

Waves Light and Heat

Jacob Denson

May 2, 2023

Table Of Contents

1	Physical Examples	2
1.1	Simple Harmonic Motions	2
1.2	Superposition	4
1.3	Dampening	7
1.4	Coupled Oscillators	9
1.5	Continuous Oscillation	11
1.6	Fourier Analysis	12
2	Optics	13
2.1	The Finite Speed of Light	13
2.2	Refraction	14
2.3	Huygens	15
2.4	Waves vs Particles	17

Chapter 1

Physical Examples

1.1 Simple Harmonic Motions

So many physical behaviours can be studied from the perspective of oscillation and vibration. The common feature of this phenomenon is *periodicity*, a pattern repeating itself indefinitely. Almost always these oscillations are sinusoidal, because the physical behaviour reduces (at least approximately) to the study of the *simple harmonic oscillator* $\ddot{x} = -\omega_0^2 x$, and its higher dimension variants, in which the force on an object is proportional to its spatial displacement from equilibrium. The general solution of such an equation is

$$x = A \sin(\omega_0 t + \phi) = C_1 \sin(\omega_0 t) + C_2 \cos(\omega_0 t).$$

This equation appears so often that the terms in the equation have special names:

- A is the *amplitude*.
- ϕ is the *phase shift* or *phase lag*.
- ω_0 is the *angular frequency* of oscillation.
- The frequency is then $f = \omega_0/2\pi$.

In the case of simple harmonic motion, the frequency and period of the oscillation is fixed; it is independent of the initial position and velocity of the object.

Most phenomena in physics is *nonlinear*, and thus does not precisely follow simple harmonic oscillation. Nonetheless, simple harmonic oscillation is often a useful approximation of many physical behaviours. Given an arbitrary equation

of the form $\ddot{x} = F(x)$ with $F(0) = 0$ and $F'(0) < 0$, solutions to this equation with initial conditions close to the origin, and starting with small velocity are closely approximated by solutions of the equation $\ddot{x} = F'(x_0)x$, which is a simple harmonic oscillation. Let us consider some physical examples:

- The motion of an object held on a *linear spring* follows *Hooke's Law*

$$m\ddot{x} = -kx.$$

It oscillates at an angular frequency of $\sqrt{k/m}$.

- The motion of a pendulum of length l , if θ denotes the angle of the pendulum relative to the vertical, and g the gravitational force, follows the equation

$$\ddot{\theta} = -(g/l) \sin \theta.$$

For $|\theta| \ll 1$, we have $\ddot{\theta} \approx -(g/l)\theta$, and thus we expect that, for small angles, the pendulum to oscillate at a frequency of $\sqrt{g/l}$.

- Consider a piston of mass M and surface area A holding down a fixed number of N moles of air particles by gravity. Boyle's law tells us that the pressure P pushing up on the lever is equal to NRT/V , where R is the ideal gas constant, T is the temperature of the air inside the piston, and V is the volume of the space inside the piston. Now if H is the height of the lever relative to the ground, then $V = AH$. If we assume the force pushing down on the lever is roughly independent of H with some magnitude F , i.e. the combined force of gravity pushing down on the piston, and the atmospheric pressure, then our equation of motion becomes

$$M\ddot{H} = AP_{\text{up}} - F = NRT/H - F.$$

This is a nonlinear equation in H . But near the equilibrium point $H_{\text{eq}} = NRT/F$, the equation behaves approximately like a simple harmonic oscillator. If we let $H_{\Delta} = H - H_{\text{eq}}$, then we find that for $|H_{\Delta}| \ll 1$, we have

$$M\ddot{H}_{\Delta} \approx \frac{-NRT}{H_{\text{eq}}^2} H_{\Delta} = \frac{-F^2}{NRT} H_{\Delta}.$$

The piston thus oscillates near H_{eq} at a frequency of $F/(NRTM)^{1/2}$.

One useful thing to notice is that simple harmonic motion is just the projection of a point in a plane travelling about a circle of radius A at an angular frequency of ω_0 . This is closely related to the *complexification* of the study of oscillation. A general solution of the equation $\dot{x} = -\omega_0^2 x$ over the complex numbers is an equation of the form

$$x = Ae^{i(\omega_0 t + \phi)} = C_1 e^{i\omega_0 t} + C_2 e^{-i\omega_0 t}.$$

Since our differential equation is linear, with real coefficients, the real part of any complex solution to an ordinary differential equation will continue to solve the equation, so there is really no more generality in the study of complex solutions. They are just a notational convenience, and have a computational advantage over the use of sines and cosines in that they are eigenfunctions for the derivative operator.

1.2 Superposition

Another useful tool in the study of linear oscillators is the principle of *superposition*, i.e. that the linear combination of solutions to the equation is a solution. One solution of the harmonic oscillator is $x = \cos(\omega_0 t)$, with $x_0 = 1$, and $\dot{x}_0 = 0$. Another solution is $x = \omega_0^{-1} \sin(\omega_0 t)$, with $x_0 = 0$, and $\dot{x}_0 = 1$. The principle of superposition tells us that the general solution of the harmonic oscillator with initial conditions x_0 and \dot{x}_0 is

$$x = x_0 \cos(\omega_0 t) + \frac{\dot{x}_0}{\omega_0} \sin(\omega_0 t).$$

Let us dwell on some phenomena which result from the principle of superposition.

We see the phenomenon of *interference*. Namely, the sum of two waves travelling at the same frequency and whose amplitudes have the same sign results in a wave travelling at the same frequency, but with a larger amplitude. Thus we witness *constructive interference*. A simple example of this phenomenon follows from the calculation

$$\sin(t) + \sin(t + \phi) = (1 + \cos(\phi)) \sin(t) + \sin(\phi) \sin(t + \pi/2).$$

For $\phi = 0$, we get

$$\sin(t) + \sin(t) = 2 \sin(t),$$

i.e. a wave travelling at twice the amplitude of the two original waves, so we see constructive interference. If $\phi = \pi$, then we find that

$$\sin(t) + \sin(t + \pi) = 0,$$

i.e. the waves perfectly cancel one another out, and we get deconstructive interference. One may witness this phenomenon if we consider a microphone placed between two speakers playing the same sound at the same time. The microphone will then hear that sound at twice the volume. On the other hand, if we move the microphone closer to one speaker, and thus further from the other, then the phase difference of the two speakers relative to the microphone changes, with one speaker eventually cancelling the other out.

We can also consider the superposition of N waves travelling at the same frequency but with regularly spaced phase delay by some quantity δ . Thus we consider

$$\sum_{k=0}^{N-1} \sin(\omega_0 t + k\delta).$$

This is the imaginary part of the sum of the N vectors

$$X = \sum_{k=0}^{N-1} e^{i(\omega_0 t + k\delta)} = \sum_{k=0}^{N-1} X_k.$$

Note that the points $\{X_0, \dots, X_{N-1}\}$ form consecutive vertices of a polygon. Consider the unit vector $X_A = e^{i\omega}$, where ω is equal to the average angle of the vectors $\{X_k\}$, i.e.

$$\omega = \frac{1}{N} \sum_{k=0}^{N-1} (\omega_0 t + k\delta) = \omega_0 t + \frac{N-1}{2} \delta.$$

For $0 \leq i < (N-1)/2$, each of the vectors $X_i + X_{N-i-1}$ lies in the direction of X_A , and if N is odd, then $X_{(N-1)/2} = X_A$. Thus the sum must also point in the direction of X_A . The magnitude clearly does not depend on ω_0 , so we may as well calculate the magnitude of

$$\left| \sum_{k=0}^{N-1} e^{ik\delta} \right| = \frac{|e^{Ni\delta} - 1|}{|e^{i\delta} - 1|} = \frac{2 \sin(N\delta/2)}{2 \sin(\delta/2)} = \frac{\sin(N\delta/2)}{\sin(\delta/2)}.$$

Thus we conclude that

$$X = \frac{\sin(N\delta/2)}{\sin(\delta/2)} X_A,$$

and so

$$\sum_{k=0}^{N-1} \sin(\omega_0 t + k\delta) = \frac{\sin(N\delta/2)}{\sin(\delta/2)} \sin\left(\omega_0 t + \frac{N-1}{2}\delta\right).$$

For $\delta \ll 1/N$, we have

$$\frac{\sin(N\delta/2)}{\sin(\delta/2)} \approx N - \frac{N(N+1)(N-1)}{24}\delta^2 + O((N\delta)^3)$$

Thus, if the waves all roughly have the same frequency, then the amplitude of the sum will be proportional to N . However, as δ increases beyond this range the deconstructive interference between the different phases causes the overall amplitude to decrease. When $\delta = 2\pi/N$, there is complete desconstructive interference.

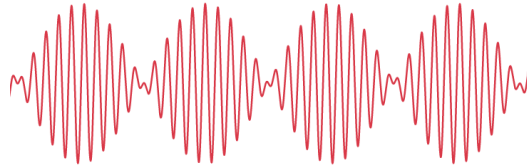
We also have the phenomenon of *beats*. Consider the cosine addition law

$$\cos(a) + \cos(b) = 2 \cos\left(\frac{a+b}{2}\right) \cos\left(\frac{a-b}{2}\right).$$

This tells us, for instance, that

$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2 \cos\left(\left(\frac{\omega_1 + \omega_2}{2}\right)t\right) \cos\left(\left(\frac{\omega_1 - \omega_2}{2}\right)t\right).$$

If $|\omega_1 - \omega_2| \ll \omega_1 + \omega_2$, then the sum of the two waves will oscillate at the average frequency of ω_1 and ω_2 , except that this oscillation will be modulated by a low frequency signal at an angular frequency of $(\omega_1 - \omega_2)/2$.



If one were to place two tuning forks together vibrating at the same amplitude, and frequencies 300 and 302 Hertz, then one would hear a wave with twice the amplitude and with frequency 301 Hertz, but modulated by a signal travelling at 1 Hertz.

1.3 Dampening

Now consider the *damped* harmonic oscillator

$$\ddot{x} = -\omega_0^2 x - b\dot{x}.$$

The values $\xi \in \mathbb{C}$ that produce a solution $x = e^{i\xi x}$ to this equation satisfy $\xi^2 = \omega_0^2 + ib\xi$, so that

$$\xi = \frac{ib}{2} \pm \sqrt{\omega_0^2 - b^2/4}.$$

Provided that the oscillator $b < 2\omega_0$, so that the equation is *lightly damped*, we see that solutions to this equation are of the form

$$x = Ae^{-(b/2)t} \cos(\omega_1 t + \phi),$$

where solutions continue to oscillate at an angular frequency of

$$\omega_1 = \sqrt{\omega_0^2 - b^2/4},$$

slightly smaller than the frequency if there was no dampening, but now also with an exponential dampening factor of $e^{-(b/2)t}$. On the other hand, if $b > 2\omega_0$, so the oscillation is *over damped*, then the solutions to the equation are of the form $x = c_1 e^{-\delta_1 t} + c_2 e^{-\delta_2 t}$, where $0 < \delta_1, \delta_2 < b/2$ are solutions to the equation

$$\delta = b/2 \pm \sqrt{b^2/4 - \omega_0^2}.$$

In particular, at the *critical damping point* with $b = 2\omega_0$, the solution is $x = (c_1 + c_2 t)e^{-(b/2)t}$, which is the amount of damping which results in the fastest return to equilibrium.

On the other hand, now let us add a *driving force*, i.e. consider the harmonic oscillator

$$\ddot{x} = F_D(t) - \omega_0^2 x - b\dot{x},$$

for some fixed driving force F_D . Let us consider a simple driving force, e.g. $F_D(t) = \cos(\omega_D t)$. It is *only* a frequency of ω_D that is being added over time, whereas the dampening force causes all frequencies to decay over time. Thus we might expect that if this equation has a *steady state solution*, i.e. a solution to which other solutions decay over time, then we might expect this solution to oscillate at a frequency ω_D . If we consider a function of the form $x(t) = A \cos(\omega_D t + \phi)$, then the equation becomes

$$A(\omega_0^2 - \omega_D^2) \cos(\omega_D t + \phi) = \cos(\omega_D t) + b\omega_D A \sin(\omega_D t + \phi).$$

Using the addition formulas for cosine and sine, and the independence of the functions $t \mapsto \cos(\omega_D t)$ and $t \mapsto \sin(\omega_D t)$, we conclude that we must have

$$A[(\omega_0^2 - \omega_D^2) \cos(\phi) - b\omega_D \sin(\phi)] = 1$$

and

$$A[b\omega_D \cos(\phi) + (\omega_0^2 - \omega_D^2) \sin(\phi)] = 0$$

Since $A = 0$ does not give a solution to the problem, we must have $A \neq 0$, and so we conclude that

$$b\omega_D \cos(\phi) + (\omega_0^2 - \omega_D^2) \sin(\phi) = 0.$$

This means that there exists $\lambda \neq 0$ such that $\cos(\phi) = \lambda(\omega_0^2 - \omega_D^2)$, and $\sin(\phi) = -\lambda b\omega_D$, where

$$\lambda^2 = \frac{1}{(\omega_0^2 - \omega_D^2)^2 + (b\omega_D)^2}$$

In particular,

$$\tan(\phi) = \frac{-b\omega_D}{\omega_0^2 - \omega_D^2}.$$

Plugging this into the first equation gives that

$$A = \frac{1}{\lambda} \frac{1}{(\omega_0^2 - \omega_D^2)^2 + (b\omega_D)^2} = \frac{1}{\sqrt{(\omega_0^2 - \omega_D^2)^2 + (b\omega_D)^2}}.$$

Thus we have found a steady state solution to the equation. For a fixed ω_0 , this steady state will have the highest amplitude when

$$\omega_D = \max \left(0, \sqrt{\omega_0^2 - b^2/2} \right).$$

This is the *resonant frequency* for the equation. The phase lag for the resonant frequency will satisfy

$$\tan(\phi) = -(2/b) \max \left(0, \sqrt{\omega_0^2 - b^2/2} \right),$$

so that there will be a slight lag relative to the driving force.

1.4 Coupled Oscillators

Finally, we consider the case of multidimensional harmonic oscillators. The simplest examples are the uncoupled oscillators, for instance, describing the x and y motion of a pendulum,

$$\ddot{\theta}_1 = -(g/l)\theta_1 \quad \text{and} \quad \ddot{\theta}_2 = -(g/l)\theta_2.$$

We can in this case just solve each equation separately, and we thus find that the solutions both oscillate at the same angular frequency $\sqrt{g/l}$. However, the *overall motion* need not be periodic, for instance, if the relative phase shift between the two motions is irrational. But if the phase shift is rational, written in simplest form p/q , then the overall motion will be periodic with angular frequency $1/q$. More general uncoupled harmonic oscillators given by

$$\ddot{x} = -\omega_1^2 x \quad \text{and} \quad \ddot{y} = -\omega_2^2 y$$

can oscillate separately in each direction with differing frequencies. The resulting motion can again be periodic or aperiodic. In case the solution is periodic, the curve of the motion is called a *Lissajous figure*.

Now we consider an example where the two motions are *coupled*. Consider two pendula of length l , attached by a linear spring of constant k , which is at equilibrium when both pendula are perfectly vertical, and thus at a distance D from one another. If θ_1 and θ_2 are the deviations from the vertical, and we assume the centre of the first pendulum lies at the origin, then the positions of the two pendula are at $(l \sin \theta_1, -l \cos \theta_1)$ and $(D + l \sin \theta_2, -l \cos \theta_2)$ respectively. Thus if we define

$$d(\theta_1, \theta_2) = \sqrt{(D + l(\sin \theta_2 - \sin \theta_1))^2 + l^2(\cos \theta_2 - \cos \theta_1)^2},$$

to be the distance between the two points, then we conclude that the equations of motion are

$$\ddot{\theta}_1 = -(g/l) \sin(\theta_1) + (k/m)(d(\theta_1, \theta_2) - D)$$

and

$$\ddot{\theta}_2 = -(g/l) \sin(\theta_2) - (k/m)(d(\theta_1, \theta_2) - D)$$

Near the equilibrium point, i.e. for small θ_1 and θ_2 , we can consider the approximate equations of motion

$$\ddot{\theta}_1 \approx -(g/l)\theta_1 + (lk/m)(\theta_2 - \theta_1) = -(g/l + lk/m)\theta_1 + (lk/m)\theta_2$$

and

$$\ddot{\theta}_2 \approx -(g/l)\theta_2 - (lk/m)(\theta_2 - \theta_1) = -(g/l + lk/m)\theta_2 + (lk/m)\theta_1.$$

To find the solutions to this equation, we rely on the symmetry of the problem. Let $\Sigma = \theta_1 + \theta_2$ and $\Delta = \theta_1 - \theta_2$. Then

$$\ddot{\Sigma} \approx -(g/l)\Sigma \quad \text{and} \quad \ddot{\Delta} \approx -(g/l + 2kl/m)\Delta.$$

These are two uncoupled equations of harmonic motion. Thus Σ oscillates at a frequency $\sqrt{g/l}$, and Δ oscillates at a frequency $\sqrt{g/l + 2kl/m}$.

This gives rise to two qualitatively distinct solutions to the equation. The first is where $\Delta = 0$ at all times, in which case

$$\theta_1 = A \cos\left(\sqrt{g/l} \cdot t + \phi\right) \quad \text{and} \quad \theta_2 = A \cos\left(\sqrt{g/l} \cdot t + \phi\right)$$

for some amplitude A and phase shift ϕ . The spring does nothing, since it always lies at equilibrium. The next equation occurs when $\Sigma = 0$, in which case

$$\theta_1 = A \cos\left(\sqrt{g/l + 2k/m} \cdot t + \phi\right) \quad \text{and} \quad \theta_2 = -A \cos\left(\sqrt{g/l + 2k/m} \cdot t + \phi\right).$$

In this case, the spring constant causes the two pendula to oscillate at a faster rate as they are pulled away and towards one another. These are *normal modes* to the equation, in the sense that all components of the equation are constant multiples of one another, and these are *all normal modes*, up to a constant multiple. The first mode is called the *sum mode*. The second, the *difference mode*. All solutions to this equation are given by a superposition of the two normal modes.

Let us consider an example of this superposition calculation. Suppose we start by lifting one of the modes, e.g. so that $\theta_1(0) = 0$, and $\theta_2(0) = D$ for some $D > 0$. Then $\Sigma = \theta_1 + \theta_2$ has initial position D and velocity zero, and so by the equations above, we conclude that

$$\Sigma = D \cos\left(\sqrt{g/l} \cdot t\right)$$

Similarly, we calculate that

$$\Delta = D \cos\left(\sqrt{g/l + 2k/m} \cdot t\right).$$

Thus

$$\theta_1 = \frac{\Sigma + \Delta}{2} = (D/2) \left(\cos \left(\sqrt{g/l} \cdot t \right) + \cos \left(\sqrt{g/l + 2k/m} \cdot t \right) \right)$$

and

$$\theta_2 = \frac{\Sigma - \Delta}{2} = (D/2) \left(\cos \left(\sqrt{g/l} \cdot t \right) - \cos \left(\sqrt{g/l + 2k/m} \cdot t \right) \right).$$

In particular, we can see from these equations that there is a beat phenomenon at play here, namely

$$\theta_1 = D \cos \left(\left(\frac{\sqrt{g/l} + \sqrt{g/l + 2k/m}}{2} \right) t \right) \cos \left(\left(\frac{\sqrt{g/l + 2k/m} - \sqrt{g/l}}{2} \right) t \right).$$

If $g/l \gg 2k/m$, so that the sum and difference oscillate at roughly the same rate, then we conclude that

$$\theta_1 = D \cos \left(\frac{g^{1/2}}{l^{1/2}} \cdot t \right) \cos \left(\frac{kl^{1/2}}{2mg^{1/2}} \cdot t \right).$$

Thus the energy is transferred from one mass over to the other over. This is the first example of a *travelling wave*.

1.5 Continuous Oscillation

TODO: Motivation for wave equation as obtained by a limit of coupled oscillators.

Let us begin with the wave equation on a clamped spring, i.e. the equation

$$\frac{\partial^2 u}{\partial t^2} = (1/v^2) \frac{\partial^2 u}{\partial x^2}$$

where $u : [0, L] \times \mathbb{R}_+ \rightarrow \mathbb{R}$, and $u(0, t) = u(L, t) = 0$ for all $t \in \mathbb{R}$. The normal modes for this equation are solutions of the form $u(x, t) = f(x)g(t)$. The equation then becomes

$$f(x)g''(t) = (1/v^2)g(t)f''(x)$$

or equivalently,

$$\frac{f''(x)}{f(x)} = v^2 \frac{g''(t)}{g(t)}.$$

It follows that there is a common value γ such that $f'' = -\gamma f$ and $g'' = (-\gamma/v^2)g$. If $\gamma > 0$, $\gamma = \lambda^2$, then $f(x) = \sin(\lambda t + \phi)$. The condition that $f(0) = f(L) = 0$ implies that $\phi = 0$, and $\lambda = (\pi n/L)$ for some integer $n \geq 0$. We therefore have modes of the form

$$u(x, t) = A \sin((\pi n/L)x) \cos((\pi n/Lv)t + \phi).$$

If $\gamma < 0$, then $\gamma = -\lambda^2$, and the conditions that $f'' = \lambda^2 f$ with $f(0) = f(L) = 0$ are then impossible to satisfy. Thus there does not exist any solutions corresponding to this value. Thus the description above gives all normal modes to the wave equation for a clamped string. The solution have temporal period $\omega = 2Lv/n$, and wavelength $\lambda = 2L/n$. The wave number k is then $2\pi/\lambda = \pi n/L$, and counts the number of radians of a wave in a given unit of space.

1.6 Fourier Analysis

Joseph Fourier's main contribution to analysis were two simple principles: all solutions to the wave equation are the superposition of normal modes, and a method to find the superposition given fixed initial conditions. Thus we expect a general solution of the wave equation on a clamped string is of the form

$$u(x, t) = \sum_{n=0}^{\infty} A_n \sin\left(\frac{\pi n}{L} \cdot x\right) \cos\left(\frac{\pi n}{Lv} \cdot t + \phi_n\right).$$

Chapter 2

Optics

How do we explain *refraction* and *diffraction*, and the *finite speed of light*.

2.1 The Finite Speed of Light

Before the 1700s, it was a very difficult problem to determine the longitude of a ship at sea, and in 1598 Philip III of Spain had offered a prize if a scientist could find a reliable method to solve this task. We note that the longitude is the position on Earth relative to the Earth's axis of rotation, and thus measuring this longitude is equivalent to measuring the *time of day* on Earth, provided one has some other time keeping device which can keep track of the time of day at another fixed point on earth. Galileo proposed establishing this time by measuring the times of the eclipses of Jupiter, a method not significantly improved upon until the invention of mechanical clocks two centuries later. The difficulty of performing astronomy on a boat made this technique impractical at sea, but it could work on land. Italian astronomer Cassini began making tables predicting when an eclipse would be visible from a particular location. As director of the new Royal Observatory in Paris, Cassini sent astronomer Jean Picard, together with his assistant Ole Römer, to Tycho Brahe's observatory, based on the island of Uraniborg near Copenhagen, to measure the times of eclipses of Jupiters moons.

Ole Römer was observing Io, one of the innermost moons of Jupiter, when he noticed an interesting phenomenon. The plane of orbit of Io about Jupiter is very close to the plane of orbit of Jupiter about the sun, which means that, in each orbit of Io about Jupiter, which occurs about once every 42.5 hours, an eclipse occurs as Io enters the shadow of Jupiter relative to the sun. Entering the shadow, Io

disappears, what is called an *immersion*. Exiting the shadow, Io reappears, in what is called an *emergence*. It is not possible from the perspective of earth to see the immersion and emergence of a single eclipse, because the Earth is closer to the sun than Jupiter, and so Jupiter will block the view of Io. If the sun lies directly between the Earth and Jupiter, neither the immersion nor the emergence will be visible. This is called the *point of opposition*.

For about four months after the point of opposition, it is possible to view the emergences of Io, while for about four months before, it is possible to view the immersions. For the remaining five to six months of the year, one cannot see the eclipses at all because the sun is too close to Jupiter.

The key phenomenon Rømer observed was that the time between eclipses was *not constant* from the perspective of the earth. He was fairly sure the time that Io took to orbit Jupiter should be constant. The fact that the time between eclipses grew as Jupiter moved *away* from the earth, and shrank as Jupiter moved *towards* the earth, allowed Römer to believe this phenomenon was the result of the *finite speed of light*. Römer's theory was controversial, and he never convinced Cassini to completely accept it. But it gained support of Christiaan Huygens and Isaac Newton, and the phenomenon's further use in explaining Stellar Abberation, given by James Bradley in 1729, lead to increasing credence in the phenomenon.

2.2 Refraction

Greatly prized by all man is the diamond and many are the joys which similar treasures bring, such as precious stones and pearls... but he who, on the other hand, prefers the knowledge of unusual phenomena to these delights, he will, I hope, have no less joy in a new sort of body, namely, a transparent crystal, brought to use from Iceland, which perhaps is one of the greatest wonders nature has produced.

In 1669, Danish professor of mathematics and medicine, Erasmus Bartholinus, managed to get ahold of some crystals, obtained in Iceland, from a sailor. These crystals we now call *Iceland Spar*. When viewing small objects through these crystals, Bartholinus found that these objects appeared to duplicate. Bartholinus

discovered why this was the case, a result of *refraction*, or the redirection of light as it passes through a medium. Such a procedure was previously known, the law being described by Dutch mathematician and physicist Willebrord Snel van Royen in the early sixteen hundred, what we now call *Snell's Law*. But Bartholinus had found a medium which, when a ray of light passes through it, splits apart into two separate rays. One of the rays resembles the usual one given by Snell's Law, which Bartholinus called the *ordinary ray*. The other, the *extraordinary ray*. This is the cause for an object to appear to double when viewed through Iceland Spar.

2.3 Huygens

Around the time when Römer was discovering the finite speed of light, Huygens had taken ill, and had returned to the Hague from Paris, staying there for two years to recover. It was during this recovery that Huygens formulated the *wave theory of light*. Twelve years later, he published an account of this theory in his famous book *Traité de la Lumière*. This book was the first attempt to find a common explanation for the different properties of light that had been discovered. Huygen's showed that the essential features of light: it's tendency to travel in straight lines, it's behaviour under refraction, both simple and double, it's behaviour under reflection, and the finite speed of propagation of light, all follow from simple assumptions about the wave-like properties of light.

Let us now study Huygen's main principle:

Each particle in which a wave spreads, ought not to communicate it's motion only to the next particle which is in the straight line drawn from the luminous point, but that it also imparts some of it necessarily to all the others which touch it and which oppose themselves to it's movement. So it arises that around each particle there is made a wave of which the particle is it's common centre.

How do rays of light arise from this principle? The primary pulse of light at a given time is equal to the superposition of the secondary pulses of light

In terms of our modern conception, Huygen's theory is rudimentary. Huygen's was familiar with how sound travelled, i.e. via the vibration of matter. So Huygen's postulated that light travelled in the same way. But earlier experiments due to Torricelli had shown that light was transmitted even in a space free of matter. So Huygen's had to postulate the existence of an imperceptible medium, *ether*, a

collection of tiny particles, and that light transmits through the ether by the collisions of this ether. Huygen's had previously formulated a theory of collision in mechanics, and showed that, if light travelled in the way postulated, it would lead to light travelling in a finite speed.

Another key insight into the behaviour of light was the concept of a *wavefront*. Huygen's explained his theory by stating that light is emitted from a source of light in an expanding

, spherical wave. Each point on this sphere becomes a new source of light, emitted in a sphere around itself, though this emittance interacts with the other wavefronts that have been emitted, i.e. the emittance back towards the source of light destructively interfering with the smaller spherical waves that have been emitted from the source moments after the initial emittance. And emittances transverse to the wavefront also interact destructively with the transverse emittances from the other points on the concentric sphere. This latter property gives an explanation for why light tends to travel in straight lines. Under the assumption that light travels at different speeds through other mediums than ether, Huygens was able to provide an explanation for Snell's Law given his theory. He was *also* able to explain the problem of double refraction by postulating that, when refraction through Iceland Spar, light produces both *ordinary*, spherical emittances, but also *extraordinary*, ellipsoidal emittances.

The success of the wave theory of light was so successful that the *corpuscular* theory of light, i.e. that light travels as small particles through space, fell out of favor for several centuries, until it began to be studied again in the 20th century.

Huygen's theory was expanded upon by British physician Thomas Young, who noticed that light travelling from two different sources interacts with one another, destructive interference occurs in a way that cannot be explained using Huygen's theory of propagation. Namely, Huygen had noticed that, when a ray of light was diffracted into two rays using Iceland Spar, if each of the rays of light is incident to another crystal of Iceland Spar, the same phenomenon does *not occur*. Thomas Young explained this by postulating that light oscillated in a direction *transverse* to their direction of propagation, in contrast to sound waves, which oscillate *longitudinally*. TODO: WHY DOES THIS EXPLAIN POLARIZATION. This theory was later confirmed by Maxwell, who gave an explanation of light as emerging from transverse oscillations of electric and magnetic fields.

Seventeen years after Young, French engineer and scientist Augustin Fresnel synthesized Huygen's principle with the principle of interference. He was able to explain the phenomenon of *diffraction*, discovered by Grimaldi in 1660, in which the path of some rays of light deviate from straight lines when they touch the

edges of apertures and boundaries of bodies.

2.4 Waves vs Particles

Broadly speaking, we can divide the theories of light into two categories: *corpuscular*, or *emission* theories of light characterized light as a sequence of individual particles subject to forces from material bodies. *Wave theorists* characterized light as a spreading disturbance through a universally present medium they called the ether. The wave theory began to dominate in the 1800s. Buchwald suggests this was not due to the failure of the corpuscular theory in explaining new phenomena, but instead because the wave theory of light was able to give quantitative results where the corpuscular theory failed, which was highly valued at the time. By the 1830s, the corpuscular theory had almost vanished from modern publications.

Many scientists found it very difficult to think in terms of 'wavefronts', i.e. abandoning the intuitive ideas of isolated rays of light.