Revision Guide for Chapter 7

Contents

Revision Checklist

Revision Notes

Fermat's least time principle	. 4
Interference of photons	. <u>5</u>
Quantum behaviour	<u> 5</u>
Photon	. 6
Electron diffraction	. 7
Intensity	. 8
Phase and phasors	. 9
Probability	<u>10</u>
Superposition	10
Diffraction	11
Interference	12
Path difference	14
Accuracy and precision	14
Systematic error	16
Uncertainty	16

Summary Diagrams

A path contributes an arrow	<u>18</u>
Finding probabilities	19
Mirror: contributions from different paths	20
Photons and refraction	21
Focusing photons	22
Restricting photons	23
Evidence for photons	24

Revision Checklist

Back to list of Contents

I can show my understanding of effects, ideas and relationships by describing and explaining:

how phasor arrows come to line up for paths near the path that takes the least time			
Revision Notes: Fermat's least time principle			
Summary Diagrams: <u>A path contributes an arrow</u> , <u>Finding probabilities</u>			
how phasor arrows 'lining up' and 'curling up' account for straight-line propagation, reflection, refraction, focusing, diffraction and interference (superposition) of light			
Revision Notes: Fermat's least time principle, interference of photons			
Summary Diagrams: <u>Mirror: contributions from different paths</u> , <u>Photons and refraction</u> , <u>Focusing</u> <u>photons</u> , <u>Restricting photons</u>			
that the probability of arrival of a quantum is determined by graphical addition of arrows representing the phase and amplitude associated with each possible path			
Revision Notes: <u>quantum behaviour</u> , <u>photon</u>			
Summary Diagrams: <u>A path contributes an arrow</u> , Finding probabilities			
evidence for random arrival of photons			
Revision Notes: <u>photon</u>			
evidence for the relationship $E = hf$			
Summary Diagrams: Evidence for photons			
evidence from electron diffraction that electrons show quantum behaviour			
Revision Notes: electron diffraction			

I can use the following words and phrases accurately when describing effects and observations:

frequency, energy, amplitude, phase, superposition, intensity, probability	
Revision Notes: intensity, phase and phasors, probability, superposition	
path difference, interference, diffraction	
Revision Notes: diffraction, interference, path difference	

I can interpret:

diagrams illustrating how paths contribute to an amplitude

Summary Diagrams: <u>A path contributes an arrow</u>, <u>Finding probabilities</u>, <u>Mirror: contributions</u> <u>from different paths</u>, <u>Photons and refraction</u>, <u>Focusing photons</u>, <u>Restricting photons</u>

I can calculate:

the energy of a photon using the relationship E = hf

the de Broglie wavelength of an electron using the relationship $\lambda = h/mv$

Revision Notes: electron diffraction

Summary Diagrams: Evidence for photons

I can show my ability to make better measurements by:

measuring the Planck constant h

Revision Notes: accuracy and precision, systematic error, uncertainty

I can show an appreciation of the growth and use of scientific knowledge by:

commenting on the nature of quantum behaviour

Revision Notes: guantum behaviour

Revision Notes

Back to list of Contents

Fermat's least time principle

Fermat had the idea that light always takes the 'quickest path' – the path of least time. You see below a number of paths close to the straight line path from source to detector. A graph of the time for each path has a minimum at the straight line path.





Near the minimum the graph is almost flat. This is a general property of any minimum (or maximum). That is, near the minimum the times are all almost the same.

The amount by which a photon phasor turns along a path is proportional to the time taken along the path. Thus, for paths near the minimum all the phasors have turned by more or less the same amount. They are therefore all nearly in phase with one another. They 'line up', giving a large resultant amplitude. This is the reason why Fermat's idea works. Only for paths very close to the path of least time is there a large probability for photons to arrive. The photons try all paths, but all except the paths close to the least-time path contribute very little to the probability to arrive.

The idea explains photon propagation in a straight line, reflection and refraction.

Back to Revision Checklist

Interference of photons

If light from a narrow source is passed through a pair of closely spaced slits onto a screen, a pattern of interference fringes is seen on the screen. Photons have two paths to the screen, and must be thought of as trying both. There is a phasor quantity (amplitude and phase) associated with each path. Since the paths are nearly equal in length the magnitude of the amplitudes for each path is similar, but the phases differ.

The phasor for a path rotates at the frequency of the light. The phase difference between two paths is proportional to the path difference.

At points on the screen where the phasors have a phase difference of half a turn, that is 180°, dark fringes are observed because the phasors added 'tip to tail' give zero resultant. Where the phasors are in phase (zero or an integer number of turns difference) there are bright fringes. The intensity on the screen is proportional to the square of the resultant phasor.





Back to Revision Checklist

Quantum behaviour

Quantum behaviour can be described as follows:

- 1. Particles are emitted and absorbed at distinct space-time events.
- 2. Between these events there are in general many space-time paths.
- 3. The presence of all possible space-time paths influences the probability of the passage of a particle from emission to absorption.
- 4. Each path has an associated amplitude and phase, represented by a rotating phasor arrow.
- 5. The phasor arrows for all possible paths combine by adding 'tip to tail', thus taking account of amplitude and phase.
- 6. The square of the amplitude of the resultant phasor is proportional to the probability of the emission event followed by the absorption event.

A photon, although always exchanging energy in discrete quanta, cannot be thought of as travelling as a discrete 'lump' of anything. Photons (or electrons) arriving at well-defined places and times (space-time events) are observable. But their paths between emission and detection are not well-defined. Photons are not localised in time and space between emission and absorption. They must be thought of as trying all possible paths, all at once.

In the propagation of photons from source to detector across an empty space, the probability of arrival of photons anywhere but close to the straight line from source to detector is very low. This is because, not in spite of, the many other possible paths. The quantum amplitudes for all these paths add to nearly zero everywhere except close to the straight line direction.

As soon as the space through which the light must go is restricted, by putting a narrow slit in the way, the probability for photons to go far from the direction of straight line propagation increases. This is because the cancelling effect of other paths has been removed.

The net effect is that the narrower one attempts to make the light beam, the wider it spreads.

Back to Revision Checklist

Photon

Electromagnetic waves of frequency f are emitted and absorbed in quanta of energy E = h f, called photons.

Photons are quantum objects, exhibiting quantum behaviour. They are emitted and absorbed at random. Their intensity is given by the probability of arrival. But this probability is the square of a phasor amplitude found by combining phasor arrows for all possible paths from emission to detection. In this sense, the photon cannot be thought of as localised on any particular path from emitter to detector. Rather, photons 'try all paths'.

For a point source of photons emitting energy at a rate W, the number of photons per second emitted by the source = W / h f since each photon carries energy h f.

Random arrival of photons

The random nature of the arrival of photons is most easily seen using high energy gamma ray photons, which can be heard arriving randomly in a Geiger counter.

The pictures below illustrate the random arrival of photons. They are constructed as if made by collecting more and more photons to build up the picture. Where the picture is bright the probability of arrival of a photon is high. Where it is dark, the probability is low. You can see how the random arrival, governed by these probabilities, builds up the final picture.



Emission of photons from atoms

When an electron moves from a higher to a lower energy level in an atom, it loses energy which can be released as a photon of electromagnetic energy. Since the energy of a photon = h f, then if an electron transfers from an energy level E_2 to a lower energy level E_1 , the energy of the photon released = $h f = E_2 - E_1$.

In this way, the existence of sharp energy levels in atoms gives rise to sharp line spectra of the light they emit.

Back to Revision Checklist

Electron diffraction

Electron diffraction is the diffraction of a beam of electrons by a regular arrangement of atoms.

Possible paths for electrons being scattered by successive layers of atoms differ in length, and so in the phase of the associated phasor. The phasors for paths going via successive layers of atoms only combine to give a large amplitude in certain directions.

If the quantum behaviour of a free electron is thought of as associated with a wave motion, the wavelength of the waves is the de Broglie wavelength

 $\lambda = \frac{h}{p}$

where *p* is the momentum of the electron and *h* is the Planck constant.



Electron diffraction

Back to Revision Checklist

Intensity

The intensity of a wave is the energy per second carried by the waves and incident normally on unit area of surface.

The SI unit of intensity is the joule per second per square metre (J s⁻¹ m⁻²) which is the same as the watt per square metre (W m⁻²)

The intensity of radiation from a point source varies with distance from the source in accordance with the inverse square law, provided the radiation is not absorbed by the substance it travels through.



The inverse square law

Consider a point source that radiates energy at a rate of *W* joules per second. At distance *r* from the source in an non-absorbing substance, all the radiation from the source passes through the surface of a sphere of area $4\pi r^2$, where the source is at the centre of the sphere. Hence the intensity *I* = the energy per second incident on unit area of the sphere = $W / 4\pi r^2$.

The intensity of a wave is proportional to the square of its amplitude. A single particle oscillating in simple harmonic motion at frequency *f* with an amplitude *A* has a maximum speed of $2\pi f A$ and therefore a maximum kinetic energy of $\frac{1}{2}m(2\pi f A)^2$. Thus the intensity is proportional to the square of the amplitude.

Relationships

$$I = \frac{W}{4\pi r^2}$$

Back to Revision Checklist

Phase and phasors

'Phase' refers to stages in a repeating change, as in 'phases of the Moon'.

The phase difference between two objects vibrating at the same frequency is the fraction of a cycle that passes between one object being at maximum displacement in a certain direction and the other object being at maximum displacement in the same direction.

Phase difference is expressed as a fraction of one cycle, or of 2π radians, or of 360° .

Phasors are used to represent amplitude and phase in a wave. A phasor is a rotating arrow used to represent a sinusoidally changing quantity.

Suppose the amplitude *s* of a wave at a certain position is $s = a \sin(2\pi ft)$, where *a* is the amplitude of the wave and *f* is the frequency of the wave. The amplitude can be represented as the projection onto a straight line of a vector of length *a* rotating at constant frequency *f*, as shown in the diagram. The vector passes through the +*x*-axis in an anticlockwise direction at time t = 0 so its projection onto the *y*-axis at time *t* later is $a \sin(2\pi ft)$ since it turns through an angle $2\pi ft$ in this time.

Phasors can be used to find the resultant amplitude when two or more waves superpose. The phasors for the waves at the same instant are added together 'tip to tail' to give a resultant phasor which has a length that represents the resultant amplitude. If all the phasors add together to give zero resultant, the resultant amplitude is zero at that point.



Generating a sine wave

Back to Revision Checklist

Probability

Probability has to do with uncertainty, with randomness and with quantum effects. Probability is a measure of the chance of one of a number of possible things happening.

Random events, such as the emission of an alpha particle from a radioactive nucleus, are more, or are less, likely to happen. The probability of emission in a short time interval can be estimated from the number of emissions taken over a long period of time.

The probability of the random arrival of a photon at a point in a beam of light is proportional to the intensity of the light. The intensity is proportional to the square of the classical wave amplitude, or in quantum theory, to the square of the resultant phasor amplitude for all possible paths.

Back to Revision Checklist

Superposition

When two or more waves meet, their displacements superpose.

The principle of superposition states that when two or more waves overlap, the resultant displacement at a given instant and position is equal to the sum of the individual displacements at that instant and position.

In simple terms, where a wave crest meets another wave crest, the two wave crests pass through each other, forming a 'super crest' where and when they meet. If a wave trough meets another wave trough, they form a 'super trough' where they meet. In both cases, the waves reinforce each other to increase the displacement momentarily. If a wave crest meets a wave trough, the waves cancel each other out momentarily.

An example of superposition is the interference pattern produced by a pair of dippers in a ripple tank, as shown below.

Interference



Back to Revision Checklist

Diffraction

Diffraction is the spreading of waves after passing through a gap or past the edge of an obstacle.

The spreading increases if the gap is made narrower or if the wavelength of the waves is increased.

Diffraction



Monochromatic light passing through a single narrow slit produces a pattern of bright and dark fringes. Intensity **minima** are observed at angles θ given by the equation $d \sin \theta = n\lambda$, where d is the gap width, n is a positive integer and θ is the angle between the incident direction and the direction of diffraction.



phasors add to zero

If the distance across the gap is taken to be a large number of equally spaced point sources, **1**, **2**, **3**, etc, the phasor due to **1** will be a certain fraction of a cycle behind the phasor due to **2**, which will be the same fraction behind the phasor due to **3** etc. The resultant phasor is therefore zero at those positions where the tip of the last phasor meets the tail of the first phasor.

The path difference between the top and bottom of the slit is $d \sin \theta$. If this path difference is equal to a whole number of wavelengths, $n\lambda$, and if the last and first phasors join tip to tail minima occur when

 $d\sin\theta = n\lambda$

For small angles, sin $\theta = \theta$ giving an angular width 2λ / d for the central maximum.

Back to Revision Checklist

Interference

When waves overlap, the resultant displacement will be equal to the sum of the individual displacements at that point and at that instant (if the waves superpose linearly).

Interference is produced if waves from two coherent sources overlap or if waves from a single source are divided and then reunited.



Interference using sound waves can be produced by two loudspeakers connected together to an oscillator. If you move about in the waves overlap you will detect points of reinforcement (louder) and of cancellation (quieter).

Another way to produce interference is to divide the waves from one source and then recombine them. The diagram below shows this being done for microwaves, sending part of the wave along one path and part along another. The receiver gives a minimum response when the paths differ by half a wavelength.



Division of amplitude

Other examples of interference include the 'blooming' of camera lenses, and the colours of oil films and soap bubbles.

Path difference

The path difference between two waves will determine what happens when they superpose.

If the path difference between two wavefronts is a whole number of wavelengths, the waves reinforce.

If the path difference is a whole number of wavelengths plus one or minus half of a wavelength, the waves cancel.

The importance of a path difference is that it introduces a time delay, so that the phases of the waves differ. It is the difference in phase that generates the superposition effects.

Back to Revision Checklist

Accuracy and precision

A measurement is accurate if it is close to the true value. A measurement is precise if values cluster closely, with small uncertainty.

A watch with an accuracy of 0.1% could be up to five minutes astray within a few days of being set. A space probe with a trajectory accurate to 0.01 % could be more than 30 km off target at the Moon.

Think of the true value as like the bullseye on a target, and measurements as like arrows or darts aimed at the bullseye.

Uncertainty and systematic error

Think of measurements as shots on a target. Imagine the 'true value' is at the centre of the target



An accurate set of measurements is like a set of hits that centre on the bullseye. In the diagram above at the top, the hits also cluster close together. The uncertainty is small. This is a measurement that gives the true result rather precisely.

On the left, the accuracy is still good (the hits centre on the bullseye) but they are more scattered. The uncertainty is higher. This is a measurement where the average still gives the true result, but that result is not known very precisely.

On the right, the hits are all away from the bullseye, so the accuracy is poor. But they cluster close together, so the uncertainty is low. This is a measurement that has a systematic error, giving a result different from the true result, but where other variations are small.

Finally, at the bottom, the accuracy is poor (systematic error) and the uncertainty is large.

A statement of the result of a measurement needs to contain two distinct estimates:

- 1. The best available estimate of the value being measured.
- 2. The best available estimate of the range within which the true value lies.

Note that both are statements of belief based on evidence, not of fact.

For example, a few years ago discussion of the 'age-scale' of the Universe put it at 14 plus or minus 2 thousand million years. Earlier estimates gave considerably smaller values but with larger ranges of uncertainty. The current (2008) estimate is 13.7 ± 0.2 Gy. This new value lies within the range of uncertainty for the previous value, so physicists think the estimate has been improved in precision but has not fundamentally changed.

Fundamental physical constants such as the charge of the electron have been measured to an astonishing small uncertainty. For example, the charge of the electron is $1.602\ 173\ 335 \times 10^{-19}$ C to an uncertainty of $0.000\ 000\ 005 \times 10^{-19}$ C, better than nine significant figures.

There are several different reasons why a recorded result may differ from the true value:

1. **Constant systematic bias**, such as a zero error in an instrument, or an effect which has not been allowed for.

Constant systematic errors are very difficult to deal with, because their effects are only observable if they can be removed. To remove systematic error is simply to do a better experiment. A clock running slow or fast is an example of systematic instrument error. The effect of temperature on the resistance of a strain gauge is an example of systematic experimental error.

2. **Varying systematic bias**, or drift, in which the behaviour of an instrument changes with time, or an outside influence changes.

Drift in the sensitivity of an instrument, such as an oscilloscope, is quite common in electronic instrumentation. It can be detected if measured values show a systematic variation with time. Another example: the measured values of the speed of light in a pipe buried in the ground varied regularly twice a day. The cause was traced to the tide coming in on the nearby sea-shore, and compressing the ground, shortening the pipe a little.

3. Limited resolution of an instrument. For example the reading of a digital voltmeter may change from say 1.25 V to 1.26 V with no intermediate values. The true potential difference lies in the 0.01 V range 1.25 V to 1.26 V.

All instruments have limited resolution: the smallest change in input which can be detected. Even if all of a set of repeated readings are the same, the true value is not exactly equal to the recorded value. It lies somewhere between the two nearest values which can be distinguished.

4. Accidental momentary effects, such as a 'spike' in an electrical supply, or something hitting the apparatus, which produce isolated wrong values, or 'outliers'.

Accidental momentary errors, caused by some untoward event, are very common. They can often be traced by identifying results that are very different from others, or which depart from a general trend. The only remedy is to repeat them, discarding them if further measurements strongly suggest that they are wrong. Such values should never be included in any average of measurements, or be used when fitting a line or curve.

5. **Human errors**, such as misreading an instrument, which produce isolated false recorded values.

Human errors in reading or recording data do occur, such as placing a decimal point wrongly, or using the wrong scale of an instrument. They can often be identified by noticing the kinds of mistake it is easy to make. They should be removed from the data, replacing them by repeated check observations.

6. **Random fluctuations**, for example noise in a signal, or the combined effect of many unconnected minor sources of variation, which alter the measured value unpredictably from moment to moment.

Truly random variations in measurements are rather rare, though a number of unconnected small influences on the experiment may have a net effect similar to random variation. But because there are well worked out mathematical methods for dealing with random variations, much emphasis is often given to them in discussion of the estimation of the uncertainty of a measurement. These methods can usually safely be used when inspection of the data suggests that variations around an average or a fitted line or curve are small and unsystematic. It is important to look at visual plots of the variations in data before deciding how to estimate uncertainties.

Back to Revision Checklist

Systematic error

Systematic error is any error that biases a measurement away from the true value.

All measurements are prone to systematic error. A systematic error is any biasing effect, in the environment, methods of observation or instruments used, which introduces error into an experiment. For example, the length of a pendulum will be in error if slight movement of the support, which effectively lengthens the string, is not prevented, or allowed for.

Incorrect zeroing of an instrument leading to a **zero error** is an example of systematic error in instrumentation. It is important to check the zero reading during an experiment as well as at the start.

Systematic errors can change during an experiment. In this case, measurements show trends with time rather than varying randomly about a mean. The instrument is said to show **drift** (e.g. if it warms up while being used).

Systematic errors can be reduced by checking instruments against known standards. They can also be detected by measuring already known quantities.

The problem with a systematic error is that you may not know how big it is, or even that it exists. The history of physics is littered with examples of undetected systematic errors. The only way to deal with a systematic error is to identify its cause and either calculate it and remove it, or do a better measurement which eliminates or reduces it.

Back to Revision Checklist

Uncertainty

The uncertainty of an experimental result is the range of values within which the true value may reasonably be believed to lie. To estimate the uncertainty, the following steps are needed.

1. Removing from the data **outlying** values which are reasonably suspected of being in serious error, for example because of human error in recording them correctly, or because of an unusual external influence, such as a sudden change of supply voltage. Such values should not be included in any later averaging of results or attempts to fit a line or curve to relationships between measurements.

- 2. Estimating the possible magnitude of any systematic error. An example of a constant systematic error is the increase in the effective length of a pendulum because the string's support is able to move a little as the pendulum swings. The sign of the error is known (in effect increasing the length) and it may be possible to set an upper limit on its magnitude by observation. Analysis of such systematic errors points the way to improving the experiment.
- Assessing the **resolution** of each instrument involved, that is, the smallest change it can detect. Measurements from it cannot be known to less than the range of values it does not distinguish.
- 4. Assessing the magnitude of other small, possibly random, unknown effects on each measured quantity, which may include human factors such as varying speed of reaction. Evidence of this may come from the spread of values of the measurement conducted under what are as far as possible identical conditions. The purpose of repeating measurements is to decide how far it appears to be possible to hold conditions identical.
- 5. Determining the combined effect of possible **uncertainty** in the result due to the limited resolution of instruments (3 above) and uncontrollable variation (4 above).

To improve a measurement, it is essential to identify the largest source of uncertainty. This tells you where to invest effort to reduce the uncertainty of the result.

Having eliminated accidental errors, and allowed for systematic errors, the range of values within which the true result may be believed to lie can be estimated from (a) consideration of the resolution of the instruments involved and (b) evidence from repeated measurements of the variability of measured values.

Most experiments involve measurements of more than one physical quantity, which are combined to obtain the final result. For example, the length L and time of swing T of a simple pendulum may be used to determine the local acceleration of free fall, g, using

$$T=2\pi\sqrt{\frac{L}{g}}$$

so that

$$g=\frac{4\pi^2 L}{T^2}.$$

The range in which the value of each quantity may lie needs to be estimated. To do so, first consider the resolution of the instrument involved – say ruler and stopwatch. The uncertainty of a single measurement cannot be better than the resolution of the instrument. But it may be worse. Repeated measurements under supposedly the same conditions may show small and perhaps random variations.

If you have repeated measurements, 'plot and look', to see how the values vary. A simple estimate of the variation is the spread = $\pm \frac{1}{2}$ range.

A simple way to see the effect of uncertainties in each measured quantity on the final result is to recalculate the final result, but adding or subtracting from the values of variables the maximum possible variation of each about its central value. This is pessimistic because it is unlikely that 'worst case' values all occur together. However, pessimism may well be the best policy: physicists have historically tended to underestimate uncertainties rather than overestimate them. The range within which the value of a quantity may reasonably be believed to lie may be reduced somewhat by making many equivalent measurements, and averaging them. If there are *N* independent but equivalent measurements, with range *R*, then the range of their average is likely to be approximately *R* divided by the factor \sqrt{N} . These benefits are not automatic, because in collecting many measurements conditions may vary.

Summary Diagrams

Back to list of Contents

A path contributes an arrow



Finding probabilities

Many path arrows can be summed to give an amplitude. The square of the amplitude gives a probability.



Mirror: contributions from different paths Phasors for paths near the middle of the mirror line up. Phasors for paths away from the middle curl up.



Back to Revision Checklist

Photons and refraction

Light refracted at a surface takes the path of least time.

Refraction – explorations through a surface					
S					
	Place the source, detector and surface.	Light appears to travel more slowly below the surface, so we reduce the speed of the exploring phasor. The frequency is unchanged.			
S B B B	Choose a photon frequency and define a characteristic set of paths going via the surface.	The trip time is calculated in two parts: above and below the surface. The phasor spins at the same frequency. The time taken determines the angle through which it has turned.			
	Explore each path by moving a phasor along the path. Start with a fresh phasor each time and record the final arrow. Record these arrows in order.				
4 - A	Obtain and square the amplitude to find the chance that a photon ends up at this detector.	Refraction occurs – quantum mechanics says that there is a large chance that the photon be found at the detector. Most of the final amplitude comes from paths just to the right of the straight line path; paths close to the path of least time.			
near least time path far from least time path	Explore more paths to get more arrows, a clearer picture and greater accuracy.	The pattern is clear. Most of the amplitude comes from the paths close to the path that takes least time, only a little from those far out.			

Focusing photons

Curving a mirror shortens paths going via the edges. When all paths to the focus take the same time, photons are very likely to arrive there.



Back to Revision Checklist

Restricting photons



Evidence for photons

Here are three pieces of evidence, each supporting the view that light carries energy in certain sized packets.

Light emitted by LEDs

The frequency of the emitted light is compared with the potential difference needed to emit light of that colour.



Light ejecting electrons from metals

The frequency of the absorbed light is compared with the potential difference needed to stop electrons emitted from clean metal surfaces.



Light emitted or absorbed by atoms

The frequency of the light is compared with the possible energy levels within atoms.



Back to Revision Checklist

Back to list of Contents