The Discrete Fourier Transform

and its threat to modern day cryptography.

Maximilian Powers
International School of Nice

(Extended Essay)

(Wordcount:)

(Dated: October 15, 2018)

Contents

I. Introduction

a. History In 1807, Joseph Fourier wrote the paper *Mémoire sur la propagation de la chaleur dans les corps solide*, which explains how sinusoids can be used to model heat flow, and was the first known method to solve the heat equation, a partial differential equation. However its uses expanded far beyond that, and many of the transforms developed from Fourier Series are used everyday in electrical, sound, and structural engineering, to compression, and most recently in solving the factorisation of large numbers problem, and therefore breaking most of modern day public key encryption.

b. Summary. In this extended essay, the concepts of Fourier Series will be touched on first, and will end by representing the function e^x between $-\pi$ and π as a sum of sines and cosines a long with its proof of convergence as more and more sines and cosines are added. Then the Fourier Transform will be introduced which attempts to approximate aperiodic functions, such as e^x, globally. The Discrete Fourier Transforms takes in discontinuous data at regular time intervals, and is a specific case for the Fourier Transform. An explanation of how it works will be introduced mainly through visualisation of the roots of unity, as well as an example. Lastly the factorisation problem will be introduced, and a method using the Discrete Fourier Transform adapted to quantum computers will be discussed, named Shor's Algorithm.

c. Motivation. The idea of representing all functions, well most under Lebesgue's restrictions, by a combination of circular frequencies and amplitudes, is one of the most elegant descriptions of mathematics I have ever encountered. While mathematics is strongly driven by rigour, and logic, I think it is important that the beauty of mathematics is still appreciated, and to see the duality between abstraction and illustration is unique to the Fourier Series, not solely because of its widespread uses but also because of the simplicity of the idea.

II. Fourier Series.

Before deriving the Fourier Series, it is important to understand the conditions upon it is built on, both periodicity and piecewise functions are a pre requisit for a Fourier Series.

i. Addition of Periodic Signals.

When combining two periodic signals, defined as,

Definition 1. Periodic Functions A function f is said to be periodic on a set F with period t if and only if,

$$f(x+t) = f(x) \quad x \in F \tag{1}$$

Now the notion of periodicity has been established, the concept of signal addition can be postulated. Suppose there is a function f with the fundamental period, the lower value for the frequency, t_1 and another function g with the fundamental period of t_2. If $t_1 + t_2$ have a lowest common multiple t_3 then $f + g$ is periodic with frequency t_3. This is not always the case as in when $f = \cos(x)$ and $g = -\cos(x)$. Globally there are also exceptions wherein the resulting function is not periodic, such as $\cos(x) + \cos(\pi)$. Therefore it can be postulated that,

$$f_1 + f_2 = f_3, \quad [f_1, f_2] \mid f_3,\ f_1, f_2, f_3 \in R^+ \tag{2}$$

For example the signal $f(t) = 4.8\sin(\pi(t+3)/6) + 5.1$ to model the tides and another periodic function $g(t) = 2.4\sin(\pi/2(t+1))$ models the discharge of a dam into the sea. The resulting water level frequency will be given by,

$$LCM(\frac{(t+3)\pi}{6}, \frac{(t+1)\pi}{2}) = 6 \tag{3}$$

Therefore the resulting function f_3 will have frequency $6 \cdot 2 = 12$. More generally if a function with frequency 1Hz is taken, then a group of functions will have the same frequencies,

$$1\text{Hz} + n\text{Hz} = 1\text{Hz}, \quad n \in \mathbb{Z}^+ \tag{4}$$

$$2\text{Hz} + 2^n\text{Hz} = 2\text{Hz}, \quad n \in \mathbb{Z}^+ \tag{5}$$

$$xHz + x^n Hz = xHz, \quad n \in \mathbb{Z}^+, x \in \mathbb{C} \tag{6}$$

Now this can be expanded to the addition of many signals such that

$$\sum_{n=1}^{\infty} nHz = 1Hz, n \in \mathbb{Z}^+ \tag{7}$$

Now the only thing that remains is to find the amplitudes of each of these signals, under the constraint that each successive signal must be a multiple of the previous' frequency. For example to represent the square wave function $f(t)$ with a period of 1, the function can be represented as,

$$f(t) = \sum_{n=0}^{\infty} \sin(2\pi nt) \tag{8}$$

When $n = 0$ there will be no function, however at as n increments in 1's, the final function will have period 1. However for each of these successive sine functions, there must be an associated magnitude. This will be constructed using the orthogonal property of sine waves. A Fourier Series does not have to model a continuous function, in fact it was proven that $f(x)$ need only be piecewise continuous.

ii. Piecewise Continuous Functions.

A function $f(x)$ is said to be piecewise continuous if it can be partitioned into a finite group $[a_1, a_2, ..., a_i]$ wherein each interval is continuous. Hence a function such as $f(x) = \frac{1}{x}$, $\{-1 < x < 1\}$ would not work. This condition can be defined more rigorously as,

$$\lim_{\epsilon \to 0+} f(x + \epsilon) = f(x + 0), \quad \epsilon > 0 \tag{9}$$

$$\lim_{\epsilon \to 0-} f(x - \epsilon) = f(x - 0), \quad \epsilon > 0 \tag{10}$$

iii. Motivation

The fundamental motivation of Fourier Series was solve the general Heat Equation, a partial differential equation that had no known general solution. The concept was to represent a function as a superposition of sine and cosines and to write the solution as a linear combination of the function. For example take the function $f(x) = \ln(x)$, it could be approximated locally via a Taylor Expansion, however Fourier Series allow for a global approximation as $n \to \infty$. The rigorous definition of the Fourier Series was only rigorously postulated 100 years after the initial concept and lead to many advances in Real Analysis such as Hilbert Spaces and Lebegues Integrals. Now more generally For example take this saw tooth wave, how could one globally approximate it using a combination of sine and cosines.

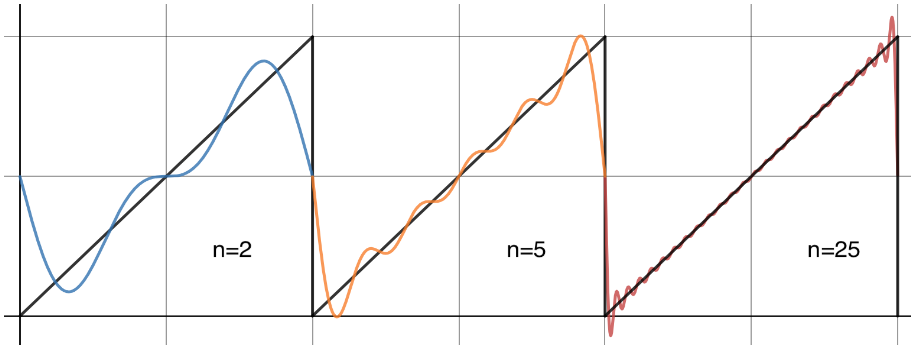


Figure 1: Fourier Approximated sawtooth wave at $n = \{2, 5, 25\}$.

This is a very powerful tool and the theory behind, and framework will now be developed.

iv. The Fourier Series

The reasoning behind Fourier Series is that for a given function f is in $\mathbb{R}^n$,

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \tag{11}$$

To break this down, the approximation will be useless if the function does not lie on x axis, hence a constant must be used often denoted a_0. The second term represents the co-sinuoidal part of the function, often used to represent even signals, wherein $f(x) = f(-x)$ while the third term represents the sinusoidal part, which simplifies the approximation of odd functions. If the function $f(x)$ is neither even or odd then both sinusoidal and co-sinusoidal components must be used. The coefficients, a_n and b_n give information on that importance of each frequency, and when the coefficient is plotted against the frequency, it is very easy to find the most important frequencies, essentially allowing for data compression of noise, frequencies with low coefficients, which is a common use of Fourier Series. It can also be said that the above approximation is a periodic one, and hence assume some inherent periodic nature of our function $f(x)$ which is not true for all functions. However, it can approximate discontinuous functions, explained through the Gibbs Phenomenon, as at limit of the series, ∞ can be assumed to be the same, as seen in Figure 1 when $n = 25$ the function jumps down quickly.

v. Fourier coefficients

There is no known way to decompose periodic functions into polynomials. However, taking the integral of the function will give rise to the solution,

$$\int_0^{2\pi} f(x)\mathrm{d}x = \int_0^1 a_0 + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx)\mathrm{d}x \tag{12}$$

Now decomposing this the integral of a_0 is $a_0 \cdot 1$, which is essentially just the integral of the function $f(x)$ while the remain terms cancel to 0 following the property of periodicity, visualised by.

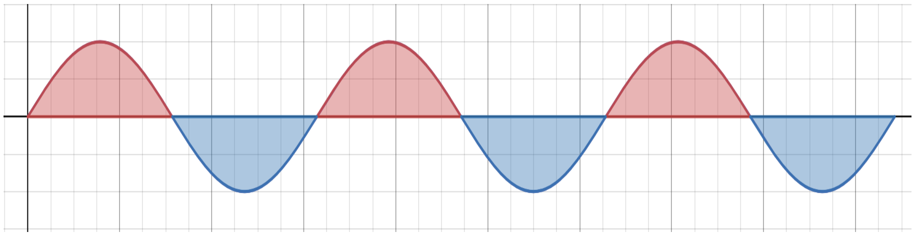


Figure 2: Area beneath a sine wave, or a periodic function, over a multiple of its period, in this case 2π is always 0.

$$\int_0^{2\pi} f(x) \cdot \cos(nx)\mathrm{d}x = \sum_{n=1}^{\infty} a_n \int_0^{2\pi} \cos(nx)\cos(nx)\mathrm{d}x = 0 \tag{13}$$

Therefore it can be said that,

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\mathrm{d}x \tag{14}$$

This can further be generalised to the case with a period of T,

$$a_0 = \frac{1}{2T}\int_T f(x)\mathrm{d}x \tag{15}$$

Essentially, a_0 represents the average value of the function, now when dealing with the a_n and b_n, the property of orthogonality must be introduced.

Definition 2. Orthogonal. This term comes from vector spaces wherein if two vectors are perpendicular, hence their dot product is 0, they are considered to be orthogonal. However for function this means that when 2 different function are multiplied together as an inner product, which is defined as

$$\langle a_n, b_n \rangle = \sum_{i=1}^{n} a_i b_i \tag{16}$$

And when they are then passed under a definite integral they are equal to 0, but when computed with itself, are equal to some value other than 0. Explaining more simply, imagine a set of particles and another set of antiparticles, now when combined they annihilate leaving nothing, but when mixed with each other they're is mass. Now more formally take the two function ϕ_m, ϕ_n and they are orthogonal under some interval $[a, b]$. Therefore,

$$\begin{aligned}\langle \phi_n, \phi_m \rangle &= \int_a^b \phi_m \cdot \phi_n \mathrm{d}x \\ &\rightarrow 0 \iff n \neq m \\ &\rightarrow ||\phi_n^2|| \iff n = m\end{aligned}$$

Where $||\vec{v}||$ is the square norm of a vector, which is just an absolute value but for vectors so that they're positive in all directions. Now take another function $f(x)$, such that

$$\begin{aligned}f(x) &= \sum_{n=0}^{\infty} C_n \phi_n \\ \langle f(x), \phi_m \rangle &= \sum_{n=0}^{\infty} C_n \langle \phi_n, \phi_m \rangle \\ &= \sum_{n=0}^{\infty} C_n \int_a^b \phi_m \cdot \phi_n \, \mathrm{d}x,\end{aligned} \tag{17}$$

Which can be simplified using the aforementioned property,

$$\int_a^b \phi_n \cdot \phi_m \, \mathrm{d}x = 0 \iff n \neq m \tag{18}$$

Leaving the following when $\phi_n = \phi_m$.

$$\langle f(x), \phi_m \rangle = C_n(\phi_n \cdot \phi_n) \tag{19}$$

$$C_n = \frac{\langle f(x), \phi_m \rangle}{\phi_n \cdot \phi_n} = \frac{\langle f(x), \phi_m \rangle}{\|\phi_n^2\|} \tag{20}$$

This principle is then used with sinusoidal functions throughout this chapter, and it is this key property that underlies all of Fourier Series and Transforms.

From the aforementioned definition,

$$\int_0^{2\pi} f(x)\mathrm{d}x = \int_0^{2\pi} a_0 + \sum_{n=1}^{\infty} a_n \cdot \cos(nx) + b_n \cdot \sin(nx)\mathrm{d}x \tag{21}$$

Now to prove sinusoidal orthogonality recall,

$$\cos(A \pm B) = \cos(A)\cos(B) \mp \sin(A)\sin(B) \tag{22}$$

$$\cos(A+B) + \cos(A-B) = \cos(A)\cos(B) - \sin(A)\sin(B) + \cos(A)\cos(B) + \sin(A)\sin(B) = 2\cos(A)\cos(B) \tag{23}$$

$$2\int_0^1 \cos(2\pi nx)\cos(2\pi mx)\mathrm{d}x = \int_0^1 \cos(2\pi(m+n)x)\mathrm{d}x + \int_0^1 \cos(2\pi(m-n)x)\mathrm{d}x \tag{24}$$

$$\sin(A+B)+\sin(A-B) = 2\sin(A)\cos(B) \tag{25}$$

Theorem 1. The function $f(x)$ is orthogonal to the summation $\cos(mx)$ and $\sin(mx)$ in the form of a Fourier Series.

Proof. To prove orthogonality first prove the integral is equal to 0 when $n \neq m$. The function $f(x)$ under an inner product argument of $\cos(2\pi mx)$, and then of $\sin(2\pi mx)$. Using the identity in Equation 23,

$$\int_0^{2\pi} f(x)\cdot\cos(mx) = \int_0^{2\pi} a_0\cdot\cos(mx) + \sum_{n=0}^{\infty} a_n\cos(nx)\cos(mx) + b_n\sin(nx)\cos(mx)\mathrm{d}x \tag{26}$$

Breaking this down, the first term will always be equal to 0 regardless of m following,

$$\int_0^{2\pi} a_o\cos(mx)\mathrm{d}x = a_0(-\sin(mx)\Big|_0^{2\pi} = 0 \tag{27}$$

The last term of II v will cancel to 0 using Equation 25,

$$\sum_{n=1}^{\infty} b_n\int_0^{2\pi}\sin(nx)\cos(mx)\mathrm{d}x = \sum_{n=1}^{\infty} b_n(\int_0^{2\pi}\sin((n+m)x)\mathrm{d}x + \int_0^{2\pi}\sin((n-m)x)\mathrm{d}x) \tag{28}$$

Using Figure 2 the integral of a periodic function, in this case sine, is equal to 0 so long as it's taken between it's period, in this case 0 → 2π. As for when $n = m$, the first sinusoid behaves normally while $\sin(0) = 0$, which still gives 0. Now that all the other terms of Equation , apart from the second, which can now be solved as by using Equation 23,

$$\sum_{n=1}^{\infty} a_n\int_0^{2\pi}\cos(nx)\cos(mx)\mathrm{d}x = \sum_{n=1}^{\infty} b_n(\int_0^{2\pi}\cos((n+m)x)\mathrm{d}x + \int_0^{2\pi}\cos((n-m)x)\mathrm{d}x) \tag{29}$$

All the terms will cancel to 0, apart from when $n = m$ as illustrated by, □

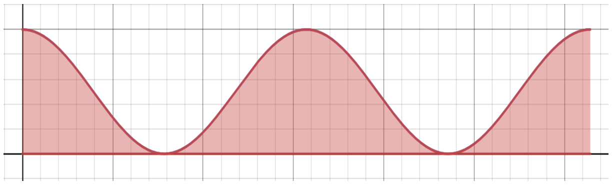


Figure 3: The area beneath $\int_0^{2\pi}\cos(nx)\cos(nx)\mathrm{d}x$ is not equal to 0.

Formally this can be proved by using,

$$\cos^2(x) = \frac{1-\cos(2x)}{2} \tag{30}$$

Then from then solving for the integral when $n = m$,

$$\begin{aligned} a_n \int_0^{2\pi} \cos(nx)\cos(nx)\mathrm{d}x &= \sum_{n=1}^{\infty} a_n \int_0^{2\pi} \frac{1-\sin(2xn)}{2}\mathrm{d}x \\ &= a_n \cdot \frac{1}{2}[x + \frac{1}{2n}\sin(2nx)]\Big|_0^{2\pi} \\ &= a_n \cdot \frac{1}{2}[2\pi + 0 - 0 - 0] \\ &= a_n \cdot \pi \end{aligned} \tag{31}$$

Therefore,

$$a_n \pi = \int_0^{2\pi} f(x)\cos(nx)\mathrm{d}x \tag{32}$$

Giving,

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\mathrm{d}x \tag{33}$$

The exact same can be applied to the sine part of Equation II v, using the fact that $f(x)$ is also orthogonal to $\sin(mx)$, hence

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\mathrm{d}x \tag{34}$$

vi. Fourier Approximation of e^x.

In order to illustrate the power of the Fourier Series, the function e^x will be approximated as a sum of sines and cosines instead.

$$f(x) = e^x \tag{35}$$

To approximate this function between $-\pi$ and π, first compute a_0 recalling Equation 15, when the period is equal to 2π.

$$\begin{aligned} a_0 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^x \mathrm{d}x \\ &= \frac{1}{\pi}\left[e^x\right]_{-\pi}^{\pi} \\ &= \frac{e^{\pi} - e^{-\pi}}{\pi} \\ &= 3.68 \end{aligned} \tag{36}$$

Hence the average of the function between those values is ≈ 3.68. Now for the sums, using Equation 34 for b_n gives,

$$b_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^x \cdot \sin(nx)\mathrm{d}x \tag{37}$$

Applying integration by parts wherein $u = e^x,\ v' = \sin(nx)$

$$b_n = -e^x \cdot \frac{1}{n}\cos(nx) - \int -e^x \cdot \frac{1}{n}\cos(nx)\mathrm{d}x \tag{38}$$

Applying integration by parts again, where $u = e^x,\ v' = \cos(nx)$ and using $w = \displaystyle\int_{-\pi}^{\pi} e^x \cdot \sin(nx)\mathrm{d}x$.

$$w = -e^x \cdot \frac{1}{n}\cos(nx) - \Bigg(-\frac{1}{n}\bigg(e^x \cdot \frac{1}{n}\sin(nx) - \int e^x \cdot \frac{1}{n}\sin(nx)\mathrm{d}x\bigg)\Bigg) \tag{39}$$

Therefore,

$$w = -e^x \cdot \frac{1}{n}\cos(nx) - \Bigg(-\frac{1}{n}\bigg(e^x \cdot \frac{1}{n}\sin(nx) - \frac{1}{n}w\bigg)\Bigg) \tag{40}$$

By isolating w, the equation simplifies to,

$$\begin{aligned} w &= -e^x \cdot \frac{1}{n}\cos(nx) + e^x \cdot \frac{1}{n^2}\sin(nx) - \frac{1}{n^2}\cdot w\mathrm{d}x \\ w(1+\frac{1}{n}) &= e^x\frac{1}{n}(-\frac{1}{n}\cos(nx) + \frac{1}{n}\sin(nx) \\ w &= \frac{e^x(-n\cos(nx)+\sin(nx)}{n^2+1} \end{aligned} \tag{41}$$

Then computing the boundaries, and multiplying by $1/2\pi$,

$$b_n = \Bigg(\frac{e^\pi\sin(n\pi) - e^\pi n\cos(n\pi)}{2\pi n^2 + 2\pi} - \frac{\sin(-\pi n) - n\cos(-\pi n)}{2\pi e^\pi n^2 + 2\pi e^\pi}\Bigg) \tag{42}$$

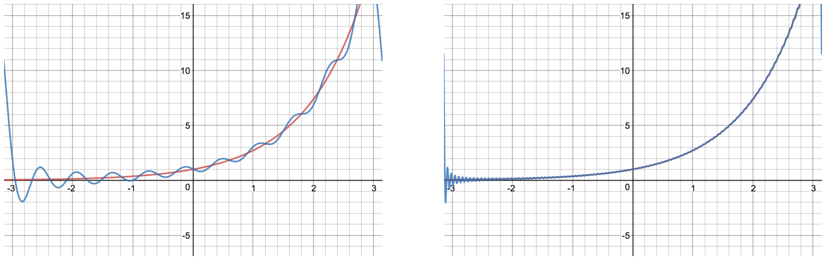
The same process can be applied to get a_n, giving,

$$a_n = \Bigg(\frac{e^\pi n\sin(n\pi) + e^\pi\cos(n\pi)}{2\pi n^2 + 2\pi} - \frac{n\sin(-\pi n) + n\cos(-\pi n)}{2\pi e^\pi n^2 + 2\pi e^\pi}\Bigg) \tag{43}$$

Now the Fourier Approximation has been established and as $n \to \infty,\ f(x) = \mathcal{F}(x)$ when $-\pi \le x \le \pi$. A Fourier Series of degree n will be denoted by $\mathcal{F}_n(x)$ for a function $f(x)$

$$\begin{aligned} e^x = 3.68 + \sum_{n=1}^{\infty}\Bigg(\frac{e^\pi n\sin(n\pi) + e^\pi\cos(n\pi)}{2\pi n^2 + 2\pi} - \frac{n\sin(-\pi n) + n\cos(-\pi n)}{2\pi e^\pi n^2 + 2\pi e^\pi}\Bigg)\cos(nx) \\ + \sum_{n=1}^{\infty}\Bigg(\frac{e^\pi\sin(n\pi) - e^\pi n\cos(n\pi)}{2\pi n^2 + 2\pi} - \frac{\sin(-\pi n) - n\cos(-\pi n)}{2\pi e^\pi n^2 + 2\pi e^\pi}\Bigg)\sin(nx) \end{aligned} \tag{44}$$

Which is visualised by,



(a) Approximation at $n = 10$ (b) Approximation at $n = 100$

Figure 4: An approximation of e^x using Fourier Series.

vii. Complex Fourier Series Derivation.

The aforementioned derivation of the coefficients for the Fourier Series of e^x were tedious to calculate, and required 3 different coefficients. In order to simplify this, the complex Fourier series is used instead, recalling Euler's Formula,

$$\cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{1}{2}e^{i\theta} + \frac{1}{2}e^{-i\theta} \tag{45}$$

$$\sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i} = -\frac{1}{2}ie^{i\theta} + \frac{1}{2}ie^{-i\theta} \tag{46}$$

Now substituting into Equation 11,

$$f(x) = a_0 + \sum_{n=1}^{\infty} \frac{a_n}{2}\left(e^{inx} + e^{-inx}\right) + \sum_{n=1}^{\infty} \frac{b_n}{2i}\left(e^{inx} - e^{-inx}\right) \tag{47}$$

$$f(x) = a_0 + \sum_{n=1}^{\infty} e^{inx}\left(\frac{a_n - ib_n}{2}\right) + \sum_{n=1}^{\infty} e^{-inx}\left(\frac{a_n + ib_n}{2}\right) \tag{48}$$

This is still a complex conjugate, hence a real number, but in order to turn all terms into one notice that,

$$\sum_{n=1}^{\infty} e^{-inx}\left(\frac{a_n + ib_n}{2}\right) = \sum_{n=-\infty}^{-1} e^{inx}\left(\frac{a_{-n} + ib_{-n}}{2}\right) \tag{49}$$

Therefore taking,

$$C_n = \frac{a_n + ib_n}{2} \text{ and } C_{-n} = \frac{a_{-n} + ib_{-n}}{2} \tag{50}$$

All the term can be put into one, where a_0 represent $n = 0$,

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{inx} \tag{51}$$

Following this simplification,

$$\begin{aligned} f(x) \cdot e^{-imx} &= \sum_{n=-\infty}^{\infty} C_n e^{inx} e^{-imx} \\ f(x) \cdot e^{-imx} &= \sum_{n=-\infty}^{\infty} C_n e^{(n-m)ix} \\ \int_0^{2\pi} f(x) \cdot e^{-imx} dx &= \sum_{n=-\infty}^{\infty} C_n \int_0^{2\pi} e^{(n-m)ix} dx \end{aligned} \tag{52}$$

The last part follows the trait of orthogonality and hence,

$$\int_0^{2\pi} e^{(n-m)ix)} dx = [x \cdot e^{(n-m)ix)}]_0^{2\pi} = 2\pi e^{(n-m)\cdot 2\pi i} - 0 = 0 \tag{53}$$

Using Euler's Identity $e^{n\cdot 2\pi i} = 0$. Now when $m = n$,

$$\int_0^{2\pi} e^0 dx = [x]_0^{2\pi} = 2\pi - 0 = 2\pi = T \tag{54}$$

Where T is period. Using this for Equation 52,

$$\begin{aligned}\int_0^{2\pi} f(x)\cdot e^{-imx}dx &= C_n\int_0^{2\pi} e^0 dx\\ C_n &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\cdot e^{-inx}dx\\ C_n &= \frac{1}{2T}\int_T f(x)\cdot e^{-in\pi x/T}dx\end{aligned} \tag{55}$$

Hence only one coefficient needs to be calculated.

viii. Complex Fourier Approximation of e^x.

Using the complex form of Fourier Approximation of e^x should be much easier as it only requires 1 coefficient, hence 1 integral. Therefore using 55, the coefficient can be calculated as follows.

$$\begin{aligned}C_n &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^x\cdot e^{-inx}dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{x(1-i\cdot n)}dx\\ &= \frac{1}{2\pi}\left[\frac{e^{x(1-i\cdot n)}}{1-i\cdot n}\right]_{-\pi}^{\pi}\\ &= \frac{1}{2\pi}\left[\frac{e^{\pi-\pi i\cdot n)}}{1-i\cdot n}\right]-\left[\frac{e^{-\pi+\pi i\cdot n)}}{1-i\cdot n}\right]\\ &= \frac{1}{2\pi\cdot 1-in}(e^{\pi}e^{-i\pi n}-e^{-\pi}e^{i\pi n})\\ &= \frac{1}{2\pi\cdot 1-in}(e^{\pi}\cdot(-1)^n-e^{-\pi}\cdot(-1)^n) = \frac{(-1)^n}{2\pi\cdot 1-in}(e^{\pi}-e^{-\pi})\end{aligned} \tag{56}$$

Then multiplying by the conjugate of $1-in$, which is $1+in$ gives,

$$C_n = \frac{(-1)^n\cdot(1+in)}{2\pi\cdot(1-in)(1+in)}\cdot(e^{\pi}-e^{-\pi}) \tag{57}$$

$$C_n = C_{-n} = \frac{(-1)^n(1+in)(e^{\pi}-e^{-\pi})}{2\pi(1+n^2)} \tag{58}$$

Therefore the Complex Fourier Series of e^x is given by,

$$e^x = \sum_{n=-\infty}^{\infty}\frac{(-1)^n(1+in)(e^{\pi}-e^{-\pi})}{2\pi(1+n^2)}\cdot e^{inx} \tag{59}$$

The calculations were a lot more straightforward, and there was only one integral, so from now on when dealing with a Fourier Series it will be taken in complex form. From now on a more rigorous definition of the Fourier Series will be used, wherein $f(x)$ is defined in the interval $(-T,T)$ and determined outside the interval by $f(x+2T)=f(x)$ which is given by,

$$f(x) = \sum_{n=-\infty}^{\infty} C_n e^{\frac{n\pi x}{L}} \tag{60}$$

Therefore,

$$C_n = \frac{1}{T}\int_{-T}^{T} f(x)^{\frac{n\pi x}{L}}dx, \quad n=0,1,2,... \tag{61}$$

There hasn't been a proof for why a Fourier Series converges to a function on its interval $(-T, T)$ as $n \to \infty$, this will now be explained.

III. Convergence of Fourier Series

In this chapter the proof for convergence will focus on 2π functions as opposed to $2L$ functions. This function is defined on $[a, b]$ and is piecewise continuous. The function is considered piecewise smooth if $f'(x)$ is also continuous between n_{i-1} and n_i. As the function is not fully continuous throughout the real or complex domain, the fundamental theorem of Calculus which needs to be adjusted to account for the redundancies of points a and b.

$$\lim_{x_a-} f(x) = f(a) \text{ and } \lim_{x_b+} = f(b) \tag{62}$$

Hence,

$$\lim_{x_a-} f(x) - \lim_{x_b+} f(x) = \int_b^a f'(x)\mathrm{d}x \tag{63}$$

Then

$$\begin{aligned} \lim_{k\to\infty} \int_b^a f(x)\sin(xk)\mathrm{d}x = 0 \\ \text{and } \lim_{k\to\infty} \int_b^a f(x)\cos(xk)\mathrm{d}x = 0 \end{aligned}$$

i. Proof

Recall that the Fourier Series is a global approximation of a function, as opposed to a Taylor series which is a local approximation.

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi nx) + b_n \sin(2\pi nx) \tag{64}$$

Which can also be written to its Nth degree as which are known as Fourier Partial sums.

$$f_N(x) = \frac{a_0}{2} + \sum_{n=1}^{n=N} a_n \cos(2\pi nx) + b_n \sin(2\pi nx) \tag{65}$$

The full Fourier sum is

$$f_N(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt + \sum_{n=1}^{n=N} \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)(\cos(nt)\cos(nx) + \sin(nt)\sin(nx))dt \tag{66}$$

Which simplifies to

$$f_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)[\frac{1}{2} + \sum_{n=1}^{N} \cos(n(x-t))dt \tag{67}$$

Following $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$.
If we substitute $s = x - t$ and assume f is 2π periodic

$$f_N(s) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-s)[1 + 2\sum_{n=1}^{n=N} \cos(ns)]ds \tag{68}$$

From this we can extract a trigonometric series, which called the Dirichlet Kernel,

$$D_N(s) = 1 + 2\sum_{n=1}^{N} \cos(ns) \tag{69}$$

Which has the following properties for $N, \forall s \in \mathbb{N}$

Remark. $D_N(s)$ is a 2π-periodic function.

Remark. $D_N(0) = 2N + 1$.

Lemma 1. $|D_N(s)|\leq 2N + 1$.

Proof. Since max$(cos(x))$ occurs at $x = 0$, at all other values $cos(x)$ is smaller. Formally this can be written as

$$\forall x \neq 0, y = 0 \tag{70}$$

□

Lemma 2. $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} D_N(s)ds = 1$.

Proof.

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(s)ds &= \frac{1}{2\pi}\int_{-\pi}^{\pi} 1 + 2\sum_{n=1}^{N}\cos(ns)ds \\ &= \frac{1}{2\pi}(s - 2\sum_{n=1}^{N}\sin(ns))|_{-\pi}^{\pi} \\ &= \frac{1}{2\pi}(2\pi + 2\sum_{n=1}^{N}\sin(n\pi) - \sin(n\pi)ds \\ &= \frac{1}{2\pi}(2\pi) \\ &= 1\end{aligned}$$

□

Lemma 3. $D_N(s) = \dfrac{\sin(N + 1/2)s}{\sin(s/2)}$.

Proof. This can be proved via some trigonometric manipulation. Firstly suppose $n \in \mathbb{N}$ & $s \neq k\pi$ for any $k \in \mathbb{Z}$

$$D_N(s) = 1 + 2\sum_{n=1}^{N} \cos(ns) \tag{71}$$

Then remembering cosine is even and sine is odd:

$$2\sum_{n=1}^{N}\cos(ns) + 1 = \sum_{n=-N}^{N} \cos(ns) \text{ I could show a geometric proof here} \tag{72}$$

$$\sum_{n=-N}^{N} \sin(ns) = 0 \tag{73}$$

Hence

$$\begin{aligned} D_N(s) &= \sum_{n=-N}^{N} \cos(ns) + \frac{\cos(s/2)}{\sin(s/2)} \sum_{n=-N}^{N} \sin(ns) \\ &= \frac{1}{\sin(s/2)} \sum_{n=-N}^{N} (\sin(s/2)\cos(ns) \\ &+ \cos(s/2)\sin(ns)) \\ &= \frac{1}{\sin(s/2)} \sum_{n=-N}^{N} \sin((n+1/2)s) \end{aligned} \tag{74}$$

Following $\sin(A)\cos(B) + \cos(A)\sin(B) = \sin(A+B)$.

Which is a telescoping sum, as in a sum that cancels itself with previous terms leaving a single fixed value.

$$D_N(s) = \frac{\sin((N+1/2)s)}{\sin(s/2)} \tag{75}$$

Going back to the Dirichlet Kernel, which can be defined in two domains following the 2nd trait of the Dirichlet

$$D_N(s) = \begin{cases} \dfrac{\sin((N+1/2)s)}{2\pi\sin(s/2)} & \text{if } s \neq 0, \pm 2\pi, ... \\ \\ \dfrac{2N+1}{2\pi} & \text{if } s = 0, \pm 2\pi, .. \end{cases} \tag{76}$$

Another interesting property of the Dirichlet is

$$\int_{-\pi}^{0} D_N(s)ds = \int_{0}^{\pi} D_N(s)ds = \frac{1}{2} \tag{77}$$

□

Now the Fourier Series can be proven to be convergent as $n \to \infty$. Since $D_N(s)$ is even the integral can broken into 2,

$$\begin{aligned} f_N(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_N(t) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{0} f(x-t) D_N(t) dt + \frac{1}{2\pi} \int_{0}^{\pi} f(x-t) D_N(t) dt \\ &= \frac{1}{2\pi} \int_{0}^{\pi} (f(x-t) + f(x+t)) D_N(t) dt \end{aligned} \tag{78}$$

Following the 4th and 5th trait of the Dirichlet kernel,

$$\begin{aligned} f_N(x) - s &= \frac{1}{2\pi} \int_{0}^{\pi} (f(x-t) + f(x+t) - 2s) D_N(t) dt \\ &= \frac{1}{2\pi} \int_{0}^{\pi} \frac{(f(x-t) + f(x+t) - 2s)}{t} \cdot \frac{t}{\sin(t/2)} \sin((N+1/2)t) dt \end{aligned} \tag{79}$$

And following the final identity in the aforementioned introduction, $f_N(x) - s \to 0$ which shows that $f_N(x)$ will converge to a point s since $t/\sin(t/2)$ is bounded to the interval $(0, \pi)$ which is a finite value. This formula is used by computers to find the the Fourier Series as it does not require any formulas for the Fourier coefficients, and instead directly calculating each consecutive term.

IV. The Fourier Transform

Now the Fourier Transforms allows engineers, scientists, programmers, and many other fields of profession from the time domain to the frequency domain of non-periodic functions, or functions of period ∞. Considering the Fourier Series assumes that the period is $2T$, what would be the result if $T \to \infty$. Firstly some slight adjustments need to be make to the variables, firstly take the function $f(t)$ with period T, its fundamental frequency ω_0 can be written as $\omega_0 = 2\pi/T$ and its harmonics can now be written in terms of ω giving,

$$f(t) = \sum_{n=-\infty}^{\infty} C_n e^{in\omega t} \qquad C_n = \frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-in\omega t} \tag{80}$$

Now as $T \to \infty$, the fundamental frequency ω becomes very small, it will be denoted $\mathrm{d}\omega$. To combat this take another variable n that approaches ∞ at the same are as ω approaches 0 such that it gives a finite frequency ω, therefore taking $C_n = C(\omega)\mathrm{d}\omega$ which conserves the integral as the discrete spectrum turns into a continuous one.

$$C(\omega)\mathrm{d}\omega = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} f(t)e^{-i\omega t}\mathrm{d}t \tag{81}$$

Substitution the aforementioned identities,

$$C(\omega)\mathrm{d}\omega = \frac{\mathrm{d}\omega}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-\omega t}\mathrm{d}t \to C(\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{-\omega t}\mathrm{d}t \tag{82}$$

Since n is now continuous, the summation for $f(t)$ in Equation 80 can be turned into an integral,

$$f(t) = \int_{-\infty}^{\infty} C(\omega)\mathrm{d}\omega e^{i\omega t} = \int_{-\infty}^{\infty} C(\omega)e^{i\omega t}\mathrm{d}\omega \tag{83}$$

The Fourier Transform however uses $2\pi C(\omega)$ instead of $C(\omega)$ and use $F(\omega)$ as the Fourier Transform therefore giving,

$$F(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\mathrm{d}t \qquad f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\mathrm{d}\omega \tag{84}$$

This is an incredibly powerful tool, and while it is hard to show an example of exactly what happens when a signal in time is turned into its representation in frequency, examples such as

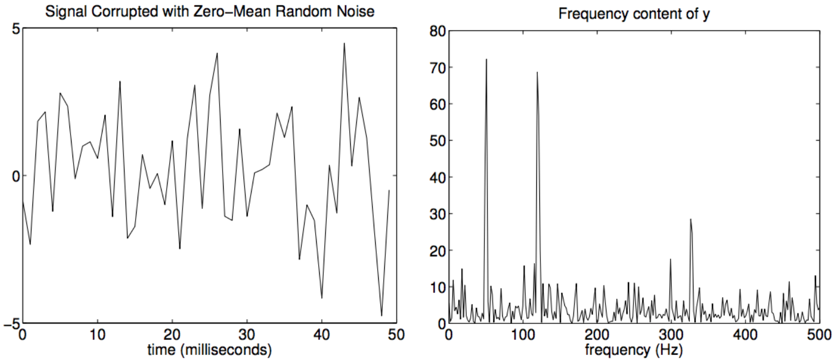


Figure 5: How white noise is often composed mainly of only a small combination of frequencies, that can then be targeted by sound engineers to create noise cancelling headphones.

V. The Discrete Fourier Transform.

While Fourier series are useful in solving differential equations, and global approximations of other functions, the true applications of the Fourier series lie in the Fourier Transform. A basic Fourier Transform is shown in Section **??**, but in this essay the more important Discrete Fourier Transform will be explored, not only because of it's numerous application in modern day acoustics such as Schorr's Prime Factorisation algorithm, infrared spectroscopy, Siri, and MRI scans. From understanding atmosphere's of distant exoplanets to creating personal assistant, the DFT, or it's more computationally efficient FFT is the most widely used Digital Signal processing tool, which is why it's listed in the top 100 most used algorithms. The DFT requires the signal to be discrete, as well as being finite in duration. In this section, N will be the number of discrete impulses, each of amplitude x.

i. Complex Geometric Series

Lemma 4. Take the finite series $S_n(z^n)$ where $z^n \in \mathbb{C}$. It's sum will be $\dfrac{1-z^N}{1-z}$.

Proof. Now writing this out,

$$\begin{aligned} S_n(z^n) &= \sum_{n=0}^{N-1} z^n = 1 + z + z^2 + ... + z^{N-1} \\ zS_n(z_n) &= z + z^2 + z^3 + ... + z^{N-1} + z^N \\ S_n(z^n) - z^n S_n(z_n) &= 1 - z^n \\ S_n(z^n) &= \frac{1-z^N}{1-z} \end{aligned} \tag{85}$$

□

ii. Proving Complex Sinusoidal Orthogonality.

To prove orthogonality, the integral of two or more waves with different frequencies is equal to 0.

$$\langle A_1 \sin(w_1 t), A_2 \sin(w_2 t)\rangle \tag{86}$$

Where $w_1 \neq w_2$. Exact orthogonality only occurs when the frequency f_s divided by the sampling rate, N is equal to the fundamental frequency, f_k

$$f_k = k\frac{f_s}{N} \tag{87}$$

Where $k = 0, 1, 2, 3, ..., N-1$ and $f_s = \dfrac{1}{T}$. These signals are then processed in the complex plane corresponding to f_k,

$$s_k(n) = e^{jw_k nT} \tag{88}$$

Where $w_k = f_k \cdot 2\pi = k\dfrac{2\pi}{N}f_s$.

$$W_N^k = e^{jk2\pi f_s T/N} \tag{89}$$

$$W_N^k = e^{jk2\pi/N} \tag{90}$$

$$[W_N^k]^N = e^{jk2\pi N/N} = e^{jk2\pi} = 1 \tag{91}$$

Regardless of k, the $[W_N^k]^N$ will always be equal to 1 since k can only be an integer, following Euler's formula:

$$e^{j2\pi k} = \cos(2\pi k) + j\sin(2\pi k), \tag{92}$$

Since $\sin(2\pi\mathbb{I}) \equiv 0$ & $\cos(2\pi\mathbb{I}) \equiv 1$.

iii. The Orthogonality of the DFT Sinusoids.

To show that the DFT are also completely orthogonal, take a signal $s_k(n)$.

$$s_k(n) = e^{jw_k nT} = e^{j2\pi kn/N} = [W_N^k]^N, \; n = 0, 1, 2, ..., N-1 \tag{93}$$

Now for the $k = 0,$ to the $N-1$ sinusoidal.

$$\begin{aligned}\langle s_k, s_l \rangle &= \sum_{n=0}^{N-1} \frac{s_k(n)}{s_l(n)} \\ &= \sum_{n=0}^{N-1} e^{j2\pi kn/N} e^{-j2\pi ln/N}\end{aligned} \tag{94}$$

$$\sum_{n=0}^{N-1} e^{j2\pi n(k-l)/N} = \frac{1 - e^{j2\pi n(k-l)}}{1 - e^{j2\pi n(k-l)/N}} \tag{95}$$

The last equation uses Lemma 4. Since the denominator is non-zero, and the numerator is 0, this proves.

$$\langle s_k, s_l \rangle \Leftrightarrow k \neq l \tag{96}$$

Now when $s_k = s_l$,

$$\sum_{n=0}^{N-1} e^{j2\pi \frac{n}{N}(k-k)} = \sum_{n=0}^{N-1} e^0 = N \tag{97}$$

iv. Visual intuition of DFT

As shown in Equation 91, W_N^k are what is known as the roots of unity.

Definition 3. Roots of Unity. Otherwise known as *de Moivre* numbers, is a complex number that gives 1 when raised to some integer n. Hence it satisfies the equation,

$$z^n = 1, \quad n \in \mathbb{Z} \tag{98}$$

These are often taken as complex numbers such that,

$$\exp\left(\frac{2k\pi i}{n}\right) = \cos(\frac{2k\pi}{n}) + i\sin(\frac{2k\pi}{n}), \quad k = 0, 1, \ldots, n-1 \tag{99}$$

The formula for solving this equation is therefore given as,

$$U_n = \{e^{2k\pi j/n} \mid k \in \{0, 1, \ldots, n-1\}\} \tag{100}$$

Where U_n is the set of values that are roots of unity. If one of the roots of unity are known, all others can be calculated following that,

$$x^n \in U_n, \quad x^k = 1 \tag{101}$$

The sum of the roots of unity, as seen in Equation 95 is 0 whenever whenever k is not divisible by l. It is often easier to visualise these as points around a complex unit sphere give be, All the values in 6 when raised to the power of 8

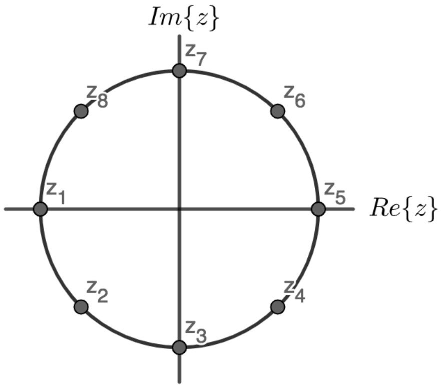


Figure 6: The roots of unity of a unit sphere where $n = 8$.

will equal to 1, and there sum is 0, the center of the circle as they are equivalently placed. Now expanding this to $n = 2, 3, 4, 5, 6$ the roots of unity are visualised as,

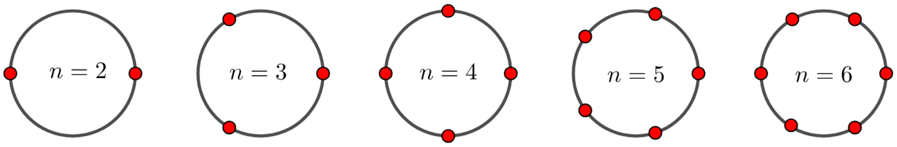
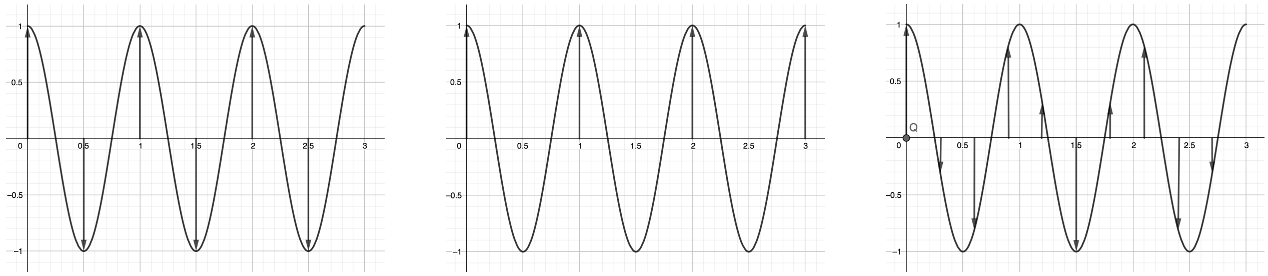


Figure 7: The roots of unity for $n = 2, 3, 4, 5, 6$

In the case where $s_k = s_l$ from Equation 97 the summation is not equal to 0. It is simpler to illustrate it with the curve $\cos(2\pi nx)$. Then using a constant sampling rate, T_s, of the height of the wave at this point, shown by



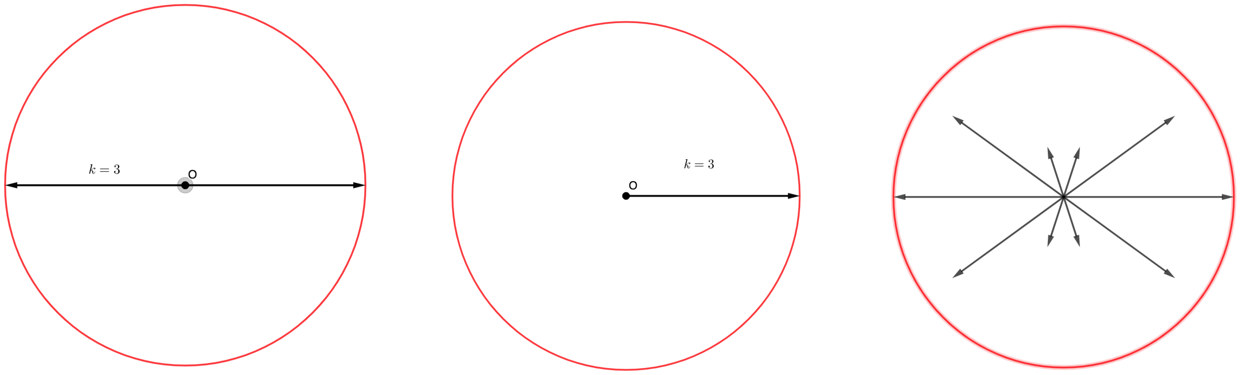
(a) When $T_s = 0.5$

(b) When $T_s = 1$

(c) When $T_s = 0.3$

Figure 8: Height of cosine wave at equally space points.

It may be obvious that the vectors in (a) and (c) will add to 0, and (b) will not, however plotting each respective sample around a circle wherein the angle of separation is given by $2\pi \cdot T_s$ In figure (b) it is clear that the total sum of the vectors is not equal to 0, and therefore the trait of orthogonality is true as when the sampling rate does not equal the frequency of the sinusoid, the resulting value is 0. The symmetry between odd and even frequencies and sampling rates is what is utilised by the Fast Fourier Transform, which is much more efficient to compute.



(a) When $T_s = 0.5$ (b) When $T_s = 1$ (c) When $T_s = 0.3$

Figure 9: The cosine wave plotted at different sampling rates, k is the number of vectors overlayed.

v. The Transform.

The Discrete Fourier Transform (DFT) is equivalent to the aforementioned Fourier Transform, but it is regardless of time and takes discrete signals instead, given as $f(t) = f[0], f[1], \ldots, f[k], \ldots, f[N-1]$ with a sampling rate of T_s. Therefore this translation from,

$$F(i\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{102}$$

Wherein each impulse $f[k]$ can be considered to have an area $f[k]$, which means the integral exists only at discrete sample points,

$$\begin{aligned} F(i\omega) &= \int_{0}^{(N-1)T} f(t)e^{-i\omega t}dt \\ &= f[0]e^{-i\cdot 0} + f[1]e^{-i\cdot 1} + \ldots + f[k]e^{-i\cdot k} + \ldots + f[0]e^{-i\cdot\omega(N-1)T} \\ &= \sum_{k=0}^{N-1} f[k]e^{-i\omega kT} \end{aligned} \tag{103}$$

Since the data is treated as periodic, the signal treats $f(0) \to f(N-1)$ the same as $f(N) \to f(2N-1)$ which is similar to the initial Fourier Series. Therefore since it is treated as periodic, the fundamental frequency (one cycle per $1/NT$Hz) the definition of ω can be change from Equation **??** such that,

$$\omega = 0, \frac{2\pi}{NT}\cdot 1, \frac{2\pi}{NT}\cdot 2, \ldots, \frac{2\pi}{NT}\cdot n, \ldots, \frac{2\pi}{NT}\cdot(N-1) \tag{104}$$

This modifies Equation 103,

$$F[n] = \sum_{k=0}^{N-1} f[k]e^{-i2\pi nk/N} \quad n = 0, 1, \ldots N-1 \tag{105}$$

Therefore $F[n]$ is the DFT of the sequence $f[k]$. As this involves many summations, it is often easier to write it in it's matrix form,

$$\begin{bmatrix} F[0] \\ F[1] \\ F[2] \\ \vdots \\ F[N-2] \\ F[N-1] \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & W & W^2 & W^3 & \dots & W^{N-1} \\ 1 & W^2 & W^4 & W^6 & \dots & W^{N-2} \\ 1 & W^3 & W^6 & W^9 & \dots & W^{N-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W^{N-1} & W^{N-2} & W^{N-3} & \dots & W \end{bmatrix} \cdot \begin{bmatrix} f[0] \\ f[1] \\ f[2] \\ \vdots \\ f[N-2] \\ f[N-1] \end{bmatrix} \tag{106}$$

$$W = \exp\left(\frac{-in2\pi}{N}\right) \tag{107}$$

$$\begin{aligned} W^{2N} &= \exp\left(\frac{-in2\pi}{N}\right) \cdot \exp\left(\frac{-in2\pi}{N}\right) \\ &= \sqrt[N]{\exp\left(-in2\pi\right)} \cdot \sqrt[N]{\exp\left(-in2\pi\right)} \\ &= \sqrt[N]{1} \cdot \sqrt[N]{1} \\ W^{2N} &= 1 \end{aligned} \tag{108}$$

As seen in Equation 106 it is very tedious to compute the DFT of a signal by hand, and is almost always done using a computer using a Fast Fourier Transform (a geometric interpretation of the DFT outlined in the Appendices Section Y). However it is easier to visual this with an example, take the sinusoid,

$$f(t) = 2 + 2\cos(2\pi t) + 4\cos(4\pi t) \tag{109}$$

The first frequency is 1 and the second is 2, therefore in our samples of the signal the amplitudes of the frequencies of 1 and 2 should be highest. Then using a sampling rate of 4 times per second, from $t = 0$ to $t = 3/4$, hence taking points A_1, B_1, C_1, D_1. Illustrate by,

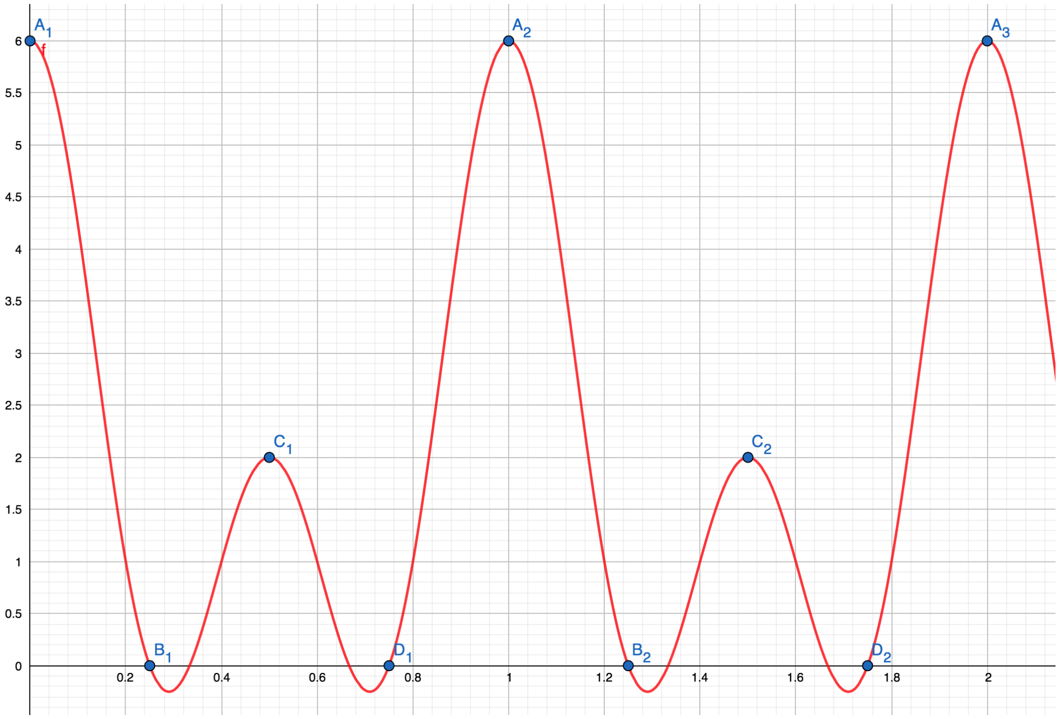


Figure 10: The function $f(t)$ samples at 3 points between $t = 0$ and $t = 1$

Therefore, by using $T_s = 4$, the sampling rate per second

$$\begin{aligned} t &= k \cdot T_s \\ k &= \frac{t}{4} \end{aligned} \tag{110}$$

Giving,

$$f[k] = 2 + 2\cos(k\pi/2) + 2\cos(k\pi) \tag{111}$$

Now expanding into it's matrix form,

$$\begin{bmatrix} f[0] \\ f[1] \\ f[2] \\ f[3] \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 0.5 \\ 0 \end{bmatrix} \tag{112}$$

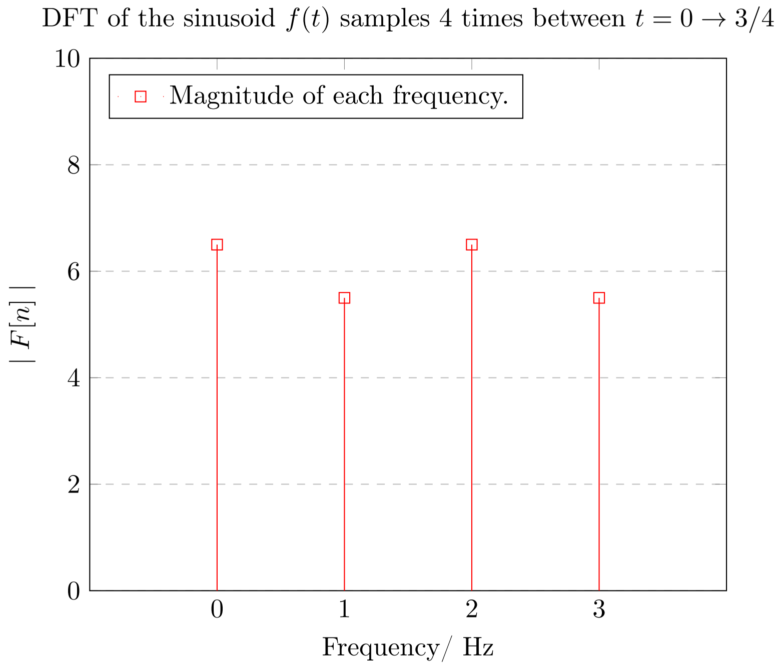
Then using Equation 105 where $N = 4$,

$$F[n] = \sum_{0}^{3} f[k] \exp\left(\frac{-i2\pi nk}{2}\right) = \sum_{0}^{3} f[k](-j)^{nk} \tag{113}$$

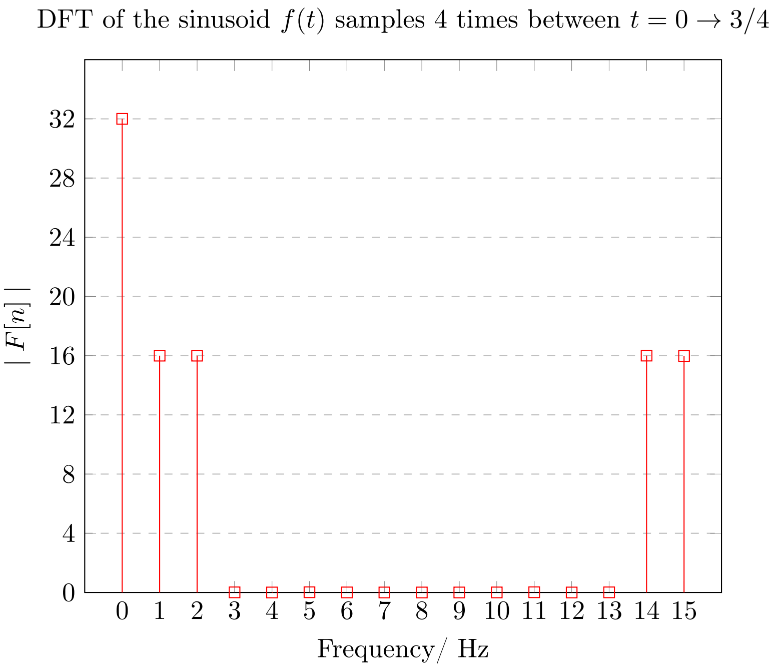
In it's matrix form it is written as,

$$\begin{bmatrix} F[0] \\ F[1] \\ F[2] \\ F[3] \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{bmatrix} \cdot \begin{bmatrix} f[0] \\ f[1] \\ f[2] \\ f[3] \end{bmatrix} = \begin{bmatrix} 6.5 \\ 5.5 \\ 6.5 \\ 5.5 \end{bmatrix} \tag{114}$$

If a complex number were to come up in the result of the DFT, or a negative number, simply take the absolute value of the result, which will give,



While it is easy to try and check for any patterns in the analysis, there is simply not enough data to make any useful analysis. Now taking a smaller sampling rate, giving higher amounts of data, there should be more data to make deductions from. Think of this as a magnification of the current data, giving,



The value of 0 comes from the fact that the number of samples, 16, when multiplied by the constant 2 gives 32, and if the task were to reverse engineer our signal then 32/2 would be taken as the constant . Which is what should be expected following the prediction initially made of having distinct frequencies at 1 and 2. As for the values of 14 and 15, this should be ignored following the Nyquist Sampling Limit, which is too complicated to derive for this essay, but the result states that only half of the sampled frequencies have meaning following the traits of modular arithmetic. Therefore any frequencies above 8 should be ignored.

vi. Inverse Discrete Fourier Transform

The process of going from a signal, to it's coefficients has been formalised before, but as all the aforementioned transforms, it works in both ways. Hence using Equation 105, and solving for $f[k]$ gives,

$$f[k] = \frac{1}{N}\sum_{0}^{N-1} F[n] \exp\left(\frac{j2\pi nk}{N}\right) \tag{115}$$

Wherein $1/N$ is used to normalised the signal, and is the complex conjugate of the matrix. In some cases this is taken as $1/\sqrt{N}$, and the proof for this is outside of the scope of this extended essay and is just assumed. For the rest of the essay how ever, the normalised IDFT will taken in the form,

$$C_n = \frac{1}{\sqrt{2N}}\sum_{n=0}^{N-1} F[n] \exp\left(\frac{j2\pi nk}{N}\right) \tag{116}$$

Where C_n is the new coefficient, which allows for the identity,

$$\sum_{n=0}^{\infty} \mid C_n \mid^2 = \int_0^{2N} f(x)dx \tag{117}$$

Wherein $f(x)$ is a real valued continuous function, the proof for this is presented in the Appendices Section Z.

vii. Discrete Fourier Transform of a Periodic Sequence

Since the function $a^r \equiv 1 \mod N$ has the property of periodicity, hence

$$f(m+r) = f(x),\ \forall m \tag{118}$$

The Discrete Fourier Transform can be used to find the periodicity, r of the sequence. Take r as a divisor of $2N$ therefore,

$$r = \frac{2N}{k},\ k \in \mathbb{Z} \tag{119}$$

Therefore if n, from the Equation 105, is not a multiple a k, the signals will cancel each other following the roots of unity, but if they are then they will be equal to the periodicity r,

$$F[n] = \begin{cases} \dfrac{k}{\sqrt{2N}} \displaystyle\sum_{m=0}^{\frac{k}{2N}-1} f(m) \exp\left(\frac{inm\pi}{N}\right) & \text{iff } n = k \cdot c, \\ 0 & n \neq k \cdot c \end{cases} \tag{120}$$

Wherein c is an integer. This means that if n is not a multiple of k then the coefficient becomes 0, whereas if it is the coefficient is non zero. The smallest of these multiples is the period of the sequence $a^r \equiv 1 \mod N$.

VI. Quantum Computing

In the modern data "big data" age, security and confidentiality is essential for communication. The most well known application of pure mathematics to this area, RSA cryptography, is reliant on the prime factorisation problem. As in take two primes p, q, multiply them together to get n, it is then assumed that,

$$p \cdot q = n \tag{121}$$

is a simple calculation for a computer, however the reverse process is "hard", factorising n into p, q, and cannot be solved in polynomial time.

Definition 4. Polynomial Time. In Complexity Theory, functions are given to illustrate how the time to compute, or verify, the problem is in function of its size. A problem is said to be "hard" if it cannot be solved in polynomial time, in this case the factorisation falls within that domain using a classical computer. O is commonly used to denote this, hence RSA $\neq O(n^k)$ using a classical computer

However, there is an algorithm that can reduce the problem of prime factorisation into 4 steps, the second of which will require a quantum computer and Shor's algorithm to find in polynomial time to find those two primes, with the only knowledge of N.

i. How to break RSA cryptography

In this section, a simple algorithm will be outlined wherein the problem of factorisation can be broken down into 4 steps.

Step 1. The first step is to pick a any number smaller than N and then check to make sure a and N are relatively prime.

Definition 5. Relatively prime. Two integers are said to be coprime integers, or relatively prime, if they're greatest common denominator is 1, for example 14 and 15 are relatively prime as they are only commonly divisible by 1 where 14 and 21 are divisible by 7.

So to check this, take the GCD of a and N such that,

$$\text{GCD}(a, N) = 1 \tag{122}$$

Which can be computed quickly via the Euclidean Algorithm, outlined in the appendix.

Step 2. The period of a number a under modular arithmetic N is simply to what power does,

$$a^r \mod N = 1 \tag{123}$$

In the case of $N = 35$ and $a = 8$, $8^4 = 1$ hence $r = 4$. This is a hard problem, known as the discrete logarithm problem, and is where a quantum computer using Shor's algorithm will be used. To make sure the remaining steps are still valid, r must be an even integer and,

$$a^{\frac{r}{2}} + 1 \not\equiv 0 \tag{124}$$

If either of these steps fail, then a new a needs to be chosen, but this will only happen half of the time.

Step 3. Now since it is known that,

$$a^r - 1 \equiv 0 \mod N \tag{125}$$

However, a $0 \mod N$ is simply it is a multiple of some values, k, giving,

$$a^r - 1 = N \cdot k \tag{126}$$

Then substituting $N = p \cdot q$, and since r is an even number it can be rewritten as,

$$(a^{\frac{r}{2}} - 1)(a^{\frac{r}{2}} + 1) = (p \cdot q) \cdot k \tag{127}$$

Step 4. Now since $a^{1/2} + 1 \not\equiv 0$ it can be said that,

$$p = \text{GCD}(a^{\frac{r}{2}} - 1, N) \tag{128}$$

$$q = \text{GCD}(a^{\frac{r}{2}} + 1, N) \tag{129}$$

Since p will divide the first term of Equation 127, and q the second, giving k since it is a multiple of N and p and q together form N.

However finding r in Step 2 is again computationally inefficient with an exponential time complexity, $O(x^n)$. However, as previously discussed the period of a discrete signal can be found using a Discrete Fourier Transform and quantum computers can find this very efficiently using a principle known as superposition.

ii. Introduction to Quantum Information Theory.

While the computer science part will not be explained fully in this essay, it is useful to understand some of the mathematical properties that can be used to apply to real world application, such as Shor's Algorithm. There main difference between a quantum computer and a normal computer is that the latter operates with strings binary, ones and zeros, and use logic gates to manipulate the bits. On the other hand quantum computers use qubits, and the just like the normal computer it is written in zeros and ones, and outputs zeros and ones. However, while a quantum computer is running its qubits can be in infinitely many superpositions between 0 and 1.

Definition 6. Superposition. This is the term coined by theoretical physicists to describe how quantum particles can be in an infinite amount of states until measured. The metaphor of a wave is often used, as waves can add destructively, and constructively in order to create a new wave, and this has infinitely many possibilities.

Following the definition of superposition, it can be said that the output of the system is only determine once an observer looks at it, as shown by the Schrodinger's Cat Experiment. Now more formally to illustrate how qubits behave, take 2 qubits, there are 4 possible basic states,

$$|x, y\rangle = \begin{matrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{matrix} \tag{130}$$

Now for example take the first state, and apply a quantum gate, α, and another β now the qubits are in a superposition, ignore the square roots.

$$|00\rangle \xrightarrow{\alpha} \sqrt{1/2} \cdot |01\rangle + \sqrt{1/2} \cdot |10\rangle \xrightarrow{\beta} \sqrt{1/2} \cdot |01\rangle + \sqrt{1/6} \cdot |10\rangle + \sqrt{1/3} \cdot |11\rangle \tag{131}$$

Hence $|01\rangle$ has probability $1/2$, $|10\rangle$ has probability $1/6$ and $|11\rangle$ has probability $1/3$ under the quantum gates α then β. Now when the observer looks at the qubit, it will collapse into one of those states according to its probability, 01 half of the time, 10 one sixth etc. Now for a more geometric interpretation, and the intuition behind the square roots.

iii. Mathematical intuition of Quantum Computers.

The aforementioned definition of a qubit is easier to understand but not completely true. Now take a unit circle, for a normal computer a line L can only be formed between the point $0 \to 0$ to represent the bit 0, or $0 \to 1$ for bit 1. The square root is used as following the Pythagorean Theorem $1 = \sqrt{p_0^2 + p_1^2}$ where p_0 is the probability of state 0, and p_1 is the probability of being in state 1. This can be visualised by, In Figure 11, the domain and range were restricted to the positive real numbers, but the same properties can be applied to the entire circle. Now for two qubits, it will take a 4 dimensional vector, as shown in Equation 130, to show allow possible combinations. The quantum gate transformation in Equation 131 can now be written in vector form as follows,

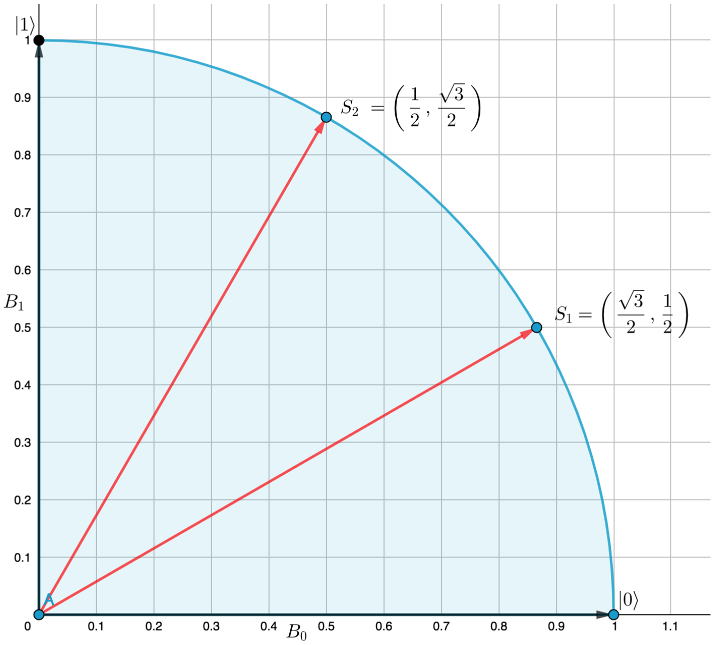


Figure 11: For a normal computer, only the vector B_0 and B_1 would be available, and the probability of each would be to 1. However for a supercomputer there is an infinite amount of possibilities between $1, 0$ and $0, 1$, for example S_1 and S_2. Now in terms of there probabilities, $S_1 : p_0 = 3/4, p_1 = 1/4$ and $S_2 : p_0 = 1/4, p_1 = 3/4$.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \xrightarrow{\alpha} \begin{bmatrix} 0 \\ \sqrt{1/2} \\ \sqrt{1/2} \\ 0 \end{bmatrix} \xrightarrow{\beta} \begin{bmatrix} 0 \\ \sqrt{1/2} \\ \sqrt{1/6} \\ \sqrt{1/3} \end{bmatrix} \xrightarrow{\Omega} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{132}$$

Where Ω is the act of observing the quantum state, in this case it turned out to be $|0100\rangle$ which is the most probable. This however is a lot harder to visualise as it will be pointing to a location on a 4-Dimensional unit sphere. This can be expanded to using 2^N quantum state, which will point to a sphere of dimension 2^N, and a quantum gate moves the point around the sphere. This transformation can be described by a unitary matrix denoted $\mathbb{U}$, further explained in the Appendices Section X, wherein a transformation on the 4 previously described quantum states is equal to,

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ \sqrt{1/2} & 0 & -\sqrt{1/2} & 0 \\ -\sqrt{1/2} & 0 & \sqrt{1/2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \sqrt{1/2} \\ \sqrt{1/2} \\ 0 \end{bmatrix} \tag{133}$$

$$\mathbb{U} \cdot \mathbf{b} = \alpha(\mathbf{b}) \tag{134}$$

Hence this unitary matrix, multiplied by our base state $b = |1000\rangle$ is equal to the quantum gate α. This unitary matrix can be changed in function of what the gate is. Now while it is true that the output of a gate is in 2^N different states, the probabilities of those are written in 2^N complex dimensions, wherein the probabilities are the absolute value of that output. For example take the unitary matrix,

$$\mathbb{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & i & 0 & 0 \\ \sqrt{1/2} & 0 & -\sqrt{1/2} & 0 \\ -\sqrt{1/2} & 0 & \sqrt{1/2} & 0 \\ 0 & 0 & 0 & i \end{bmatrix} \tag{135}$$

Then then applying that to $\mathbf{b} = |0101\rangle$, and using matrix multiplication, gives,

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & i & 0 & 0 \\ \sqrt{1/2} & 0 & -\sqrt{1/2} & 0 \\ -\sqrt{1/2} & 0 & \sqrt{1/2} & 0 \\ 0 & 0 & 0 & -i \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} |i| \\ 0 \\ 0 \\ |-i| \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \tag{136}$$

While it is often said the main advantage of quantum computing is parallelism, as in it can run many tasks in one computation, however this isn't entirely true as the quantum computer does these computation exponentially faster than a normal computer in regards to each individual computation. This can reduce the problem of checking whether each number between 1 and n is a factor of n, from the initial problem described in Equation 121 to having each qubit check one of those numbers, and then combine together so they are essentially computing the factorisation problem in 2 steps.

iv. Tensor Product

The Tensor Product is another operation on a superposition of states, and uses the notation $\otimes$. It follows the distributive property of vector addition. Given the two vectors $\mathbf{a}_i$ and $\mathbf{b}_i$,

$$\bigotimes_{i=1}^{n} (\mathbf{a}_i|0\rangle + \mathbf{b}_i|1\rangle) = \mathbf{a}_1|0\rangle \otimes \left(\bigotimes_{i=2}^{n} \mathbf{a}_i|0\rangle + \mathbf{b}_i|1\rangle \right) + \mathbf{b}_1|1\rangle \otimes \left(\bigotimes_{i=2}^{n} \mathbf{a}_i|0\rangle + \mathbf{b}_i|1\rangle \right) \tag{137}$$

It is much easier to understand this using an example,

$$(\mathbf{a}_1|0\rangle + \mathbf{b}_1|1\rangle) \otimes (\mathbf{a}_2|0\rangle + \mathbf{b}_2|1\rangle) = \mathbf{a}_1\mathbf{a}_2|00\rangle + \mathbf{a}_1\mathbf{b}_2|01\rangle + \mathbf{b}_1\mathbf{a}_2|10\rangle + \mathbf{b}_1\mathbf{b}_2|11\rangle \tag{138}$$

v. Quantum Fourier Transform

In Section V the concept of the Discrete Fourier Transform were described. Now using the properties of a quantum superposition, adjustments can be made to the DFT to allow for quicker computation. In the QFT, there transformation is down on the amplitudes of the quantum states within a quantum superposition. For a set of computer superpositions $|\phi\rangle$, each of which denoted j, the transformation maps,

$$|\phi\rangle = \sum_{j=0}^{N-1} |j\rangle \mapsto \sum_{j=0}^{N-1} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \exp\left(\frac{-2\pi ijk}{N}\right)|k\rangle \tag{139}$$

Therefore each of the states $|j\rangle$ are mapped by,

$$|j\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \exp\left(\frac{-2\pi ijk}{N}\right)|k\rangle \tag{140}$$

The first equation is going from the sum of j quantum states wherein the $1/\sqrt{n}$, or the probability, of each quantum state is transferred to a new value. This while the quantum superposition stays unobserved wherein the transformation is just passing the qubits through quantum gates.

vi. Shor's Algorithm

To find the correct primes p and q is simply a matter of filtering out all the noise, to reveal the period of $a^r \mod n$, as seen in Equation 123. To illustrate this mathematically the probability of each qubit being a factor is given as,

$$\frac{1}{\sqrt{n}}|10\rangle + \frac{1}{\sqrt{n}}|11\rangle + \ldots + \frac{1}{\sqrt{n}}|n\rangle \tag{141}$$

Wherein each state has probability $1/n$. Now when an observer goes to measure that state, a quantum computer without Shor's algorithm is just as good as a normal computer (it would just be a random guess). In a sense, Shor's algorithm just amplifies the states, giving them a higher probability, so that the noise all the random states are filtered out. Now suppose n has period r, therefore it can be said that,

$$r \approx \frac{1}{\sqrt{\mathcal{N}}}|1\rangle + \frac{1}{\sqrt{\mathcal{N}}}|10\rangle + \ldots + \frac{1}{\sqrt{\epsilon}}|r\rangle + \ldots + \frac{1}{\sqrt{\mathcal{N}}}|N\rangle \tag{142}$$

Where $\epsilon \ll 1$, $\mathcal{N} \gg 1$, $\epsilon + \mathcal{N}(N-1) = 1$. through destructive and constructive interference. As previously discussed in Subsection V vii that the Discrete Fourier Transform (DFT) is an inefficient way of finding the period of a function. However when using the Quantum Fourier Transform from Subsection VI v this can be reduced to a polynomial-time solvable problem, and hence it is efficient. In this case it will be with the knowledge of n, the large number to factorise, and a.

Step 1. Take a register L of length x wherein $n^2 < 2^x < 2n^2$, each qubit is set to $|0\rangle$ and another register R large enough to hold n.

Step 2. Set L to a superposition, as described in Equation 141 for all numbers between 0 and $2^q - 1$ using a quantum gate R_j for each qubit with index j.

Step 3. For all number $r \in L$ calculate $a^r \mod n$ and add this to the register R, this can be done simultaneously for a quantum computer. Now in the registers L and R there is the superposition,

$$\frac{1}{\sqrt{2^q}} \sum_{r=0}^{2^x-1} |r\rangle \otimes |a^r \mod n\rangle \tag{143}$$

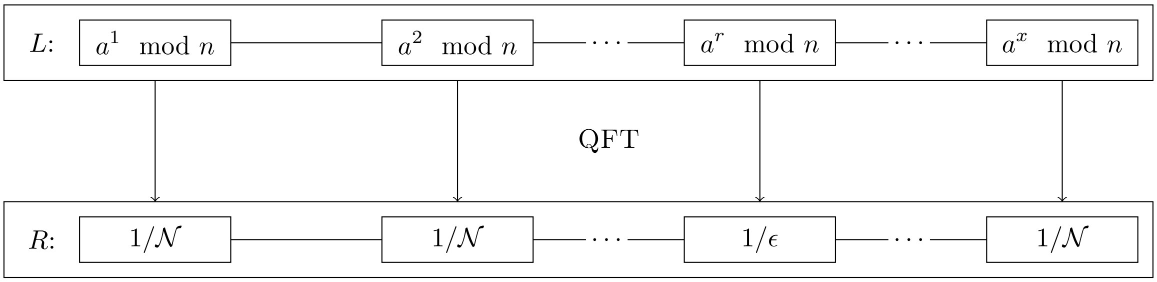
Step 4. Apply the Quantum Fourier Transform to register L such that,

$$|r\rangle = \frac{1}{\sqrt{2^x}} \sum_{c=0}^{2^x-1} \exp\left(\frac{2i\pi rc}{2^x}\right)|c\rangle \tag{144}$$

Therefore,

$$L \mapsto \frac{1}{\sqrt{2^x}} \sum_{r=0}^{2^x-1} \frac{1}{\sqrt{2^x}} \sum_{c=0}^{2^x-1} \exp\left(\frac{2i\pi rc}{2^x}\right)|c\rangle \otimes |a^r \mod n\rangle \tag{145}$$

What this step is doing can be illustrated as follows,



So that when the registry R is looked at there is a high probability that the period r comes out, which can be checked very quickly with a classical computer.

While the mathematics behind why this works requires advanced knowledge of complex analysis there is a simple geometric proof using the roots of unity. Using the example $n = 7$, and $a = 2$, it is fairly obvious that the period is 3 following, While this may be obvious for larger numbers its a lot harder, but when applying the QFT to each each

$2^1 \mod 7$	$2^2 \mod 7$	$2^3 \mod 7$	$2^4 \mod 7$	$2^5 \mod 7$	$2^6 \mod 7$	$2^7 \mod 7$
2	4	1	2	4	1	2

element in L an interesting property arises. To illustrate this place a point on the $(1, 0)$ of a unit circle, and then space equally around the circle such that it forms the N roots of unity. Recalling Figure 7 and Figure 9 it can be modified such that for each consecutive term, rotate the vector between the origin and the point, and when it falls on the value 1, record the position of the vector. Visually, for a circle with 3 roots of unity it corresponds to,

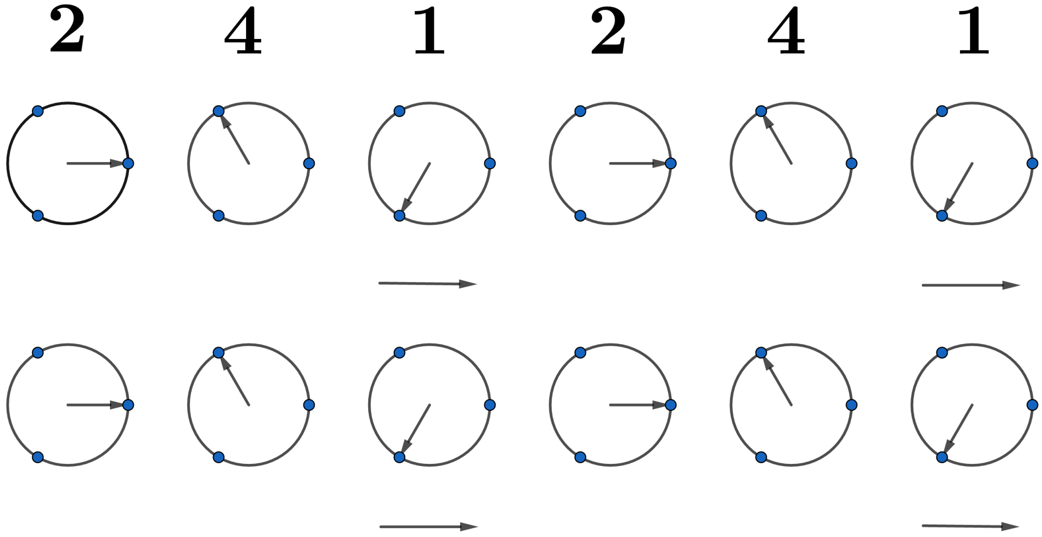


Figure 12: Finding the period of the sequence $2^n \mod 7$ using the Quantum Fourier Transform. Each vector at 1 adds constructively when using 3 roots of unity, whereas for all other non multiples of 3 will become 0 following Figure 9. Therefore the period of the sequence is will be found as 3.

This will give a spike of probabilities a long the multiples of r, and it is then for a classical computer to check whether it works or not. If the process fails, then it is simply a matter of trying again however in most of the cases, it should not take beyond a few tries before the solution is revealed, and it has time complexity of $O((\log N)^2(\log\log N)(\log\log\log N))$ which is computationally efficient which is much quicker than the current quickest time for prime factorisation, using the general number field sieve algorithm (GNFS) $O(e^{1.9(\log N)^{1/3}(\log\log N)^{2/3}})$ which can be visualised this graph,

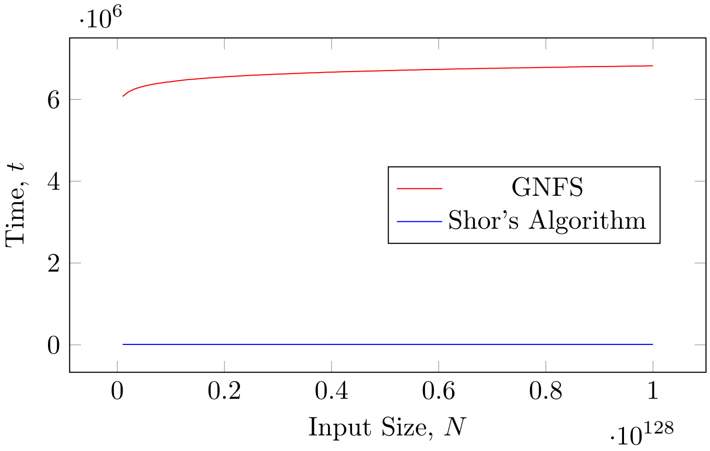


Figure 13: $2^{256} \approx 10^{77}$ is considered the minimum requirement for regular RSA encryption. Theoretically a typical computer would take around 10^6 seconds to solve (11 and a half days), whereas Shor's factorisation algorithm takes around an hour.

VII. Further Investigation

Throughout this essay, the Discrete Fourier Transform has been explained, derived, and applied. It is clear there is a lot more depth to Fourier Series, whether it be in modelling heat, to real analysis, or my own applications to machine learning, and I intend on pursuing this into university. There are also many different properties of Fourier Series that are used to prove other theorems, such as the Nyquist Sampling Limit as part of Information Theory or Hillbert Spaces and Non-Euclidean Geometry which were crucial in deriving General Relativity. Quantum Computers have also seen quite the exposure recently with Google's Bristle Cone quantum computer, which has 72 stable qubits, compared to the 4000 qubits and 100 million gates required for Shor's Algorithm to factor large numbers ($2^{128}+$). There is clearly much more research to be done in the experimental part of quantum information theory. It is to be noted however the beauty in such an old method, Fourier Analysis, being used in applications that would of seemed outlandish to Joseph Fourier.

[SUBJECTS TO TAKE FROM MAIN AND PUT IN APPENDIX ADD HERE]