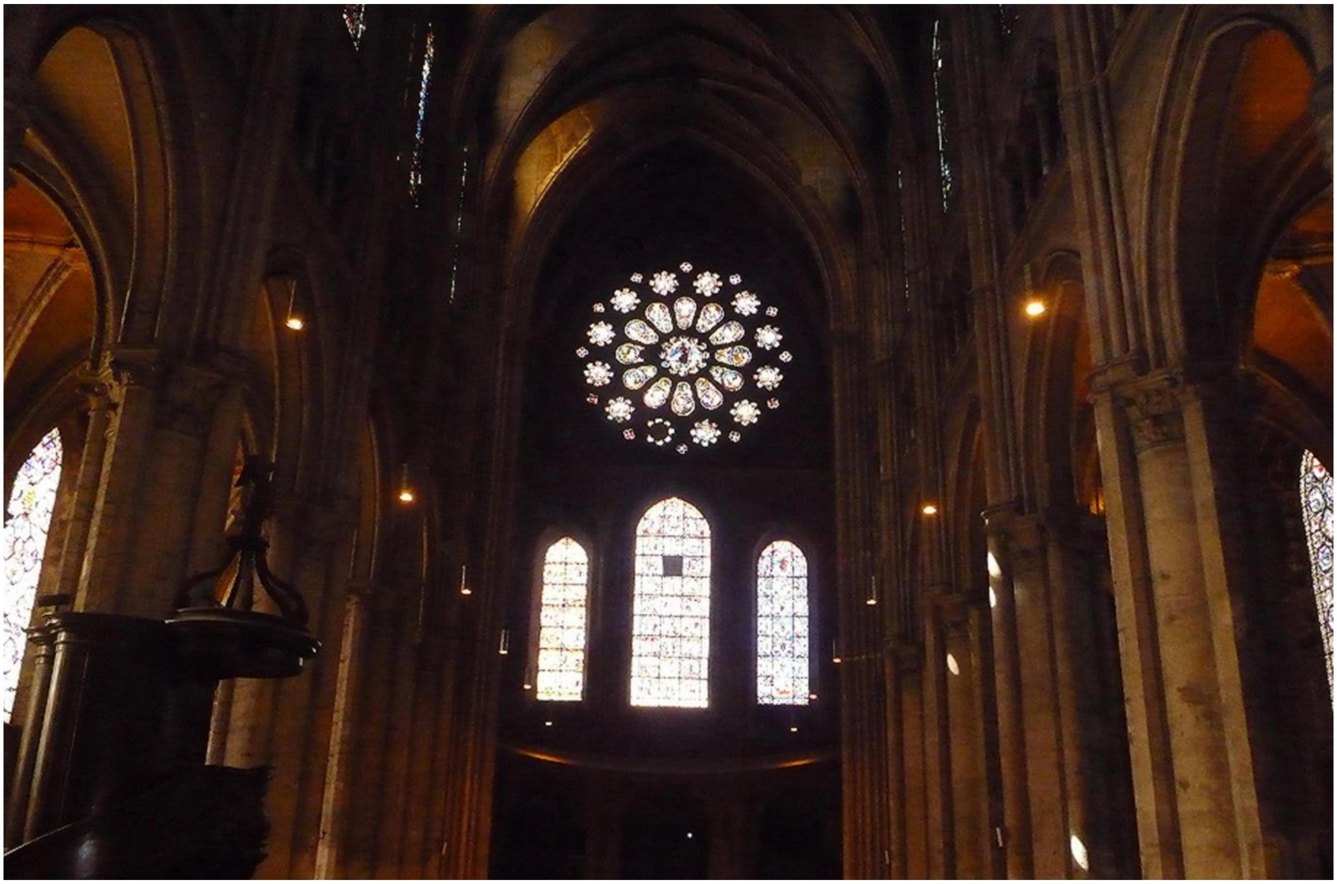
Nowhere is the interplay between harmony and light more manifest than in the design of Chartres' three rose windows, architectural devices which only became practical as a result of the innovations in Gothic design. The southern rose consists of a central medallion, surrounded by concentric rings of triangular lozenges, circles, quatrefoils and semicircles. It is typical of the geometric style, in its use of the quatrefoil, and displays a careful restraint, governed by the controlled geometry of each radiating segment. The northern rose is stylistically slightly different, replacing the ring of circles by squares, constructed in such a way as to form the corners of a larger square, rotated about the centre in three consecutive positions.

[30] There is some debate as to whether this church was responsible for the design of the crypt, which has an apsidal section with elongated radiating chapels, similar to the chevet form. In (Strachan, 2003), Gordon Strachan attributes the crypt and its current design to the Carolingian period, circa 858, whereas in (Ball, 2009) Philip Ball ascribes it to the invention of a Romanesque architect Beranger, circa 1028. These contradictory claims are based on archaeological and documentary evidence, respectively. The author is suspicious of both of these approaches, as they ignore the natural evolution of geometric styles obtained by considering a series of buildings and their developments.

[31] Also the number to which successive ratios of the Fibonacci sequence converge.

90. North rose window, Chartres Cathedral, Chartres, France.

The western rose is a central medallion, surrounded by lozenges, rosettes and quatrefoils, again typical of the geometric, in its profusion of circular forms.



91. West rose window, Chartres Cathedral, Chartres, France.

All three windows are beautiful, because of the pure harmony of their geometrical proportions, and the way in which mathematics and form are merged. There is an underlying cosmic symbolism of the rose, as representing both order and ornament, continuing the Neoplatonic philosophy inherent at Chartres, during this period. Perhaps, the most attractive feature of their style is the way that light and the eye of the observer, is focused equally about the window, an effect that the later Rayonnant style, with its emphasis on more centralized forms, was to change. One can appreciate how these concentric, wave like, patterns were motivated by earlier architectural features such as the ambulatory, a sort of half rose in design, and emulated, particularly, in the outer medallions of the north rose. Chartres labyrinth, in the nave, constructed about 1194, also develops the theme of wave like paths emanating from its centre. Nigel Pennick, in (Pennick, 1990), interprets the route of the labyrinth as symbolizing the passage from life to death.

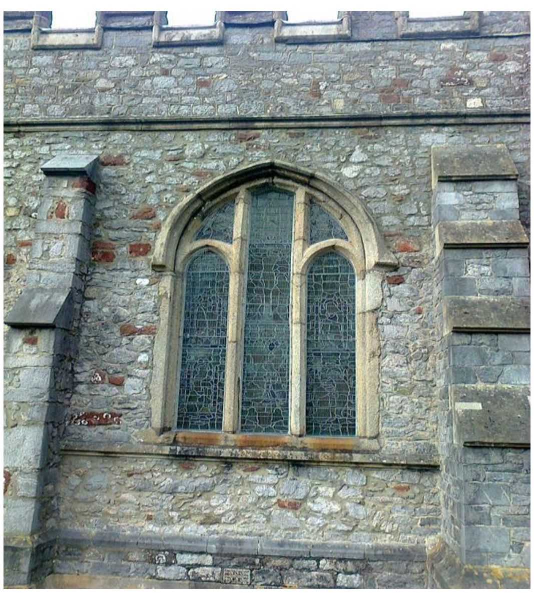
92. Labyrinth, Chartres Cathedral, Chartres, France.

The later Gothic style in France continued to draw on the influence of the Romanesque. One should, perhaps, draw attention to the cathedral of Saint Urbain in Troyes, 1262-1286, in this context, partly because of the rose window there, which we have already mentioned, and partly because of the interior, which could be considered harmonious, due to the limited number of elements, which are used repeatedly.

93. Nave of Reims Cathedral, Reims, France.

The harmonious designs found in the quincunx patterns of the Cosmati artists, and French Romanesque art, are closely linked with the idea of focusing light, visible in the rose window of the Cathedral at Troyes, France. The representation of light as a wave, allowed medieval artists to convey this last notion, in an alternative form, which predates the rose window. The following windows, from the churches of Bere Ferrers and Chudleigh in Devon, and Kilkhampton in Cornwall;


94. East and south windows of St. Andrews church, Bere Ferrers, Devon.




95. South windows of St. Martin and St. Mary's church, Chudleigh, and St. James the Great church, Kilkhampton, Devon.

demonstrating this idea, date from the Norman period, and are almost certainly derived from the French Romanesque, after 1148, when many French crusader knights returned to England after the Siege of Damascus, ending the Second Crusade. These designs predate the use of rose windows in the French geometric style, and employ the idea of analyzing the spatial forms of waves, rather than focusing light, as evidenced by the differing periods of three and four. An obvious analogy is with musical composition and the analysis of sound in terms of meter. This provides some evidence that musical understanding, at this time, was more closely linked to visual art. More harmonious designs, in the style of Cosmati artists, can be found at Bere Ferrers;

96. East window of St. Andrew's church, Bere Ferrers, Devon.

These differ from the use of harmonious patterns employed in the chevet designs of the earlier French Romanesque, with the juxtaposition of three forms in a triangular composition, rather than a radiating design around a central circle. The idea of focusing light towards an interior, as occurred later in the composition of rose windows in France, is, in some sense, derived from these ideas developed in England, and the chevet forms of the French Romanesque. The analysis of waves on a line, is replaced by an analysis of waves on a circle. This will be an important geometric theme in the following chapter.

One cannot leave the subject of the French Romanesque, without considering the importance that sculpture plays throughout its development, particularly after 1000. Found, usually, around the portals of churches, it lends a distinctive tactile quality to its art. Geometry continues to play an important role, in the way that two dimensional designs, such as foliage patterns, are interwoven with three dimensional figures. A good example are the capitals of the collegiate church of Saulieu, from the first half of the 12th century.

In English Romanesque art, circa 1000-1200, sculpture also plays an important role. One might mention the doorway of Stanley Pontlarge church in Gloucestershire, with its "solar paneled" exterior,

97. Doorway of St. Michael's church, Stanley Pontlarge, Gloucestershire.

the foliaged doorway of Kilpeck, Herefordshire,

98. Doorway of St. Mary and St. David's church, Kilpeck, Herefordshire.

the spiral columns in the church at Kempton,

99. St. Mary's church, Kempley, Gloucestershire.

or the classically friezed columns of Pershore, that we noted in Chapter 4.

The Romanesque style is often warmer and livelier than its Norman contemporaries. The wall frescoes of Stoke Orchard church, depicting a braided dragon form, were designed by travelling artists on the pilgrimage route to Santiago de Compostela.

100. St. James the Great church, Stoke Orchard, Gloucestershire.

In Italy, also, during the same period, the Romanesque evolved far beyond its initial incarnation. We have already mentioned the beautiful rose window at Tuscania, with its distinctive Lombardian influence. The abbey church of San Antimo, from the 12th century, is a reaction to the French style, with its characteristic chevet form, see (Stalley, 1999).

101. Abbey of St. Antimo, Tuscany, Italy.

Although this style never assumed great importance, it can be found occasionally in certain details, for example, the rose decoration of the bishop’s palace in Tuscania.

102. Window of the Archbishop's Palace, Tuscania, Tuscany, Italy.

The Romanesque style tended to be more dynamic, due to its Lombardian roots. An interesting example is the facade of the cathedral of Anagni,

103. Facade of Anagni Cathedral, Anagni, Lazio, Italy.

with its decoration of braiding patterns.

104. Facade detail, Anagni Cathedral, Anagni, Lazio, Italy.

It could also follow the classically restrained forms of the Roman style, in, for example, the facade of the cathedral at Civita Castellana.

105. Civita Castellana Cathedral, Civita Castellana, Viterbo, Italy.

Buildings, such as the clearstorey of San Benedetto in Subiaco, with its inventive architecture, defy a conventional classification.

106. Clearstorey of the Abbey of Santa Scholastica, Subiaco, Lazio, Italy.

The motif of the angel, holding a fiery orb, recalls some of the previous images of the Lamp of Heaven that we have considered.

44. Clearstorey detail, Abbey of Santa Scholastica, Subiaco, Lazio, Italy.

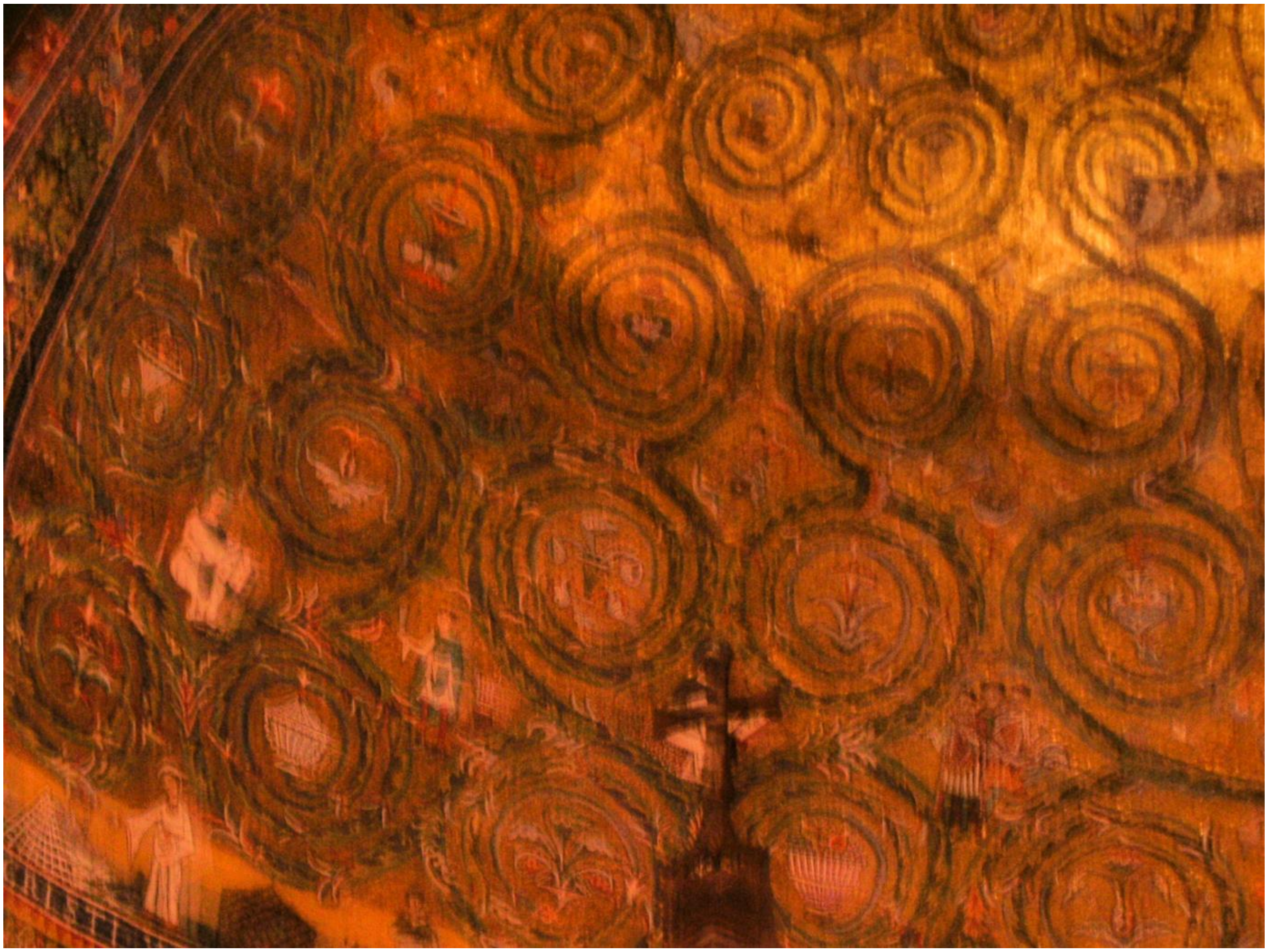
There was some limited reaction to the later French geometric style, in the elegant arcades of the Palazzo dei Papi, in Viterbo, from the middle of the 13th century, but, as with earlier French ideas, the form never assumed great prominence.

45. Palazzo dei Papi, Viterbo, Italy.

The importance of dynamism in Italian Romanesque is also a legacy of its Byzantine heritage, lasting an enormous period of time, from 400 until 1300. The mosaics in Santa Maria Maggiore date from the 5th century, and include beautiful spiral motifs of flowers and foliage, possibly attributed to the master mosaicist Jacopo Torelli, in 1295.

46. Mosaic in the church of Santa Maria Maggiore, Rome, Italy.

The same spiraling motifs can also be found in the Capella San Zeno, of 822, in Santa Prassede, and in the depiction of the Tree of Life, from the 12th century, in San Clemente.



47. Mosaic from the church of San Clemente, Rome, Italy.

The more considered geometry underlying the dynamic of Romanesque art in Italy, is brought to a climax in the work of the Cosmati artists, that I wish to now consider in more detail. In their work, we see especially, a departure from the more static conception of harmony, partly through the need to interpret the Christian image of the resurrection. Before giving specific examples of their work, let me give a brief description of the history of the Cosmati artists. The Cosmati were a family, headed by Cosmatus, who is known to have had four sons, Jacobus, Petrus, Iohannes and Deodatus, all of whom continued in their father's profession as marble workers or "marmorani romani", as they were also known. However, there were other contemporary artists such as Laurentius, Vassallettus and Drudus de Trivii, working in the 13th century, and artists from the 12th century such as the Paulus and Rainerius families, who were also known as the Cosmati, due to the similarity in the style and execution of their work. The Cosmati were more than just designers of pavements, they were also architects, responsible for many of the campaniles attached to Roman churches, such as Santa Maria in Cosmedin and San Giorgio in Vellabro, according to ([32]), as well as some of the most beautiful cloisters in Italy, pictures of which we have already seen, earlier in this chapter. They also designed triumphal arches, choirs, screens, porticoes, ambones, ciboria, candelabra tabernacoletti, episcopal thrones, altars and tombs, a good overview of this type of work can be found in ([33]). According to a recent survey in ([34]) , 91 percent of genuine

---

[32] Venturi, "A Short History of Italian Art", Macmillan(1926)

[33] Hutton, "The Cosmati-The Roman Marble Workers of the 12th and 13th Centuries"

[34] Pajares Ayuela, "Cosmatesque Ornament, Flat Polychrome Geometric Patterns in Architecture"

Cosmatesque production is found in Central Italy, and 46 percent can be found in Rome. Some notable exceptions are in Sicily and Westminster Abbey in London. The underlying historical factors which governed such a huge output in this area of Italy are complex and interesting in their own right.

I wish to consider, primarily, the geometric designs used by the Cosmati in their pavements. This study is complicated by the fact that some of the original pavements are missing, either as a result of structural damage over the centuries or extensive restorations during The Renaissance and the 19th century. In some exceptional cases, the designs of the restored pavements are known to be nothing like the original and, in other rare cases, there are even conflicting opinions as to whether the pavements have been restored at all, ([35]). A good survey on the authenticity of Cosmati pavements can be found in ([36]). However, the pavements that I will consider are mostly original, and, at worse, preserve their original geometric designs, ([37]).

Although many variations exist, there are a number of basic motifs which recur frequently in Cosmati designs. The most important of these are the guilloche and the quincunx. As we saw, the guilloche design can be found in Santa Maria in Trastevere, Rome. It consists of two interweaving threads which run along the central section of the nave. This braiding image is repeated in the architectural device of twisting columns which we saw in the pictures of Cosmati cloisters, and also in a number of Paschal candelabra. This makes it clear that the image is symbolic of the resurrection of Christ. It can also be seen as representing the image of The River of Life, from Revelations, an observation that is made in ([38]). The quincunx design can be seen in Santa Maria in Cosmedin, Rome, and also at Westminster Abbey, see the above pictures. It consists of a central roundel and four smaller surrounding eyelets, around which a single thread is intertwined. Recalling Pajares Ayuela description of the symbology of the quincunx, earlier in the chapter. It signifies a 3-dimensional reality, the tetramorph, consisting of the four hybrid beasts, which represent the evangelists, surrounding either Heaven or the symbolic representation of Christ as a circle. This symbolism was, as we have seen, used by Romanesque artists in the form of wheel rose windows, for example at Tuscania, in which the central circle is replaced by a wheel, symbolic of Ezekiel's vision from the Old Testament.

Aside from the symbolism, in both the guilloche and quincunx designs, we can see a diversity of aesthetic expression. Whereas the guilloche conveys a sense of femininity, fluency and twisting, the quincunx is more concerned with the classical idea of harmony. In both cases, there is also a suppressed aesthetic of light, the designs are projections of 3-dimensional forms, manifested by the fact that the Cosmati artists left spaces at the points where the represented curves intersect. The aesthetic is further suppressed by the use of dark colours for the marble roundels.

In the following design, from San Lorenzo Fuori le Mura,

[35] For example, at Santa Maria in Cosmedin, according to Pajares Ayuela, see previous citation, the pavement is original but according to "Guiseppe Massimi, again see previous citation, the upper part was transferred to the presbytery sometime during the 18th or 19th centuries.

[36] (Glass, 1980), "Studies on Cosmatesque Pavements".

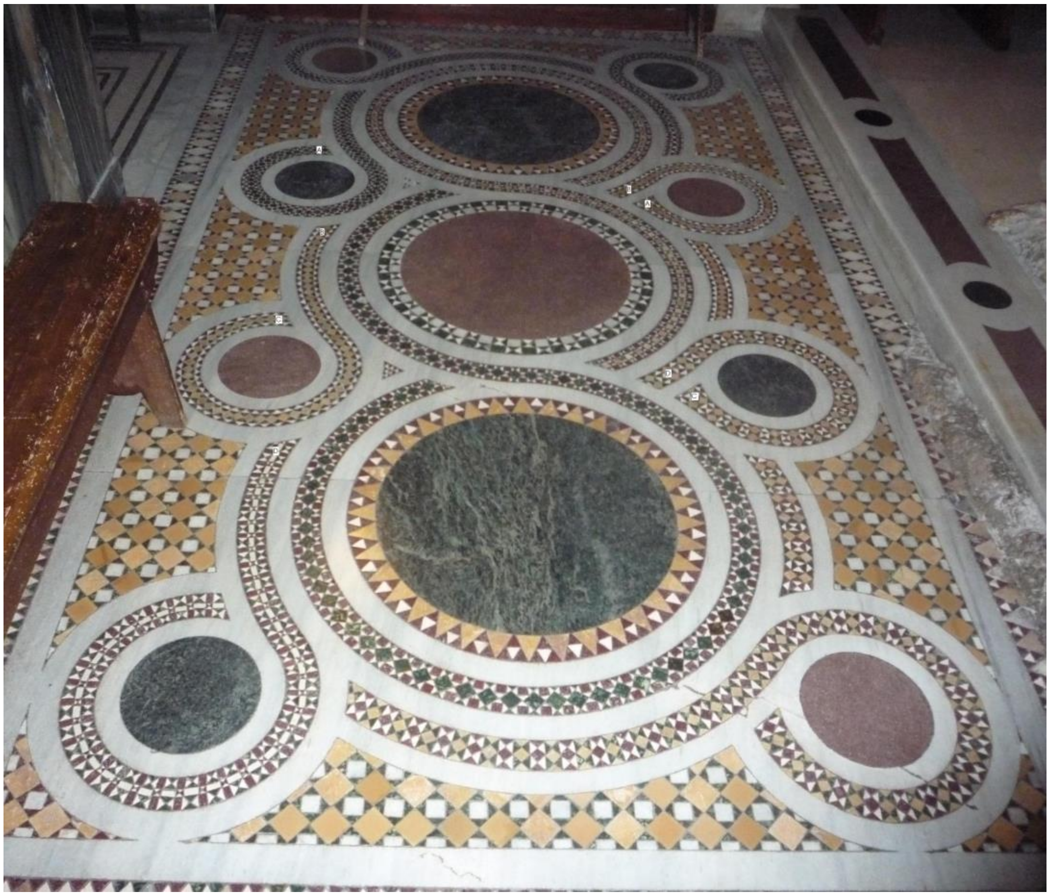
[37] In cases of doubt as to the authenticity of the geometric design of a pavement, I take the view that, unless the whole pavement was replaced, which is unlikely, the original design was maintained.

[38] (Williams, Italian Pavements, Patterns in Space, 1997)

48. Pavement, church of San Lorenzo Fuori le Mura, Rome, Italy.

we see a combination of the quincunx and guilloches. The aesthetic effect is one of dynamism, not accomplished by the quincunx or the guilloche form on its own. This aesthetic appears as a more intensified form of the feminine aesthetic of the guilloche, and clearly breaks with the rules of classical harmony. Its historical origins can probably be found in the numerous spiraling designs of Byzantine mosaics found in Rome.

Another unusual variant can be found in the Sacro Speco, Subiaco;



49. Pavement in the Sacro Speco, Subiaco, Lazio, Italy.

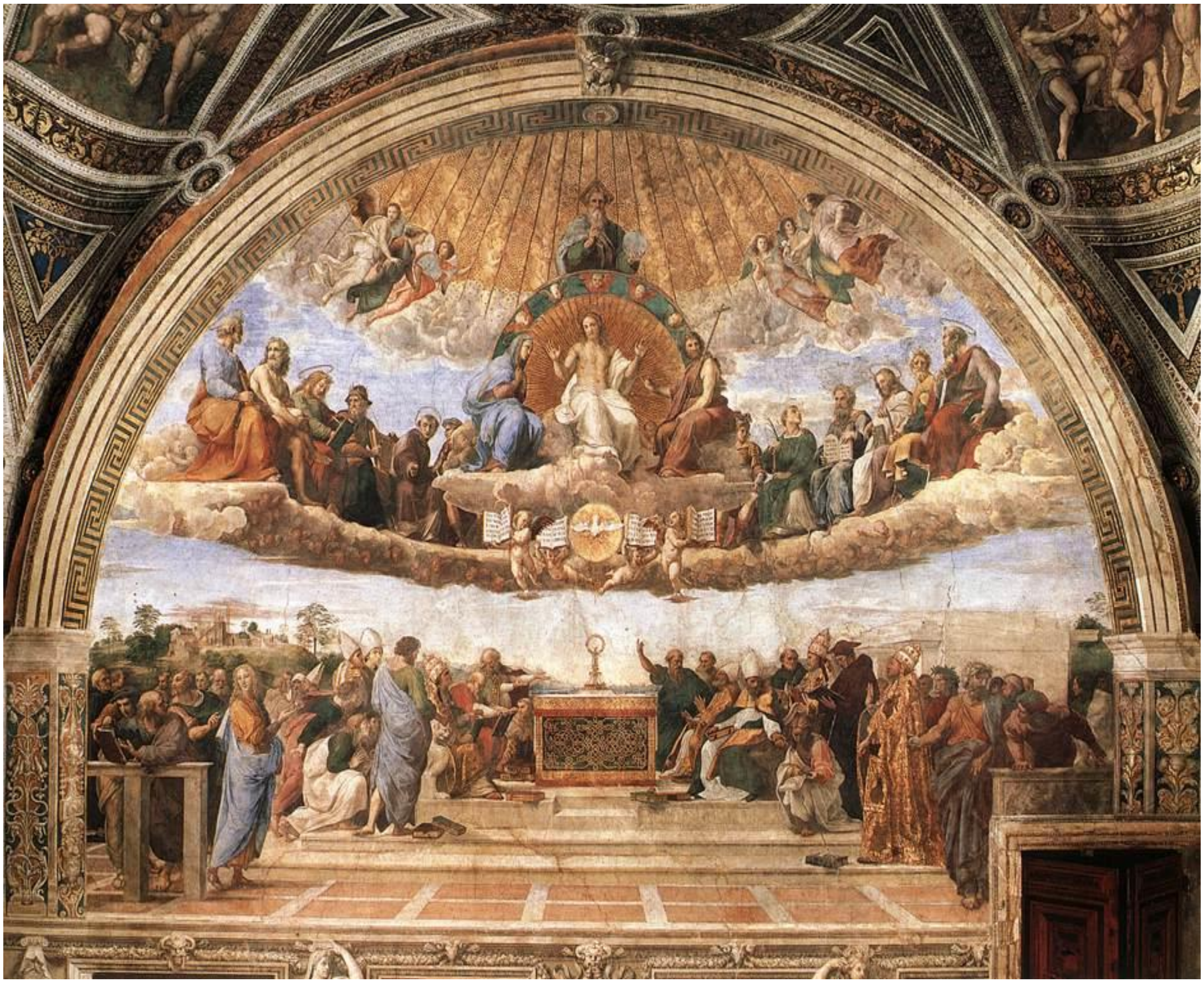
in which the basic quincunx form is cleverly switched, to create a continuous thread which runs the complete exterior of the three central roundels, ([39]). Such designs show that the Cosmati were not just artists, but were also interested in creating geometric puzzles. They demonstrate a balance between aesthetics and geometry, which, I believe, is not paralleled in our modern understanding of the subject. Understanding this synthesis in more detail is obviously the concern of this book.

We have already discussed Raphael's depictions of the event of the resurrection, but, undoubtedly influenced by Cosmati work, he was also an innovative practitioner of the geometry of curves. One can detect at least three types in his paintings, which I will refer to as classical, inflexion and spiraling. The classical theme finds its highest expression in “The School of Athens”;

[39] The direction of movement can be seen from the attached labels in the picture.

50. The School of Athens by Raphael, Raphael Stanze, Vatican, Rome, Italy.

This was commissioned as part of the decoration for the papal apartments of Julius 11, on the upper floor of the Vatican palace. The theme represented is that of the school of philosophy and geometry, an open spacious building decorated with statues of the gods. In the centre stand the two giants of Greek philosophy, Plato and Aristotle. Plato holds the Timaeus in his left arm, and points towards heaven, the realm of ideas or forms, while Aristotle, holding the Ethics, gestures towards the earthly domain, where “man is the measure of all things”. Other figures include Pythagoras, surrounded in the left foreground by a group of students, and Euclid, explaining the laws of geometry to his pupils on a black slate with a compass. Socrates can also be seen, next to Plato, rapt in discussion with a group of Athenian citizens. The composition, perhaps reflecting the rationality of the depicted disciplines, is that of harmony and order. The temple is viewed in perfect perspective, vanishing directly behind the central figures of Aristotle and Plato. The grouping of the figures is on two carefully constructed levels. Raphael is careful to connect these levels with the central figure, sprawled on the steps. Each action and gesture of the figures is carefully counterbalanced, Aristotle’s with Plato’s, the sprawling figure with the figure writing in the foreground, on the marble projection. The composition is reminiscent of “The Disputa”, also painted for the papal apartments, two years earlier, which is also constructed on two separate levels.



51. The Disputa by Raphael, Raphael Stanze, Vatican, Rome, Italy.

However, Raphael's technique seems to have developed in the School of Athens as the figures are deployed more dynamically and interact with each other, without disturbing the harmony of the whole. This notion of dynamic harmony is a defining feature of the High Renaissance period of Italian art, that culminates in the work of Raphael and that of his contemporary Michelangelo.

The inflexional geometry is beautifully expressed in "The Galatea".

52. The Galatea by Raphael, Villa Farnesina, Rome, Italy.

This was commissioned by the enigmatic patron Augustus Chigi for his villa on the banks of the river Tiber, The Villa Farnesina. The scene depicts Galatea's train, giving themselves up to the pleasures of love, while she herself looks heavenward, moving through the water on a shell ship, powered by a pair of dolphins. Again, we find the same dynamic controposto movements of Galatea's entourage and the sea monsters, each carefully balanced, the foreground figure facing us, while her companion, in the background, looks away. The bow-wielding putti serve to focus attention on Galatea herself, in a way geometrically similar to the device of a rose window. The figure of Galatea herself traces a lovely inflexion, drawn from the tip of her water chariot, whose movement is dynamically balanced by the spiraling motion of the figure on her left. Raphael's range of technique seems to surpass even his previous masterpiece “The School of Athens”, and represents one of his few forays into the subject of pagan mythology. A final painting worth considering, in relation to the use of spirals in Raphael’s art is that of “The Expulsion of Heliodorus”.

53. The Expulsion of Heliodorus by Raphael, Vatican, Rome, Italy.

This was again created for the Vatican palace, slightly later than “The School of Athens”, and depicts a scene from the Second Book of Maccabees, there appeared in the temple;

“ a horse that was finely adorned, on which sat a terrible rider”

and

“ two young men, who were beautiful and strong, who stood on either side of Heliodorus and rained blows upon him, so that he fell unconscious to the ground”.

The harmony in this scene is relatively simple compared to previous paintings, there are two, relatively unconnected groups, in the left and right foreground of the painting. However, the right foreground group is of interest, due to the dynamic spiral which dominates, emanating from the prostrate figure of Heliodorus, through the right arm of his companion, the swirl of the rider’s cloak and ending in the pointed figure of the beautiful youth in the foreground.

Raphael’s interest in geometry can even be seen in the development of his Madonna compositions. Early examples such as the Madonna and Child with Saints, 1502, and the Diotalevi Madonna, 1503, demonstrate his familiarity with the classical geometric compositions based on a circle and the balanced Greek capital form. In

1505, he experiments with a more dynamic composition, the Christ child of the Terranuova Madonna, exhibiting an inflexionary pose.

54. The Terranuova Madonna by Raphael, Staatliche Museen, Berlin, Germany.

This type of composition was fully harmonized in the supreme Madonna dell Seggiola of 1514;

55. Madonna dell Seggiola, Pitti Palace, Florence, Italy.

Raphael also experimented with linear compositions of the form described in our analysis of the sublime; examples being the Madonna del Prato, 1506, which employs triangles and diagonals, the Madonna of Loreto, 1509-1510, exhibiting a square composition and the Alba Madonna, 1511, a sophisticated use of parallel lines, see (Hartt, 2006).

7. The Transfiguration and the Right Hand of the Father

We will briefly consider how the images of The Transfiguration and the Right hand of the Father, have been employed in the visual arts. Raphael's final painting of the Transfiguration was commissioned in 1520.

56. The Transfiguration by Raphael, Vatican Museum, Rome.

We will not be concerned with the evident sense of spatial harmony, characteristic of his style, here, but note, instead, an interest in the aesthetic of colour, which, perhaps, is derived from the work of contemporary Venetian artists. The passage from Matthew, which we discussed above, describes how Jesus' clothes became white as light. This is clearly reflected in the use of a radiant purple, pink and white haze which surrounds the transfigured figure. The use of colour, in particularly that of purple and pink in the garments of the surrounding onlookers, suggests a new intuitive sense of balance. It is interesting to compare this composition with an acknowledged master colourist, Titian, in his later depiction of the Transfiguration, of 1534;

57. The Transfiguration by Titian, church of San Salvador, Venice, Italy.

Here, the description in Matthew, of how Jesus' face shone like the sun, is emphasized by the prominent use of yellow and red, in the halo surrounding Christ, and the clothes of his disciples. Again, there is a sense of colour balance in the painting, even though the forms and shapes are slightly discordant, a topic we will consider below shortly.

Titian, who worked mainly in Venice, was a contemporary of the artist Tintoretto, who we discussed above. He is famous for the use of red in his paintings, and is renowned for developing an aesthetic of colour, at the height of his career, between 1516 and 1550, before becoming a portraitist in his final years, 1550-1576. His painting of the Assumption of the Virgin, (1516-1518), in the Frari church, Venice;

58. The Assumption of the Virgin by Titian, Frari church, Venice, Italy.

demonstrates an interest in a dynamic colour scheme, particularly red and gold, as well as a slight asymmetry, for example, in the right hand of God, who floats above the ascending female figure. This aesthetic of asymmetry, which, I will argue, is closely interrelated to Titian's use of colour, can, perhaps, be understood, by consideration of the history of depictions of God, perhaps inspired by the passage from Mark above, on Jesus sitting at the right hand of the Father.

The following painting, "God Inviting Christ to sit on the Throne at his Right Hand", by de Grebber, 1645, demonstrates this image directly.

59. God Inviting Christ to Sit on the Throne at his Right Hand by de Grebber, Museum Catharijneconvent, Utrecht, Holland.

It is hard to find many remaining depictions of this exact subject in medieval art. However, I would speculate that many medieval images of Christ, in the act of creation or ascension, Christ and the Virgin, raising their hands in blessing, or even simple lyrical gestures, are influenced by this idea, as for example, in the following pictures, of the Creation of Heaven and Earth (medieval manuscript), a fresco from Saint Clement de Taull, Catalonia, the Ascension, by Jacopo di Cione, 1370-1371, the Virgin and Child, (Umbrian artist, 13th century), Vision of the Blessed Clare of Rimini, 1340, by Francesco da Rimini, and the Virgin and Child Enthroned, with narrative scenes, (1263-1264), Margerito d'Arezzo.

60. Creation of Heaven and Earth, medieval manuscript.

61. Fresco in Saint Clement de Taull, Catalonia, Italy.

62. The Ascension by di Cione, and The Virgin and Child, by an Umbrian artist, National Gallery, London.

123. Vision of the Blessed Clare of Rimini by Francesco da Rimini and The Virgin and Child Enthroned by Margerito d'Arezzo, National Gallery, London.

The following painting of "God the Father with his Right Hand Raised in Blessing", by Girolamo da Libri, from 1500, is a further, post-medieval, variation on this theme.

124. God the Father with his Right Hand in Blessing by Girolamo da Libri, National Gallery of Art, Washington, USA.

There are a few medieval depictions of the Transfiguration, for example by Duccio, (1311);

125. The Transfiguration by Duccio, National Gallery, London.

Again, one notes a simple colour balance, perhaps inspired by this biblical description of this image.

The sense of asymmetry, perhaps originally considered in depictions of Jesus Sat at the Right Hand of the Father, can be found in Titian's "Madonna of the House of Pesaro", (1519-1526) and "The Vendramin Family", (1550-1560);

126. Madonna in the House of Pesaro by Titian, Frari church, Venice, Italy.

127. The Vendramin Family by Titian, National Gallery, London.

In the second painting, the raised altar, creates a sense of asymmetry, in comparison with the carefully balanced groups of three boys, at either side. This effect is copied and enhanced in the earlier first painting, by the positioning of the red flag of the Pesaro family. The characteristic use of Titian red, begins to combine the two aesthetic ideas of colour and asymmetry, in this case, first, reinforcing the sense of asymmetry in form, but, secondly, harmonizing the composition by averaging the use of red with yellow, in an intuitively coordinated spectrum.

The Aesthetic of Asymmetry in Medieval Art and Architecture.

The use of what I will refer to as “asymmetric fragments”, that is additional pieces, decorating a more conventional form, can be found in a number of medieval church windows in Devon. The following Curvilinear designs, from St. Andrew’s church in Bere Ferrers , developing the sense of harmony, found in the window we previously considered, and St. Petroc’s church in Lydford, will, for reasons we will discuss below, originate from the Norman period, (c1180).

128. Interior and exterior views of east window in St. Andrews church, Bere Ferrers, Devon.

129. South window of St. Petroc's church, Lydford, Devon.

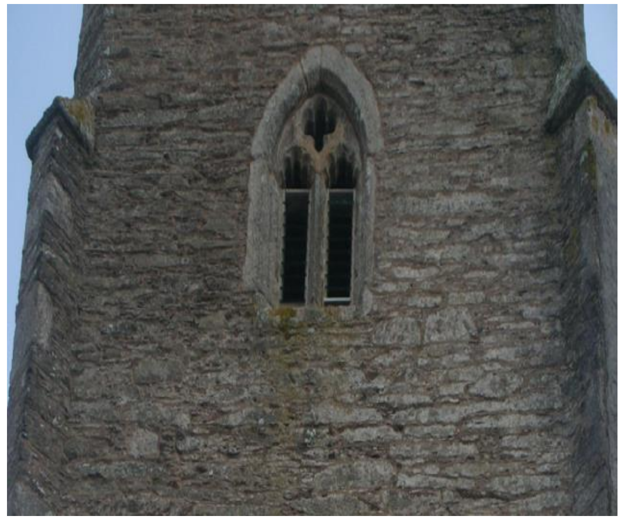
These fragments recur in the following windows, again from St. Andrew's church, and from St. Winifred's church in Branscombe, which can be easily dated from the period of the English response to the French Geometric style, (c1240-1280), ([40]), owing to the use of the quatrefoil.

130. South windows of St. Andrews church, Bere Ferrers, and St. Winifred's church, Branscombe, Devon.

Similar distinctively "spade-shaped" fragments can be found at St. Petroc's church, Lydford, from the Lancet period, (c1180-1275), and St. Werburgh's church, Wembury, which develop a style, particularly common in

[40] Although it is possible that the window at St. Winifred's church is a Gothic revival copy of a previous one dating from this period.

Devon and Cornwall churches, that I will refer to as "Cusp", (c1280), the origins of which I will discuss in Chapter 8, ([41]).



131. South and west windows of St. Petroc's church, Lydford, and St. Werburgh's church, Wembury, Devon.

I believe that the source of these fragments can be found in designs from churches in Spain, in particular, the development of the horseshoe arch. The origination of this form, which differs from the Roman semicircular arch, in that the ends are extended below the diameter, can be found in the Mosque at Cordoba, which was later converted into a Christian church, in 1236. The design was adopted by the Visigoths of Northern Spain, who converted to Catholicism in 587, the first example being from the church of San Juan de Banos, in Banos de Cerrito, Palencia, of 661;

[41] See (Bradt, 2010) for a guide to churches in this area.

132. Church of San Juan de Banos, Banos de Cerrito, Palencia, Spain.

The design can also be found in the Church of San Pedro de la Nave, Campillo, (670-680);

133. Church of San Pedro de la Nave, Campillo, Zamora, Spain.

where characteristic semicircular columns, probably of Roman origin, are also employed, and, at the church of El Salvador, Toledo;

134. Church of El Salvador, Toledo, Spain.

The Visigothic use of the form culminates in this design from the church of Santiago de Penalba, near Leon, (c 930), in which the lower ends are further elongated;

135. Church of Santiago de Penalba, Leon, Spain.

However, the design can still be found, as late as the 13th century, in the Church of San Roman, Toledo;

136. Church of San Roman, Toledo, Spain.

Variations on the theme of the horseshoe arch can be found at the church of Santa Maria la Blanca, in Toledo, originally built as a synagogue for Jewish church, at the end of the 12th century, the arch is canopied into a three-tier structure;

137. Church of Santa Maria la Blanca, Toledo, Spain.

A further variation is to point the apex of the horseshoe arch, an example which can again be found at Santa Maria La Blanca, in Toledo;

138. Windows of the church of Santa Maria la Blanca, Toledo, Spain.

As we discussed in Chapter 4, the construction of the pointed arch form became a mainstream feature of Cistercian architecture, within England and France, from the end of the eleventh century, and coincides with a pivotal moment in the Reconquista in Spain, the Battle of Toledo, 1085. The conjunction of the horseshoe and pointed arch can be seen, there, at the Puente de San Martin, from this period;

139. Puente de San Martin, Toledo, Spain.

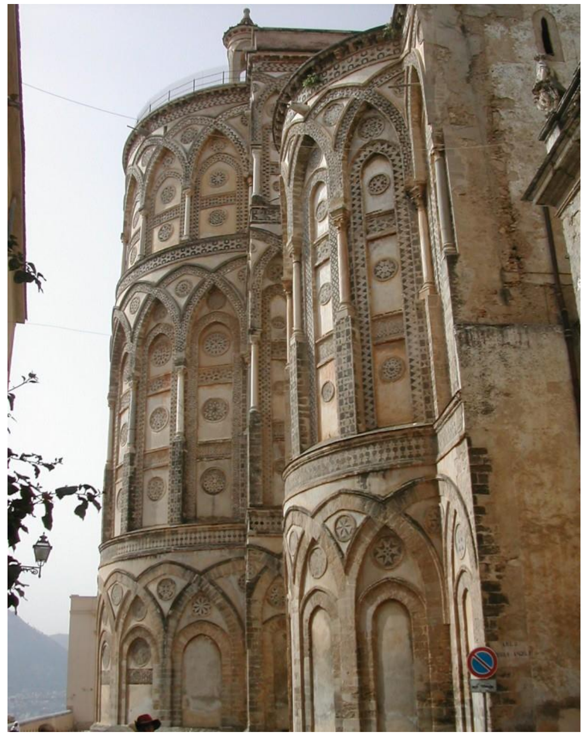
I would suggest that this is the origin of the spade-like fragments discussed above. These fragments owe much to the Jewish and Muslim communities living in Spain, but, due to the innovation in horseshoe arches, at the end of the Visigothic period, the development of the pointed arch form in Gothic art, seems to be primarily due to Christian architects.

It remains to examine the question of their asymmetrical use. The churches of San Tome, San Sebastian, San Andres, San Pedro Martir, San Domingo el Antiguo, Santiago del Arrabal and San Vincente , in Toledo, all date from around 1085, and are distinguished by the appearance of bell towers and cupolas. The tierceron pattern at the mosque of Cristo de la Luz,

140. Mosque of Cristo de la Luz, Toledo, Spain.

and the use of copulas along a wall at Santiago del Arrabal , are reminiscent of the pavement designs and the apse of Monreale, Sicily;

141. Church of Santiago del Arrabal, Toledo, Spain.

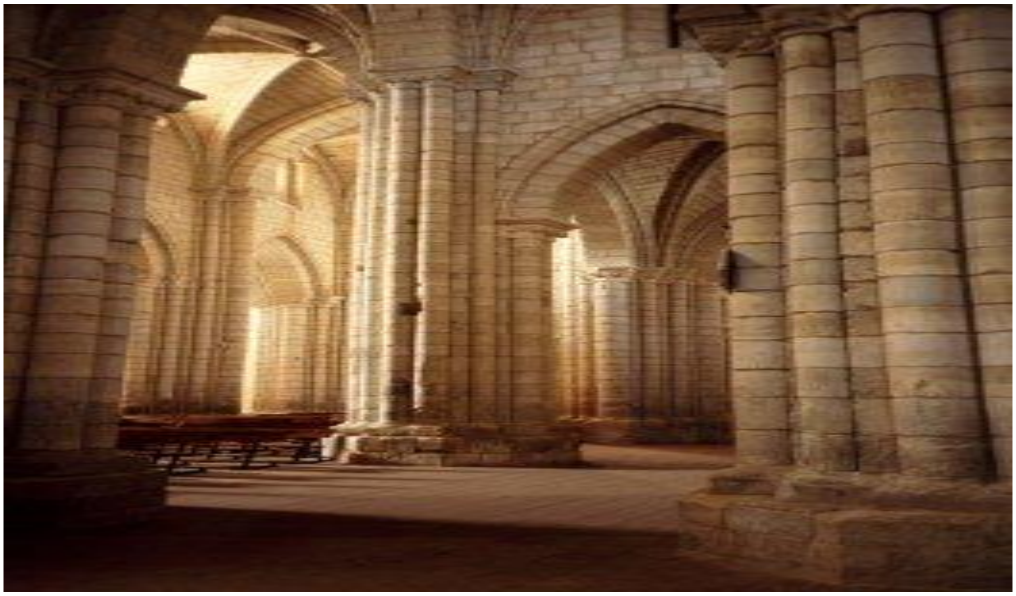


142. Monreale Cathedral, Palermo, Sicily.

suggesting a strong Norman influence. The positioning along one side of a wall, gives a sense of asymmetry, in comparison to the chevet deigns found in France. The example, at San Domingo el Antiguo , with the cupola conjoining a bell tower;

143. Church of San Domingo el Antiguo, Toledo, Spain.

reinforces this idea. The construction of the church of Santa Maria La Blanca, in 1180, roughly coincides with the Templar church of the same name, at Villacazar de Sirga;



144. Church of Villacazar de Sirga, Villacazar de Sirga, Palencia, Spain.

owing to the almost identical high clustered shafts, found at Pershore Abbey, which we discussed in Chapter 4. As we noted above, that these spade like fragments were added to a three light composition at Bere Ferrers from about 1138, it seems reasonable to place the origin of the asymmetric design, considered above, at the end of the Norman period, (c1180).

More examples, possibly constructed by the Knights Templar, can be found in Zamora. The cupola from Santa Maria Magdalena and the belltower of San Juan de Puerta Nueva make interesting comparisons with the designs from Toledo. The use of a single rather than repeated cupola and a flatter tower is prominent.

145. Churches of Santa Maria Magdalena and San Juan de Puerta Nueva, Zamora, Spain.

There are numerous Romanesque churches in Zamora. The reader is invited to consider the cupolas found also at Santa Maria La Nueva, Santa Maria La Horta, Santiago Caballeros and San Claudio de Olivares, as well as the belltowers from Santiago del Burgo, San Pedro y San Ildefonso, San Juan de Puerta Nueva, San Vincente, Santa Maria la Horta and San Cipriano.

My aim in the following chapter, is, as before, to consider the geometry behind the aesthetic ideas which we have considered.

8. RAPHAEL AND THE COSMATI GEOMETRY OF CURVES

TRISTRAM DE PIRO

In this chapter, we will look more closely at the aesthetic ideas discussed above, and see how they can be used in geometry. The use of spirals and inflexions in the artwork of Raphael, together with the braiding patterns of Cosmati artists, suggests an intelligence behind the depiction of curves in the plane, which, I will argue, is extremely useful in the study of the theory of braids, a field of study in algebraic curves connected with topology and the notion of "fundamental groups". We also examined the understanding of harmony in medieval architecture, with particular reference to the chevet form. I will argue these ideas are also extremely important in the representation of algebraic curves, using the notions of class and genus, a field pioneered by the German mathematician Julius Plucker. Finally, in the medieval conception of focusing light, either in the form of rose windows, or in the earlier devices of Romanesque windows, through the analysis of curves in terms of ocsillating patterns, we will find a connection with the analysis of functions on curves, into simpler harmonics, an area now referred to as Fourier analysis, after the French scientist Joseph Fourier, but possibly invented by his adviser Laplace.

In order to make these ideas more precise, I will give a brief explanation of some of the above mentioned mathematical terms. The term "fundamental group" was first used by Henri Poincare, in his paper Analysis Situs, of 1895. The idea is to understand some of the geometry of algebraic curves, through the possible one dimensional loops which can exist on their two-dimensional surfaces. Of course, there are an infinite number of loops through any given point, but any two loops, which can be continuously bridged by a path across the surface are called equivalent, see figure 1, ([1]) In the diagram, any loop passing through the point p, on the surface of the sphere S, can be contracted back to the initial point p, so any two loops are formally equivalent. In this case, the fundamental group is trivial, that is consists of a single element. In more complicated surfaces, such as the torus T, there exist loops which cannot be so contracted, see figure 2. Here, the loops γ_a and γ_b cannot be

[1] Formally, if S is a topological space and $s_0 \in S$ is a chosen base point, we define a closed loop to be a continuous map $\gamma : [0,1] \rightarrow S$ with $\gamma(0) = \gamma(1) = p$. We define two loops $\{\gamma_0, \gamma_1\}$ to be homotopically equivalent, $\gamma_0 \sim \gamma_1$, if there exists a continuous map;

$H : [0,1] \times [0,1] \rightarrow S$ with $H(0,t) = \gamma_0(t)$, $H(1,t) = \gamma_1(t)$, $H(s,0) = H(s,1) = p$. It is easily checked that $\sim$ is an equivalence relation on loops, and we define $[\gamma]$ to be the equivalence class of a loop γ. One can also define a composition of loops $\gamma_2 \circ \gamma_1$, compatibly with the equivalence relation $\sim$, and an identity class $[1]$. We then define;

$\pi_1^{top}(S,p) = \{[\gamma] : \gamma \text{ a closed loop}\}$

With the composition $\circ$, this defines a group. If S is connected, it is independent of the chosen base point p, and denoted by $\pi_1^{top}(S)$. If C is a projective algebraic curve over $\mathcal{C}$, then it can be considered as a topological space, and we can define $\pi_1^{top}(C)$.

contracted back to the initial point p, and, moreover, cannot be transformed continuously, one to the other. It is easily shown, that any loop is essentially a combination of these two, ([2]).

The geometric idea of the fundamental group connects easily with that of genus. The modern definition, which could be said to originate with the German mathematician Felix Klein, is that an algebraic curve C has geometric genus g, if, topologically, it is equivalent, ([3]) to a sphere S with g attached handles, see figure 3. However, the idea can be traced back to the earlier work of Julius Plucker, where it is mentioned in his "Theorie der Algebraischen Curven", of 1839, ([4]). For nonsingular curves C, there is an interesting relationship between the fundamental group and the genus. Namely, that, for a curve C of genus g, the fundamental group is generated by exactly $2g$ loops, ([5]). Figure 4 makes this clear, the loops γ_a^1 and γ_b^1 attached to the first handle, passing through the point p, are not equivalent, and, a similar argument applies, for each of the g handles.

The class of an algebraic curve C, which, also appears in the above work of Julius Plucker, counts the number of tangent lines, passing through a general point in the plane. This notion is illustrated in figure 5, in which the general point is represented by p, and the tangent lines (in black) of C (in brown), count the class as 8. There is a simple relationship between the class m and the degree of C, for nonsingular curves, known as the elementary Plucker formula, $m = d(d-1)$, in particular, the class of such curves is even. One can also obtain a relationship with the genus, using the degree-genus formula, ([6]).

One can refine the study of loops on a surface, by allowing an algebraic curve to have singularities. We discussed the concept of nodal curves in Chapter 5, and, will, again, in Chapter 11. A related idea is that of an inflexion. Formally, a point $p \in C$ is an inflexion if its tangent line, see the discussion in Chapter 5, has contact at least 3 with the curve, see figure 6 and Chapter 11. As we explain in Chapter 11, any algebraic curve is birational to a plane curve with at most nodes as singularities, and, for which the inflexions are distinct from the nodes. One can then restrict attention to loops or paths on such surfaces which are restricted by its geometry of nodes and inflexions. In figure 7, we have two loops, based

---

[2] More precisely, the fundamental group is generated by the 2 loops, γ_a and γ_b, subject to the single relation $\gamma_a\gamma_b\gamma_a^{-1}\gamma_b^{-1} = 1$.

[3] Two surfaces, S_1 and S_2, are said to be topologically equivalent, if, one can be continuously transformed into the another, by stretching, without tearing or glueing

[4] Plucker's definition, which, I believe, was later adopted by the Italian mathematician Francesco Severi, can be shown to be equivalent to the modern definition. The proof, assuming the topological degree-genus formula, which asserts that, for a nonsingular curve C, $g = \frac{(d-1)(d-2)}{2}$, where d is the degree of C, see Chapter 5, can be found in [20]. A direct proof of the degree-genus formula, using Severi's definition, can also be found in [20]. A direct proof of the topological genus formula can also be given using Plucker's methods, see [21]. A further important reference is [11].

[5] More specifically, the group is generated by the loops $\{\gamma_a^1, \gamma_b^1, \ldots, \gamma_a^g, \gamma_b^g\}$, satisfying the single relation;

$$((\gamma_a^1)(\gamma_b^1)(\gamma_a^1)^{-1}(\gamma_b^1)^{-1})\ldots((\gamma_a^g)(\gamma_b^g)(\gamma_a^g)^{-1}(\gamma_b^g)^{-1}) = 1$$

The proof of this result can be found in [11], see also [15].

[6] This is a simple calculation, from $m = d(d-1)$, $g = \frac{(d-1)(d-2)}{2}$, we obtain $d^2 - d - m = 0$, $d = \frac{1+\sqrt{1+4m}}{2}$, $g = \frac{m}{2} - d + 1$, $g = \frac{m+1+\sqrt{1+4m}}{2}$

at the points p and q. The first is an example of an inflexionary path, with inflexions at i_1 and i_2, these should occur at the actual inflexions of the surface. The second is an example of a nodal path, which passes through the node ν of C, ([7]).

[7] Formally, by a "nodal path", we mean a closed path $\gamma : (S^1, 0) \rightarrow (C, p)$ with the following further properties;

(a). γ is locally analytic in the sense of real manifolds.

(b). γ is smooth.

(c). γ has at most nodes as singularities, defined by *at least* one of the nodes in C, $(\dagger)$

In condition (a), we take $\mathcal{R}^4$ to be the object manifold. Then condition (b) means that $\gamma'(t) \neq \bar{0}$, for $t \in S^1$, and condition (c) means that γ is injective, with the exception of pairs $\{t_1, t_2\}$ such that $\gamma(t_1) = \gamma(t_2) = v$ where $v \in C'$ is a node and, in this case, $\{\gamma'(t_1), \gamma'(t_2)\}$ belong to the tangent planes $\{H_{\gamma_1}, H_{\gamma_2}\}$ of the branches γ_1, γ_2 centred at v. We define a "nodal-inflexionary" path to be a path $\gamma : (S^1, 0) \rightarrow (C, p)$, satisfying conditions $(a), (b)$ of $(\dagger)$, with the additional requirement;

$(c)'$ γ has at most nodes as singularities, defined *only* by the nodes of C, the *only* inflexions of γ are defined by the inflexions of C.

Here, by an inflexion $t \in S^1$ of γ, we mean that, for a generic choice of $P \in \mathcal{R}^4$, with conic projection $pr_P : C' \rightarrow \mathcal{R}^2$, t should define an inflexion of the induced curve $pr_P \circ \gamma : S^1 \rightarrow \mathcal{R}^2$, in the sense of [22], (this notion is well defined as inflexions are preserved by homographies of $P^1(\mathcal{C})$). A nodal-inflexionary path clearly reflects the "flow" of the curve C along the finitely many non-ordinary branches, hence seems to be an important object of study.

Let $S = \{v_1^1, v_1^2, \ldots, v_d^1, v_d^2, i_1, \ldots i_m\}$, where $v_j = v_j^1 = v_j^2$, for $1 \leq j \leq d$ denotes a node of C and i_j, for $1 \leq j \leq m$ denotes an inflexion of C. We call a partial ordering on S *good*, if, v_j^1 appears in the order iff v_j^2 appears in the order, and, in this case, $v_j^1 < v_j^2$. Using the natural ordering on S^1, induced by the interval $[0, 1)$, any nodal-inflexionary path determines a good partial ordering on S. One can define a geodesic nodal-inflexionary path γ, by adding the requirement (d) that γ defines a geodesic, and adapt the notion of homotopy to require that intermediate paths are geodesics. Conversely, one can ask;

(a). Given a good ordering on S, can one find a geodesic nodal-inflexionary path realising this ordering?

(b). Given homotopic geodesic nodal-inflexionary paths γ_1, γ_2, are the orderings determined by these paths the same?

(c). Given geodesic nodal-inflexionary paths determining the same ordering, are they homotopic?

(d). Is every path in $\pi_1^{top}(C)$ homotopic to a nodal-inflexionary or geodesic nodal-inflexionary path?

(e). More generally, we can allow S to have repeats, and define the length $l([\gamma]) = min\{l(\gamma_1) : \gamma_1 \sim \gamma\}$, where $l(\gamma_1) = 2d' + m'$, of a geodesic nodal-inflexionary path, where d' is the number of nodes (with repeats) and m' is the number of inflexions (with repeats). How does the length behave under composition, and, is there a bound $0 \leq 2d + m \leq k$, and a finite index subgroup $G \subset \pi_1^{top}(C)$ such that, given a path γ, can one find paths $\{\gamma_1, \gamma_2\}$, with $max\{l([\gamma_1]), l([\gamma_2])\} \leq k$, $[(\gamma \circ \gamma_1^{-1})] = [\gamma_2]$, $[\gamma_2] \in G$?

One can show that any nodal-inflexionary loop in the plane $\mathcal{R}^2$ has an even number of inflexions, see [22], however, suprisingly the number of inflexions of a plane nonsingular algebraic curve can be odd, see below. An interesting exercise is to give a convenient description of nodal loops, without inflexions, in the plane. These are classified as sources, in the sense that one can remove all the nodes from the path, by successively subtracting loops beween adjacent nodes, see [22]. In figure 8, the nodal path γ, centred at p, has two nodes, based at $\{\nu_1, \nu_2\}$. The ordering defined by the path (starting at p), is given by $\{\nu_2^1, \nu_1^1, \nu_1^2, \nu_2^2\}$. Subtracting $\{\nu_1^1, \nu_1^2\}$, we obtain $\{\nu_2^1, \nu_2^2\}$, and, then $\emptyset$. This result is useful in the study of nodal-inflexionary paths on an algebraic curve C, as defined above, and, considered in footnote 7, as the generic projection of a nodal path on C is a nodal path in the plane $\mathcal{R}^\epsilon$. The reduction of geometric thinking about algebraic curves to the plane, gives ample scope for the type of visual thinking, about inflexions and spirals, employed by artists such as Raphael and the Cosmati, which we considered in the previous chapter. A successful understanding of such paths, see particularly question (d) in footnote 7 about symmetry groups, might lead to geometric insights into the etale fundamental groups of nodal algebraic curves, and the structure of Galois actions, discussed in Chapter 5, Theorem 0.19, within the context of flash geometry discussed there. It is the author's hope that these considerations might lead to new proofs of Severi's conjecture, ([8]). This type of thinking is also fundamental in the study of braids, see figure 8, ([9]). In figure 9, we have 2 braids, (in pink and red), which are distinct, in the sense that one cannot be continuously transformed into the other. There is an interesting connection with algebraic curves. Given $C \subset P^2$, there exists finitely many points $\{q_1, \ldots, q_r\}$, for which the projection $pr_1 : C \rightarrow P^1$ is ramified, that is the line defined by $pr^{-1}(x)$, $x \in P^1$, is tangent to the curve C. Ignoring these points, at which the surface "branches" over the plane, any loop γ centred at $p \in D \subset P^1 \setminus \{pr_1(q_1), \ldots, pr_1(q_n)\}$, lifts uniquely to a path (possibly a loop) commencing at p', with $pr_1(p') = p$, and defines the string of a braid by tracing, over time, the moving shadow of γ' on a disc D', centred at $p_1 = pr_2(p')$, see figure 10. We can then obtain a representation of the fundamental group $\pi_1(D \setminus \{q_1, \ldots, q_r\}, p)$, in the braid group $Br_{n, pr^{-1}(p)}$, where $n = degree(pr_1)$, the cardinality of a typical fibre of the projection pr_1, $(*)$. Now, given a projective algebraic curve C, we can label its nodes and inflexions, together with the nonsingular ramification points of the

(f). Given positive answers to $(a), (b), (c), (d), (e)$, can one use the induced action of the symmetric group S_{2d+m} on $\pi_1^{top}(C)$ to give an analysis of the automorphism groups of certain Galois extensions, $Gal(C''/C')$, with $C \subset C' \subset C''$, $C' = Fix(G)$ and $Card(Gal(C''/C')) \geq 2d+m$?. Otherwise, can one construct invariants of C on which S_{2d+m} acts? Can one obtain an analysis of the etale fundamental group, possibly in nonzero characteristic, using the specialisation and lifting theorems, see [18]?

[8]A central idea in proving irreducibility of Severi varieties V, is that an automorphism of $[C] \in Sing(V)$ should extend to an analytic map between the sheets of V passing through $[C]$. This might involve some ideas from deformation theory, but also require insights into the realisation of this extension in terms of these Galois actions.

[9]Formally, a braid is a sequence of continuous paths $\{\gamma_1, \ldots, \gamma_n\}$, with $\gamma_i : [0,1] \rightarrow [0,1] \times D$, for $1 \leq i \leq n$, where D denotes a complex disc, such that $\{\gamma_i(\delta) : 1 \leq i \leq n\} \subset \{(\delta, p_1), \ldots, (\delta, p_n), \delta \in \{0,1\}\}$, (distinct points), $pr_2 \circ \gamma_i = Id_{[0,1]}$ and, for $t \in (0,1)$, $1 \leq i < j \leq n$, $\gamma_i(t) \neq \gamma_j(t)$. Similarly to the above description of the fundamental group, we can define a notion of isotopy for braids, namely $\overline{\gamma}_0 \sim \overline{\gamma}_1$, if there exists a continuous $H : [0,1] \times [0,1]^n \rightarrow [0,1] \times D$, with $H(0, \overline{t}) = \overline{\gamma}_0(\overline{t})$, $H(1, \overline{t}) = \overline{\gamma}_1(\overline{t})$, and $H(s, \overline{t})$ defining a braid for any $s \in [0,1]$. As above, we can define a composition of braids, compatible with $\sim$ and an identity [1]. We let $Br_{n,\overline{p}} = \{[\gamma] : \gamma \text{ a braid}\}$, which we call the braid group on n strings, associated to $\overline{p} = \{p_1, \ldots, p_n\}$.

projection $pr_1 : C \to P^1$, ([10]). One can then attempt to address some of the questions in footnote 7, by combining the theory of nodal-inflexionary paths in $\mathcal{R}^2$, see [22], with Moishezon's theory of braid monodromy, see [17], ([11]).

Nowhere is an understanding of how the aesthetics of harmony and balance can be used in the geometry of algebraic, more evident, than in the work of Julius Plucker, to whom we alluded to above. Julius Plcker was born on June 16, 1801 in Elberfeld, Germany, and died on May 22, 1868, Bonn. He was a mathematician of the highest order, making important advances in the theory of algebraic curves, and, in later life, a physicist. Plcker attended the universities in Heidelberg, Bonn, Berlin, and Paris. He obtained his doctorate in 1824, from the University of Marburg. In 1829, after four years as an unsalaried lecturer, he became a professor at the University of Bonn, where he wrote "Analytisch-geometrische Entwicklungen, The Development of Analytic Geometry, in 2 volumes, between 1828 and 1831. Here, he begins to develop the notation of projective geometry, in which lines rather than points become the fundamental elements. The idea, here, is that projective space, $P^2(\mathcal{C})$ considered as the set of lines passing through the origin in $\mathcal{C}^3$, can be written in homogenous coordinates $\{[X : Y : Z] : (X, Y, Z) \neq \overline{0}\}$, with an embedding of affine space $\mathcal{C}^2$, given by $\{[x : y : 1] : (x, y) \in \mathcal{C}^2\}$. The equation of a curve $F(x, y)$ in the affine plane can then be converted into an equation in projective space, by making the substitution $x = \frac{X}{Z}$ and $y = \frac{Y}{Z}$, and clearing coefficients. On (p97, Vol.2, Essen (1831)), he writes;

"Die Coordinaten dieses Mittelpunctes, den wir durch (y', x') bezeichnun wollen, sind folgende:

$x' = \frac{B}{A}\ y' = \frac{D}{A}$

and bases a number of calculation on this substitution. Plucker considers mainly problems concerning curves of low degree, with an exceptional interlude on Cramer's paradox, that the number of points of intersection of two higher-order curves can be greater than the number of arbitrary points that are usually needed to define one such curve, "Sobald $n = 3$, oder $n > 3$, und also;

$n^2 \geq (\frac{n(n+3)}{1.2})$

[10] Formally, let $\{\nu_1, \ldots, \nu_s, i_1, \ldots, i_t, q_1, \ldots, q_r\}$ denote these points. Any nodal inflexionary loop $\gamma \in \pi_1(D \setminus \{pr_1(\nu_1), \ldots, pr_1(\nu_s), pr_1(q_1), \ldots, pr_1(q_r)\})$, with inflexions defined by $\{pr_1(i_1), \ldots, pr(i_t)\}$, lifts to a nodal-inflexionary path γ' in C, and, if C is nonsingular, $\gamma' \subset C \setminus \{pr_1^{-1}(q_j) : 1 \leq j \leq r\}$ is a nodal-inflexionary path, with $pr_1(\gamma'(0)) = pr_1(\gamma'(1))$ then $pr_1 \circ \gamma'$ is a nodal inflexionary loop in $\pi_1(D \setminus \{pr_1(q_1), \ldots, pr_1(q_r)\})$. How does the ordering of the inflexions on the path γ effect its winding numbers around the projected ramification points, and, therefore, by $(*)$, its representation in the braid group?

[11] This requires the positioning of the ramification of nonsingular curves in terms of the intersections of lines, proved in [21], coming from Severi's conjecture.

begegnet man einer Art von paradox...", ([12]). In Chapter 3, Section 3, "Das Princip der Reciprocitat", (p259), of the same text, Plucker formulates the principle of duality, ([13]), "Die Polaren aller Puncte einer gegebenen geraden Linie gehen durch den Pol derselben...Die pole aller geraden Linien, welch durch einen gegebenen Punct gehen, liegen auf der Polaren dieses Punctes." Plucker does not give a general definition of class here, for an arbitrary algebraic curve, but uses it to count the number of "parallele Tangenten", for specific cubic curves, see p96. He is clearly experimenting with the idea at this stage, but addresses the problem with greater confidence in his "System der Analytischen Geometrie", Berlin, (1835), written after Plcker became an ordinary professor of mathematics at the University of Halle, in 1834. We now consider this text in greater detail.

Plucker seems to be interested now in the use of the methods developed above, to represent algebraic curves in the plane. He seems to be mainly interested in real algebraic curves, but many of his ideas extend immediately to the complex case. In Chapter 5, we considered, in depth, Newton's classification and analysis (with diagrams) of cubic curves, using asymptotes. Plucker continues this project in Section 3, "Allgemeine geometrische Construction der Curven dritter Ordnung" (Several geometric constructions of curves of the third order) and Section 5, "Aufzahlung der verschiedenen Arten der Curven dritter Ordnung" (List of different kinds of curve of the third order", (p221), obtaining 219 types. Here, he begins by considering types of cubic with the property that the three finite points where the three asymptotes intersect the curve lie on a straight line l, dividing the curves into groups, according to the relative position of this line with the intersection points of the asymptotes. In figure 11, (referring to p227) he considers the case where the line l intersects the triangle formed by the intersections, above and below the extreme vertices. He maintains Newton's terminology of an oval, to refer to the irreducible component (for real algebraic curves), formed within the triangle, and is addressing the question of how these "double intersections" (of the asymptotes) might help us to analyse algebraic curves. The aesthetics of harmony and balance are clearly employed here. In figure 12, a typical representation of an algebraic curve (of higher degree) is given in which three of its asymptotes are bitangents to the curve. Although Plucker doesn't introduce bitangents until later, where he makes the connection with nodal curves, by duality, the method of finding such asymptotes, for a suitable choice of coordinates, is critical, in finding representations, where the class points, are assigned in pairs, to the asymptote intersections, thus forming an oval and a facing hyperbola. This method, for nonsingular curves, is outlined in [21], Lemma 2.7, for, "harmonic" arrangements of lines, which radiate concentrically around a central polygon, see the aesthetic discussions in the previous chapter and figure 14, which is again discussed in relation to genus below. The pairs theorem is given in Lemma 2.8. Here, the use of duality is important in the technical detail of aligning the bitangent points along two

[12] Plucker is clearly aware of Bezout's theorem, which was published in 1779 by tienne Bzout, "Thorie gnrale des quations algbriques", see also [29], he also makes an interesting reference to a numerical notion of genus, observing that, (p244), $n^2-(\frac{n(n+3)}{2})=\frac{(n-1)(n-2)}{2}$, see footnote 4, however, this could be a coincidence, we will return to this point later

[13] Namely, any line in $l\in P^2(\mathcal{C})$ can be defined by an equation $aX+bY+cZ=0$, (up to multiplication by a scalar), hence, defines a corresponding point $p_l\in(P^2)(\mathcal{C})$, $p_l=[a:b:c]$. A point $p\in P^2(\mathcal{C})$ can be written as the intersection of lines $\{cX+dY+eZ:cp_1+dp_2+ep_3=0\}$, hence determines a line $l_p\subset(P^2)(\mathcal{C})$, given by $p_1X+p_2Y+p_3Z=0$. Moreover these constructions are inverses

parallel axes, see footnote 8. Plucker clearly understands the general duality map, in which a curve C is mapped to its set of tangents, forming a new curve C^*, ([14]), and the definition of class at this stage, as, on p291, he notes that "Eine gegebene Curve n. Ordnung hat zu ihrer Polar-Curve eine Curve der n(n-1) Ordnung". "A Plane curve of degree n has a dual curve of degree $n(n-1)$, Although there is no formal proof given of this statement, it is interesting to speculate, whether this correct statement, now usually referred to as the "Elementary Plucker Formula", was obtained by generalising results for curves of low degree, or, more likely, was proved but unpublished. A later geometric proof of this result was given by Severi, see Chapter 11, and reproduced in [20];

Theorem 0.1. *Let C be a plane projective algebraic curve of order n and class m, with d nodes. Then;*

$$n + m + 2d = n^2$$

Proof: See [20].

Lemma 0.2. *Let C be a projective algebraic curve, with finitely many flexes, then $Cl(C)$, as defined in Definition 4.1, is the same as $deg(C^*)$ and $deg(C)$ is the same as $Cl(C^*)$. In particular, if C has at most nodes as singularities, then;*

$$deg(C^*) = n(n-1) - 2d$$

Proof: See [20].

Plucker obtains, again without proof, the correct formula for the number $3n(n-2)$ of inflexions on a nonsingular complex curve C of degree n, $3n(n-2) = 3(m-n)$, see the discussion below;

"Nach dem 6. Paragraphen, auf den ich mich hier stillschweigund ofter bezogen habe, hat eine Curve n. Ordnung $3n(n-2)$ Wendungstangenten."

and also a formula for the number of bitangents to a real algebraic curve;

"und also auch die Anzahl jener doppeltengenten $\frac{n(n-2)(n^2-9)}{2}$"

giving his famous result of 28 bitangents, for quartic curves. This seems to be correct for generic real curves, the correct formula $\frac{n(n-2)(n^2-6)}{2}$ bitangents for complex nonsingular algebraic curves, is given in [21], footnote 5. It seems that, here, these results were genuinely obtained, from particular constructions of lower degree, as he obtains a deeper understanding of duality in his following "Theorie der Algebraischen Curven", Bonn, published in 1839. At this stage, he was a full professor of mathematics, until 1847, where he succeeded Karl von Mnchow in 1836, and, had married, in 1837, a Miss Altsttten.

[14] see [20] for a formal algebraic definition

There, he obtains, on p211, the following result, concerning the transformation of branches by duality;

"v die Anzahl der Wendungspuncte der gegebenen Curve und also die Anzahl der Spitzen ihrer Polar-Curve;

u die Anzahl der Doppeltangenten der gegebenen und also der Doppelpuncte ihrer Polar-Curve;

y die Anzahl der Spitzen der gegebenen Curve und also der Wendungspuncte ihrer Polar-Curve;

x die Anzahl der Doppelspuncte der gegebenen Curve und also der Doppeltangenten ihrer Polar-Curve.

A rigorous proof of this statement can be found in [20];

Theorem 0.3. *Transformation of Branches by Duality*

Let C be a plane projective algebraic curve, with finitely may flexes, then, if γ is a branch of C with character (α,β), such that $\{\alpha,\alpha+\beta\}$ are coprime to $char(L)=p$, the corresponding branch of C^* has character (β,α).

Proof: See [20].
The result is illustrated in figure 13. Here, we see how an inflexion corresponds to a cusp, in duality, which is formally described by a character, ([15]), and, then, how a bitangent, corresponds to a node, this also requires the definition of the duality map, see footnote 14.

Plucker also obtains immediately afterward the formulas, which are still associated to his name;

"Alsdann besteben die folgenden, durch das Princip der Reciprocitat paarwise mit einander verknupften, Gleichungen:

$m=n(n-1)-2x-3y$, (1)

$n=m(m-1)-2u-3v$, (2)

$v=3n(n-2)-6x-8y$, (3)

$y=6m(m-2)-6u-8v$, (4)

$u=\frac{n(n-2)(n^2-9)}{2}-(2x+3y)(n(n-1)-6)+2x(x-1)+\frac{9y(y-1)}{2}+6xy$ (5)

[15] This is explained in Chapter 11, and was a major preoccupation of the geometer Severi

$x = \frac{m(m-2)(m^2-9)}{2} - (2u+3v)(m(m-1)-6) + 2u(u-1) + \frac{9v(v-1)}{2} + 6uv$ (6)
"

Plucker's observation in the "System der Analytischen Geometrie", for the number of inflexions of a nonsingular curve C, follows from (3), $i = 3(m-d) = 3(d(d-1)-d) = 3d(d-2)$. A rigorous proof of this result (3), follows from the class formula, known as Plucker III' in Severi, and Plucker's earlier elementary Plucker formula above. The class formula is given in [20], along with the proofs relevant to the other claims;

Theorem 0.4. *Let C be a normal plane projective curve, not equal to a line, having at most nodes as singularities, with the convention on summation of branches given above and $\{m, n, \rho, d\}$ as defined in the previous lemmas. Then, we obtain the class formula;*

$3m - 3n = \sum_\gamma (\beta(\gamma) - 1)$ (1)

and the genus formula;

$6\rho + 3n - 6 = \sum_\gamma (\beta(\gamma) - 1)$ (2)

and the node formula;

$3n(n-2) - 6d = \sum_\gamma (\beta(\gamma) - 1)$ (3)

In particular, if C has at most ordinary flexes, and i is the number of these flexes, we obtain the class formula, referred to as Plucker III' in [35]*;*

$3m = 3n + i$ (4)

and the genus formula;

$6\rho = i - 3n + 6$ (5)

and the node formula, referred to as Plucker III in [35]*;*

$6d = 3n(n-2) - i$ (6)

Proof: See [20].

It seems unclear exactly how Plucker obtained these results. Again, Plucker formulates some notion of the genus of an algebraic curve, on p215, "Das gesuchte Maximum wird hiernach:

$z = \frac{(n-1)(n-2)}{2}$

as the maximum number of double points, an irreducible curve can possess, with a footnote to the Swiss mathematician Gabriel Cramer, (1704-1752). This is consistent with the

degree-genus formula, for a curve, of degree n, having at most d nodes as singularities, $g=\frac{(n-1)(n-2)}{2}-d$, see also footnote 4, as we always have that $g\geq 0$. There are exactly $\frac{(n-1)(n-2)}{2}$ "alcoves", in finite position, formed by a general set of n lines, see [21], Lemma 1.12. This gives a clear connection between the genus g of a complex algebraic curve C, and the number of ovals of real algebraic curves, see the attached figure 14, in which there are 5 lines $\{l_i : 1\leq i\leq 5\}$, 6 alcoves $\{a_i : 1\leq i\leq 6\}$ and $g=6$. The earlier reference by Plucker to Cramer, and his discussion of ovals, suggests that he might have been impressed by this idea from previous stages of his career, and, here, formulates, a precise geometric definition. The interest of Cramer in the dimensions of linear systems of curves, as in his paradox, and the strong connection with Plucker's formulae, seem to motivate Severi's definition of genus. Moreover, the intuition of figure 12, relating nonsingular curves to harmonic arrangements of lines, is the main idea in proving the equivalence of Severi's and Klein's later definition, again see footnote 4. As before, the author would suggest that much of Plucker's, or possibly even Cramer's, work on this subject was unpublished, but, perhaps, later obtained by Severi, who, was, in fact known to have written his major work, "Vorlesungen uber algebraische Geometrie", (1907), in German. Severi's proof of Plucker's formulae relies heavily on his geometric definition of genus, which is preserved by birational maps, see Chapter 11. It seems possible that many of Plucker's unpublished ideas were later rediscovered by Severi, but, perhaps, this would be to discredit the great achievements of this geometer and the later work of the "Italian School" in algebraic surfaces, ([16])
Plucker was also a professor of physics at Bonn, from 1847 to 1868. He produced a further important geometric work, "System der Geometrie des Raumes in neuer analytischer Behandlungsweise", (Dusseldorf), in 1846, which, refined his earlier results. However, this upset a parallel school of synthetic geometry, ([17]), associated with Jakob Steiner, and he mainly concentrated on research into physics after this date. In 1847 he began research into the behaviour of crystals in a magnetic field, establishing results central to a deeper knowledge of magnetic phenomena. At first alone and later with the German physicist Johann W. Hittorf, Plucker investigated the magnetic deflection of cathode rays, ([18]) Together they also made many important discoveries in spectroscopy, anticipating the German chemist Robert Bunsen and the German physicist Gustav R. Kirchhoff, who later announced that spectral lines were characteristic for each chemical substance. In 1862 Plucker pointed out that the same element may exhibit different spectra at different temperatures. According to his pupil J.W. Hittorf, Plucker was the first to identify the three lines of the hydrogen spectrum, which a few months after his death were recognized in the spectrum of solar radiation, and preceded the celebrated experiments of R. Bunsen and G. Kirchhoff in Heidelberg. The discreteness of spectral lines was an important motivation in the later development of quantum physics. Plucker wrote 59 papers on pure physics, published primarily in the "Annalen der Physic

[16]See [2] for a good modern survey of their results in this area.

[17]In which infinitesimals are defined algebraically as "vanishing powers", without the use of coordinates

[18] The earliest version of the cathode ray tube, now used in televisions, was invented by the German physicist Ferdinand Braun in 1897, possibly influenced by the work of the English physicist J. J. Thomson, who identified the electron on the basis of the deflection of cathode rays in both electric and magnetic fields. Plucker was influenced by the English scientist Michael Faraday, (1791-1867), with whom he corresponded, and who motivated his work on magnetism. Plucker later studied the phenomena of electrical discharge in evacuated gases. A modern text on electrodynamics is [8]. For a discussion of Maxwell's equations, which include Faraday's law of induction, and Gauss's laws for electric and magnetic fields, see [33].

und Chemie" and the "Philosophical Transactions of the Royal Society", ([19])

Following Steiners death in 1863, Plucker returned to the study of mathematics with his final work in geometry, "Neue Geometrie des Raumes gegrndet auf die Betrachtung der geraden Linie als Raumelement" (New Geometry of Space Founded on the Treatment of the Straight Line as Space Element). He died before finishing the second volume, which was edited and brought to completion by his gifted young pupil Felix Klein, who served as Pluckers physical assistant from 1866 to 1868, and whom we referred to above. Although Pluckers accomplishments were unacknowledged in Germany, English scientists appreciated his work more than his compatriots did, and, in 1868, he was awarded the Copley Medal.

Other scientists, who, I will argue, used aesthetic ideas in their work, in particular the analysis of wave and periodic motions, are the contemporary French mathematicians Pierre Simon Laplace and Joseph Fourier. Laplace was born in Beaumont-en-Auge, Normandy, in 1749. In 1765, he began to read theology at the University of Caen, where he remained for 5 years. In 1768, he published his first paper, "Recherches sur le calcul integral aux differences infiniment petites, et aux differences finies", (1766-1769, Volume 4, Melanges de Turin), which resulted in correspondence with another famous French mathematician Lagrange. Here, he advocates the use of a finite difference method to find solutions of ordinary differential equations;

"Plusieurs principes du calcul integral aux differences infiniment petites, ont egalement lieu pour les differences finies, ainsi toutte fonction de x, par ex. qui satisfera pour y^x", dans l'equation A, (differential equation omitted, see text (formulation of counterpart) e qui refermera un nombre n, de constantes arbitraires, en fera l'integrale complette. Ce principe qui est de plus grand usage dans le calcul integral aux differences infiniment petites, n'est pas d'un usage moins etendu, dans le calcul aux differences finies." (p301, Section XIV)

There is the use of integral notion, early in the text, and in Section XV, p302, he gives a definition of the finite difference operator and its iterates;

"Comme on a

$\Delta . y^x = y^{x+1} - y^x$

$\Delta^2 . y^x = y^{x+2} - 2y^{x+1} + y^x$

$\Delta^3 . y^x = y^{x+3} - 3y^{x+2} + 3y^{x+1} - y^x$

..."

[19] His interestingly named study of wave surfaces (1839) and the reflection of light at quadric surfaces, (1847), combine both physics and geometry, and are counted among his 41 mathematical papers.

In Section XV1, p303, he obtains a series solution to a difference equation, by this difference method;

"Soit propose d'integrer l'equation du premier ordre

$X^x = y^x + H^x y^{x+1}$

$y^x = p^{x-1}(A - \Sigma . [\frac{X^x}{p^{x-1}}]$"

The above passage and the simultaneous use of sum and integral notation, show that Laplace was, at an early age, comfortable with the idea of switching between step-by-step difference methods and ideas from calculus, in solving ordinary differential equations. This was, of course, due to the influence of Newton's (and Liebniz's) use of infinitesimals, rather than limits, in the definition of integration and differentiation, see Chapter 5, a point of view which was unaltered until the 19th century. As we saw there, Newton construed integrals as sums of infinitesimals, hence the sums obtained, by the difference method, of simple differential equations such as $\frac{dy}{dx} = G(x, y(x))$, $(*)$ can be interpreted immediately as integral solutions $y(x) = \int_0^x G(x', y(x'))dx'$, which solve the original differential equation by the Fundamental Theorem of Calculus. These arguments are, in some sense, unacceptable by modern standards, however, now, with the development of nonstandard analysis, it is possible to show these two viewpoints are possible simultaneously, provided that one replaces equality by "infinitely close". The interested reader can find a good discussion of the finite difference method in [25], where a number of analogues of simple calculus results are proved. The nonstandard definition of the integral, which we briefly encountered in Chapter 5, can be found in [27], together with conditions under which the nonstandard integral is infinitely close to the standard integral, Definition 3.26 and Theorem 3.28. A rigorous demonstration of $(*)$ is in [27], Theorem 4.1,([20])

In 1771, Laplace moved to Paris, having supposedly abandoned theology and become an atheist. Intent on becoming a great mathematician, he carried a letter of introduction to Jean le Rond d'Alembert. After gaining a teaching position at the Ecole Militaire, on his recommendation, he threw himself into original research until 1787. In 1778, he published "Un memoire sur le calcul integral", ([21]), and, in 1785, "Theorie des Attractions des Spheroids et de la Figure des Planetes". In this text, he considers the gravitational potential function and shows that it always satisfies the partial differential equation;

"$\nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0$"

This is still known today as Laplace's equation. Later, there is an attempt to solve the equation using power series, which we considered in Chapter 5. Real valued solutions of

[20] It can be shown that, under quite general conditions, solutions to ordinary differential equations, obtained by the difference method, will be infinitely close to the standard solution. However, there are some fairly exceptional cases. We will consider the corresponding problem for partial differential equations below.

[21] Here, he demonstrates that $\int_{\mathcal{R}} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$, predating the development of Cauchy's later complex analysis.

Laplace's equation are known as harmonic functions. In two dimensions, with variables (x, y), a function V is harmonic iff it is the real part of a complex analytic function, ([22]). The Laplacian operator occurs in a number of important partial differential equations, involving more than one spatial variable, ([23]), and can even be generalised to an operator on an arbitrary algebraic curve. Understanding the eigenvalues of this operator, that is solutions to the partial differential equation $\nabla^2 W = \lambda W$, for λ infinitely close to real values, seems, in conjunction with methods, later developed by Laplace and Fourier, to be the clue to using the finite difference method to solve partial differential equations, under quite general conditions, on surfaces, we discuss this question in greater depth below.

Laplace continued his analytical research into the solar system with his "Mechanique Cleste" published in five volumes. The first two volumes, published in 1799, contain methods for calculating the motions of the planets, determining their figures, and resolving tidal problems. In the third volume, published in 1802, (Complete Works, Tome 3), Laplace applies these methods to understand geodesic lines on the surface of the earth. On p118, he defines geodesics in terms of meridians, and rotated overlapping triangles;

"Telle est donc la propriete caracteristique de la courbe tracee par les operations geodesiques. Son premier cote, dont la direction peut etre supposee quelconque, est tangent a la surface de la Terre; son second cote est le prolongement de cette tangente, plie suivant une verticale; son troisieme cote est le prolongement du second cote, plie suivant une verticale, et ainsi de suite"

and observes that such lines are the shortest distance between two points on a surface;

"Ainsi les lignes tracees par les mesures geodesiques ont la propriete d'etre les plus courtes que l'on puisse mener sur la surface de spheroide, entre deux de leurs points quelconques;"

[22] The subject of complex analysis was developed later, by the French mathematician Augustin-Louis Cauchy, (1789-1857), who is thought to have essentially proved the famous Cauchy's Integral Theorem in 1814, publishing it entirely in "Memoire sur les integrales definies prises entre des limites imaginaires" (Memorandum on definite integrals taken between imaginary limits), which was submitted to the Acadmie des Sciences on February 28, 1825. Cauchy defined residues in "Sur un nouveau genre de calcul analogue au calcul infinitsimal" (On a new type of calculus analogous to the infinitesimal calculus), Exercices de Mathematique, vol. 1, p. 11 (1826), and proved the residue theorem in Mmoire sur les rapports qui existent entre le calcul des Residus et le calcul des Limites, et sur les avantages qu'offrent ces deux calculs dans la resolution des equations algebriques ou transcendantes (Memorandum on the connections that exist between the residue calculus and the limit calculus, and on the advantages that these two calculi offer in solving algebraic and transcendental equations), presented to the Academy of Sciences of Turin, November 27, 1831. Cauchy is sometimes associated with the introduction of limits, but, this is unclear. It seems, from these publications, that he was ambivalent, and adapted to the new perspective in his later work. His "Cours d'Analyse", (1821), uses $\epsilon - \delta$ notation and infinitesimals.

[23] Examples include the heat equation, the wave equation and Schrodinger's equation, see [16] and [6] for discussions of these equations in the plane. The concept of potential also occurs in close connection with Maxwell's equations and Navier-Stokes equations.

He finally deduces, using an infinitesimal argument, that the second derivative of the geodesic path is perpendicular to the surface, ([24]).;

"Ainsi les lignes tracees par les mesures geodesiques ont la propriete d'etre les plus courtes que l'on puisse mener sur la surface de spheroide, entre deux de leurs points quelconques;"

"on aura donc, par la nature de la verticale et en supposant que $u = 0$ soit l'equation de la surface de la Terre,

$0 = \frac{\partial u}{\partial x} d^2 y - \frac{\partial u}{\partial y} d^2 x$

$0 = \frac{\partial u}{\partial x} d^2 z - \frac{\partial u}{\partial z} d^2 x$

$0 = \frac{\partial u}{\partial y} d^2 z - \frac{\partial u}{\partial z} d^2 y$
"

An important analysis occurs on p131, with the attached figure 15. There is a clear understanding of periodic motion here, with the geodesical paths traced out by circular orbits around a sphere, and inclusion of the sine and cosine functions. Laplace states;

"Mais dans l'hypothesese de la Terre spherique on a;

$\frac{d\phi}{ds} = \frac{sin.\lambda}{cos.\psi}\ \frac{dd\phi}{ds^2} = \frac{2.sin.\lambda.cos.\lambda}{cos.\psi}.tang.\psi$

$\frac{d\psi}{ds} = cos.\lambda\ \frac{dd\psi}{ds^2} = -sin.^2\lambda.tang.\psi$"

obtaining expressions for the variation of angle with arc length. In the fourth volume, published in 1805, we obtain further applications (p515) in the study of the satellites of Saturn, Laplace develops his theory of gravitational potential to calculate the eccentricity of their orbits, which he assumes to be elliptical. There is clearly some development of the field, now known as Lagrangean mechanics, a good discussion of this topic can be found in [13]. The fifth volume, published in 1825, contains figure 16, giving a more geometric analysis, ([25])

In 1807, Joseph Fourier published his treatise "Memoire sur la propogation de la chaleur dans les corps solides", which exerted a great influence on Laplace. In Fourier's " Memoire sur la Theorie Analitique de la Chaleur", (1822), p594, which is, presumably, taken from the memoire of 1807, Fourier defines the heat equation (Section 3,p586);

[24]This is equivalent to the property that $\bigtriangledown_{\gamma'}(\gamma') = 0$, for the induced connection, obtained by an orthogonal projection of the covariant derivative $\bigtriangledown_{\overline{\gamma'}}(\overline{\gamma'})$ onto the surface S. It seems likely that Gauss was aware of this interpretation by 1825, when he published "Disquisitiones generales circa superficies curvas". Good modern references on differential geometry are [4] and, for curves and surfaces, [19].

[25]No doubt Laplace was influenced by Newton's earlier theory of planetary motion, in which he gave a derivation of Kepler's laws, in his Philosophiae Naturalis Principia Mathematica of 1687, see Chapter V.

"Or l'equation differentielle de mouvement lineaire de la chaleur est $\frac{dv}{dt} = \frac{k}{C.D}\frac{d^2v}{dx^2}$, et si l'on ecrit $\frac{kt}{C.D}$ au lieu de t, on a $\frac{dv}{dt} = \frac{d^2v}{dx^2}$"

and employs a series solution involving sine functions;

"Nous employons en premier lieu l'expression suivante:

$v = \sum e^{-i^2t} sin(ix)\alpha_i$, $(**)$, ([26]).

en designant par α_i une fonction inconnue du temps t qui contient aussi l'indice i, ([27]);"

Laplace, perhaps motivated by his work we have considered on planetary motion, recognised that Fourier's method of series for solving the heat equation could only apply to a limited region of space as the solutions were periodic. Such a view seems reasonable, as periodic solutions on an infinite line would be a physically unrealistic model of diffusion, and a periodic solution reflects the geometry of a bounded interval. In figure 17, we see how the sine function, restricted to a bounded interval $[-\pi, \pi]$ can be wrapped around a circle, to provide a new continuously differentiable function. This property turns out to be not only physically realistic, ([28]), but mathematically essential, in the sense that the Fourier series of such functions converge uniformly. In this case, a complete analysis of a function in terms of its harmonics can be obtained, see figure 18, ([29]). A proof using infinitesimals is given in [26], ([30]);

Theorem 0.5. *For $g \in C^\infty(\mathcal{S})$, there is a non standard proof of the uniform convergence of its Fourier series.*

Proof: The idea is to obtain the result first using infinitesimals $\frac{1}{\eta}$, the summation being up to η . This turns out to be relatively easy, by the finite difference method, ([31]), without

[26]This is basically correct, the general periodic solution, with values in $\mathcal{C}$, on $\mathcal{R}_{\geq 0} \times [-\pi, \pi]$ is given by, $v(x, t) = \sum_{n=-\infty}^{\infty} A_n e^{-n^2t} e^{-inx}$, hence the real solutions are given by $\sum_{n=-\infty}^{\infty} a_n e^{-n^2t} sin(nx) + b_n e^{-n^2t} cos(nx)$

[27]Fourier makes a miscalculation, here, by requiring that the terms α_i depend on time t

[28]It turns out that solutions to the closely related wave equation with fixed endpoints have this property, see [30]

[29]The harmonics, α_i in Fourier's notation, refer to the coefficients of the sine/cosine terms in the series, the $sin(ix)$ terms. A formal definition of these can be found in [26], Definition 0.1 and Remarks 0.3. The definition of Fourier coefficients in terms of infinitesimals, infinitely close to the standard Fourier coefficients, is given in Definition 0.8. It seems likely that both ideas were again used simultaneously in the early development of the subject.

[30]The standard proof of this result can be found in [36], based on Dirichlet's memoir of 1829, "Sur la convergence des series trigonometriques qui servent a representer une fonction arbitraire entre des limites donnees". According to [37], Fourier had originally claimed that any periodic function could be written as a series of sines and cosines, this being true only if one construes equality in the measure theoretic sense of "almost everywhere"

[31]This could be classed as a result in the theory of fast or discrete Fourier transforms, see [14], there is some evidence that an analogue of the result could be found in Gauss's "Theoria Interpolationis Methodo

imposing any particular restrictions on g;

$g_\eta(x) = \frac{1}{2}\sum_{m\in\overline{\mathcal{Z}}_\eta} \hat{g}_\eta(m) exp_\eta(\pi i x m)$ (*)

The continuously differentiable property is required to show that the tail terms of this series converge like $\frac{1}{m^2}$, hence rapidly enough to ensure that the total of the summands with infinitely large indices is infinitesimal. It follows that the infinitesimal series is infinitely close to the standard series.

It is not clear whether a proof along these lines was known at this time, though it does seem more intuitive than Dirichlet's later result of 1829,([32]) However, as we will now see, Laplace, in his "Memoire sur divers points dAnalyse", (Journal de lEcole Polytechnique, Tome VIII., 229-265, Oeuvres Completes, XIV, 178-214, (1809)) approached the question of finding solutions that diffused indefinitely in space, using his earlier method of finite differences, and obtained the correct result. On p189, he defines the heat equation using both the method of finite differences and as a standard partial differential equation;

"L'equation aux differences finies,

$\Delta^2 y_{x,x'} = \Delta' y_{x,x'}$

se change dans une equation aux differences infiniment petites, en y substituant $\frac{\partial}{\partial x}$ et $\frac{\partial}{\partial x'}$, au lieu des caracteristiques Δ et Δ', (Memoires de l'Academie des Sciences, 1779), et en y changeant $y'_{x,x'}$ en y, on a

$\frac{\partial^2 y}{\partial x^2} = \frac{\partial y}{\partial x'}$

noting, on p188, the step by step method of solution;

"L'equation precedente aux differences partielles donne

$y_{x,x'+1} = y_{x,x'} + \Delta^2 y_{x,x'}$.

and obtains, on p190, the solution;

$y = \frac{1}{\sqrt{\pi}}\int e^{-z^2}\phi(x+2z\sqrt{x'})dz$, (*), ([33])

Nova Tractata", which was published posthumously in Volume 3 of Gauss's collected works, appearing from 1863 to 1871, but predated Fourier's work of 1807.

[32]It is claimed in [1], that Fourier began to tackle the problem, in the style that Dirichlet later used, within a later work of 1822, see below, but the evidence involves limits which can't be found in the original text, though they may just have been introduced at this time, see 22.

[33]Here, $\phi(w)$ denotes the boundary condition $y(0,w)$. The solution is usually now written in the form $\frac{1}{\sqrt{4\pi x'}}\int_{\mathcal{R}} e^{\frac{-(x-w)^2}{4x'}} g(w)dw$. The two expressions are easily seen to be the same by making the substitution $w = x + 2z\sqrt{x'}$

In order to understand how Laplace arrives at this answer, we need to consider his contemporary paper "Memoire sur les Approximations des Formules qui sont Fonctions de Tres Grands Nombres et sur Leur Application aux Probabilites" (Memoire de l'Academie des Sciences, 1809). On p336, he notes;

"Considerons le cas de $i=\frac{1}{2}$; on aura pour l'integrale de equation (u)

$\phi(r,n)=\int\frac{1}{\sqrt{x}}e^{\frac{-x^2}{6}}(acosrx+bsinrx)dx$, [34]

a et b etant deux constantes arbitraires, et l'integrale etant prise depuis x nul jusqu'a x infini"

The expression on the right is essentially the transform which he is still known for, in this case applied to the function $\frac{1}{\sqrt{x}}e^{\frac{-x^2}{6}}$, [35]. On p332, Laplace gives the expression;

$\phi'(r,n)=\sqrt{\frac{3}{2\pi}}e^{\frac{-3r^2}{2}}[1-\frac{30}{n}(1-6r^2+3r^3)+\ldots]$

There is a similarity here, with the calculation, for a transform named after Fourier, [36], namely that $\mathcal{F}(K)=e^{-tr^2}$, where $K(t,x)=\frac{1}{\sqrt{4\pi t}}e^{\frac{-x^2}{4t}}$, the function appearing in Laplace's earlier solution $(*)$. This calculation, using Fourier transforms, is an essential step in recovering Laplace's solution from the reduction of the original partial differential equation to a set of ordinary differential equations indexed by the space step x, [37]. Laplace clearly knew how to solve the simpler 1 variable equations from $(\dagger)$, using the difference method, which he presented in 1768, indeed, on p328 of his approximation memoir, he gives the step by step solution;

"on a donc cette equation aux differences partielles finies et infiniment petites

$(p)\phi(r',n-1)-\phi(r'',n-1)=\frac{3}{\sqrt{n}}\phi'(r,n)$, [38]

The final step, to obtain Laplace's solution, requires the inversion of the Fourier transform, [39]. Laplace is carrying out a parallel calculation with the Laplace transform, which is more difficult to assess, as a simple inversion theorem is not known, [40]. Laplace's concern with

[34]The reader should not confuse the function of 2 variables, $\phi(r,n)$, here, with the function of 1 variable ϕ given above.

[35]The precise definition is $\mathcal{L}(f)(r)=\int_0^\infty f(x)e^{-ixr}dx$

[36] We define the Fourier transform by $\mathcal{F}(f)(r)=\int_{\mathcal{R}} f(x)e^{-ixr}dx$; there is a clear connection with the Laplace transform, in that the domain of integration is changed from the real line $\mathcal{R}$ to the half line $(0,\infty)$, although in modern treatments r is usually treated as a complex variable for the Laplace transform.

[37]Namely, if f solves the heat equation, $f_t-f_{xx}=0$, $f(0,w)=g(w)$, then $d\frac{\mathcal{F}(f)}{dt}+x^2\mathcal{F}(f)=0$, $(\dagger)$ $\mathcal{F}(f)(0,x)=\mathcal{F}(g)(x)$, for each $x\in\mathcal{R}$.

[38]The standard solution of $(\dagger)$ is given by $A(x)e^{-tx^2}$, where $A(x)=\mathcal{F}(g)(x)$

[39]The method of solving the heat equation using the difference method, and its relation with the inversion theorem for Fourier transforms can be found in [23]

[40]The interested reader can find the author's result for the inversion of Laplace transforms in [24]

the heat equation arose mainly from considerations in probability, we will return to this point below. Inspired by his memoir, Joseph Fourier wrote the "Thorie du mouvement de la chaleur dans les corps solides", a prize essay, deposited with the Institut on September 28, 1811, and published in Memoires lAcadmie Royale des Sciences de lInstitut de France, years 1819, 1820, 185-556. Unfortunately, this work is difficult to trace (it appears in Fourier's complete works, Tome1), however, the contents are reproduced in " Memoire sur la Theorie Analitique de la Chaleur", (1822). In Ch(IX) Section 1,342-371, (p428-429), Fourier addresses the solution of the heat equation on an infinite line;

"et l'on aura

$\frac{du}{dt} = k\frac{d^2u}{dt^2}$"

obtaining the solution;

"Il en resulte que la valeur de u pourra etre exprimee ainsi:
$u = \int dqQ.cos.qxe^{-kq^2t}$
"

Here, Fourier writes the solution as the (inverse) Fourier transform applied to the function e^{-kq^2t}. The reader should observe the similarity between this expression and the one he obtained in (**), replacing the sum by an integral, and the coefficients, by the (transform) of the boundary condition. This analogy is an important feature of their approach using infinitesimals, indeed, the inversion theorem for Fourier transforms is, from a nonstandard perspective, exactly analogous to the convergence of Fourier series. The following result is taken from [27], the reader can compare this with the previous theorem;

Theorem 0.6. *For $g \in S(\mathcal{R})$, the Fourier Inversion Theorem holds and admits a non standard proof.*

Proof: Again, one obtains the result using infinitesimals $\frac{1}{\eta}$, the nonstandard integral converting to a sum up to η^2.
$g_\eta(x) = \frac{1}{2}\int_{\overline{\mathcal{R}}_\eta} \hat{g}_\eta(r)exp_\eta(\pi ixr)d\lambda_\eta(r)$

The condition that $g \in S(\mathcal{R})$, meaning rapid decay at infinity, analogous to the previous continuous differentiability condition, is required to show that the transform $\hat{g}_\eta(r)$ also decays like $\frac{1}{r^2}$, hence rapidly enough to ensure that the tail of the integral above infinite values is infinitesimal. It follows that the integral of the nonstandard transform is infinitely close to the integral of the standard transformation.

The result is illustrated in figure 19. A single application of the transformation alters the dispersion and amplitude of the function, and a second application alters just the amplitude, the initial function is essential preserved by two applications of the transformation, ([41]).

[41] Precisely, for the definition in footnote 36, $(\mathcal{F})^2(g)(x) = 2\pi g(-x)$, the double transformation is a constant multiple of the mirror image of the original function.

Consideration of the text shows that Fourier arrived at the result by analogy with his previous paper. It seems likely that Laplace and Fourier saw that their two different integral solutions to the heat equation, implied an inversion property for a particular type of function, ([42]), which might imply a general inversion theorem for (Fourier) transforms. On p561, Chapter IX, Fourier states the Inversion Theorem;

"Il faut maintenant, dans le second membre de l'equation;

$fx = \frac{1}{2\pi}\int d\alpha f\alpha \int dpcos.(px - p\alpha)$", ([43])

However, the proof is unclear. It is interesting to speculate whether Laplace and Fourier had unpublished work on this result, using infinitesimals, before Dirichlet's paper. As we noted above, Fourier believed that any function could be decomposed this way, and observes on p558;

"La fonction fx acquiert en quelque sorte, par cette transformation, toutes les proprietes des quantites trigonometriques; les differentiacions, les integrations et la sommation des suites s'appliquent ainsi a des fonctions generales de la meme maniere qu'aux fonctions trigonometriques exponentielles."

We have seen, in the nonstandard proof, that, at the level of infinitesimals, Fourier was correct, and the inversion theorem holds for any function. However, restrictions on the functions are required when recapturing the classical result, ([44]).

Fourier's considerations are guided by his geometric intuition of the propagation of heat. In figure 20, we see his visualisation of how an initial symmetric temperature profile, on part of an infinite line, propagates to an asymmetric function, in order to equalise the zero temperature on the rest of the line;

"Nous considerons d'abord le premier cas, qui est celui ou a chaleur se propage librement dana la ligne infinie dont une partie ab a recu des temperatures initiales quelconques; tous les autres points ayant la temperature initiale o. Si l'on eleve en chaque point de la barre l'ordoonee d'une courbe plane qui represente la temperature actuelle de ce point, on voit qu'apres une certaine valeur du temps t, l' etat du solide est exprime par la figure de'la courbe. Nous designerons par $v = Fx$ l'equation donnee qui correspond a l'etat initial, et nous supposons d'abord pour rendre le calcul plus simple que la figure initiale de Ia courbe, est composee de deux parties symetriques, en sorte que l' on a la condition $Fx = F(-x)$."

This type of visual thinking, in terms of symmetric and asymmetric functions, guides his considerations on series. Fourier observes on p254, see also figure 21, that any function can

---

[42] $\mathcal{F}^{-1}(\frac{e^{\frac{-x^2}{4x'}}}{\sqrt{4\pi x'}}) = e^{-x'x^2}, \mathcal{F}(e^{-x'x^2}) = \frac{e^{\frac{-x^2}{4x'}}}{\sqrt{4\pi x'}}$

[43] Again, this is nearly correct, if f is real-valued, $f(x) = \frac{1}{2\pi}\int_{\mathcal{R}} d\alpha f(\alpha)\int_{\mathcal{R}} dpcos.(px - p\alpha) + \frac{1}{2\pi}\int_{\mathcal{R}} d\alpha f(\alpha)\int_{\mathcal{R}} dpsin.(px - p\alpha)$, Fourier makes a similar mistake in his definition of Fourier series.

[44] Classical proofs of the inversion theorems and convergence of Fourier Series can be found in [12], with further developments in [10]

be decomposed into a symmetric and asymmetric component,

"La fonction ϕx, developpee en cosinus d'arcs multiples, est representee par une ligne formee de deux arcs egaux places symetriquement de part et d'autre de l'axe des y, dans l'intervalle de $-\pi$ a $+\pi$ (voy. fig. 11); cette condition est exprimee ainsi $\phi x = \phi(-x)$. La ligne qui represente la fonction $\psi(x)$ est au contraire formee dans la meme intervalle de deux arcs opposes, ce qu'exprime l'equation

$$\psi x = -\psi(-x)$$

Une fonction quelconque Fx, representee par une ligne tracee arbitrairement dans l'intervalle de $-\pi$ a $+\pi$, peut toujours etre partagee en deux fonctions telles que ϕx et ψx."

later developing a seperate treatment of these in terms of cosines and sines,([45]).

As well as considering the heat equation on an infinite line, Fourier reformulates it on the solid sphere, (Ch V,283-289), the solid cylinder (Ch VI,290-305), the rectangular prism (Ch VII,306-320), and the solid cube (Ch VIII, 321-341). Fourier uses the Laplacian notation we discussed above, to describe the equation in this case;

$$\frac{\partial v}{\partial t} = \frac{K}{C.D}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right)$$

and obtains some of the solutions using separation of variables, and reduction to the one dimensional case. The problem of solving the heat equation on surfaces or algebraic curves has always been relevant, and general solutions to this problem are still unknown, ([46]). The general idea for solving partial differential equations (depending on space and time) on surfaces (algebraic curves), and, I think, Fourier and Laplace understood this intuitively, is to determine the eigenvalues for the corresponding Laplacian, ([47]). The exponential terms e^{inx}, with $n \in \mathcal{Z}$, or e^{itx}, with $t \in \mathcal{R}$, are eigenfunctions of the 1-dimensional Laplacian operator $\frac{d^2}{dx^2}$ on a bounded interval or an infinite line, and most of the properties of and proofs around Fourier series or transforms rely on this fact. Fourier himself observes this point on p557;

"On aura donc, en designant par i un nombre entier quelconque

$$\frac{d^{2i}}{dx^{2i}} fx = \pm \int d\alpha f\alpha \int dp p^{2i} cos.(px - p\alpha)$$

[45] see also [30] and [31], in which the uniform convergence of cosine/sine series to such functions is discussed, and an asymmetric reflection method is used to solve the wave equation.

[46] One of the major obstacles to scientists working in quantum physics was finding solutions to the related Schrodinger's equation on surfaces such as the sphere, cube or torus. This would be necessary in order to determine the correct orbits for charged particle paths, by relating the discrete frequencies arising from the generalised Fourier series, we explain this further below, and those arising from classical calculations.

[47] There is a general result, known as the Hodge Theorem, which says that the eigenvalues of the Laplacian are discrete (with finite multiplicity) and any smooth function can be decomposed into a series, with the eigenfunctions, corresponding to these eigenvalues, replacing the classical periodic sine and cosine terms. However, the proof of uniform convergence is unknown.

On ecrit le signe superieure lorsque i est pair, et le signe inferieur lorsque i est impair. On aura en suivant cette meme regle relative au choix du signe

$\frac{d^{(2i+1)}}{dx^{(2i+1)}}fx = \mp \int d\alpha f\alpha \int dpp^{2i+1} sin.(px - p\alpha)$"

With the eigenfunctions in place, one can then develop the series or transform method for solving the equation as we have discussed above, exactly for the 1 dimensional case. The problems of uniform convergence and inversion can be handled analogously using the nonstandard method,([48]).

---

[48] The idea being to investigate the discrete set of standard eigenvalues $\{\lambda_i : i \in I\} \subset \mathcal{R}$, and smooth solutions for Laplace's equation $\Delta f = \lambda f$ on a projective algebraic curve C. Using the method of [5], (nonstandard solutions to ordinary differential equations), one can find an (internal) set of distinct nonstandard eigenvalues $\{\mu_i : i \in {}^*I\} \subset {}^*\mathcal{C}$, such that;

(a). Given $i \in I$, there exists $i' \in {}^*I$, with ${}^\circ\mu_{i'} = \lambda_i$.

(b). $Card(A_i) = mult(\lambda_i)$, where $A_i = \{i' \in I : {}^\circ\mu_{i'} = \lambda_i\}$

As the operator ${}^*\Delta$ is almost symmetric, on an internal space V, which includes the lifts $\overline{f}$ of smooth functions $f \in C^\infty(C)$, the corresponding eigenfunctions $\{f_i : i \in {}^*I\}$ are almost orthogonal and form a basis for V. Using the Gramm-Schmidt orthogonalisation procedure, and the method of [26], to control the decay rate of the nonstandard Fourier coefficients, one can show that, for $f \in C^\infty(C)$;

$\overline{f} \simeq \sum_{i \in {}^*I} a_i f_i,\ a_i = \int_C f(x,y)\overline{f_i(x,y)} d\mu(x,y)$

and, taking standard parts, that the Fourier series $f = \sum_{i \in I''} b_i g_i$, where $\{g_i : i \in I''\}$ are the corresponding standard eigenfunctions, taking into account multiplicity, is uniformly convergent.

Using this method, we can define the nonstandard Fourier transform $\hat{\overline{f}} : Sp({}^*\Delta) \rightarrow {}^*\mathcal{C}$, by $\hat{\overline{f}}(\mu_i) = a_i$, to obtain a nonstandard inversion formula, for arbitrary functions;

$\overline{f}(x,y) = \int_{Sp({}^*\Delta)} \hat{\overline{f}}(\mu)\overline{g(\mu,x,y)} d\lambda(\mu)$

where $g(\mu,x,y)$ parametrises the nonstandard eigenfunctions, and $d\lambda(\mu)$ is a counting measure on $Sp({}^*\Delta)$. One can generalise this formula in the case when C is not compact, using the method of [25]. Taking standard parts, we aim to obtain the inversion formula for smooth functions on algebraic curves;

$f(x,y) = \int_{Sp(\Delta)} \hat{f}(\lambda) h(\lambda,x,y) d\rho(\lambda)$

where $h(\lambda,x,y)$ parametrises the standard eigenfunctions, including multiplicity. We can use this technique, combined with the method in [23], to solve partial differential equations on algebraic curves, such as the heat and wave equation, or Schrodinger's equation, see [28] and [33]. Generalising this method, one might approach outstanding questions such as the Atiyah-Singer Index Theorem for noncompact algebraic curves, or provide an alternative approach to the existing theory for projective curves, see [34]. This might involve insights into the geometrical representations of genus g curves we discussed above. The decay rate of Fourier coefficients on an algebraic curve C satisfying certain periodic relations (possibly obtained from the fundamental group) is also addressed by the Ramanujan-Petersson and Weil conjectures, although in simple cases, an ordinary analysis seems to suffice, see [32]. Interest in the Weil conjectures can be traced back to Gauss's "Disquisitiones Arithmeticae", 1798, the modern proof can be found in [7]

Although not a fervent Christian, in his later years, like Newton, Laplace remained curious about the question of God. According to [37], he frequently discussed Christianity with the Swiss astronomer Jean-Frdric-Thodore Maurice, and told Maurice that "Christianity is quite a beautiful thing" and praised its civilizing influence. According to Hahn, "Nowhere in his writings, either public or private, does Laplace deny God's existence." Indeed, he is known to have made the following touching remark to his son, in 1809, "Je prie Dieu qu'il veille sur tes jours. Aie-Le toujours prsent ta pense, ainsi que ton pre et ta mre." (I pray that God watches over your days. Let Him be always present to your mind, as also your father and your mother). Like Laplace, Fourier had an early acquaintance with theology, training for the priesthood and entering the Benedictine abbey of St Benoit-sur-Loire, in 1787, which is one of the surviving examples of the chevet form in France. He became a teacher at the Benedictine college, cole Royale Militaire of Auxerre, where he had studied, in 1790. Plucker's religious views are unknown, although the 5-petalled rose window in the cathedral of Cologne, near to where he was born and worked, may have been a source of inspiration in his work.

References

[1] A History of Analysis, ed. Hans Niels Jahnke, American mathematical Society, (2003).

[2] Complex Algebraic Surfaces, A. Beauville,, LMS Student Texts 34, (1996).

[3] Braids, Links and Mapping Class Groups, J.S. Birman, Princeton University Press, (1975).

[4] Riemannian Geometry: A Modern Introduction, I. Chavel, CUP, (2006).

[5] Loeb Measures in Practice: Recent Advances, Nigel Cutland, Lecture Notes in Mathematics, Springer, (2001).

[6] Partial Differential Equations, L.C.Evans, American Mathematical Society, (2010).

[7] Etale Cohomology and the Weil Conjectures, E.Freitag, R. Kiehl, Springer-Verlag, (1998).

[8] Introduction to Electrodynamics, D. Griffiths, Addison-Wesley ,(1998).

[9] Riemannian Geometry, S. Gallot, D. Hulin, J. Lafontaine, Springer, (2004).

[10] An Introduction to Harmonic Analysis, Y. Katznelson, CUP, (2004).

[11] Complex Algebraic Curves, Frances Kirwan, LMS Student Texts 23, (1992).

[12] Fourier Analysis, T.W. Korner, CUP, (1989).

[13] A First Course in Mechanics, M. Lunn, OUP, (1991).

[14] A Nonstandard Analysis Approach to Fourier Analysis, W.A.J Luxemburg, Studies in Logic and the Foundations of Maths, Volume 69, vol 69, North-Holland, Amsterdam, (1972).

[15] Algebraic Topology, C.R.F. Maunder, Dover, (1996).

[16] Quantum Mechanics, Volume 1, Albert Messiah, Elsevier, (1961).

[17] B. Moishezon, M. Teicher, Braid Group Techniques in Complex Geometry, I: Line Arrangements in CP2, Contemporary Mathematics, Volume 78, (1988).

[18] Lectures on an Introduction to Grothendieck's Theory of the Fundamental Group, J. Murre, Tata Institute, (1967).

[19] Elementary Differential Geometry, B. O'Neill, Academic Press, (1966).

[20] A Theory of Duality for Algebraic Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2009).

[21] A Theory of Harmonic Variations, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[22] A Note on Inflexions of Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[23] Solving the Heat Equation using Nonstandard Analysis, Tristram de Piro, available at http://www.magneticstrix.net, (2013).

[24] An Inversion Theorem for Laplace Transforms, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[25] A Simple Proof of the Fourier Inverion Theorem using Nonstandard Analysis, Tristram de Piro, available at http://www.magneticstrix.net, (2013).

[26] A Simple Proof of the Uniform Convergence of Fourier Series using Nonstandard Analysis, Tristram de Piro, available at http://www.magneticstrix.net, (2013).

[27] Applications of Nonstandard Analysis to Probability Theory, Tristram de Piro, available at http://www.magneticstrix.net, (2013)

[28] Solving Schrodinger's Equation Using Nonstandard Analysis, Tristram de Piro, available at http://www.magneticstrix.net, (2014)

[29] A Nonstandard Bezout Theorem for Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2006).

[30] A Simple Proof of the Uniform Convergence of Fourier Series in Solutions to the Wave Equation, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[31] A Note on Convergence of Fourier Series, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[32] Decay rates of Cusp Functions, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[33] Non Standard Analysis and Physics, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[34] The Laplacian on a Riemannian Manifold, Steven Rosenberg, CUP, (1997).

[35] Trattato di geometria algebrica, 1: Geometria delle serie lineari, Francesco Severi, Zanichelli, Bologna, (1926).

[36] Fourier Analysis, An Introduction, Elias Stein and Rami Shakarchi, Princeton Lectures in Analysis, (2003).

[37] Joseph Fourier, Wikipedia Article, http://www.en.wikipedia.org/wiki/Joseph-Fourier, (2014)

9. TITIAN AND THE GEOMETRY OF COLOUR

TRISTRAM DE PIRO

In this fragment, which continues the previous chapter, we consider how the aesthetics of asymmetry and colour, which we found in some of Titian's paintings, can, again, be used to guide the analysis of functions, this time, in terms of randomly evolving patterns and averaging effects, within the framework of probability theory. We will continue with Laplace's work as a motivating example. Although, as we have seen, Laplace's early published work of 1771 started with solving differential equations and the finite difference method, by 1774 he was already starting to think about the mathematical and philosophical concepts of probability and statistics, and, in 1774, published his first paper on the subject, "Memoire sur la probabilite des causes par les evenements", "Savants etranges 6", (pp621-656), Oeuvres 8, p27-65. In Section 2, p29, Laplace formulates the following principle with an application;

"PRINCIPE. -Si un evenement peut etre produit par un nombre n de causes differentes, les probabilites de l'existence de ces causes prises de l'evenement sont entre elles comme les probabilites de l'evenement prises de ces causes, et la probabilite de l'existence de chacune d'elles est egale a la probabilite de l'evenement prise de cette cause, divisee par la somme de totes les probabilites de l'evenement prises de chacune de ces causes.

La question suivante eclaircira ce principe, en meme temps quelle en fera voir l'usage:

Je suppose que l'on me presente deux urnes A et B, dont la premiere contienne p billets blancs et q billets noirs, et la seconde contienne p' billets blancs et q' billets noirs; je tire de l'une de ces urnes (j'ignore de laquelle) $f+h$ billets, dont f sont blancs et h sont noirs; on demande, cela pose, quelle est la probabiliteque l'urne dont j'ai tire ces billets est A ou qu'elle est B.

En supposant que cette urne soit A, la probabilite d'en tirer f billets blancs et h billets noirs est

...

Soit K cette quantite; si l'on suppose maintenant que l'urne dont j'ai tire les billete est B, la probabilites d'en tirer f billets blancs et h billete noirs se determinera en changeant, dans K, p et q en p' et q'; soit K' ce quo devient alors cette expression. Cela pose, les probabilites que l'urne dont j'ai tire les billete est A ou B sont entre elles, par le principe enonce ci-dessus, comme K est a K'; la probabilite que cette urne est A est egale a $\frac{K}{K+K'}$, et celle qu'elle est B est egale a $\frac{K'}{K+K'}$".

Here, Laplace explains the idea of Bayes' Theorem, ([1]), and how the probability of an event is determined by its conditional probabilities on a set of mutually exclusive causes, ([2]). He illustrates this concept with the problem of drawing white and black tickets from two urns A and B, concluding that if the probability of drawing f white tickets and h black tickets from A is K, the the probability of drawing f white tickets and h black tickets from B is K', then the probability that the the urn was A, given the draw of f white tickets and h black tickets, is $\frac{K}{K+K'}$, assuming an equal probability of drawing from urn A and from B, ([3]). In Section 5, p42, Laplace defines the probability density function, illustrating his definition with figure 23, and recognises that the notion of independence between observations, see figure 22, amounts to taking the product of the probability density functions;

"Probleme III.- Determiner le milieu que l'on doit prendre entre trois observations donnees d'un meme phenomene.

Solution.- Representons le temps par une droite indefinie AB (fig. 1), et supposons que la premiere observation fixe l'instant du phenomee au point a, la seconde au point b et la troisieme (figure 22) au point c;

...

nous representerons l'equation par celle-ci: $y = \phi(x)$. Or voici les proprietes de cette courbe:

1° Elle doit etre partagee en deux parties entierement semblables par la droite VR, car il est tant aussi probable que l'observation s'ecartera de la verite a droite comme a gauche: (figure 23)

2° Elle doit avoir pour asymptote la ligne KP, parce que la probabilite que l'observation s'eloigne de la verite a une distance infinie est evidemment nulle;

3° L'aire entiere de cette courbe doit etre egale a l'unite, puisqu'il est certain que l'observation tombera sur un des points de la droite KP. Supposons maintenant (fig. 1) que le veritable instant du phenomene soit au point V, a la distance x du point a, la probabilite que les trois

---

[1]Named after the English mathematician, Thomas Bayes, 1701-1761, specifically, that;

$$P(A|B)P(B) = P(B|A)P(A)$$

for events A and B.

[2]$P(A) = P(A|B_1)P(B_1)+\ldots+P(A|B_n)P(B_n)$, for mutually exclusive and exhaustive events $\{B_1, \ldots, B_n\}$

[3]Specifically, denoting the event of drawing f white tickets and h black tickets by C, and the events of drawing from urn A and urn B by A and B, we have, if $P(C|A) = K$, $P(C|B) = K'$, then;

$$P(A|C) = P(C|A)\frac{P(A)}{P(C)} = P(C|A)\frac{P(A)}{P(C|A)P(A)+P(C|B)P(B)} = \frac{K.\frac{1}{2}}{\frac{K}{2}+\frac{K'}{2}} = \frac{K}{K+K'}$$

by the two parts of Laplace's proposition and the assumption on equal probability.

observations a, b et c s'ecarteront aux distances Va, Vb et Vc sera;

$\phi(x)\phi(p-x)\phi(p+q-x)$."

The intuitive idea behind independence is that the probability of an event A is unaffected by the occurrence (or not) of an event B. Laplace naturally extends this idea to continuous random variables, which encode the possible (continuous) set of values of an event. Laplace is able to formalise his intuition, due to his command of the analytic method, using infinitesimals. ([4]).

In 1778, Laplace continued his research into probability with the "Memoire sur les probabilites", (Memoires de lAcademie des Sciences 1778,(1781), p227-237, Oeuvres 9,p383-485). Here, in Section 18, p423, he formulates a continuous version of the urns problem, finding the probability x of being a boy, given a sample of p boys and q girls, from a binomial distribution, with probability x, and, assuming the distribution of probabilities is uniform, as $\frac{x^p(1-x)^q dx}{\int_0^1 x^p(1-x)^q dx}$, ([5]);

"Soient x la possibilite de la naissance d'un garcon et $1-x$ celle de la naissance d'une fille; la probabititu que, sur $p+q$ enfants, il y aura p garons et q filles, sera, comme on l'a vu dans l'article precedent, egale a $\lambda x^p(1-x)^q$; or, si l'on regarde x comme une cause particulicre de

[4]The formal definition of independence for events, is that $P(A \cap B) = P(A)P(B)$, or, $P(A|B) = P(A)$. The modern definition of a probability density function f_X for a random variable X can be found in [5], p60, so that the probability $P(a \leq X \leq b) = \int_a^b f(x)dx$. As Laplace notes in 3°, we then have that $P(-\infty \leq X \leq \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$, and, in 2°, $lim_{n\rightarrow\infty} \int_{|x|\geq n} f(x)dx = 0$, (a formal proof of this point follows from the integrability of the pdf and the dominated convergence theorem). A definition of a joint pdf, $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$, for variables $\{X_1,\ldots,X_n\}$, can be found in [5], p78, with Laplace's generalisation of independence that $f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{X_1}(x_1)\ldots f_{X_n}(x_n)X_n$ appearing on p81 (Theorem 6B). A nonstandard formulation of independence, with its relation to the standard definition, can be found in [8], chapter 7, Definition 7.1 and Theorem 7.2, see also [1]

[5]Laplace construes the upper integral in the text as varying indefinitely between limits in the interval $[0,1]$

cet evenement, $\frac{\int x^p(1-x)^q dx}{\int x^p(1-x)^q dx}$ sera, par l'article XV, la probabitite de cette cause, ... ", ([6]).

Laplace use his method of finite differences to calculate $\int x^p(1-x)^q dx$, comparing the two expressions (λ), (p424), and (γ), (p433);

"Cette suite est, dans les differences tinies, ce qu'est la suite (λ) de l'article XVIII dans les differences infiniment petites."

In his "Mmoire sur divers points dAnalyse", of 1809, (Journal de lcol Polytechnique, Tome VIII., 229-265, 1809, Oeuvres Completes, XIV, 178-214), which we considered above, Laplace develops the calculus of generating functions in Section 1, "Sur le calcul des fonctions generatrices". These represent functions as the coefficients of power series, see Chapter 5, and are required for his work in probability. He makes the important observation on p187;

"Soit u une fonction des deux quantites t et t', et concevons qu'en la developpant dans une serie ordonnee par rapport aux puissances de t et de t', $y_{x,x'}$ soit le coefficient de $t^x t'^{x'}$ dans cette serie; u sera la fonction generatrice de $y_{x,x'}$; $u[(\frac{1}{t}-1)^2-(\frac{1}{t'}-1)]$ sera la fonction generatrice de $\Delta^2 y_{x,x'} - \Delta' y_{x,x'}$, la caracteristique Δ etant relative a la variable x, et la caracteristique Δ' a la variable x'."

relating generating functions to the method of solving the heat equation, by finite differences.

At about the same time, Laplace finished his "Memoire les Approximations des Formules qui sont Fonctions de Tres Grands Nombres et sur leur Application aux Probabilites", (Memoires de lInstitut de France, 1st Series, T. X, year 1809, (1810), pp. 383389. Oeuvres completes XII, pp. 301353), in which he proves the Central Limit Theorem. Laplace introduces the idea of a characteristic function for a discrete random variable, in Section 6, p322, observing that the corresponding function for n independent variables, can be obtained by

[6]Formally, we can generalise the principle of Laplace's first paper to continuous distributions;

$p(x|y)p(y) = p(y|x)p(x)$ (Bayes's Theorem)

$p(y) = \int p(y|x)p(x)dx$ (Conditioning on Events)

to obtain $p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)dx}$. Then, letting $p(y,z|x)$ denote the probability of there being y boys and z girls, in a sample of size $p+q$, given the probability x of a boy, $p(x)$ the distribution of x, and $p(x|y,z)$ the probability of being a boy, given y boys and z girls, in a sample of size $p+q$, we have that;

$p(p,q|x) = \lambda x^p(1-x)^q$, where ($\lambda = C_p^{p+q}$), $p(x) = 1_{[0,1]}$

and

$p(p,q) = \int_0^1 \lambda x^p(1-x)^q dx$. Then;

$p(x|p,q) = \frac{p(p,q|x)p(x)}{\int_0^1 p(p,q|x)p(x)dx} = \frac{\lambda x^p(1-x)^q}{\int_0^1 \lambda x^p(1-x)^q dx} = \frac{x^p(1-x)^q}{\int_0^1 x^p(1-x)^q dx}$

taking the $n'th$ power;

"Cela pose, representons par $\phi(\frac{s}{i+i'})$ la probabilite de l'erreur s pour chaque observation, et considerons la fonction

$$\phi(\tfrac{-i}{i+i'})e^{-i\omega\sqrt{-1}} + \phi(\tfrac{-(i-1)}{i+i'})e^{-(i-1)\omega\sqrt{-1}} + \ldots \phi(\tfrac{0}{i+i'}) + \ldots + \phi(\tfrac{i'-1}{i+i'})e^{(i'-1)\omega\sqrt{-1}} + \phi(\tfrac{i'}{i+i'})e^{i'\omega\sqrt{-1}}$$

En elevant cette fonction a la puissance n, le coefficient de $e^{r\omega\sqrt{-1}}$ dans le developpement de cette puissance sera la probabilite que la somme des erreurs de n observations sera r, ... ", ([7]). Using the infinitesimal method, he can also use this definition for a continuous random variable, defining, on p323, the nonstandard versions $\frac{q'}{h}$ and k' of the expectation and variance;

"q' est la abcisse correspondante a l'ordonnee du centre de gravite de l'aire de la courbe"

...

$\frac{k'=\int \frac{(x'-q')^2}{h^2}\phi(x}{h)}$", ([8]), and, obtaining an expression for the characteristic function on $p324$;

"$e^{-\frac{k'}{2k}nt^2}(1 + Ant^4 + \ldots)$"

On $p324$, Laplace uses the inversion theorem, which, in the previous section, I have suggested was known by him at this time;

"Si, conformement a l'analyse de l'article IV, on multiplie la fonction (o) par $2cosl\omega$, le terme independant de ω dans le produit exprimera la probabilite que la somme des erreurs sera ou $nq - l$ ou $nq + l$"

to obtain the probability distribution of the sum $X_1 + \ldots + X_n$, and, then, on p325, finds the normal probability distribution of $\frac{X_1+\ldots+X_n}{\sqrt{n}}$, for infinite n;

"Si l'on multiplie cette fonction par dl, en integrant on aura la probabilite que la somme des erreurs sera comprise dans les limites $nq \pm l$ ou $nq \pm (i + i')r\sqrt{n}$; or on a;

$dl = (i + i')dr\sqrt{n}$;

cette probabilite sera donc

[7] Formally, the characteristic function of a random variable X, is given by $c_X(t) = E(e^{itX})$, where E denotes the expectation or mean value. For n independent random variables, with the same probability distribution, $c_{X_1+\ldots+X_n}(t) = E(e^{it(X_1+\ldots X_n)}) = [E(e^{itX})]^n$, as $E(Y_1 \ldots Y_n) = E(Y_1) \cdot \ldots \cdot E(Y_n)$, for n independent random variables $\{Y_1, \ldots, Y_n\}$. The corresponding property for n independent nonstandard random variables can be found in Lemma 3 of [9].

[8] The nonstandard definitions of expectation and variance can again be found in Lemma 3 of [9]

$\frac{2}{\sqrt{\pi}}\sqrt{\frac{k}{2k'}}\int e^{\frac{-k}{2k'r^2}}dr$", ([9]) Laplace announced his result to the Academie in the same year.

In the same paper, Section 1, p305, Laplace considers the question of of the probability distribution of the average inclination of n independent satellite orbits, defining the distribution of the independent sum, ([10]);

"et l'on demande la probabilite que l'inclinaison moyenne de n orbites sera comprise dans des limites donnees"

...

Cela pose, nommons $t, t_1, t_2, \ldots$ les inclinations des n orbites, et supposons leur somme egale a s, nous aurons

$t + t_1 + \ldots t_{n-1} = s$".

After deriving the formula (a), on p307, for this defined distribution, Laplace recognises some connection with the generatrice functions used in his "Mmoire sur divers points dAnalyse", and obtains the probability distribution, on p317 for the orbit inclinations, presumably by solving the associated nonstandard difference heat equation;

"Si l'on neglige les termes de l'ordre $\frac{1}{n}$, l'integrale $\frac{2}{\sqrt{\pi}}\int dse^{-t^2}$ ou $\frac{3}{\sqrt{\pi}}\sqrt{\frac{3}{2}}\int dre^{\frac{-3}{2}}r^2$ exprime la probabilite que la somme des inclinations des orbites sera comprise dans les limites $\frac{h}{2}-\frac{rh}{2\sqrt{n}}$ et $\frac{h}{2}+\frac{rh}{2\sqrt{n}}$."

a particular case of his later general formula in the central limit theorem. Laplace is, here, and in his proof of the central limit theorem, generalising the result that Brownian motion can be obtain from a random walk, with coin tossings as the individual steps, ([11]). Figure 23 clarifies this idea. The first part of the diagram illustrates a possible path $R(x_0, t)$, obtained as the "average" of discrete independent observations $\{\omega_1, \ldots, \omega_n\}$ over time, the randomness of the motion resulting from the assumption of independence. For discrete times t_i, with a fixed number of observations, the possible values $R(x, t_i)$, corresponding to the final values of possible paths of duration t_i, display an increasingly random and irregular pattern, but, as each segment splits over successive times, the total area is preserved; in this sense, there is one rather than two degrees of freedom, for the partition, at each step. The second part shows the continuous analogue of this process, known as Brownian motion, obtained over continuous time t, after using infinitesimally small time steps, with a possible paths $W(x_0, t)$ and time slices $W(x, t_i)$. An important property is that "almost every" path

[9] Laplace's nonstandard proof of the central limit theorem is very similar to the accepted standard proof, see p130 of [5]. A statement of the nonstandard version is given in Theorem 7.4 of [8], see also [1], though its proof uses the standard formulation of the result.

[10] Here, as in the proof of the central limit theorem, Laplace construes the average of a sample of n elements, by dividing through with $\sqrt{n}$

[11] See Definition 7.7 of [8]. A nonstandard proof of this last result can be found in [8], Theorem 7.8, based on [1].

is continuous but nowhere differentiable. The averaging behaviour continues, the defining charateristic of a more general class of processes called martingales, ([12]). These, in fact, originated as betting strategies in France, during the 18th century, ([13]), and it seems reasonable that Laplace knew of their existence, ([14]). The fact that martingales, arising from gambling, are equivalent to processes with the above averaging behaviour follows from Theorem 0.5 and Theorem 0.13 of [9];

Theorem 0.1. *Any martingale X is representable as a stochastic integral, and, conversely, every stochastic integral is equivalent to a martingale.*

Proof: The idea is to lift a standard martingale X to a nonstandard martingale $\overline{X}$, and show that for infinite numbers $\{\eta,\nu\}$, with $\eta = 2^{\nu}$;

$$\overline{X}_t(x) = \sum_{j=0}^{[\nu t]} c_j(t,x)\omega_j(x)$$

where the functions c_j depend only on the previous information. Under certain technical restrictions, one can then show that the right hand side sum corresponds to a stochastic integral, defined in terms of Brownian motion, rather than the steps of a random walk. The converse involves showing that such sums have the required averaging behaviour, using the fact that $E(\omega_j|F_{j-1}) = 0$, for F_{j-1} representing the past information; that is each ω_j is obtained as an innovation from 0 at time t_{j-1}. Observe the similarity with Theorems 1.5 and 1.6 from the previous section, ([15]).

Laplace's alternative derivation of the distribution, using the heat equation, is a precursor to a general method of finding the (time-evolving) probability distributions of martingales, or more complex processes, ([16]), and investigation of his papers could still lead to interesting new insights into the subject. Indeed, in the subsequent "Memoire sur les Integrales Definies

[12]See Footnote 2 and Definition 0.7 of [9] for the standard and nonstandard definitions

[13]Based on a fair game with independent, identically distributed random variables, $\{\omega_1,\ldots\omega_n\}$, a martingale strategy is to bet on ω_n at time n as a function e_n of the previous results $\{\omega_1,\ldots,\omega_{n-1}\}$, the payoff $Y_n = e_1\omega_1 + e_2(\omega_1)\omega_2 + \ldots + e_n(\omega_1,\ldots,\omega_{n-1})\omega_n$ is known as a martingale. The continuous analogues of such payoffs are known as stochastic integrals, denoted by $Y_t = \int e(t,x)dW_t$

[14]Indeed, in his later "Essai Philosophique sur les Probabilites", (1814), within the chapter "Applications du Calcul des Probabilites Des Jeux", (p40), Laplace formulates the idea behind martingale strategies for fair games;

"Que l'on jette dans une urne, cent numeros depuis un jusqu'a cent, dans l'ordre de la numeration, et qu'apres avoir agite l'urne, pour meler ces numeros, on et tire un; il est clair que si le melange a ete bien fait, les probabilites de sorte des numeros, seront les memes. Mais si l'on craint qu'il n'y ait entre elles, de petites differences dependentes de l'ordre suivant lequel les numeros ont ete jetes dans l'urne; on diminuera considerablement ces differences, en jetant dans une seconde urne, ces numeros suivant leur ordre de sortie de la premier urne, et en agitant ensuite, cette seconde, pour meler ces numeros. Une troisieme unrne, une quatrieme, etc., diminueraient de plus ces differences, deja insensibles dans la seconde urne"

[15]A standard proof of this result can be found in [14], for other general results on martingales, see [16]. Viewing an algebraic curve as a probability space, one can also hope to generalise the theorem to this setting.

[16]The Fokker-Planck Theorem gives a partial differential equation $\frac{\partial p}{\partial t} = \frac{\partial^2 Dp}{\partial x^2}$ for the probability distributions $p(x,t)$ of random variables X_t, satisfying $X_t = \int_0^t \sqrt{2D(X_s,s)}dW_s$, usually abbreviated as a "stochastic differential equation", $dX_t = \sqrt{2D(X_t,t)}dW_t$. In the case that $D = 1$, we obtain the heat equation for the probability distribution of rescaled Brownian motion, as $\sqrt{2}\int_0^t dW_s = \sqrt{2}W_t$. In certain cases, when the

et leur Application aux Probabilites", (Memoires de lAcademie des Sciences, 1810 (1811), Ouevres Completes XII,pp357-412), Section 5, Laplace succesfully finds a new differential equation for the following probability distribution;

"Considerons deux urnes A et B, renfermant chacune le memenombre n de boules; et supposons que, dans le nombre total $2n$ de boules, il y en ait autant de blanches que de noires. Concevons que l'on tire en meme temps une boule de chaque urne, et qu'ensuito on mette dans une urne la boule extrait de l'autre. Supposons que l'on repete cette operation un nombre quelconque r de fois, en agitant a chaque fois les urnes pour en bien mcler les boules: et chercons la probabilite qu'apres ce nombre r d'operations il y aura x boules blanches dans l'urne A. Soit $z_{x,r}$ cette probahilite;", ([17]).

After obtaining a difference equation for this probability, he claims;

" ... $z_{x,r} = U$;

the preceding equation in the partial differences will become, by neglecting the terms of order $\frac{1}{n^2}$,

$\frac{\partial U}{\partial r'} = 2U + 2\mu\frac{\partial U}{\partial \mu} + \frac{\partial^2 U}{\partial \mu^2}$."

Laplace immediately followed his work on the Central Limit Theorem, with the "Supplement au Memoire sur les Approximations des Formules qui sont Fonctions Dde Tres Grands Nombres", (Memoires de lAcademie des Sciences, 1st Series, T. X, 1809, (1810), Oeuvres Completes T. XII. pp. 349353). In this short work, Laplace begins, on p351, by formulating a principle of maximum likelihood;

"Soient donc l la distance du point qu'il faut choisir a l'origine de la courbe des probabilites, et z l'abscisse correspondante a y et comptee de la meme origine; le produit de chaque erreur par sa probabilite, abstraction faite du signe, sera $(l - z)y$, depuis $z == 0$ jusqu'a $z == l$, et ce produit sera $(z - l)y$ depuis $z == l$ jusqu'a l'extremite de la courbe. On aura donc

integrand of a martingale representation is a function of the original martingale, or, when a stochastic equation of the above form is given, the theorem provides a method of calculating the required distributions. In formulating a nonstandard analogue of Fokker-Planck, one just needs to follow the method of [1] and [8], Definition 7.20, Remarks 7.21, in giving a nonstandard version of the stochastic integral, and replacing the differential equation with a difference equation. The reader might be interested in deciding whether the nonstandard formulation is true. An alternative nonstandard method of solving SDE's directly, in terms of the random variable, again generalising Laplace's finite difference method, can be found in [6].

[17]"We consider two urns A and B, each containing the same number n of balls; and we suppose that, in the total number $2n$ of balls, there are as many of them white as of black. We imagine that we draw at the same time a ball from each urn, and that next we put into one urn the ball extracted from the other. We suppose that we repeat this operation any number r times, by agitating at each time the urns in order to well mix the balls; and we seek the probability that after this number r of operations there will be x white balls in urn A. Let $z_{x,r}$ be this probability;"

$\int(l-z)ydz+\int(z-l)ydz$

pour la somme de tous ces produits, la premiere integrale etant prise depuis z nul jusqu'a $z=l$, et la seconde etant prise depuis $z==l$ jusqu'a la derniere valeur de z. En differentiant la somme precedente par rapport a l, il est facile de s'assurer que l'on aura

$dl\int ydz - dl\int ydz$

pour cette diffcrcntielle, qui doit etre nulle dans le cas du minimum; on adonc alors

$\int ydz=\int ydz$", ([18]). Laplace formulates the likelihood function for n independent observations on p352, using his result from above;

"Dans le cas present, on a, en faisant $x=X+z$,

$y=pp'p''\dots e^{-p^2\pi_1 X+z_1^2-p'^2\pi q-X-z_1^2-p''^2\pi q'-X-z_1^2-\dots}$

p etant egal a $\frac{a\sqrt{n}}{\sqrt{pi}}$, et par consequent exprimant la plus grande probabilite du resultat donne par les observations n; p' exprime pareillement la plus grande ordonnee relative aux observations n', et ainsi du reste: r pouvant, sans erreur sensible, s'etendre depuis -∞ jusqu'a $+\infty$, comme on l'a vu dans l'article VII du memoire cite, on peut prendre dans les memes limites, et alors si l'on choisit X de maniere que la premiere puissance de z disparaisse de l'exposant de c, l'ordonnee y correspondante a z nul divisera l'aire de la courbe en deux parties egales, et sera en meme temps la plus grande ordonnee. En effet, on a, dans ce cas,

$X=\frac{p'^2q+p''^2q'+\dots"}{p^2+p'^2+p''^2+\dots}$.

Laplace then finds the maximum of this function, to obtain a parameter estimate X, as a function of the independent observations $\{q,q',\dots,q_i,\dots,q_n:1\leq i\leq n\}$, ([19]). On p353, Laplace makes the observation;

---

[18] Laplace minimises the error, relative to the unknown true mean l_0, of a set of observations $\{z_i:1\leq i\leq n\}$, (corresponding to independent random variables $\{Y_i:1\leq i\leq n\}$), with a known probability distribution $y(z,l)$. Laplace maximizes the function $f(l)=\int_{z\leq l}(l-z)ydz+\int_{z\geq l}(z-l)ydz$, obtaining the equation $\int_{z\leq l}ydz=\int_{z\geq l}ydz$, to estimate l_0. It is important to realise that the estimators $(\hat{l})_n$ should be functions of the random variables $\{Y_i:1\leq i\leq n\}$; some work is required to show this, here, interpreting the integral as a sum over the observations. An estimator is said to be unbiased if $E(\hat{l})=l_0$, and consistent if $lim_{n\rightarrow\infty}(\hat{l})_n=l_0$; Laplace may be trying to estimate the median rather the mode here, and, for a certain class of distributions, these properties may hold.

[19] Laplace expresses the the likelihood $y_n(q_1,\dots,q_n,l)=\prod_{i=1}^n y(q_i,l)$, and solves $\frac{dy_n}{dl}=0$, to obtain the estimator X. A more efficient, but equivalent, method, is to solve $\frac{dlog(y_n)}{dl}=0$, or, equivalently, $\frac{d}{dl}(\sum_{i=1}^n logy(q_i,l))=0$, to obtain the estimators $(\hat{l})_n$. Under certain assumptions on the likelihood $y(q,l)$, the crucial one being that it has a unique local maximum, uniformly in the observation q, the estimators are consistent, see footnote 18 and [15], but, are not necessarily unbiased. The proof of consistency in [15] relies on the Strong Law of Large Numbers, a particular case of the Ergodic Theorem, see [12], and [8] for a nonstandard proof.

"La valeur precedente de X est celle qui rend un minimum la fonction

$(pX)^2 + [p'(q - X)]^2] + [p''(q' - X)]^2 + \ldots$

c'est-a-dire la somme des carres des erreurs de chaque resultat, multipliee respectivement par la plus grande ordonnee de la courbe de facilite de ses erreurs."

that the above procedure is equivalent to estimating the slope of a "line of best fit" by minimising the sum of squares of the residuals. Figure 24 illustrates the idea of least squares, [20], with a collection of data points $\{(p_i, q_i) : 1 \leq i \leq n\}$, obeying the relation $q_i = (\lambda_0 p_i + c_0) + u_i$, where the slope λ_0 and intercept c_0 are fixed, and the residuals, or error terms, u_i, are drawn from a fixed distribution, with unknown variance, ([21]). If the errors terms are drawn from a normal distribution, which is assumed here, then, as Laplace points out correctly, the method of maximum likelihood yields the same estimators $\{(\hat{\lambda})_n, (\hat{c})_n\}$ for the slope and intercept, as least squares, in particularly such estimators are consistent. These estimators are also unbiased, ([22]), and, in the final paragraph, Laplace applies his central limit theorem of the previous paper, observing;

"Ainsi cette propriete, qui n'est qu'hypothetique lorsqu'on ne considere que des resultats donnes par une seule observation on par un petit nombre d'observations, devient necessaire lorsque les resultats outre lesquels on doit prendre un milieu sont donnes chacun par un tres grandes nombre d'observations, quelles que soient d'ailleurs les lois de facilite des erreurs de ces observations.

...

la probabilite que l'erreur de resultat choisi $A + X$ sera comprise dans les limites $\pm \frac{T}{\sqrt{n}}$ sera;

$\frac{2\int dt e^{-t^2}}{\sqrt{\pi}}$"

that the estimators are asymptotically normal, ([23])

[20]Originally introduced by Legendre in 1805, but not in the context of probability

[21]Minimising $\sum_{i=1}^n u_i^2$, one obtains the estimators $(\hat{\lambda})_n = \frac{s_{p,q,n}}{s_{p,n}^2}$, where $s_{p,q,n} = \frac{\sum_{i=1}^n (p_i - \overline{p}_n)(q_i - \overline{q}_n)}{n}$, $s_{p,n}^2 = \frac{\sum_{i=1}^n (p_i - \overline{p}_n)}{n}$, $\overline{p}_n = \frac{\sum_{i=1}^n p_i}{n}$, and $\overline{q}_n = \frac{\sum_{i=1}^n q_i}{n}$ and $(\hat{c})_n = \overline{q}_n - \frac{(\hat{\lambda})_n s_{q,n} \overline{p}_n}{s_{p,n}}$, where $s_{q,n}^2 = \frac{\sum_{i=1}^n (q_i - \overline{q}_n)}{n}$ which depend only on the data $\{(p_i, q_i) : 1 \leq i \leq n\}$. Laplace obtains a similar result, but his notation is still unclear.

[22]This was shown in Gauss's contemporary astronomical paper "Theoria motus corporum coelestium in sectionibus conicis solem ambientum" of 1809, in which he also derives the normal distribution. However, interestingly, the estimator obtained using maximimum likelihood for the variance of the residuals is biased, differing from the unbiased estimator by a factor of $\frac{n-1}{n}$

[23]In the sense that $lim_{n \to \infty}(\sqrt{n}(\hat{\lambda})_n)$ follows a normal distribution. This would also be true for the least squares estimators if the error terms were independently drawn from another distribution, it is not clear whether Laplace realises this.

In 1812, Laplace continued his work in probability, by publishing his "Theorie analytique des probabilites". The first part of this text essentially gives a more mature presentation of his previous work on the theory of generatrices, with some possibly new remarks on the passage from finite to infinitely small quantities, in Chapter 2, Section 19, "Considerations sur le passage du fini a l'infinement petit". In the second part, entitled a "Theorie General des Probabilites", Laplace recalls 4 central principles of probability, used in his earlier work,([24]), which establish his Bayesian philosophy, in the theory of probability. In his "Essai Philosophique sur les Probabilites" of 1814, Laplace reiterates this principles among a more comprehensive list of 10, in the chapter "Principes generaux du Calcul des Probabilites", ([25]).

As we have seen Laplace's major work in physics was his consideration of planetary and satellite motion in the "Mechanique Celeste". Undoubtedly influenced by Fourier's work on the heat equation, he also published the important "Mmoire sur les mouvements de la lumire dans les milieux diaphanes" in 1810, (Mmoires de lAcadmie des Sciences, Ist Srie, Tome X,

[24]Namely;

"I^{er} PRINCIPE. La probabilite d'un evenement compose de deux evenements simples, est la produit de la probabilite d'un de ces evenements, par la probabilite que cet evenement etant arrive, l'autre evenement aura lieu"

"II^{e} PRINCIPE. Le probabilte d'un evenement futur, tiree d'un evenement observe, est la quotient de la division de la probabilite de evenement compose de ces deux evenemens, et determinee a priori, par la probailite de l'evenement observe, determinee pareillement a priori"

"III^{e} PRINCIPE. Si un evenement observe peut resulter de n causes differentes; leurs probabilites sont respectivement, comme les probabilites de l'evenement, tirees de leur existence; et la probabilite de chacune d'elles est une fraction dont le numerateur est la probabilite de l'evenement, dans l'hypothese de l'existence de la cause, et dont le denominateur est la somme des probabilites semblables, relatives a toutes les causes"

"IV^{e} La probabilite d'un evenement futur est la somme des produits de la probabilite de chaque cause, tiree de l'evenement observe, par la probabilite de chaque cause, tiree de l'evenement observe, par la probabilite que cette cause existant, l'evenement future aura lieu"

The first two principles with two events $A$and B being that (i). $P(A)=P(A|B)P(B)$ and (ii). $P(A|B)=\frac{P(A\cap B)}{P(B)}$

The second two principles, for an event B and exclusive events $\{A_i : 1\leq i\leq n\}$, that (iii). $P(A_i|B)=\frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n}P(B|A_j)P(A_j)}$, for $(1\leq i\leq n)$ and (iv). $P(B)=\sum_{j=1}^{n}P(B|A_j)P(A_j)$.

[25]The position of Bayesianism is usually contrasted with that of frequentism, in which events are usually construed as having a well defined, predetermined probability, rather than being continually updated by new evidence, based on Bayes's theorem. For the most part, in the modern theory, these approaches seem to have been reconciled, for example the simultaneous treatment of the true value of a parameter in a distribution versus its estimates as random variables obtained by maximum likelihood. There still appear to be some remaining issues in the interpretation of hypotheses, modern Bayesianists preferring to construe these as random variables, rather than either true or false, the latter; I believe, still being the usual approach. It is an interesting point that, although this is a question mainly of interpreting evidence, rather than any paradox in mathematics, the two interpretations can lead to very different decisions on whether to accept or reject a particular hypotheses, as in, for example, "Lindley's Paradox".

Oeuvres Completes de Laplace, Vol 12, 267-298). Clearly aware of Huygen's wave theory of light in "Traite de la Lumiere", Laplace discusses the idea of refraction in crystals, and, on p287, advocating Newton's theory of light as a particle, explains this effect in terms of the reflections from the molecular surface;

"Les memes resultats ont lieu relativement aux rayons extraordinaires; car, sans connaitre la cause de la refraction extraordinaire, on peut cependant assurer qu'elle est due a des forces attractives et repulsives qui agissent de molecule a molecule, suivant des fonctions quelconques de la distance, et qui, dans les cristaux, sont modifiees par la figure de leurs molecules integrantes, par celle des molecules de la lumiere et par la maniere dont ces molecules se presentent les unes aux autres."

He goes on to derive the heat equation for the propagation of temperature by the transmission of light molecules, quoting Newton's principle;

"On est parti, dans la theorie de l'equilibre et du mouvement de la chaleur, de ce principe, donne par Newton, savoir que la chaleur communiquee par un corps a un autre qui lui est contigu est proportionnelle a la difference de leurs temperatures"

It is interesting that Laplace was unable to incorporate his ideas from probability theory in this work, although this featured heavily in his discussion of the inclinations of orbits, ([26]).

References

[1] A Non-Standard Representation for Brownian Motion and Ito Integration, Robert M. Anderson, Israel Journal of Mathematics, Vol. 25, (1976).

[2] Statistical Inference, G. Casella, R. Berger, Thomson Learning, (2002).

[3] Macroscopic and Staistical Thermodynamics, Yi-Chen Cheng, World Scientific Publishing, (2006).

[4] Thermodynamics, E. Fermi, Dover, (2000).

[5] Probability, An Introduction, Geoffrey Grimmett and Dominic Walsh, OUP, (1986).

[6] An Infinitesimal Approach to Stochastic Analysis, Jerome Keisler, AMS, (1984).

[7] Quantum Mechanics, Volume 2, Albert Messiah, Elsevier, (1961).

[8] Applications of Nonstandard Analysis to Probability Theory, Tristram de Piro, available at http://www.magneticstrix.net, (2013)

[26] The application of ideas of probability theory to random particle motion could still be an interesting research topic. The process of reversing martingales to model such effects, and as an efficient method of propagating heat, or concentration, can be found in [10]. The use of stochastic differential equations can be found in [3], and [7] has a discussion of statistical results in physics. Older approaches ,including the laws of thermodynamics, can be found in [4], [13] and [11].

[9] A Simple Proof of the Martingale Representation Theorem using Nonstandard Analysis, Tristram de Piro, available at http://www.magneticstrix.net, (2013).

[10] Electron Bunching, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[11] Non Standard Analysis and Physics, Tristram de Piro, available at http://www.magneticstrix.net, (2014).

[12] Ergodic Theory, K. Peterson, CUP, (2012).

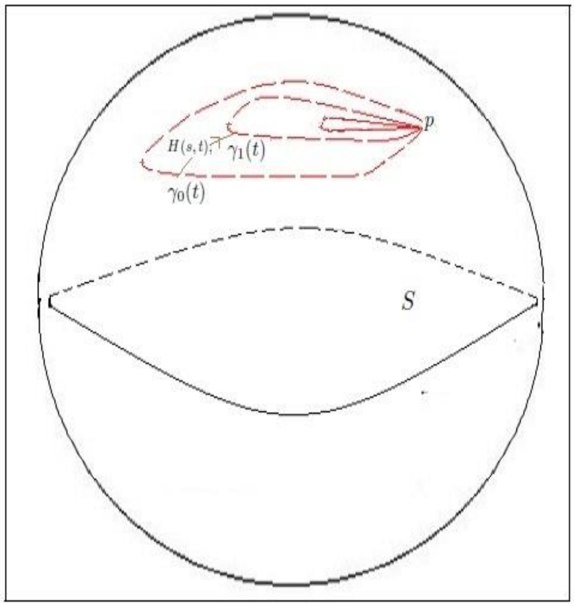
[13] Elements of Classical Thermodynamics, A.B. Pippard, CUP, (1957).

[14] Stochastic Calculus and Financial Applications, J.M. Steele, Springer, (2010).

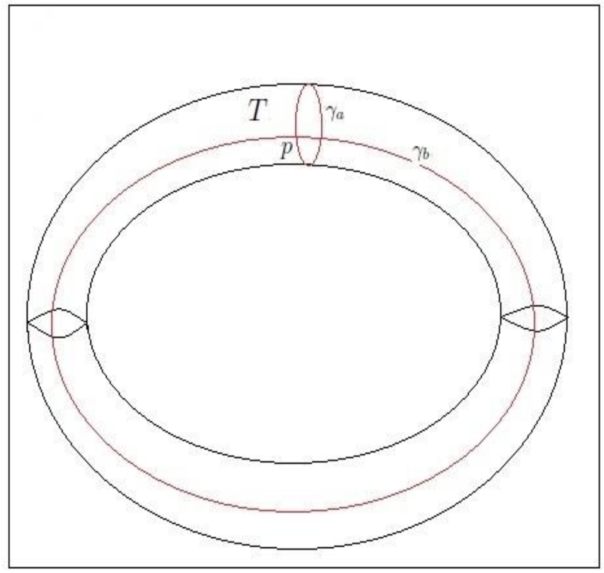
[15] Note on the Consistency of the Maximum Likelihood Estimate, Abraham Wald, The Annals of Mathematical Statistics 20, (1949) .

[16] Probability with Martingales, D. Williams, CUP, (2000).

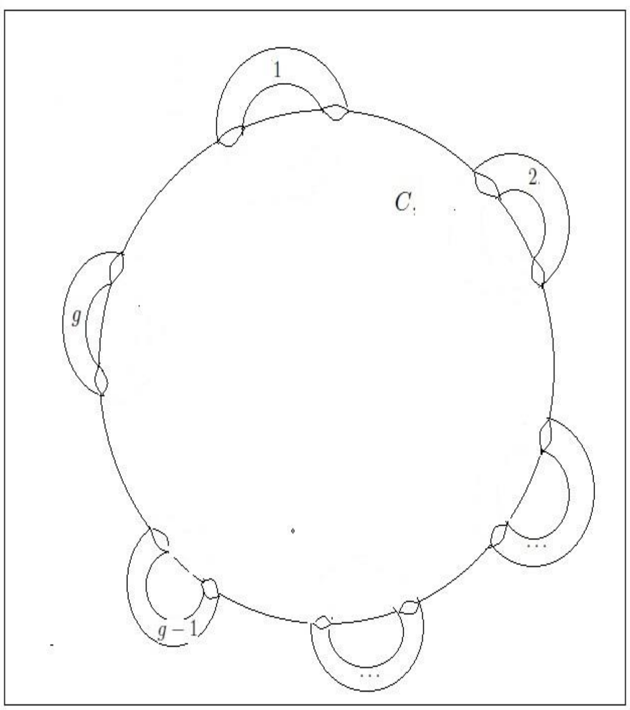
Illustrations



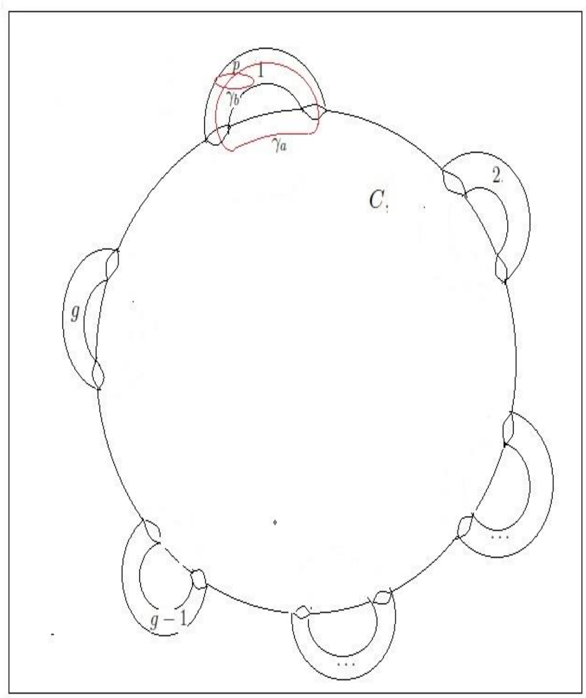
1



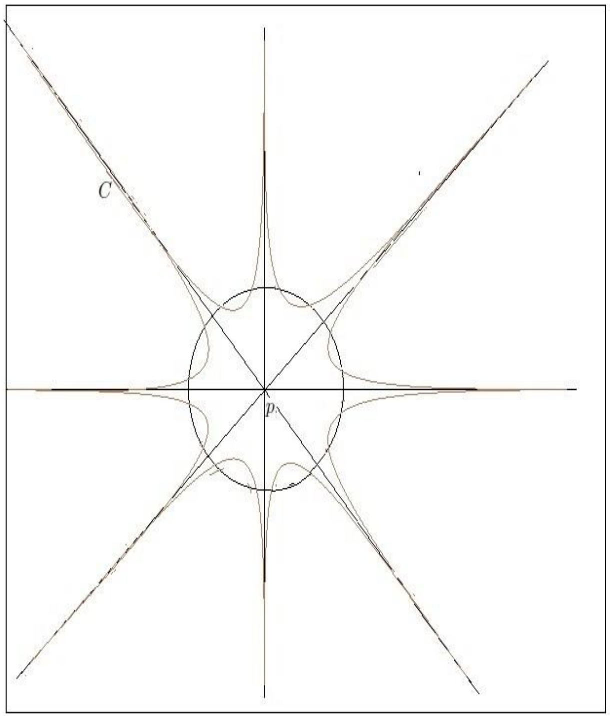
2



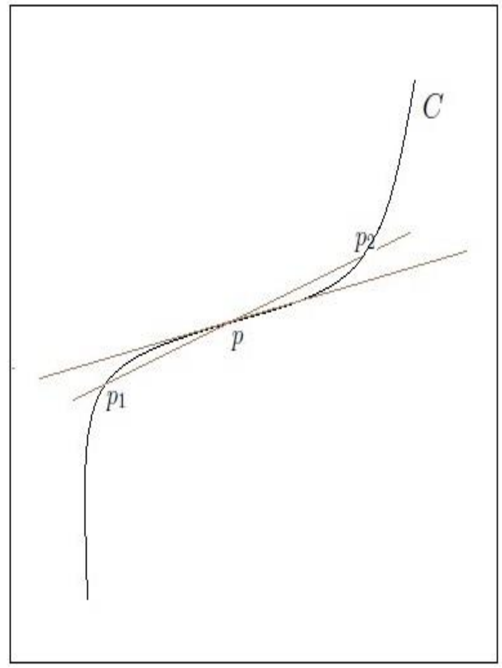
3



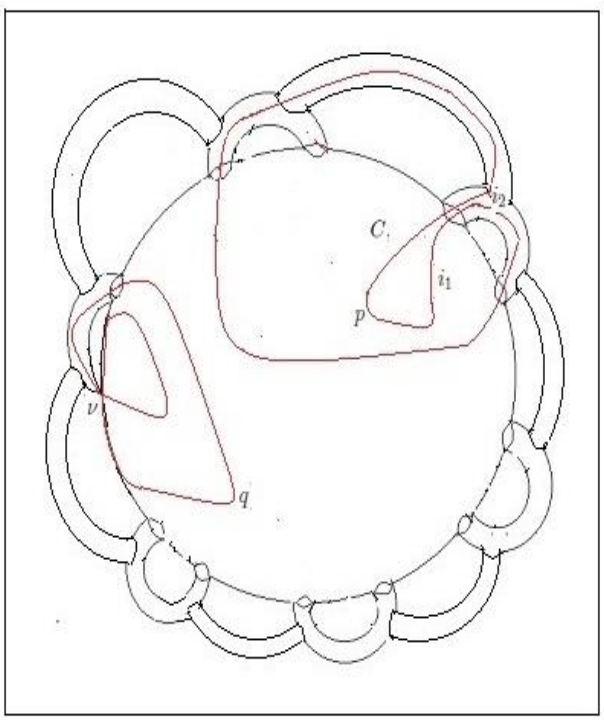
4



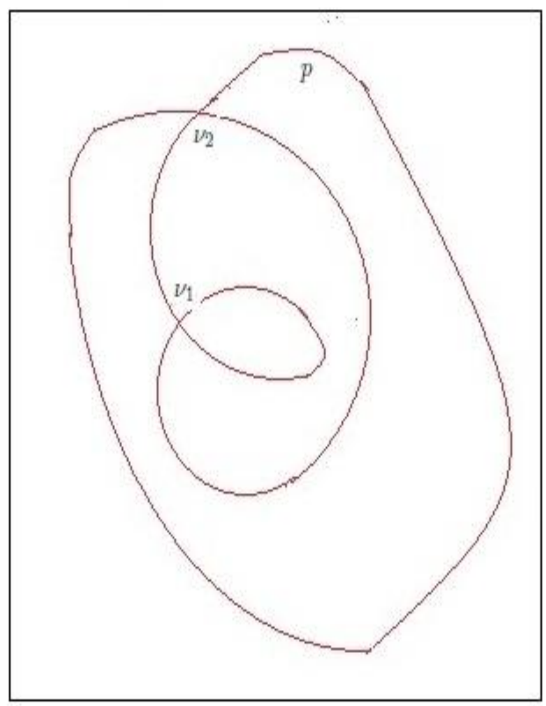
5



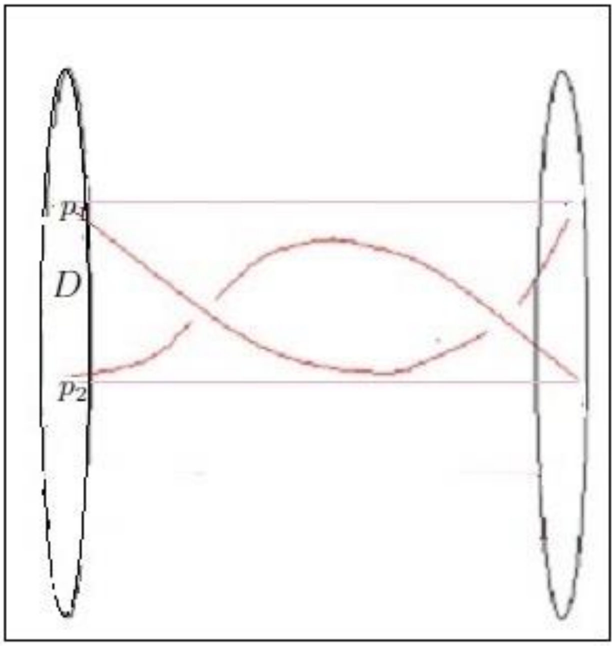
6



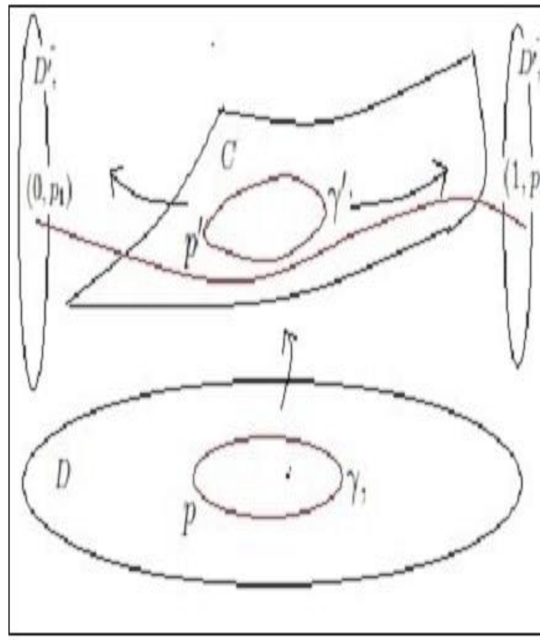
7



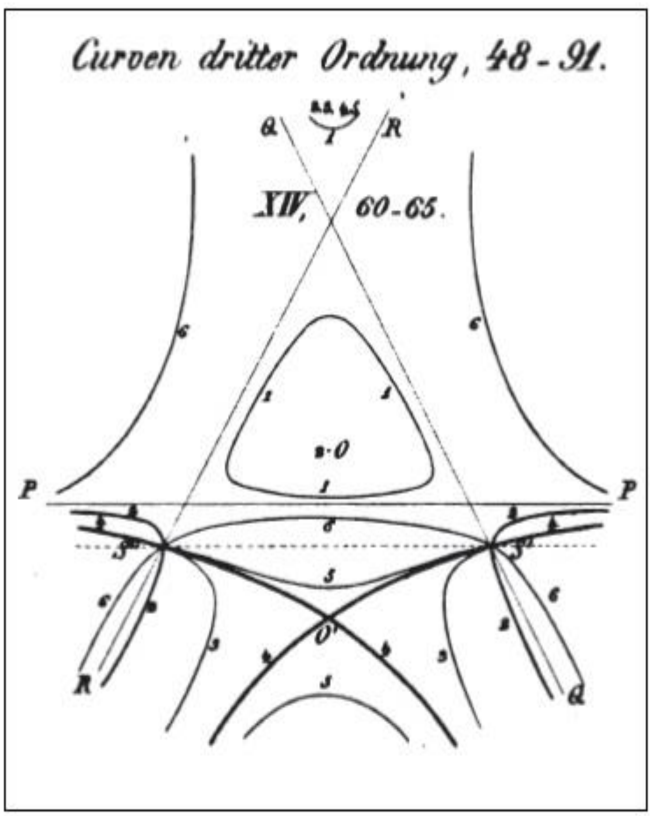
8


p1
D
p2

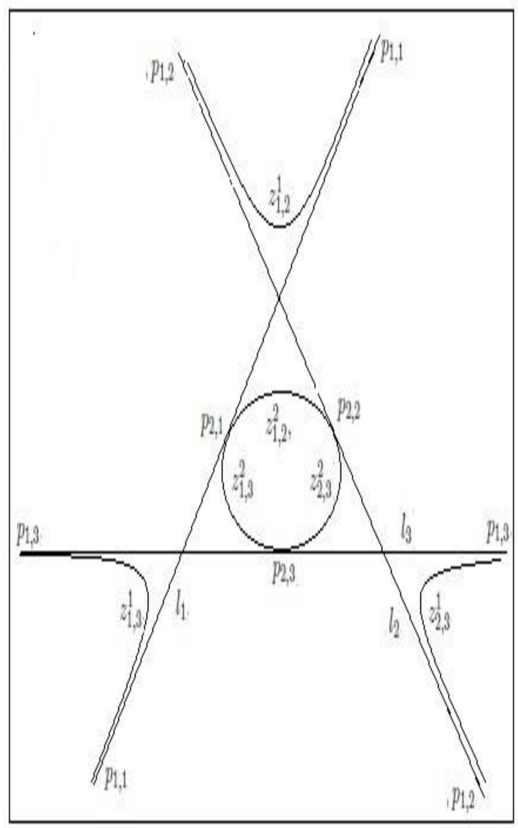
9


D'
D'
C
(0, p1)
(1, p1)
γ'
p'
D
γ1
p

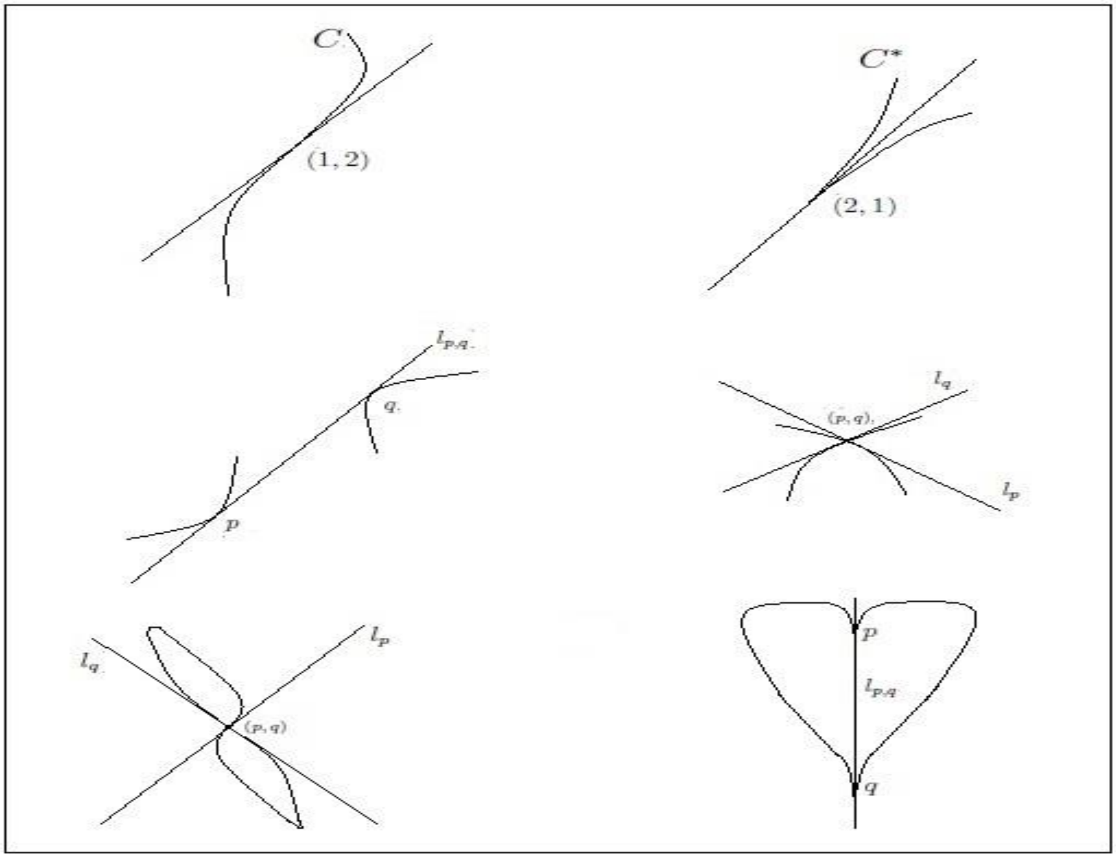
10


Curven dritter Ordnung, 48 - 91.
Q
R
XIV, 60-65.
P
P
O'
R
Q

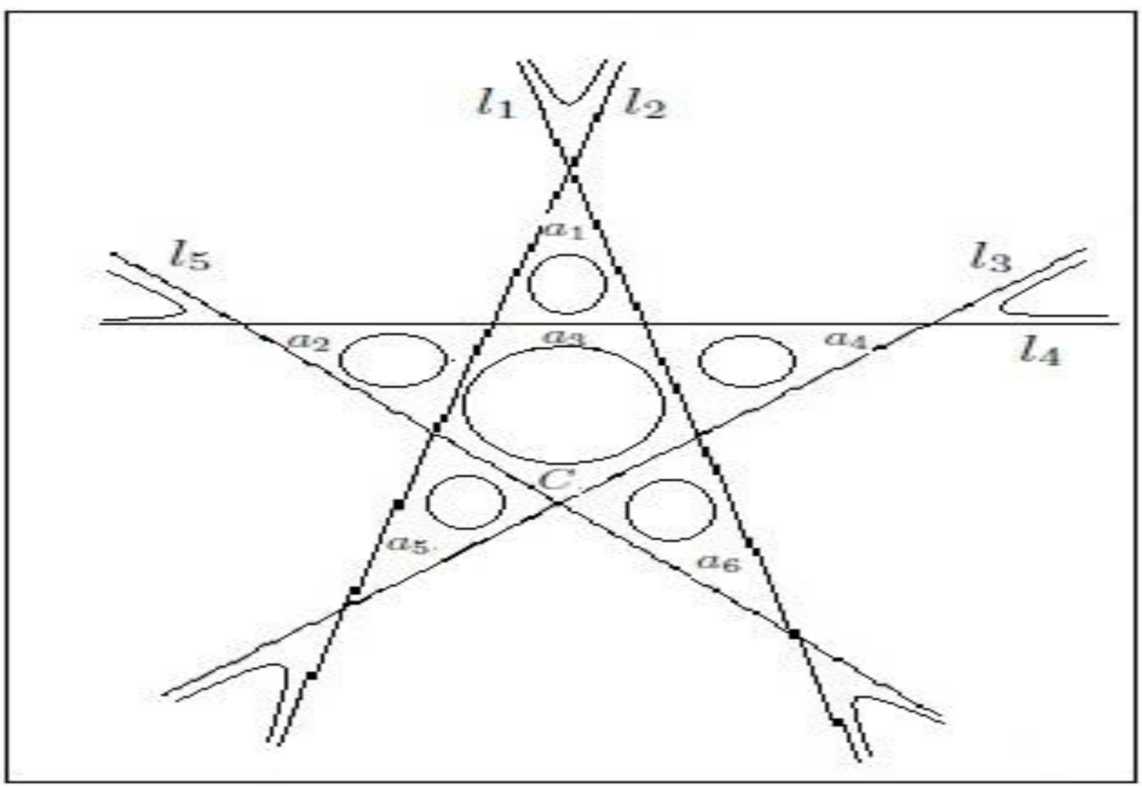
11


p1,2
p1,1
z1,2
p2,1
p2,2
p1,3
l3
p1,3
p2,3
l1
l2
p1,1
p1,2

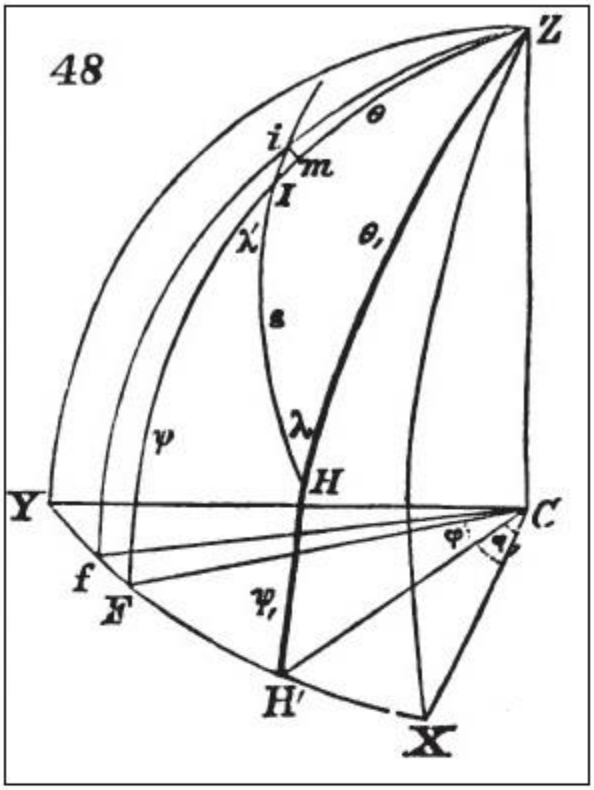
12


C
$(1,2)$
C^*
$(2,1)$
$l_{p,q}$
q
p
l_q
(p,q)
l_p
l_q
l_p
(p,q)
p
$l_{p,q}$
q

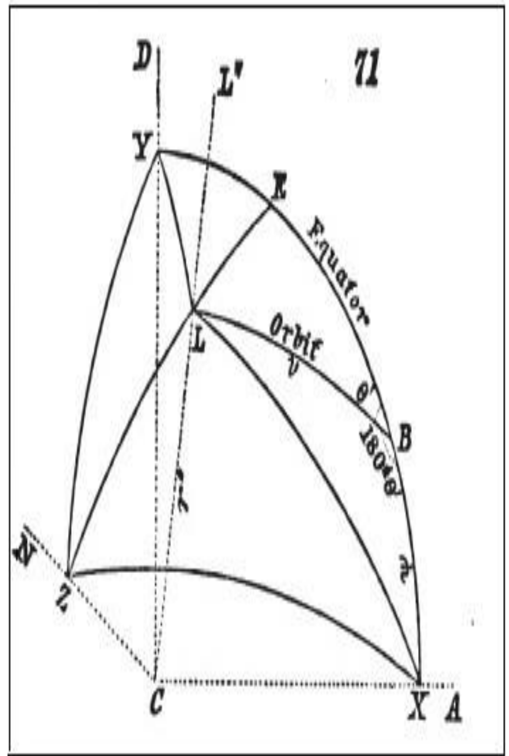
13


l_1
l_2
a_1
l_5
l_3
a_2
a_3
a_4
l_4
C
a_5
a_6

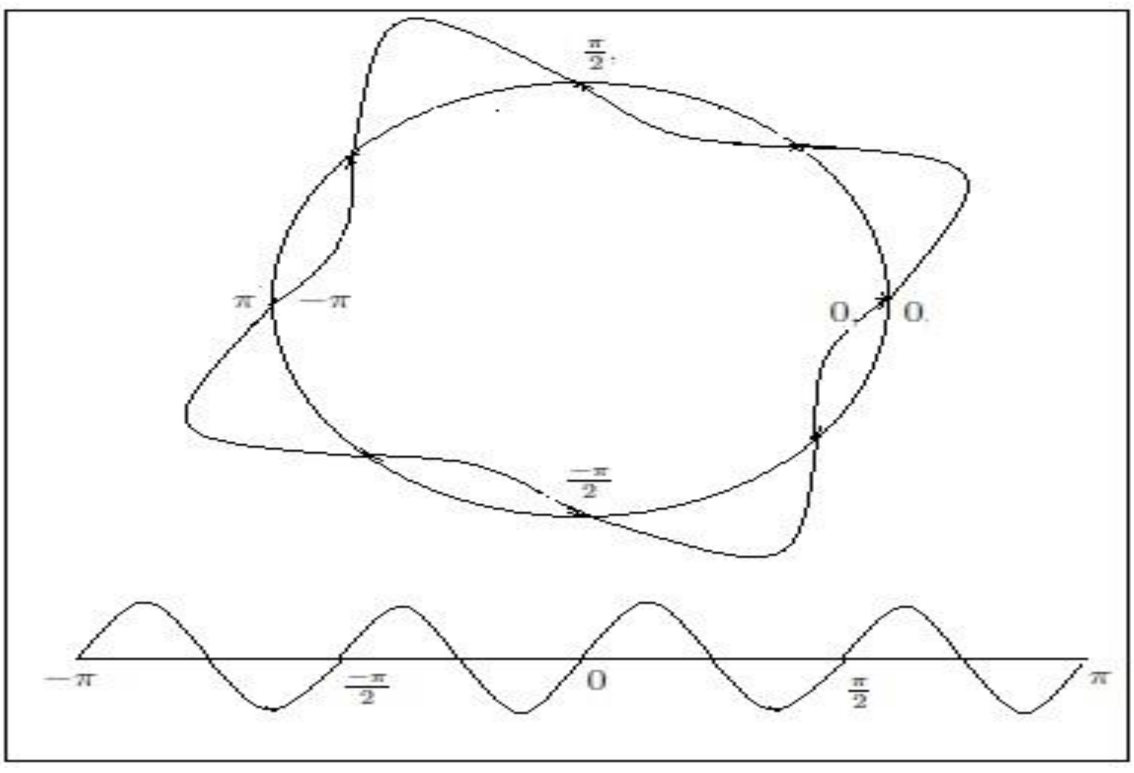
14


48
Z
i
m
θ
I
λ'
θ,
s
ψ
λ
H
Y
C
φ'
f
F
ψ,
H'
X

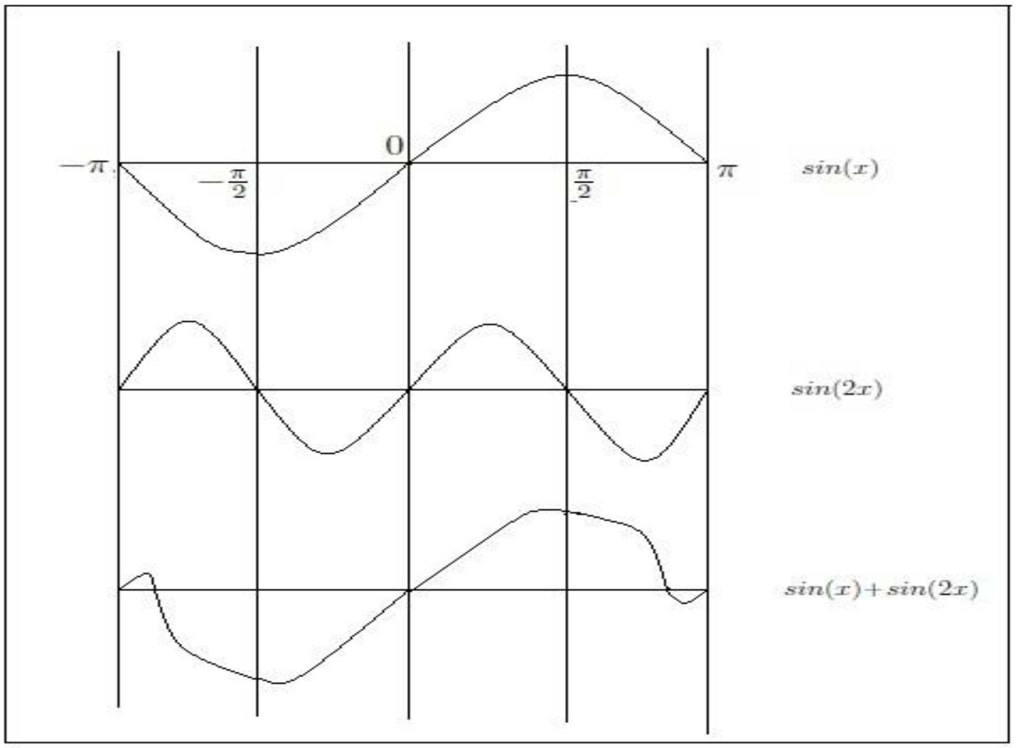
15

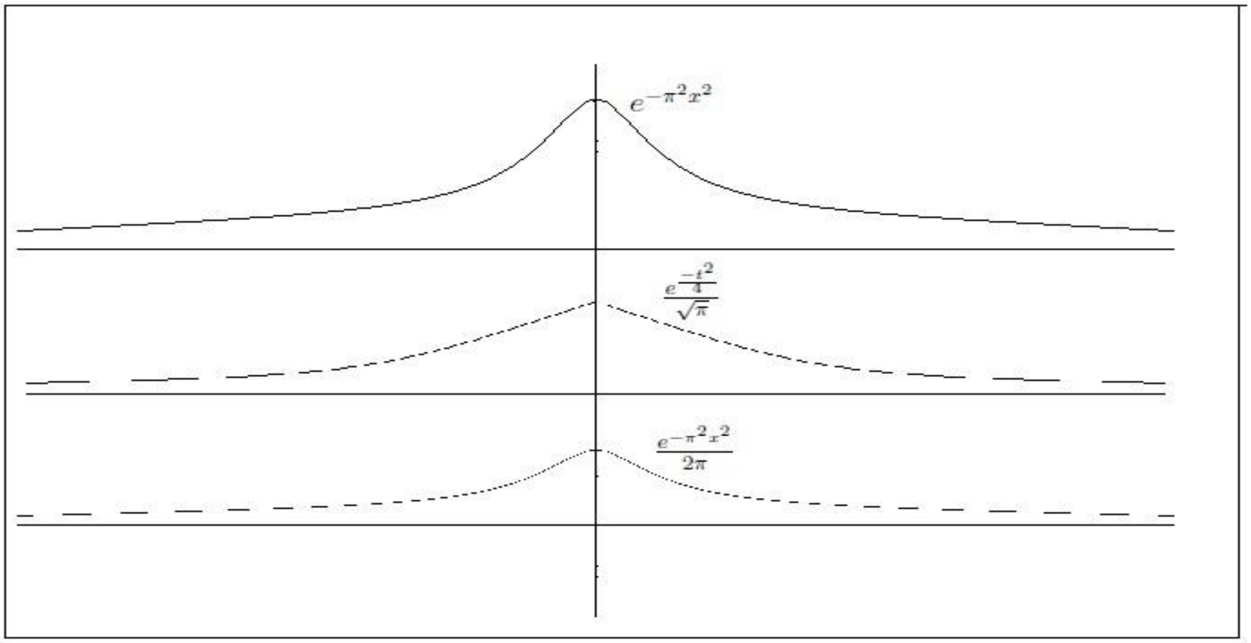

D
L'
71
Y
K
Equator
L
Orbit
ν
θ'
B
180°-θ'
N
Z
C
X
A

16

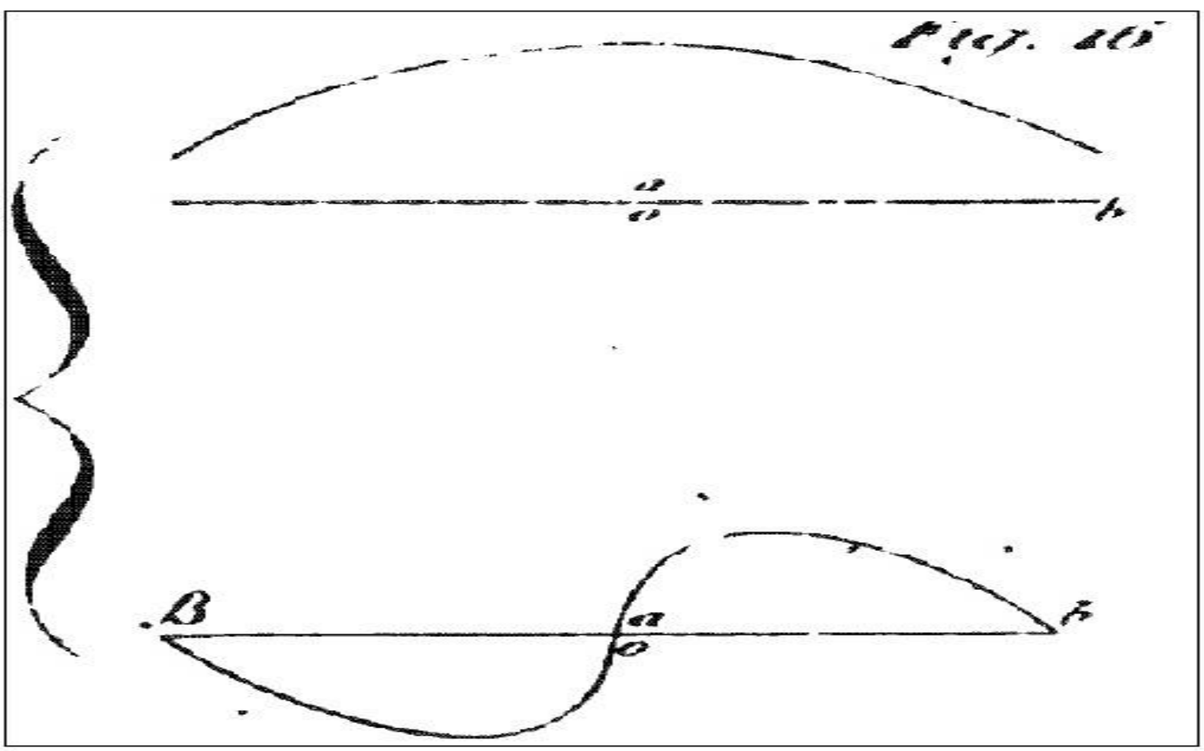

π/2
π
−π
O
O
−π/2
−π
−π/2
O
π/2
π

17

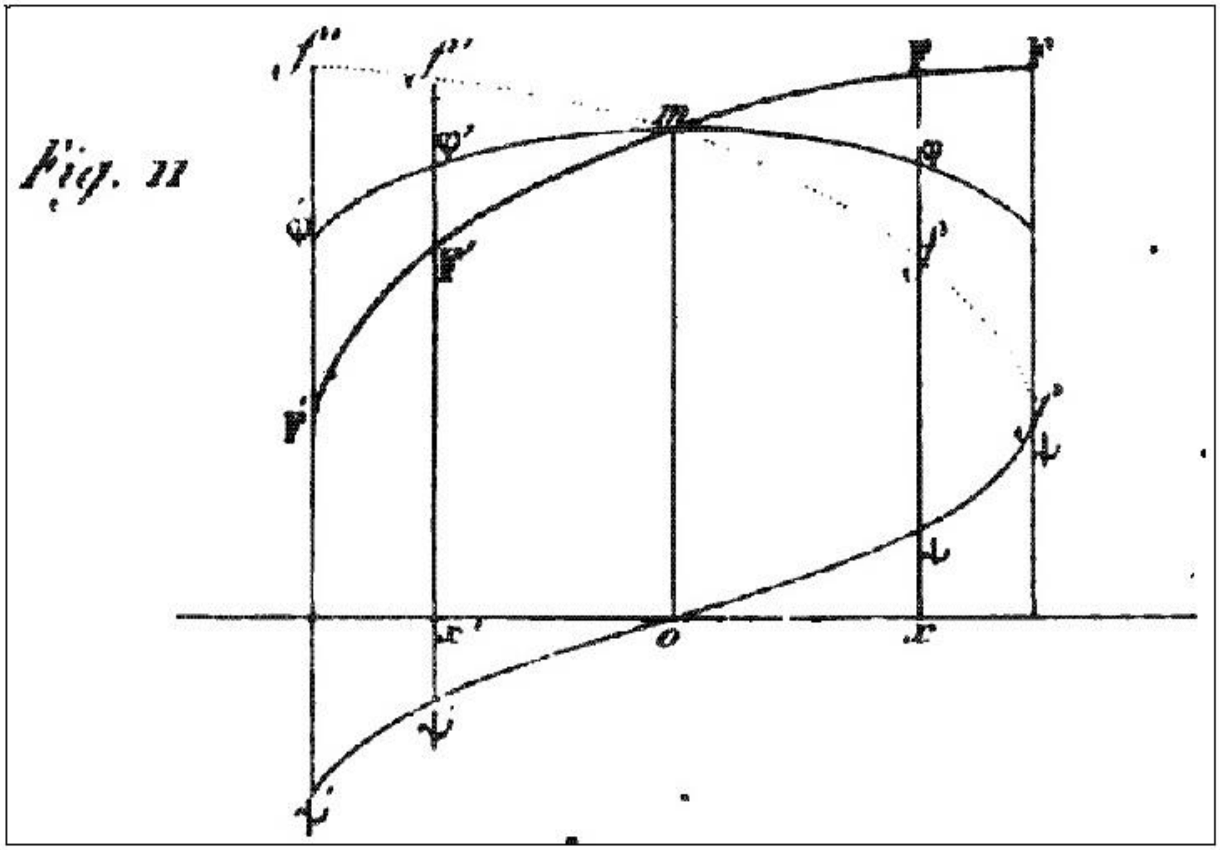

0
$-\pi$
$-\frac{\pi}{2}$
$\frac{\pi}{2}$
π
$sin(x)$
$sin(2x)$
$sin(x)+sin(2x)$



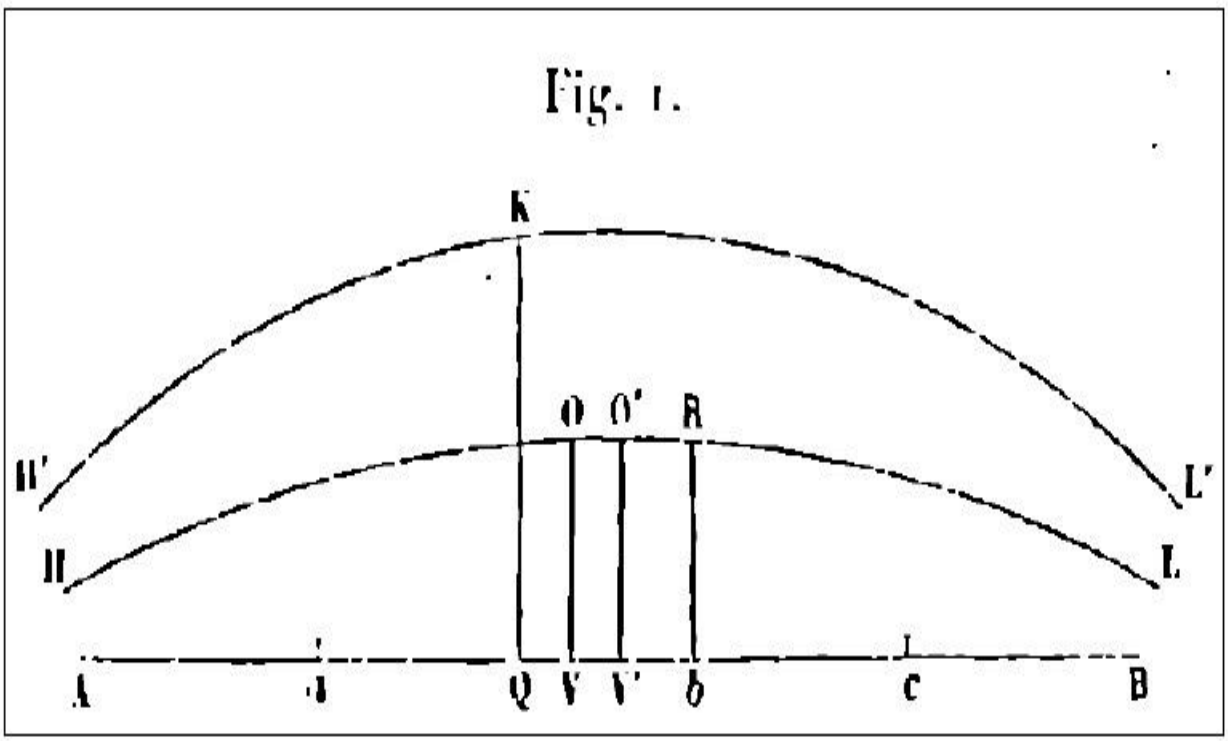
19



20



21



22

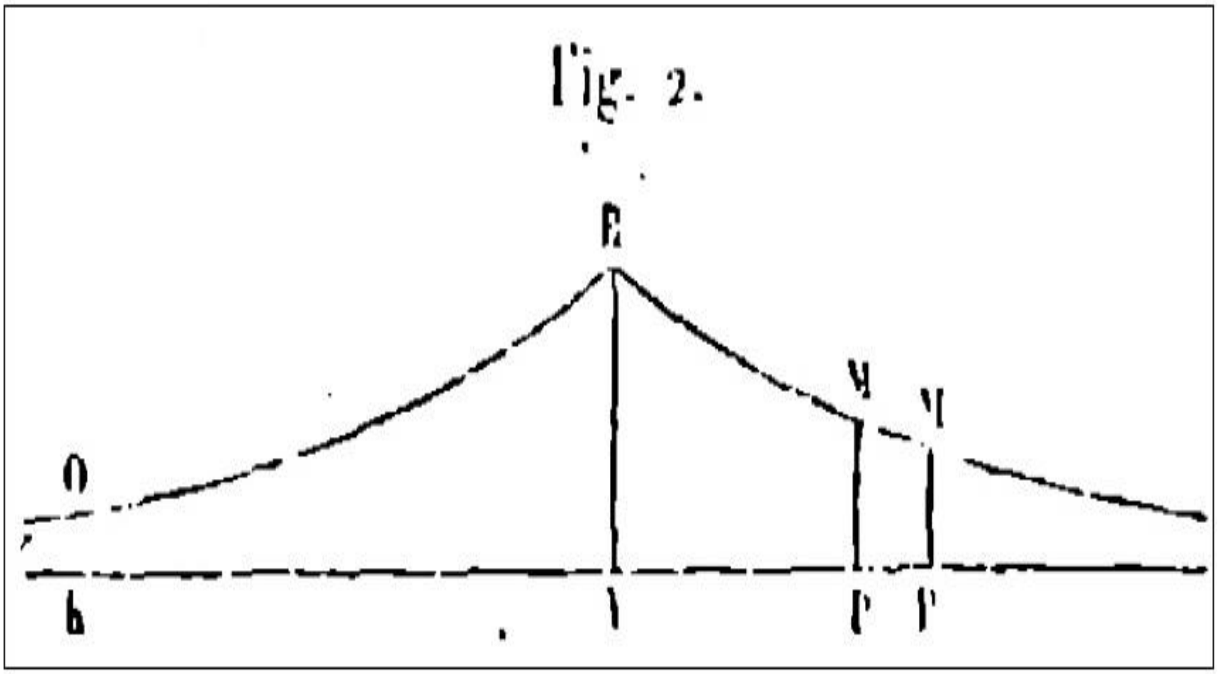
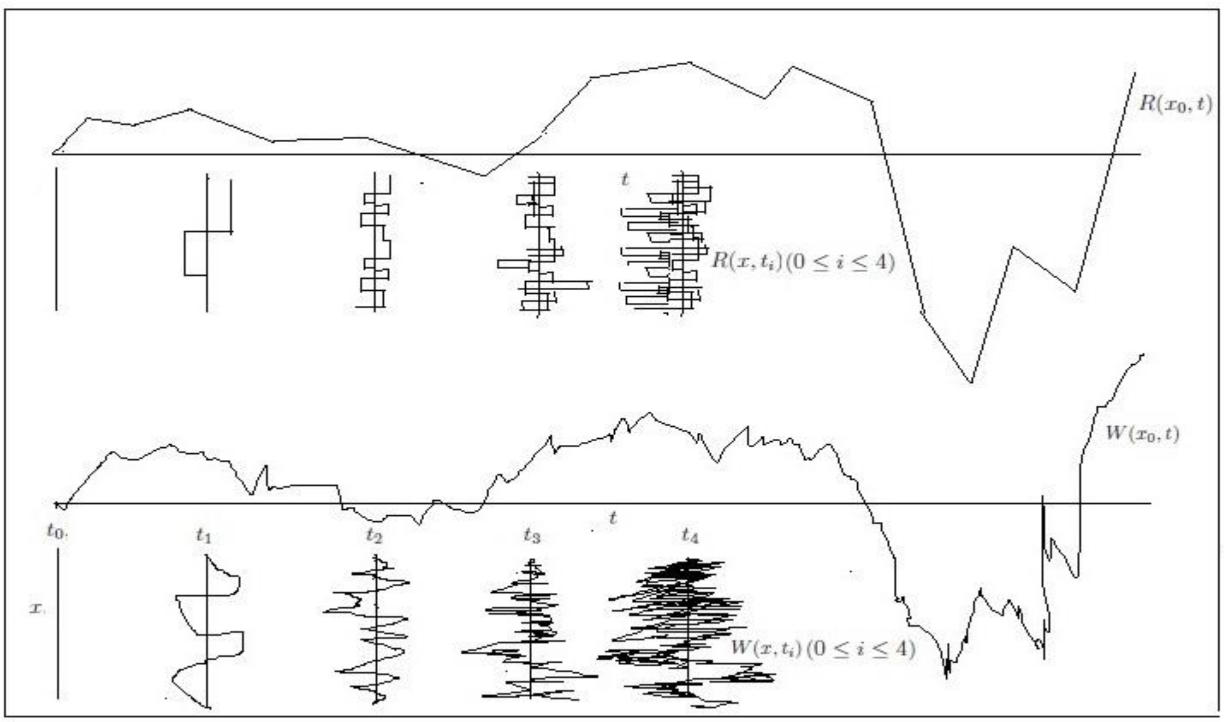
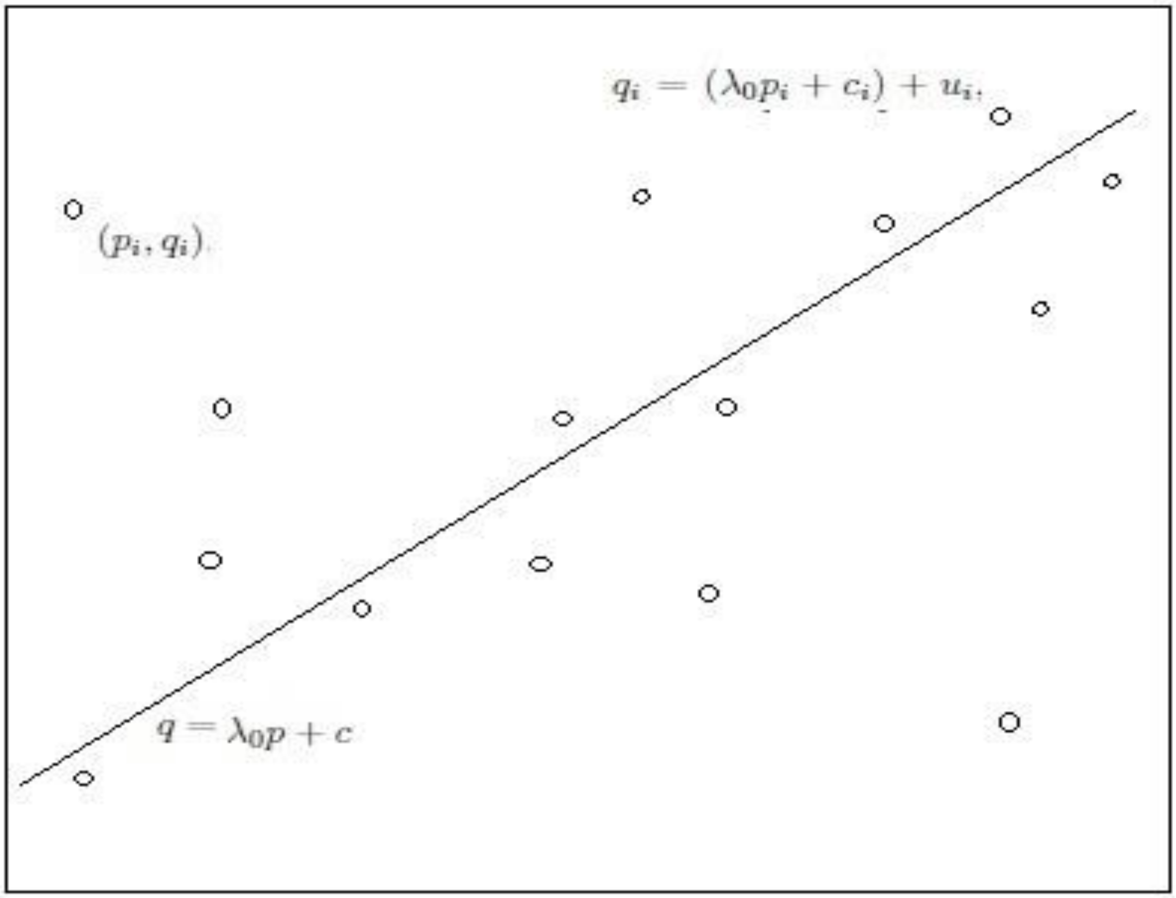


Fig. 2.
R
O
M
M
b
P
P

23


$R(x_0, t)$
t
$R(x, t_i)(0 \leq i \leq 4)$
$W(x_0, t)$
t_0
t_1
t_2
t_3
t
t_4
x
$W(x, t_i)(0 \leq i \leq 4)$

24



25

10. The Lamb and the Bright and Morning Star

We now consider how the images of The Lamb and the Bright and Morning Star, have been employed symbolically in art. Although there are few depictions of Christ, represented as the Lamb, in medieval paintings, one can find a number of depictions of this motif in medieval architecture and sculpture. The use of lambs as representing the apostles and disciples occurs in the following mosaic of the Transfiguration, which we considered above;

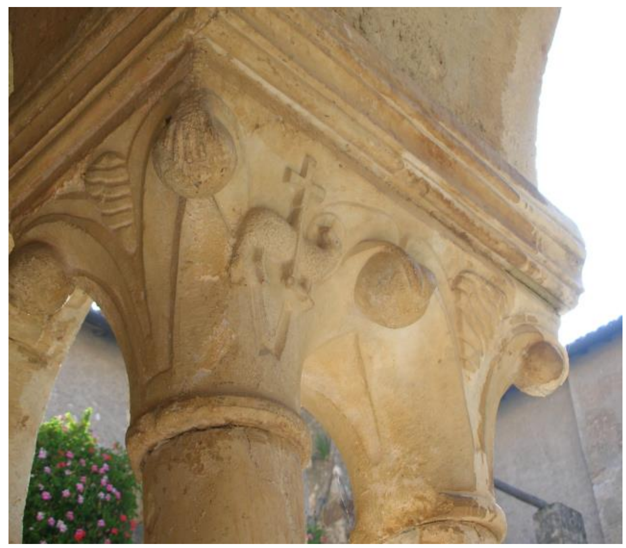
The Aesthetics of Light, Projection and Alternation in Medieval Art and Architecture.

Two good examples can be found outside the church of San Vittoria, Monteleone Sabino, and in the cloisters of the Abbey at Valviscolo;

146. Church of San Vittoria, Monteleone Sabino, Rieti, Italy.

147. Detail of San Vittoria, Monteleone Sabino, Rieti, Italy.

148. Cloister of the Abbey of Valviscolo, Latina, Italy.



149. Detail of Valviscolo Abbey, Latina, Italy.

As we already discussed, the symbol of the Lamb signifies grace and humility. In these examples, a new aesthetic of refinement and gentleness is introduced. The relationship of this image to that of the Crucifixion, which we observed, is, here, demonstrated by the Lamb carrying a cross, like a flag. As we noted, there is a spiritual strength to these motifs, the Lamb seems to be walking forward, looking back at the event of the Crucifixion, like a distant memory, and, yet, carrying the cross with a sense of purpose. This, at first, seems contradictory, but, in light of the later resurrection, as was clearly understood in the medieval age, these two images of strength and weakness are reconciled, and there is a simple balance demonstrated here. There is a sense of the suffusion of light, surrounding the cross like a halo. It is well known that the "Agnus Dei" was employed by the Knights Templar, (Ralls, 2007). This suggests that the above qualities were an essential part of the requirements to be part of this order.

It is difficult to find exact depictions of Christ as the Bright and Morning Star in medieval art, although, there is an example of a vault, decorated by stars, at Carlisle Cathedral;

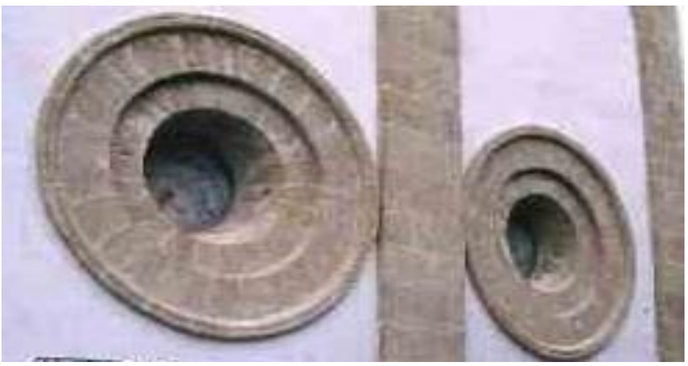
150. Carlisle Cathedral, Carlisle, Scotland.

In the medieval mind, the representation of Christ as a star, or a source of light, was achieved through the use of circular windows or oculi. These can be seen at, for example, the Cistercian Abbey of San Galgano and the church of San Michele, Arcangelo, in Bevagna;

151. Abbey of San Galgano, Tuscany, Italy.

152. Church of San Michele Arcangelo, Bevagna, Umbria, Italy.

As in the sculptural forms, depicting the Lamb, the oculus induces an aesthetic sense of balance, unity and truth; the smooth, simple shape of the circle appeals to a visual intelligence, emphasizing the clear passage of light through the window, which would be obstructed by more sophisticated aesthetic forms, involving harmony or symmetry. The origin of the oculus window is an interesting question, discussed in the above citation of Paignton Cowan. Early examples of such windows can be found in the Pantheon, Rome, and at the church of Burdj Hedar. However, a simple inspection of the Doric architecture of this last building allows us to discount the use of such forms, with a Christian symbolism in mind. A further early example is the pair of oculi found at Cefalu Cathedral in Sicily;



153. Cefalu Cathedral, Cefalu, Sicily, Italy.

According to Cowan, "the early 12th century oculi at Cefalu cathedral might originally have had some kind of tracery in wood, lead or stone". This, together with the pairing of the windows, suggests that the above representation of Christ as a source of light, was not intended. Given the importance of Cefalu in the development of Norman architecture, this could mean that they were added later. There are a number of other early examples of oculi with radiating tracery, for example at San Vincente, Cordoba, 6th century, San Salvador, Priesca, (921) and Pomposa (1028). However, although these are important in the origination of rose windows, the aesthetic device of unity is, as we have noted, not considered.

I would suggest, then, that the origination of this idea can be found in Cistercian architecture, (1180-1275), of which there are numerous examples in England and Italy. We already discussed the importance of the Cistercian influence, Chapter 4, in the origin of the pointed vault at Durham, and the aesthetic of fragmentation. The simple oculus form can be found at Canterbury Cathedral (1180) and Byland Abbey (1177);

154. Canterbury Cathedral, Canterbury.

155. Byland Abbey, Yorkshire.

which we mentioned in conjunction with the development of the Lancet style. We have seen the use of the form at San Galgano, further later regional variations, after 1250, can be found at the Abbazia di Farfa and the Upper Church of St. Francis, Assisi;

156. Farfa Abbey, Lazio, Italy.

157. Upper church of St. Francis, Assisi, Umbria, Italy.

The use of this form, together with a cross, similar to the depictions of the Lamb, (and a regional tracery design, from the Rayonnant period), occurs at the Abbey of Valviscolo, which we discussed, and, at the church in Visso;

158. Abbey of Valviscolo, Latina, Italy .

159. Collegiate church of Santa Maria, Visso, Umbria, Italy.

This combination of the aesthetics of fragmentation, in the use of intersecting lines, and unity is an important development, suggesting the influence of the Knights Templar. There is a sense of projection here, the image of the lines being clarified in a 2-dimensional space by the ambient light shining through the window. This idea is frequently repeated, in the use of the circle, together with the pointed arch form, as at the Cistercian Tintern Abbey, in Wales, and the church of St. Andrews, in Bere Ferrers, Devon;

160. Tintern Abbey, Monmouthshire, Wales.

161. North window of St. Andrews, Bere Ferrers, Devon.

Variations in the use of a cross, and a single line, can found at the Cistercian Abbey of Buckland, in Devon;

162. Churches at Buckland Abbey, Devon.

The use of circular forms and combinations of horizontal and vertical lines, in architecture, as a response to Cistercianism, also occurs frequently in Denmark. The round churches of Osterlars, Olsker,

163. Churches of Osterlars and Olsker, Bornholm, Denmark.

Nyker and Nylars, on the island of Bornholm, are known to have been constructed by the Knights Templar order, between 1250 and 1300. The characteristic use of stepped roofing and crossing designs occurs in a regional style, 1250-1400, at Vor Frue Kirke in Odense and Svendborg;

164. Vor Frue Kirke, Odense, Denmark.

165. Vor Frue Kirke, Svendborg, Denmark.

An important development in the Cistercian use of the circular form can be found at in the use of cusps, as at Pontigny Abbey, France, (c1160);

166. Pontigny Abbey, Pontigny, Burgundy, France.

This suggests that the simple circular form could have been in use within Christian France, before 1160, but I know of no remaining examples, except at the Carolingian church of St Genereux, (950), which can be easily discounted on aesthetic grounds.

The use of the cusp form in circular designs can be found at St. Anne's chapel, Malbork Castle, Poland, (1280), an important castle of the Teutonic Knights, as a response to the English Geometric style of quatrefoils, which we discussed in Chapter 3;

167. St. Anne's chapel, Malbork, Poland.

This recurs at a number of English churches in Cornwall, for example at Gulval, St.Ives and Kilkhampton;

168. Windows from the churches of St. Gulval, Gulval, All Saints, St. Ives, and St. James the Great, Kilkhampton, Cornwall.

from the Curvilinear period, (after 1280), possibly influenced by St. Anne's Chapel and the extensive use of the trefoil form in the English Decorated style. These designs introduce another cusped form in the shape of a heart, which we observed briefly in the west window of York Cathedral, which recurs again at Kilkhampton and Stoke Cannon, in Devon;

169. Windows at St. James the Great, Kilkhampton, Cornwall, and at St. Mary Magdalene, Stoke Cannon, Devon.

The reason for the recurrence of the cusped heart shaped form in Devon and, particularly, Cornwall is unclear; perhaps there is some Romantic connotation here.

The aesthetics of light and projection, which, I have suggested were highly developed in the medieval era, found their expression in painting, during the later Renaissance period. The suggestion of three dimensionality, through the modelling of light and shadow, could be attributed to the artist Giotto. The following image of an angel is taken from “Scenes from the Life of Christ (Lamentation)”, in the Arena Chapel, Padua, (1304-06);

170. Fresco at the Arena chapel, Padua, Veneto, Italy.

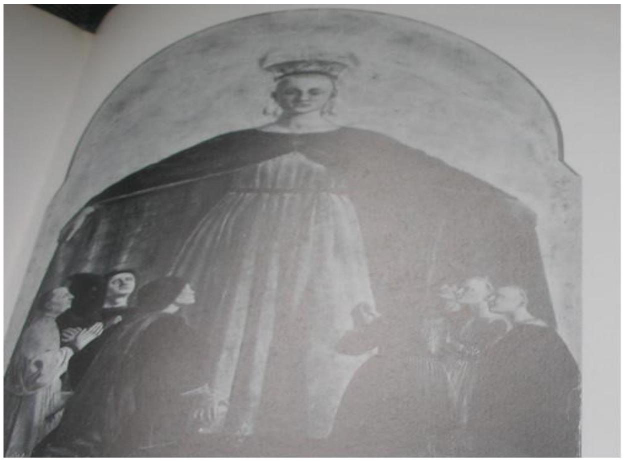
There is an understanding of light, perhaps not found in medieval painting, in the shadows cast upon the face and arms. This sense of the projection of light, through shadow, is emphasized by the deepened colour of the suggested cruciform shape, as in the window at Valviscolo. In “Lives of the Artists”, Vasari relates that when the Pope sent a messenger to Giotto, asking him to send a drawing to demonstrate his skill, Giotto drew, in red paint, a circle so perfect that it seemed as though it was drawn using a compass and instructed the messenger to give that to the Pope. Perhaps, this intuitive sense of the circular form, is what permitted the innovation of the portrayal of three-dimensionality on a surface; a decisive moment in the history of art.

This aesthetic idea of projection, was visually realized, during the Early Renaissance, by Brunelleschi, in the field of architecture, and by Piero della Francesca, in painting. The dome of the Duomo in Florence improves upon the design of the Pantheon, in Rome, by creating a hemispherical, rather than elliptically shaped surface. The construction is held in place by a single keystone, placed at the apex of the Cathedral. As we have

171. Florence Cathedral, Florence, Italy.

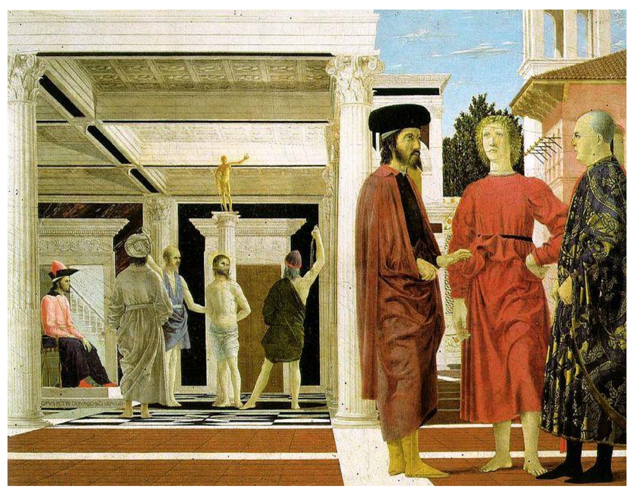
discussed in Chapter 3, the pointed vault is more effective, from the functionalist point of view. However, one has to admire the careful balancing act which was required to create this structure, that, along with the windows, conveys a sense of simplicity and grace.

We will discuss the science of perspective and projection, developed by Piero della Francesca, in greater detail, in the following chapter. For now we will consider some of his earlier paintings. In “The Madonna of Mercy”, (1445), we find a modelling of light and shadow, and, an absence of line, reminiscent of the style of Giotto, and his follower Masaccio. There is an understanding of how this creates a sense of three-dimensionality, rather than a preoccupation with the form of curves in the plane. In “The Brera Madonna” (1465), this sense of space is repeated in the construction of the hemispherical dome above the ostrich egg, hanging from the ceiling, perhaps influenced by the architectural ideas of Brunelleschi ;

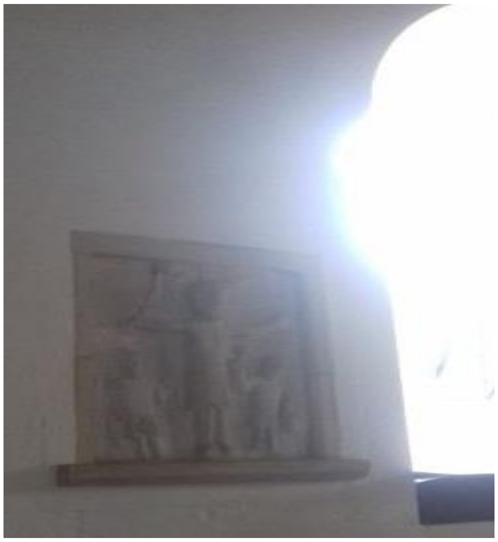
172. Madonna of Mercy and the Brera Madonna by Piero della Francesca, in the Pinacoteca, Sansepolcro, and the Pinocoteca di Brera, Milan.

In the “Discovery of the Wood of the True Cross“, (1452-7), we find the beginnings of a mathematical interest in perspective, influenced by Alberti, with whose ideas, according to Hartt, he may have become aware of, during his time in Florence, from 1439 to 1442. In this painting, Piero utilizes Alberti’s one point perspective construction, the vanishing point, traceable from the intersection of the left hand roof of the portico and the right hand beam, intersecting at a level, just below the head of the kneeling figure. Again, this construction appears in “The Flagellation”, (1459), in this case, the vanishing point, occurs half way up the central column.

173. Discovery of the Wood of the True Cross and the Flagellation by Piero della Francesca, in the church of San Francesco, Arezzo, Umbria, and the Galleria Nazionale delle Marche, Urbino, Marche, Italy.

The idea of a vanishing point and its link with conic projections, suggests the form of a triangle, in which the projection is viewed from the side. The image of the Lamb, and the associated ideas of the projection of the branches of curves onto a plane, also introduces semicircular motifs, but at a microscopic rather than macroscopic level. Many of these images can be found within the Anglo-Saxon art of England, and, to a lesser extent, Italy. The depiction of Christ in Majesty, found in an Anglo-Saxon carving at the church of the Holy Rood in Daglingworth, suggests a link with the image of the Bright and Morning Star, as well as a triangular Y-form, in that the arms are slightly raised, rather than the conventional depiction of a cross;



174. Anglo-Saxon carving of Christ in Majesty, at the Holy Rood church in Daglingworth.

The use of dogtooth, as a triangular form, is well known, and can be found in windows at Earl's Barton, Deerhurst, and doorways at St. Michael's church, Duntisbourne Rouse and St. Peter's church, Barton-on-Humber;

175. Dogtooth windows at Earl's Barton church, Northamptonshire, Deerhurst, Gloucestershire, and doorways at St. Michael's, Duntisbourne Rouse, and St. Peter's church, Barton-on-Humber.

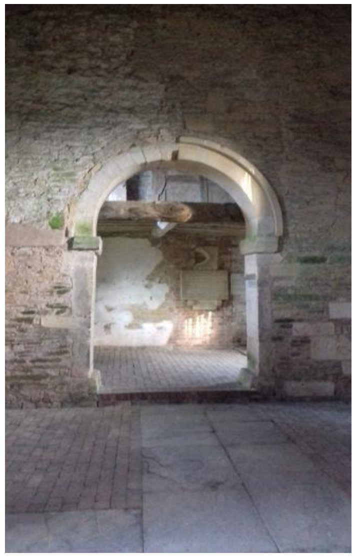
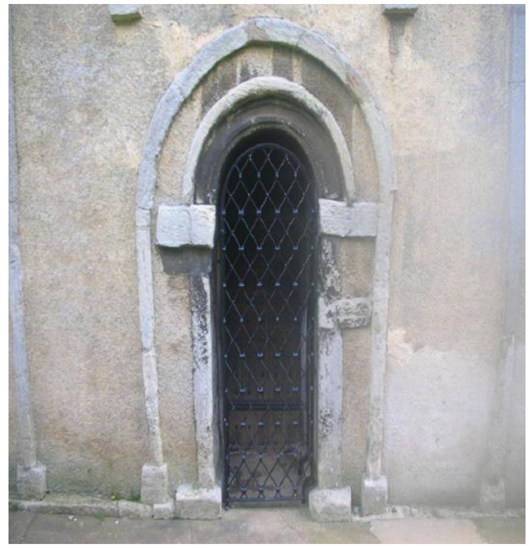
The architectural feature of a saddle back tower, with a triangular turret, is also employed at numerous churches within Gloucestershire;

176. Saddleback towers at St. Michael's, Duntisbourne Rouse, and St. Michael's church, Bagendon.

Semicircular arches and windows, significantly smaller than the Norman design, are also located in various English churches;

177. Archways at All Saint's church, Turksdean, Church of St. Mary and All Saints, Hawksworth, Church of the Holy Rood, Daglingworth, and St. Lawrence church, Bradford-on-Avon.

The development of arch supports on pillars, possibly related to the earlier use of quoins is also an Anglo Saxon innovation;

178. Quoin and archway at Odda's chapel, Gloucestershire, and archway at Earl's Barton church, Northamptonshire.

The later use of spandrels occurs in conjunction with Norman architecture. The trefoil design, introduced at Winchester cathedral around 1180, is associated with the Lancet period. These ideas are combined in Rayonnant designs such as the window at St.Mary's church in Boyton, Wiltshire. Their emphasis on a triangular geometry, could be related to the Anglo-Saxon use of dogtooth in windows and archways;



179. The use of spandrels at St. Michael and All Angels, Ledbury, and trefoil plate tracery at Winchester Cathedral.

George Zarnecki takes the view that Anglo-Saxon traditions continued in the English Romanesque style, from 1050 into the 12th century, with little influence from the Norman conquest. The church at Alatri can be dated back to 1137, and shows some aesthetic similarities with Anglo-Saxon churches in England, for example the narrow circular arches and pillars in the aisles, similar to those found at Odda's chapel, Gloucestershire, and All Saints church, Turkdean. The trefoil design can be found in the window of the façade.

180. The trefoil rose window, archways and a column at Santa Maria Maggiore church, Alatri.

The aesthetic forms evident in Anglo-Saxon architecture combine to produce an idea of alternation, the suggestion that the smooth geometries of semicircular forms and waves, can alter and change direction continuously along the vertices of linear motifs such as triangles. There is also a sense that linear forms can be associated to curves, with less of an emphasis on the standard harmonics associated to radiating and focused patterns, but still retaining a close relationship with simple circular geometries.

In Ireland, one can see the influence of lintels and simple linearity in the use of asymmetrical quadrilateral forms in doorways, dating, as Petrie notes, from the 6th-7th centuries.

181. The use of quadrilateral forms in the doorways of Our Lady's Church, Glendalough, St. Kevin's church, Reefert, Ratass church, Tralee, St. Molua's church, Killaloe, and The Sons of Nessan church, Dublin.

The characteristic square capitals of English Anglo-Saxon architecture, together with narrow semicircular archforms can be found at Cormac's chapel, in the Rock of Cashel, the west doorway at Killaloe and the east window at Ireland's eye.

182. The use of narrow semicircular archforms at Cormac's Chapel, Cashel, the west window of St. Molua's church, Killaloe, the east window at the church of The Sons of Nessan, Dublin, and square headed capitals inside Cormac's chapel.

As noted by Bell, herringbone work, again typical of the Anglo-Saxon period, occurs at St. Finian's chapel, Clonmacnoise, while the square form is repeated on smaller and larger scales in the round towers at Cashel and Clones.

11. PIERO DELLA FRANCESCA AND THE TRUE GEOMETRY OF LIGHT

TRISTRAM DE PIRO

Piero della Francesca was born in the Umbrian town of Sansepolcro, the exact year being unknown, but, it is speculated, between 1420 and 1422. There is no doubt that Piero della Francesca was an extremely accomplished mathematican, not only in geometry, but also in algebra. He wrote "Trattato d'Abaco", which deals with radicals and algebra resolving "equations above the second degree that cannot be reduced" and "De quinque corporibus regolaribus", a treatise on Platonic solids. P. Arrighi described Piero as "the greatest mathematician of his era". Vasari also explains how; "In his youth Piero applied himself to mathematics, and although when he became fifteen it was decided that he should be a painter he nevertheless always kept up his earlier studies".

As we mentioned in the previous chapter, he is known to have worked in Florence, between 1439 and 1442, where he came across the geometric constructions of Alberti, in particularly, one point perspective, from his Della pittura [On Painting] (1435). Possibly the first work involving the geometry of projection can be found in Euclid's "Ottica" (Optics), (c 300 BC). Theorem 11 states that, "Tra i piani che stanno sopra l'occhio i (piu) lontani appaiano piu in basso." (Amongst the planes which fall on the eye, the furthest away appear to be the lowest). Euclid's simple proof of this proposition is illustrated in Figure 1; the points x, y and z illustrating the diminishing height of a object as it recedes from the eye, situated at O. Although Euclid does not explicitly make this connection, it was realised, possibly by Alberti, that, as the object recedes to infinity, it converges to the point ∞, the vanishing point, marked on the baseline of Figure 1. Alberti's one-point construction can be found in Figure 2, Euclid's planar figure being represented by the planes Owz and Oxu, which intersect with P, at the vanishing point y, along the normal line between the plane P and the eye, situated at O; the lines zw and xu being perpendicular to the plane P. It seems reasonable to say that Albertian perspective was the basis for the modern subject of linear algebra. The projection, described in Figure 2, has the property that it maps a line l, either to a line l' or to a point q, in the plane P. For a suitable choice of coordinates, the projection pr described in Figure 2, has the property $pr(\alpha x' + y') = \alpha pr(x') + pr(y')$, for a scalar α, which, is the modern definition of a linear map. This is easily visualised by taking the eye O to be at infinity, y to be the origin of the coordinate system in the plane P, with the rays, defining parallel perpendiculars to the plane P. In fact, using the well known "rank-nullity" theorem, see [5] or [4], any linear map can be decomposed into a projection and a linear transformation of the plane to itself.

Towards the end of his life, Piero wrote "De Prospectiva Pingendi", on perspective for painters. Perhaps dissatisfied with the limitations of the linearity of Albertian perspective, it is innovative for being the first scientific work to deal with the laws of perspective applied to curves; indeed it is written as a series of purely geometrical constructions of curves onto a plane. "De Prospectiva Pingendi" was completed before 1482, 10 years before his death in 1492. There has been some speculation that Piero was blind when it was written. The following quotation is from [2];

"According to Vasari, the aged Piero was blind, and in the mid sixteenth century a man still lived who claimed that, as a boy, he had led Piero about Borgo San Sensepolcro by the hand. This story has been doubted, but it must contain more than a grain of truth even though in 1490, two years before his death, Piero still wrote a beautiful hand."

Irrespective of the truth of this claim, there can be no doubt that "De Prospectiva Pingendi" was written at least 5 years before this period, as one can find a record from 5 July 1487, which records that Piero appears before the notary Ser Lionardo di ser Mario Fedeli, elderly yet "sound of mind and body", see [2]. Piero della Francesca was a major influence on the geometric style of later Renaissance artists. In particular, the representation of "smooth" curves in 3-dimensions as curves in a 2-dimensional plane, having at most "nodes as singularities", is clearly reflected in the general drawing style of that period.

In figures 3 and 4, which correspond to figures 61 and 62, of "De Prospectiva Pingendi", we see the geometric idea behind Piero's method. An object, in this case, a pillar, is projected onto the plane of the canvas, corresponding to the vertical line in figure 61. This is done, by choosing an (x, y) grid of points on the object, and drawing lines connecting these points to the eye. The rendering of the object in the plane, is obtained by finding the intersections of these lines with the vertical. Piero, first, finds the intersections of the mesh points in the y-direction, and, then, repeats the construction in the x-direction. Observe that the pillar is being drawn above and to the right. In figure 62, the final result, of connecting the projected points, is shown, and, the calculation is repeated for various distances of the canvas. The interested reader can find a detailed description of this construction, in pp51-55 of the text. Note that, unlike Alberti's construction, the method easily incorporates curved figures.

The geometrical idea discovered by Piero is now referred to as the method of Conic Projections, and was highly influential on the Italian development of algebraic geometry, in the 19th century, particularly the work of Severi, Castelnuovo and Enriques. Severi was born in Arezzo, Italy, on 13 April 1879, and was the major figure of the Italian school. He wrote an article on [3] in [11], perhaps becoming interested in his work, due to the proximity of Sansepolcro to his place of birth. We can describe the construction in a more mathematical language, using the framework of algebraic curves, which we introduced in Chapter 4. If $C \subset P^3$ is an algebraic curve, ([1]), and $P \notin C$, we define $Cone_P(C) = \bigcup_{x \in C} l_{xP}$. Then, we have the following theorem;

Lemma 0.1. *Let $C' \subset P^3$ be a nonsingular projective curve, which is not contained in a plane, then, if $O \in P^3$, and H is a 2-dimensional hyperplane, with $O \notin H$, we have that $(Cone_O(C') \cap H)$ defines a projective algebraic curve C.*

[1] P^3 is the standard terminology for 3-dimensional projective space.

The proof is basically a simple consequence of the fact that $Cone_P(C')$ has dimension 2, H has dimension 2, $Cone_P(C') \neq H$, and the fact that two such surfaces intersect in a curve, the reader should look at [8] for further details. Of course, the proof is visually obvious, see figure 5, but, Piero's treatise is interesting because it provides some of the first insights into how visual and logical thinking are related.

In order to explain one of the most important mathematical results, motivated by this type of optical thinking, we require some terminology. We say that two curves C and C' are birational, if there exist open subsets $U \subset C$, $U' \subset C'$, and an isomorphism between U and U'. We introduced Newton's notion of an infinitesimal in Chapter 4. This notion was intuitively extended by Severi, in his discussion of algebraic curves, see [12], by defining an infinitesimal neighborhood $\mathcal{V}_p(C)$, of a nonsingular curve C, to be a disc of infinitesimal radius, surrounding a point p. ([2].) In [12], Severi combined these two ideas, to give an intuitive description of the local geometry of a, possibly singular, curve. Namely, given a curve C, he first showed that there exists a nonsingular curve C', birational to C, ([3]. This lack of rigour was one of the main criticisms of the Italian geometers approach.), see [12] or [8]. Secondly, given a point $p \in C$, located at a singularity, he defined the branches $\gamma_p^i(C)$, for $1 \leq i \leq n$, centred at p, to be the infinitesimal neighborhoods of the points $\{p_1, \ldots, p_n\}$, corresponding to p, on C'. That this is a good definition, is visually obvious, see figure 6, a rigorous proof can be found in [8]. Severi then defined a node p to be the origin of two branches $\{\gamma_p^1(C), \gamma_p^2(C)\}$, which intersect transversely, that is, their tangent lines are distinct, see figure 7, ([4]).

The main achievement of this visual method of proof, was to show the following theorem, the main idea being due to Severi and Castelnuovo;

Theorem 0.2. *Let $C' \subset P^w$ be a nonsingular projective curve, then C' is birational to a plane projective curve $C \subset P^2$, having at most nodes as singularities.*

Again, the geometric idea is clear, however, a clear formulation of the idea of a branch is required to find a rigorous proof, even to formulate the notion of a node correctly. The proof

[2] A more rigorous modern formulation can be found in [13]. The proof that this is equivalent to Severi's definition can be found in [6]

[3] More precisely, he showed that C' has no multiple points, that is points p with contact at least 2, (see footnote 4) for any hyperplane H passing through p. This turns out to be equivalent to the algebraic definition of nonsingularity, see Lemmas 3.2 and 4.14 of [6], Lemma 4.14 of [7]

[4] The notion of a tangent line at a branch γ of a plane curve C requires the development of Severi's notion of "contact", (contatto), between a branch γ and a line l, which I refer to as $I_{italian}(p, \gamma, C, l)$, in [8]. Namely, a line l and a branch γ have contact of order m, if, varying the line generically by an infinitesimal amount, defines m distinct intersections along the branch, see figure 8, (compare figure 2 of Chapter 4 for nonsingular points). The tangent line l_γ is defined uniquely by the requirement that the contact has order at least two. That is a good definition depends heavily on the idea of birationality, namely the intersection points are preserved by the isomorphism between open sets, and we can isolate the intersections along the branch γ, by counting them near to the corresponding nonsingular point, see figure 8 below and figure 2 of Chapter 4. That this corresponds to Newton's idea of a branch for nonsingular points, using power series, see Chapter 4, relies heavily on the method developed in [7]. This notion is also required to make the notion of tangent line for a branch precise, as, for nonsingular points, it is simple to define the tangent line algebraically, and a power series representation easily shows that it is unique. to A full description of the possible singularities of a curve C was given by Cayley, in the 19th century.

proceeds by first showing the following lemma, attributed to Castelnuovo. A rigorous proof can be found in [8].

Lemma 0.3. *Let $C' \subset P^3$ and suppose that $\{A, B\}$ are independent generic points of C'. Then if C' is not contained in a plane, the line l_{AB} does not otherwise meet the curve.*

The idea might seem clear from a visual point of view, by considering the shadow of, say, a smooth loop of thread on a surface, and revolving the thread until the shadow has no triple intersections. However, to make this idea geometrically precise seems difficult, as the loop could, in principle, be very complicated. The essence of the proof can be found in figure 9, if there existed a third point $P \in C'$, on the line l_{AB}, then projecting from P onto a hyperplane H, the tangent line l_A would lie in the plane spanned by the tangent line l_B and the line l_{AB}, $(*)$, as the projected point $pr_P(A)$ and $pr_P(B)$ coincide, and, as this point is nonsingular on $pr_P(C')$, so do the projected tangent lines $pr_P(l_A)$ and $pr_P(l_B)$. As A is generic with respect to B, this property $(*)$ holds for all but finitely many points x on the curve, (i.e replace A with x) and entails a strange symmetry. Now, projecting from B, see figure 9, all the tangent lines of the projected curve $pr_P(C')$ pass through $Q = pr_B(l_B)$. It is easily shown that the only curve with this property is a line l, which implies that the original curve C' is contained in a plane $pr^{-1}(l)$.

With this lemma shown, one can then show that the union of trisecant lines on a smooth curve $C' \subset P^3$, that is lines passing through three distinct points, on a smooth curve C' is two dimensional, $(\dagger)$. It is easily shown that every point in P^3 lies on a bisecant line of C', unless C' is contained in a plane, so the dimension of the bisecants is larger than that of the trisecants. In order to prove the theorem, one defines the following sets;

(i). The union of tangent lines to C'.

(ii).The union of bitangent chords to C', that is chords l_{ab}, for which the tangent lines l_a and l_b, at the two ends are parallel.

(iii). The union of osculatory chords of C', that is chords l_{ab}, for which the plane H of maximal contact with C' at a, passes through b.

(iv). The singular cone, which is the union of two dimensional cones $Cone_a(C')$, for the finitely many points which are "inflexions".

(v). Points lying on infinitely many bisecant lines of C'.

Similarly to $(\dagger)$, one can show that all these sets have dimension at most two. Hence, one can pick a point $O \in P^3$, avoiding all of them, as well as the union of trisecant lines. The idea is then to project the curve C' from this point O onto a hyperplane H. As O avoids (v), the projection is birational. As O avoids the trisecant lines, and C' is smooth, there are only finitely many points $\{p_1, \ldots, p_n\}$ on C', which project onto points of C, which are the origin of two branches. As O avoids (iv), these points are not inflexions. As O avoids (i), the projections of these branches do not "cusp". As O avoids (iii), and using (iv), the projections are not inflexions. As O avoids (ii), the tangent lines are distinct, so, by Severi's definition, $\{pr_P(p_1),\ldots,pr_P(p_n)\}$ are nodes. The proof is summarized in figure 10.

The purpose of the above is to demonstrate how the aesthetic ideas of projection, simple fragmentation, in the form of a cross, and, the circle, combine. The notion of a branch can be considered as a microscopic circle surrounding a point on a curve, which could be the origin of a node (the cruciform shape). The idea of projection using conics captures the idea of a macroscopic circle of light emanating from a point and projected onto a surface by shadows. This aesthetic intuition of the interplay between circles and lines, as we have seen, facilitates a deeper understanding of the geometry of curves, which a purely rational, non-visual, analysis could not provide, ([5]).

We now turn to the idea of alternation. We are given a function $f(x)$ and multiplying it by a sequence of sines $sin(kx)$ of decreasing wavelength $\frac{2\pi}{k}$. The intuition is that as k increases, the integral of $f(x)sin(kx)$ over a finite interval should approach zero. The idea, as suggested by figure 11, is that, provided f is sufficiently well behaved, the interval I can be divided into a finite sequence of r subintervals $\{I_s : 1 \leq s \leq r\}$ on which f is strictly increasing or decreasing, and positive or negative. For large enough k, the zeroes of $sin(kx)$ divide each subinterval I_s into even finer intervals $\{I_{s,t} : 1 \leq t \leq t(k,s)\}$. If we consider the integrals $\int_{I_{s,t}} f(x)sin(kx)dx$, we obtain an alternating, and strictly increasing/decreasing in modulus sequence of areas. By a simple consideration of such sequences, due to the alternation, there is some cancellation, and the sum of such areas is bounded in modulus by one term, either the integral at the beginning or end. As k increases, the interval of this integration decreases in length, and the total integral converges to zero. The fact that the series alternates is critical, there are infinite decreasing sequences which diverge, for example $\sum_{m\in\mathcal{N}}\frac{1}{m}$, whereas the alternating version $\sum_{m\in\mathcal{N}}\frac{(-1)^m}{m}$ converges to a limit close to $\frac{1}{2}$, though the exact limit is unknown. This property of alternation is peculiar to oscillating functions like sine and cosine, hence the importance of this result in Fourier analysis, a subject which we discussed above. The formal result is known as the Riemann-Lebesgue lemma, but the standard proof doesn't really employ the idea discussed above. A rigorous proof using nonstandard analysis can be found in [11];

Lemma 0.4. *Let $f \in C(-1,1)$, then $lim_{|m|\rightarrow\infty}\mathcal{F}(f)(m) = 0$.*

Riemann (1826-1866) and Lebesgue (1875-1941) were German and French mathematicians respectively, who formulated the eponymous lemma at different times according to their definitions of the integral. Riemann did this in his habilitation thesis, while Lebesgue's idea can be found in a paper from 1903, "Sur les Series Trigonometriques". The result is instrumental in the proof of Dirichlet's theorem, named after the German mathematician Peter Gustave Lejeune Dirichlet (1805-1859) who was Riemann's teacher. The theorem addresses the question as to the conditions when a Fourier series converges to a given function f. In Dirichlet's 1829 paper "Sur la convergence des series trigonometriques qui servent a

[5]The reader can find out more about the geometry of curves and surfaces embedded in three dimensional space, in [13] and [1]. A nonstandard approach to the interpretation of differentials and results in vector calculus, such as the Stokes and Divergence Theorems, can be found in [12].

representer une fonction arbitraire entre des limites donnees", Dirichlet discusses the idea of alternation, and makes the important observation that the partial sums of Fourier series;

$S(k,f)(x) = \frac{1}{2}\sum_{|m|\leq k}\mathcal{F}(f)(m)exp(\pi ixm)$

can be written as the integral;

$\frac{1}{2}\int_0^1 D_k(v)(f(x+v)+f(x-v))dv$

where $D_k(v)$ is the Dirichlet kernel, $\frac{sin(\pi v(k+\frac{1}{2}))}{sin(\frac{\pi v}{2})}$. The Riemann-Lebesgue Lemma is used to replace the Dirichlet kernel by the Fourier kernel, $F_k(v) = \frac{sin(\pi v(k+\frac{1}{2}))}{\frac{\pi v}{2}}$, noting that the difference of the denominators of these two functions is continuous on $(0,1)$, so the conditions for alternation with $sin(\pi v(k+\frac{1}{2}))$ apply as k tends to infinity. The rest of the proof involves noting that, if f is piecewise differentiable with left and right limits $\{f_x^+, f_x^-\}$, then;

$\frac{1}{2}\int_0^1 F_k(v)(f(x+v)+f(x-v))dv$

converges to $\frac{f_x^+ + f_x^-}{2}$, as k tends to infinity, where one can see the differentiability is being used to remove the terms $\frac{f(x+v)-f_x^+}{v}$ and $\frac{f(x-v)-f_x^-}{v}$. Dirichlet's theorem admits a nonstandard proof, relying on the idea of alternation, close to the spirit of his original paper, in [9]. We illustrate the idea in figure 12, where we see the Fourier series $S(k,T)(x)$ of a sawtooth function $T(x)$ converging to the midpoint at the singular endpoints. Excellent discussions of Fourier Analysis can also be found in [4] and [16], where you can find a statement of Dirichlet's Theorem.

Theorem 0.5. *If f is piecewise differentiable, then it's Fourier series $S(k,f)(x)$ converges to the average $\frac{f_x^+ + f_x^-}{2}$ as $k \rightarrow \infty$.*

References

[1] Vector Analysis, D. Bourne and P. Kendall, Taylor and Francis, (1992).

[2] Piero della Francesca, Marco Bussagli, Giunti Editore, (2010).

[3] De Prospectiva Pingendi, Piero della Francesca, c1482.

[4] Fourier Analysis, T. Korner, Cambridge University Press, (1988).

[5] Linear Algebra, S. Lipschutz, Schaum, (1996).

[6] Linear Algebra, A.O Morris, Chapman and Hall, (1982).

[7] A Non Standard Approach to the Theory of Algebraic Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2012).

[8] A Non Standard Bezout Theorem, Tristram de Piro, available at http://www.magneticstrix.net, (2002).

[9] A Non Standard Version of Dirichlet's Theorem, Tristram de Piro, available at http://www.curvalinea.net, (2020)

[10] A Theory of Branches for Algebraic Curves, Tristram de Piro, available at http://www.magneticstrix.net, (2006).

[11] Simple Proofs of the Riemann-Lebsgue Lemmas using Nonstandard Analysis, Tristram de Piro, available at http://www.curvalinea.net, (2019).

[12] Non Standard Analysis and Physics, Tristram de Piro, available at http://www.magneticstrix.net, (2006).

[13] Elementary Geometry, John Roe, OUP, (1993).

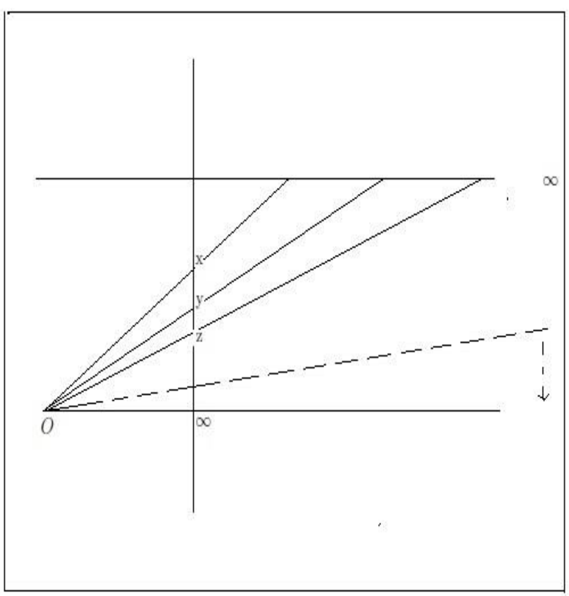
[14] A Proposito del "De Prospectiva Pingendi di Piero della Francesca", Francesco Severi, Atti e Mem. della R. Acc. Petrarca di Lettere, Arti e Scienze, 27-2, (1943).

[15] Trattato di geometria algebrica, 1: Geometria delle serie lineari, Francesco Severi, Zanichelli, Bologna, (1926).

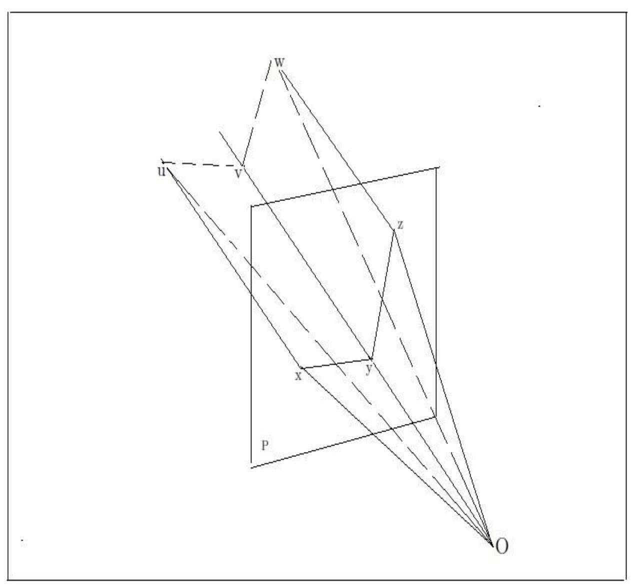
[16] Fourier Analysis: An Introduction, Princeton Lecture Series, E. Stein, R.Shakarchi, (2003).

[17] Zariski Geometries, Geometry from the Logician's Point of View, Boris Zilber, CUP, (2010).

Illustrations



1



2

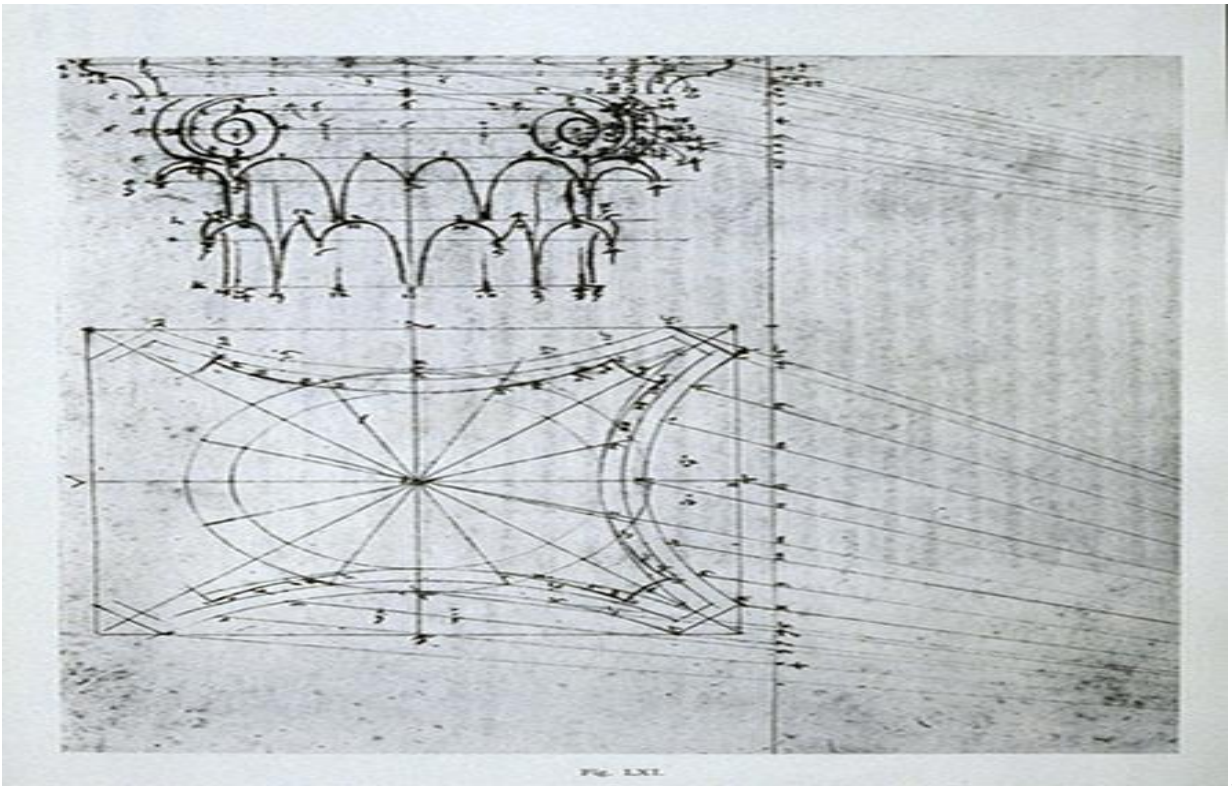
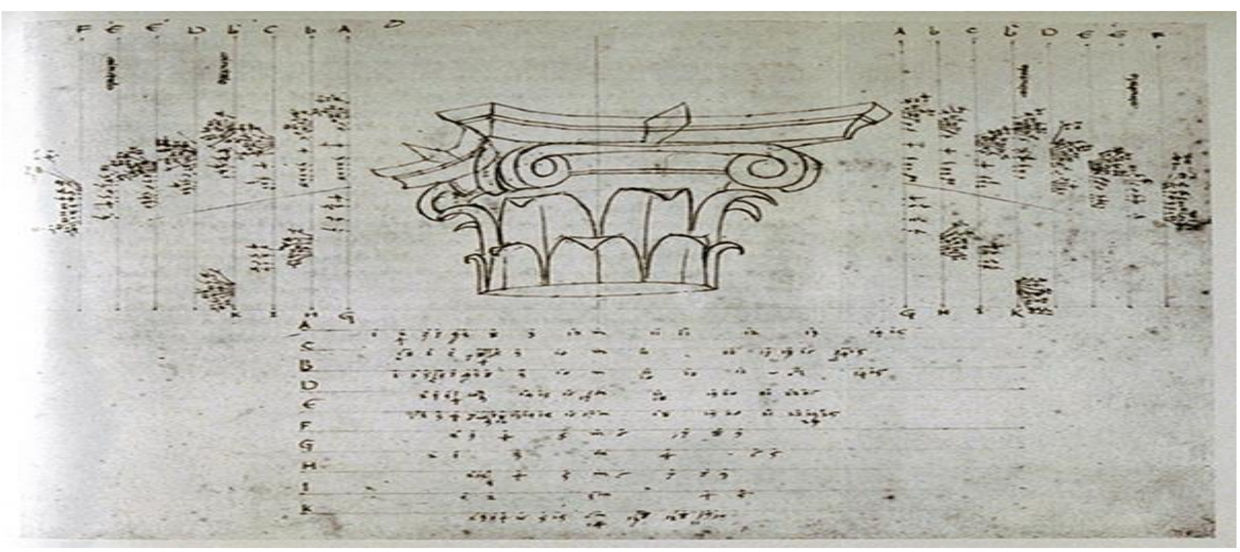
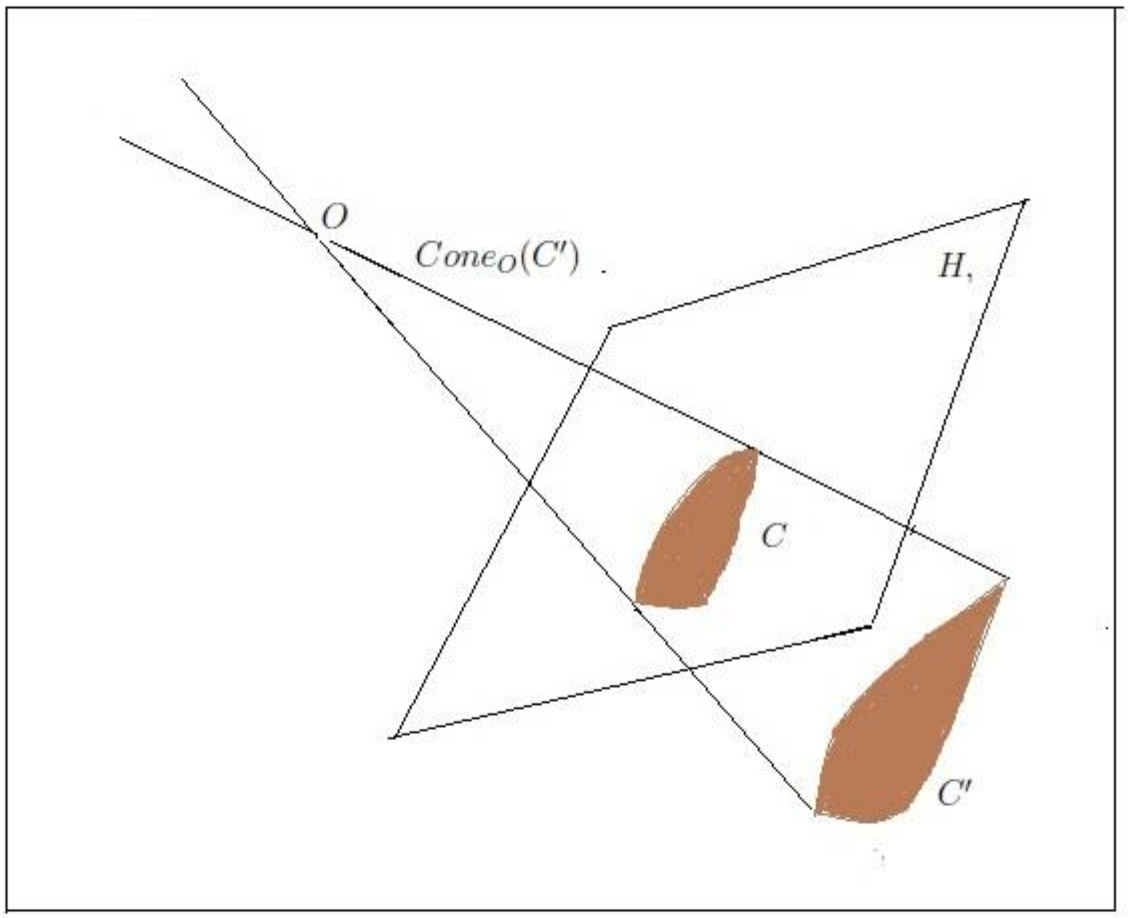


Fig. LXI

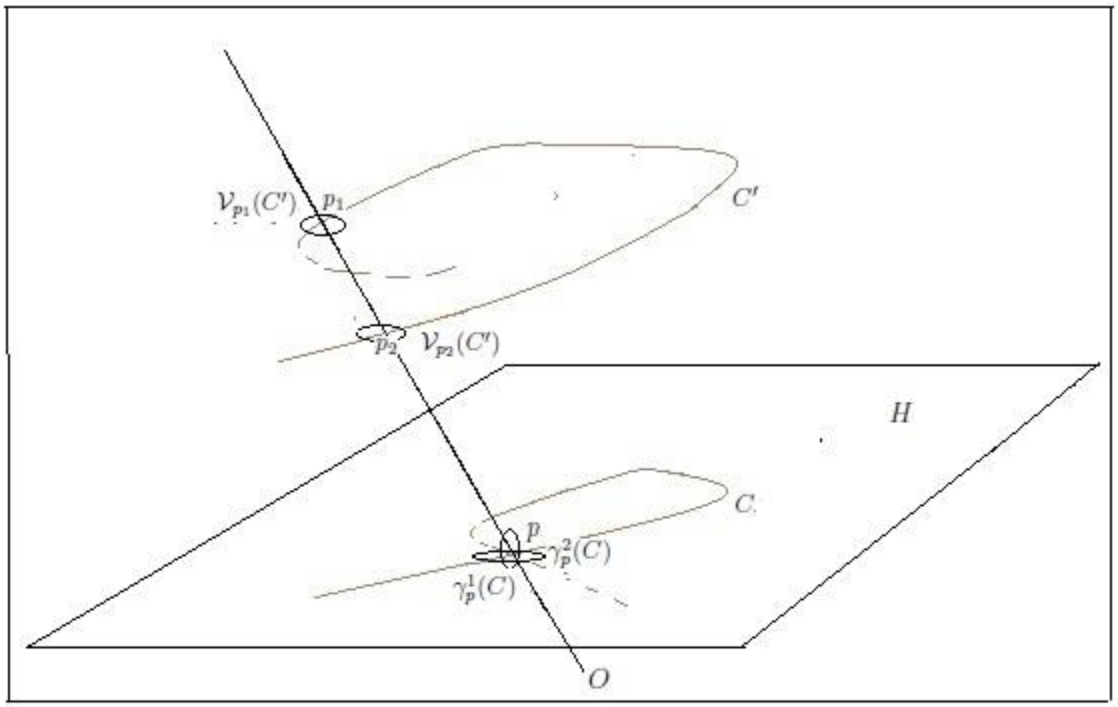
3



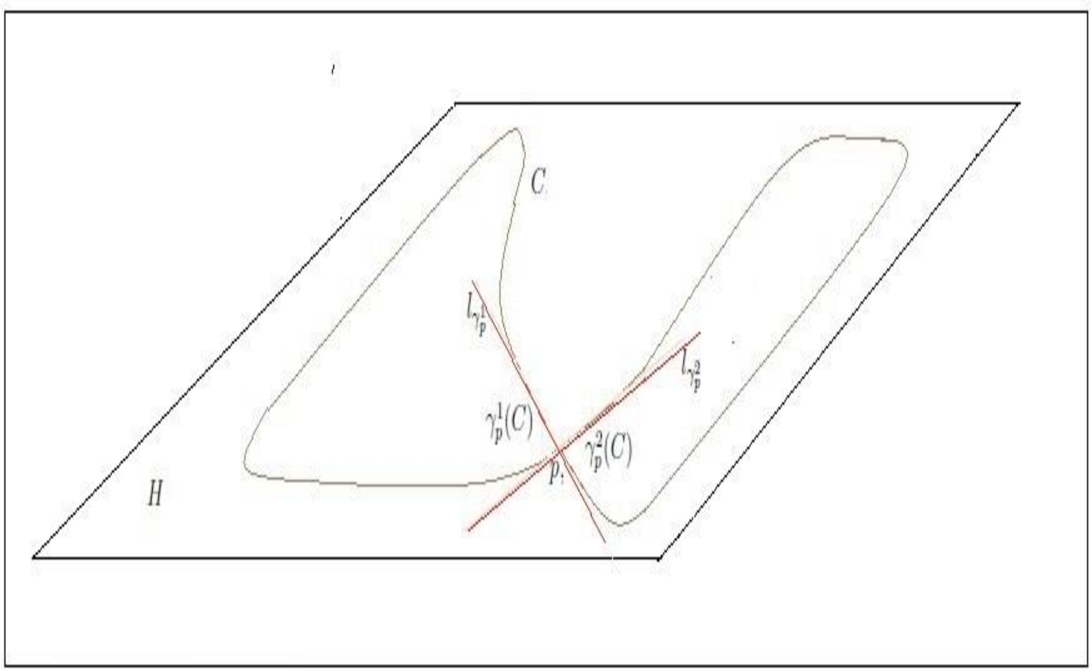
4



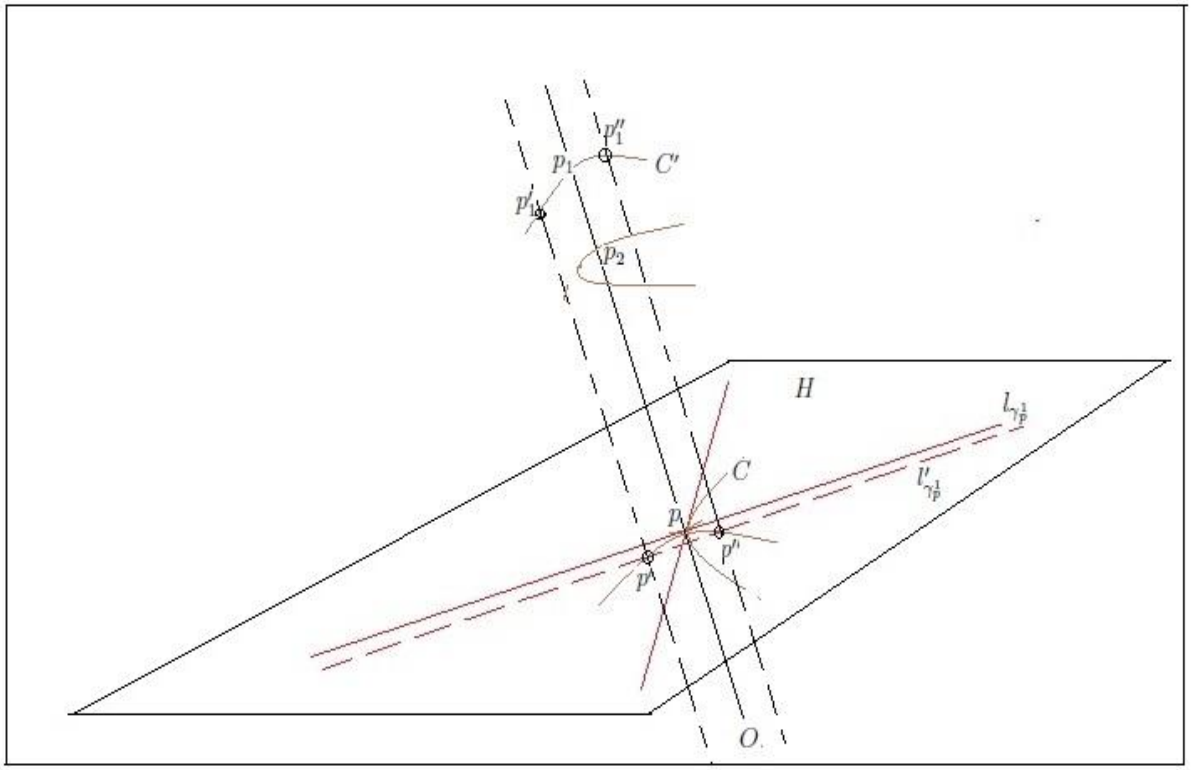
5


$\mathcal{V}_{p_1}(C')$
p_1
C'
p_2
$\mathcal{V}_{p_2}(C')$
H
C
p
$\gamma_p^2(C)$
$\gamma_p^1(C)$
O

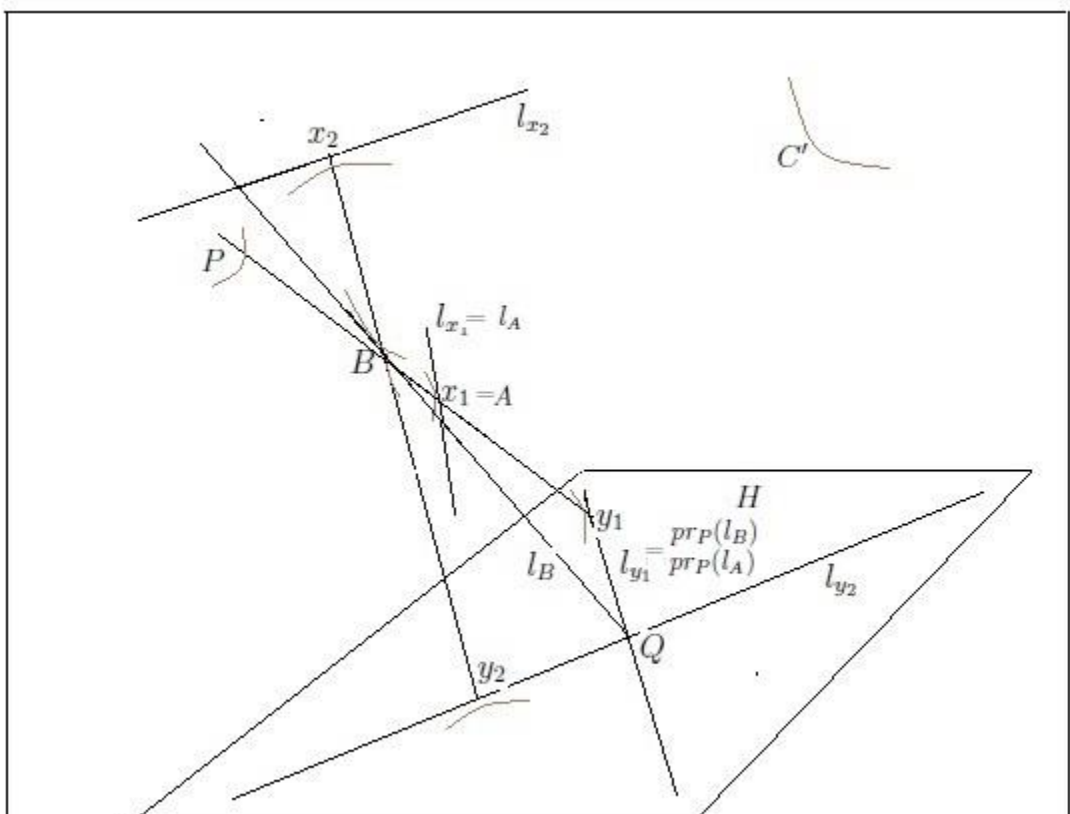
6


C
$l_{\gamma_p^1}$
$l_{\gamma_p^2}$
$\gamma_p^1(C)$
$\gamma_p^2(C)$
p
H

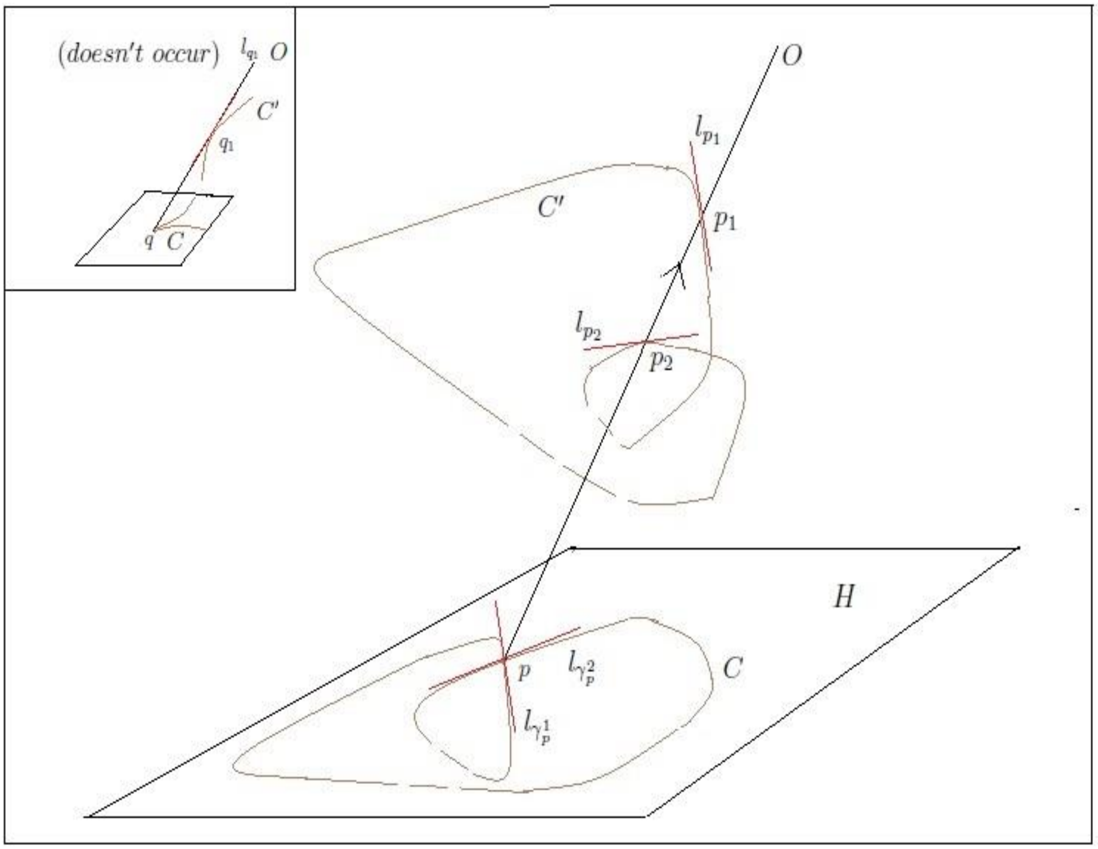
7


p_1''
p_1
C'
p_1'
p_2
H
$l_{\gamma_p^1}$
C
$l'_{\gamma_p^1}$
p
p''
p'
O

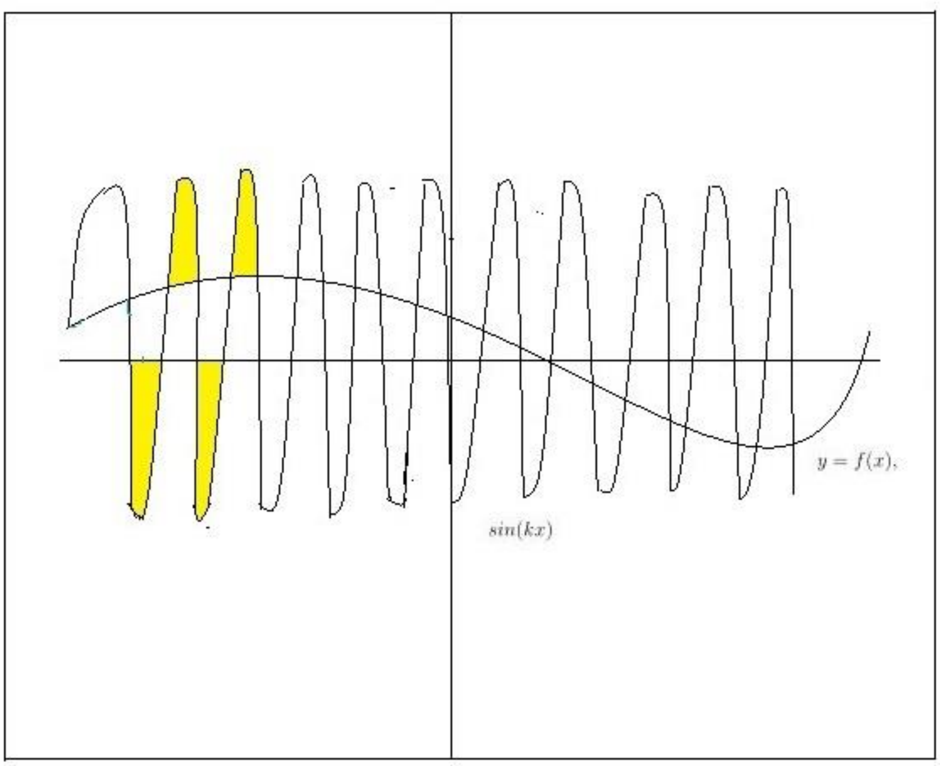
8


x_2
l_{x_2}
C'
P
$l_{x_1} = l_A$
B
$x_1 = A$
H
y_1
$l_{y_1} = pr_P(l_B) = pr_P(l_A)$
l_B
l_{y_2}
Q
y_2

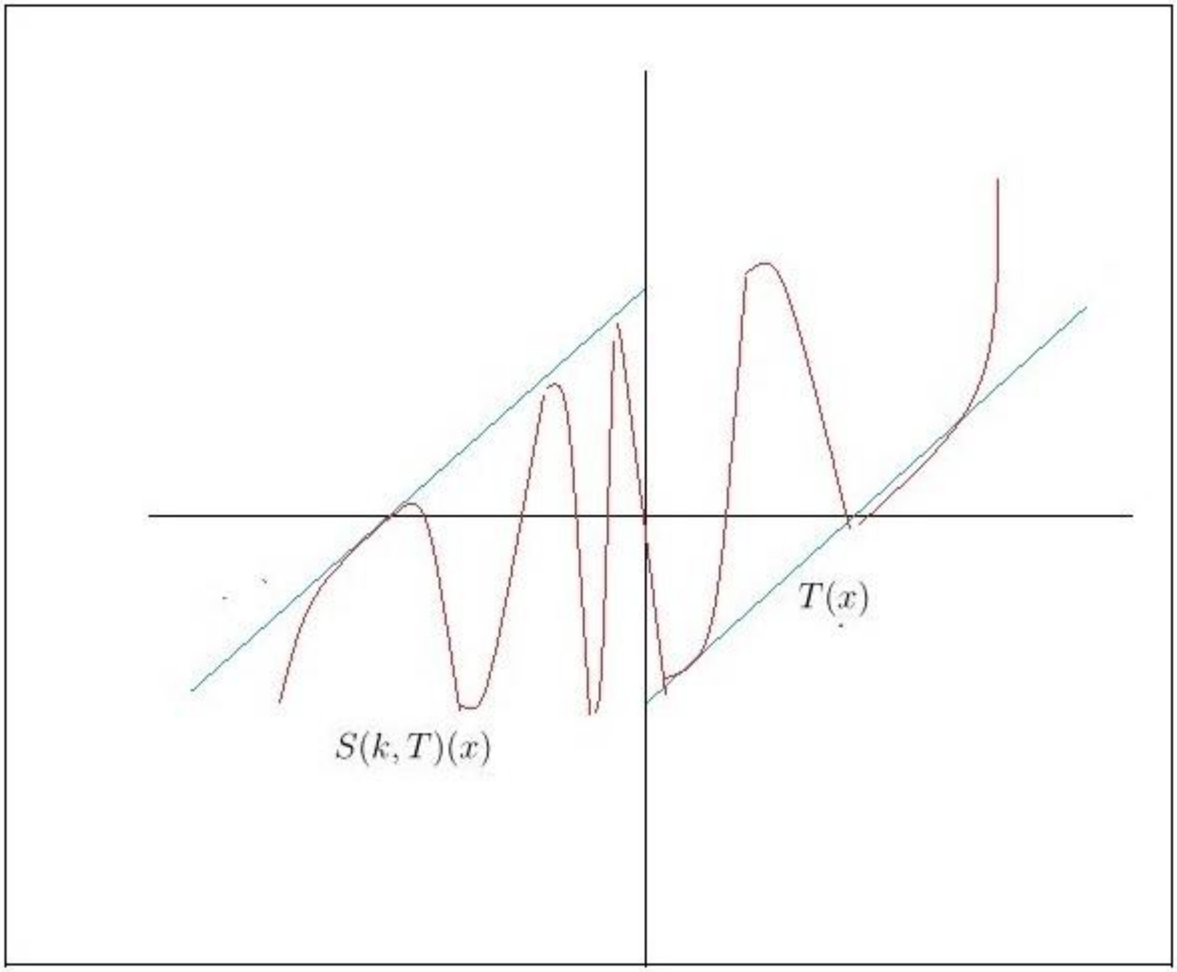
9


(doesn't occur)
l_{q_1}
O
C'
q_1
q
C
O
l_{p_1}
C'
p_1
l_{p_2}
p_2
H
p
$l_{\gamma_p^2}$
C
$l_{\gamma_p^1}$

10


$y = f(x),$
$sin(kx)$

11



12

12. Artistic Bibliography.

Arthaud, B. (1962). *Chartres et sa Cathedrale.* Paris: b. Arthaud.

Ball, P. (2009). *Universe of Stone; Chartres Cathedral and the Triumph of the Medieval Mind.* Vintage.

Bell, T. (1829). *An Essay on the Origin and Progress of Pointed Architecture,* Wakeman, Dublin.

Bilson, J. (1899). *The Beginnings of Gothic Architecture, II, Norman Vaulting in England.* Journal of Royal Institute of British Architects.

Boker, H. (1998). The Bishop's Chapel of Hereford Cathedral and the Question of Architectural Copies in the Middle Ages. University of Chicago Press.

Bony, J. (1979). *The English Decorated Style, Gothic Architecture Transformed 1250-1300.* Cornell University Press.

Bradt, H. (2010). *Slow Devon and Exmoor.* Bradt Travel Guides.

Burke, E. (1757). *Philosophical Enquiry into the Origin of our Ideas of the Sublime and Beautiful.* London: R.Dodsley.

Carl, V. C. (2008). *Romanesque Art.* Ingram Publishing Services.

Coates, N.D. *A History of the Church and Parish of Boyton.*

Cowen, P. (2005). *The Rose Window: Splendour and Symbol.* Chicago: Thames and Hudson.

Demouy, P. (1995). *Rheims Cathedral.* Paris Caisse Nationale des Monuments Historiques et des Sites.

Fergusson, P. (2010). *Rievaulx Abbey.* English Heritage Guidebooks.

Glass, D. (1980). Studies on Cosmatesque Pavements. *British Archaeological Reports*.

Glazier, R. (1889). *A Manual of Historic Ornament.* Hubbard.

Gombrich, E. H. (2006). *The Story of Art.* Phaidon Press.

Goodall, J. (2003). *Whitby Abbey.* English Heritage .

Hartt, F. (2006). *History of Italian Renaissance Art.* Prentice Hall.

Herbert, R. P. (1976). *A History of the County of Gloucester: Volume 11: Bislee and Longtree Hundreds.* British History Online.

Hilton, T. (1970). *The Pre Raphaelites.* London: Thames and Hudson.

Hubert, J. (1938). *Romanesque Art.*

Lewis, M. (2002). *The Gothic Revival.* Thames and Hudson.

Martingdale, A. (1985). *Gothic Art.* Thames and Hudson.

Massimi, G. (1953). *La Chiesa di Santa Maria in Cosmedin.* Canella.

Milner, J. (1811). *Treatise on Ecclestiastical Archiecture of England during the Middle Ages.* J. Taylor.

Nicolle, D. (2004). *Crusader Castles in the Holy Land.* Osprey Publishing.

Pajares-Ayuela, P. (2002). *Cosmatesque Ornament.* W.W Norton and Co.

Pennick, N. (1990). *Mazes and Labyrinths.* Robert Hale Ltd.

Petrie, G. (1970). *The Ecclesiastical Architecture of Ireland.* Irish University Press.

Ralls, K. (2007). *Knights Templar Encyclopaedia.* Career Press.

Rickmann, A. (1819). *An Attempt to Discriminate the Styles of English Architecture.* London: Longman, Hurst, Rees, Orme, and Brown.

Ruskin, J. (1843-60). *Modern Painters.* Smith, Elder and Company.

Ruskin, J. (1849). *Seven Lamps of Architecture.* John Wiley.

Ruskin, J. (1851-1853). *Stones of Venice.* London: Smith, Elder and Co.

Simpson, O. v. (1956). *The Gothic Cathedral.* Rotledge and Kegan Paul.

Stalley, R. (1999). *Early Medieval Architecture.* OUP.

Strachan, G. (2003). *Chartres, Scared Geometry, Sacred Space.* Floris Books.

Strafford, P. (2005). *Romanesque Churches of France.* Giles de la Mare Publishers.

Taylor, C. (1967). *Cathedrals of England.* Thames and Hudson.

Vasari. (2005). *Lives of the Artists.* Dover Fine Art: History of Art.

Williams, K. (1997). *Italian Pavements, Patterns in Space.* Anchorage Press .

Wilson, C. (1992). *The Gothic Cathedral.* Thames and Hudson.

13. Scientific Bibliography.

The reader can find the scientific references at the end of Chapters 5,8 and 9.