I. Introduction.

PV system is widely applied into our daily as an important source of renewable energy. However, the efficiency of the system is not always at its best due to many problems. One of that is due to electrical problems such as open circuit, short circuit, arc. Environment also can affect to the system such as partial shading and damaged due to the obstacle that hit the panel of scratch on the surface during installation process. To understand more about the behavior of fault and a sensor system is developed with digital signal processing to filter and extract the character of the fault for fault detection purpose.

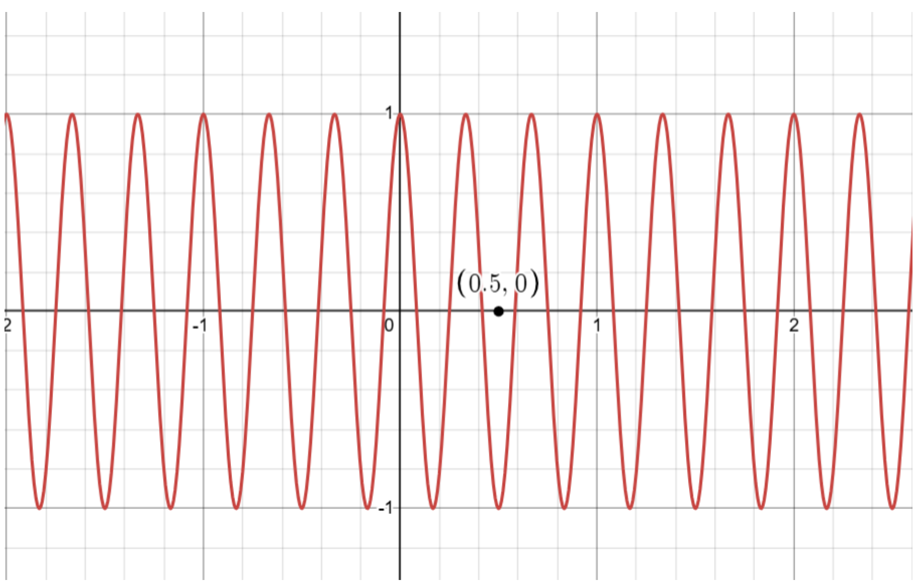
In this project, a low pass filter and FFT method is used for extract important data from collected sample of the system. Try to analyze and collect the characteristic of the fault behavior of the system.

A. DFT and Fast fourier transform

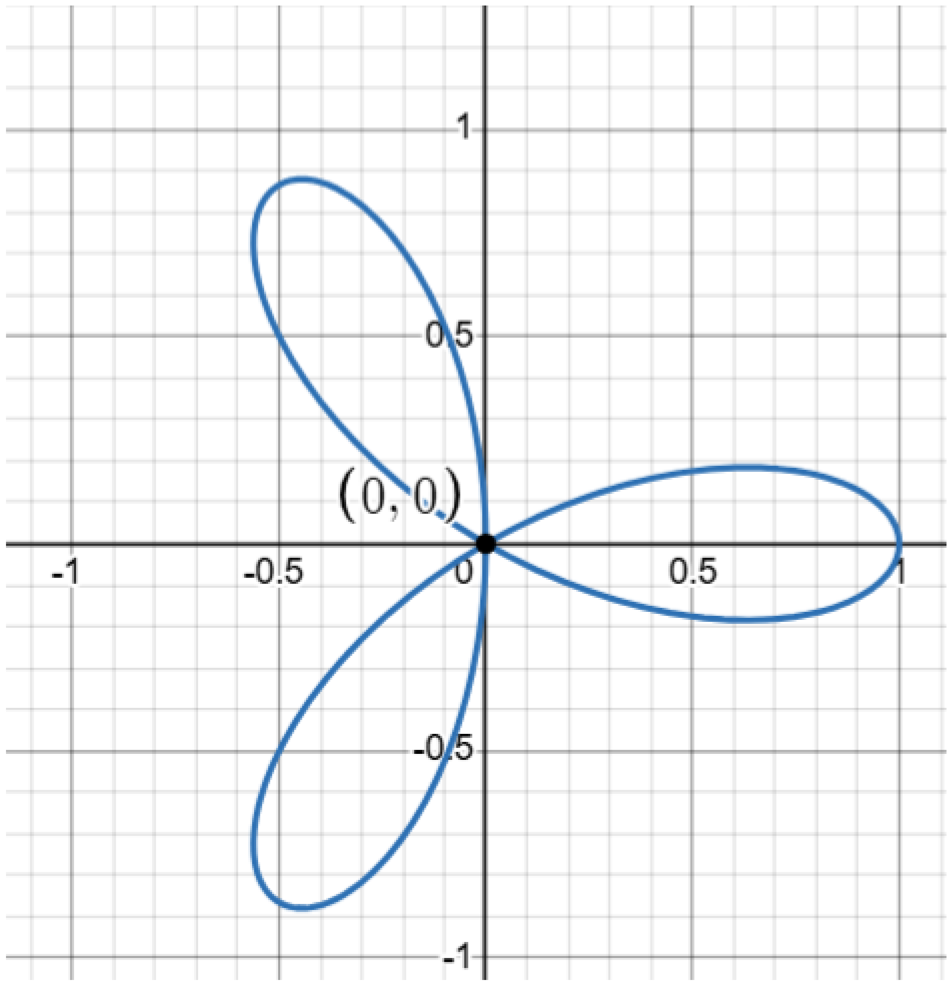
We all hear about the DFT and learn about it, so in this case, I will briefly go this. Before DFT, we should understand about the FT first: and FT is based on the idea how we can decompose or analyzing different signal in a real-life signal, or at composed signal.

So what is the idea of FT, imagen that we have a signal like this:

https://www.desmos.com/calculator/oz3m4oez5f



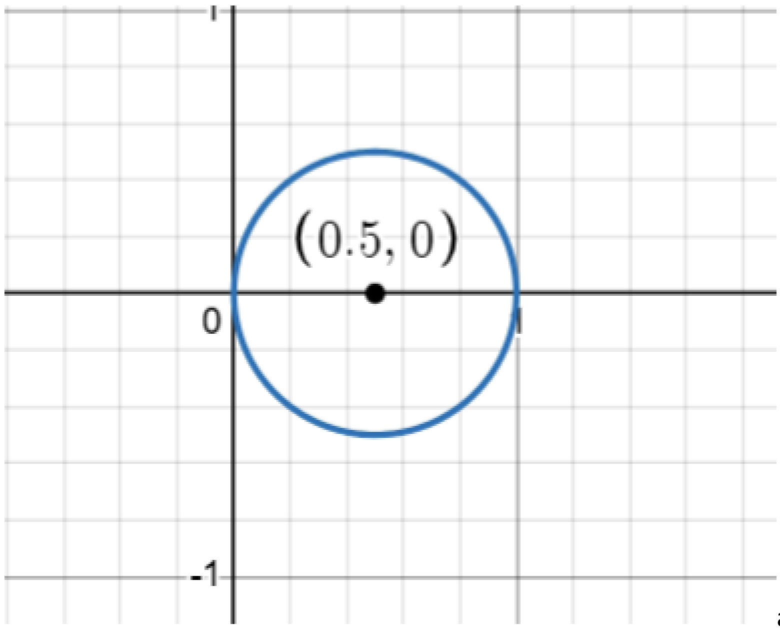
So we want to warp these signal around a circle, may be at 1 second:



So this is a very nice picture, right, we can try different graph, note that there is a black point, it is you can say that it is the center of mass with the formula

$$\left(\frac{1}{5}\int_0^5 g(t)\cos(2a\pi t)\,dt, \frac{1}{5}\int_0^5 g(t)\sin(2a\pi t)\,dt\right)$$

And when you trying to change a in this formula 😊



at a = 3

See the interesting now the center if mass move far away from the origin. And do you notice that 3 is also the frequency of the sine wave 😊

So that lead us to how we can determine or find the frequency of the signal using this technique:

Let $f(t)$ be the continuous signal which is the source of the data. Let N samples be denoted $f[0], f[1], f[2], \ldots, f[k], \ldots, f[N-1]$.
The Fourier Transform of the original signal, $f(t)$, would be

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

We could regard each sample $f[k]$ as an *impulse* having area $f[k]$. Then, since the integrand exists only at the sample points:

$$\begin{aligned} F(j\omega) &= \int_{o}^{(N-1)T} f(t)e^{-j\omega t}dt \\ &= f[0]e^{-j0} + f[1]e^{-j\omega T} + \ldots + f[k]e^{-j\omega kT} + \ldots f(N-1)e^{-j\omega(N-1)T} \end{aligned}$$

$$\text{ie.} \qquad F(j\omega) = \sum_{k=0}^{N-1} f[k]e^{-j\omega kT}$$

i.e. set $\omega = 0, \frac{2\pi}{NT}, \frac{2\pi}{NT} \times 2, \ldots \frac{2\pi}{NT} \times n, \ldots \frac{2\pi}{NT} \times (N-1)$

or, in general

$$F[n] = \sum_{k=0}^{N-1} f[k] e^{-j\frac{2\pi}{N}nk} \quad (n = 0 : N-1)$$

Let us sample $f(t)$ at 4 times per second (ie. f_s = 4Hz) from $t = 0$ to $t = \frac{3}{4}$. The values of the discrete samples are given by:

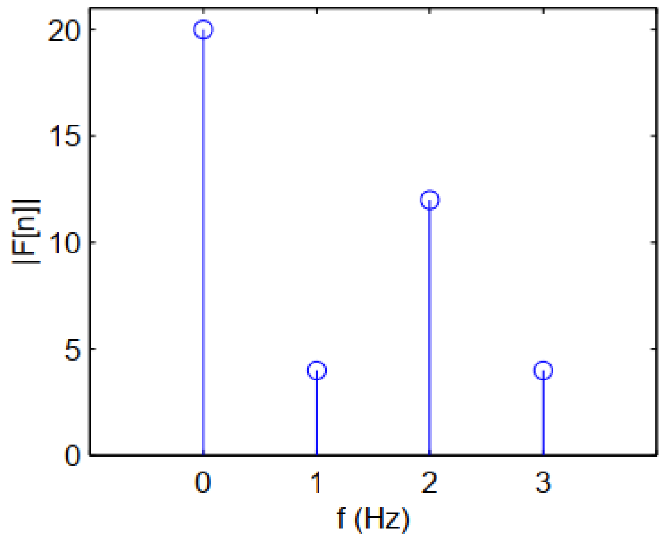
$f[k] = 5 + 2cos(\frac{\pi}{2}k - 90^o) + 3cos\pi k$ by putting $t = kT_s = \frac{k}{4}$

i.e. $f[0] = 8$, $f[1] = 4$, $f[2] = 8$, $f[3] = 0$, $(N = 4)$

$$\text{Therefore} \quad F[n] = \sum_{0}^{3} f[k] e^{-j\frac{\pi}{2}nk} = \sum_{k=0}^{3} f[k](-j)^{nk}$$

$$\begin{pmatrix} F[0] \\ F[1] \\ F[2] \\ F[3] \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{pmatrix} \begin{pmatrix} f[0] \\ f[1] \\ f[2] \\ f[3] \end{pmatrix} = \begin{pmatrix} 20 \\ -j4 \\ 12 \\ j4 \end{pmatrix}$$

The magnitude of the DFT coefficients is shown below in Fig. 7.3.



So that is for the DFT so what is about the FFT ?

Hm so based on the Wikipedia, which I read and some other DSP book. FFT is DFT but faster. In

1. **Divide-and-Conquer**:
   - The algorithm breaks down a DFT of size N into two smaller DFTs of size $N/2$ repeatedly until reaching the base case of size $N = 1$.
   - This is why it is called a **radix-2** algorithm—it splits the data into two parts at each stage.
2. **Input Reordering**:
   - The input sequence is reordered into **bit-reversed order** before processing.

The radix-2 DIT algorithm rearranges the DFT of the function x_n into two parts: a sum over the even-numbered indices $n = 2m$ and a sum over the odd-numbered indices $n = 2m+1$:

$$X_k = \sum_{m=0}^{N/2-1} x_{2m} e^{-\frac{2\pi i}{N}(2m)k} + \sum_{m=0}^{N/2-1} x_{2m+1} e^{-\frac{2\pi i}{N}(2m+1)k}$$

One can factor a common multiplier $e^{-\frac{2\pi i}{N}k}$ out of the second sum, as shown in the equation below. It is then clear that the two sums are the DFT of the even-indexed part x_{2m} and the DFT of odd-indexed part x_{2m+1} of the function x_n. Denote the DFT of the ***E***ven-indexed inputs x_{2m} by E_k and the DFT of the ***O***dd-indexed inputs x_{2m+1} by O_k and we obtain:

$$X_k = \underbrace{\sum_{m=0}^{N/2-1} x_{2m} e^{-\frac{2\pi i}{N/2}mk}}_{\text{DFT of even-indexed part of } x_n} + e^{-\frac{2\pi i}{N}k} \underbrace{\sum_{m=0}^{N/2-1} x_{2m+1} e^{-\frac{2\pi i}{N/2}mk}}_{\text{DFT of odd-indexed part of } x_n} = E_k + e^{-\frac{2\pi i}{N}k} O_k \qquad \text{for } k = 0, \dots, \frac{N}{2} - 1.$$

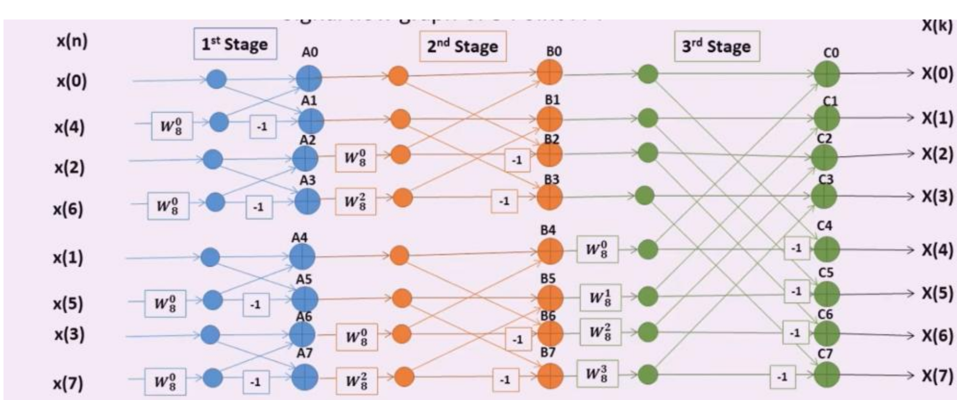
Note that the equalities hold for $k = 0, \dots, N-1$, but the crux is that E_k and O_k are calculated in this way for $k = 0, \dots, \dfrac{N}{2} - 1$ only. Thanks to the periodicity of the complex exponential, $X_{k+\frac{N}{2}}$ is also obtained from E_k and O_k:

$$\begin{aligned}
X_{k+\frac{N}{2}} &= \sum_{m=0}^{N/2-1} x_{2m} e^{-\frac{2\pi i}{N/2}m(k+\frac{N}{2})} + e^{-\frac{2\pi i}{N}(k+\frac{N}{2})} \sum_{m=0}^{N/2-1} x_{2m+1} e^{-\frac{2\pi i}{N/2}m(k+\frac{N}{2})} \\
&= \sum_{m=0}^{N/2-1} x_{2m} e^{-\frac{2\pi i}{N/2}mk} e^{-2\pi m i} + e^{-\frac{2\pi i}{N}k} e^{-\pi i} \sum_{m=0}^{N/2-1} x_{2m+1} e^{-\frac{2\pi i}{N/2}mk} e^{-2\pi m i} \\
&= \sum_{m=0}^{N/2-1} x_{2m} e^{-\frac{2\pi i}{N/2}mk} - e^{-\frac{2\pi i}{N}k} \sum_{m=0}^{N/2-1} x_{2m+1} e^{-\frac{2\pi i}{N/2}mk} \\
&= E_k - e^{-\frac{2\pi i}{N}k} O_k
\end{aligned}$$

We can rewrite X_k and $X_{k+\frac{N}{2}}$ as:

$$\begin{aligned}
X_k &= E_k + e^{-\frac{2\pi i}{N}k} O_k \\
X_{k+\frac{N}{2}} &= E_k - e^{-\frac{2\pi i}{N}k} O_k
\end{aligned}$$

An example:



- Here 4 numbers of Twiddle factors are being used
- We know that $W_N^m = e^{-j\frac{2\pi m}{N}} = cos(\frac{2\pi m}{N}) - j\,sin(\frac{2\pi m}{N})$
- $W_8^0 = e^{-j\frac{2\pi 0}{8}} = cos(\frac{2\pi 0}{8}) - j\,sin(\frac{2\pi 0}{8}) = 1$
- $W_8^1 = e^{-j\frac{2\pi 1}{8}} = cos(\frac{2\pi 1}{8}) - j\,sin(\frac{2\pi 1}{8}) = 0.7071 - j*0.7071$
- $W_8^2 = e^{-j\frac{2\pi 2}{8}} = cos(\frac{2\pi 2}{8}) - j\,sin(\frac{2\pi 2}{8}) = -j$
- $W_8^3 = e^{-j\frac{2\pi 3}{8}} = cos(\frac{2\pi 3}{8}) - j\,sin(\frac{2\pi 3}{8}) = -0.7071 - j*0.7071$

Computation of DFT using Radix-2 DIT FFT algorithm

$W_8^0 = 1$

INPUT

x(n)={1, 1, 1, 1, 2, 2, 2, 2}

1st Stage

x(n)

x(0) = 1

x(4) = 2

x(2) = 1

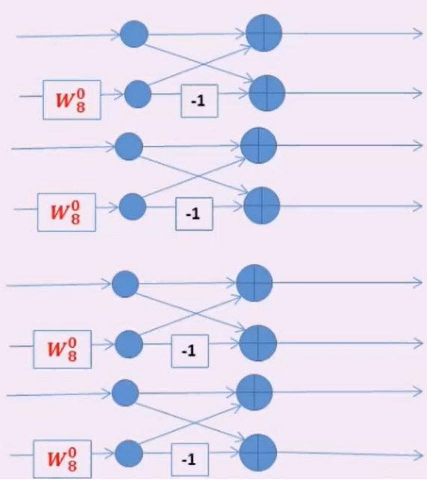
x(6) = 2

x(1) = 1

x(5) = 2

x(3) = 1

x(7) = 2



A0 = x(0) + W_8^0x(4) = 1 + 1*2 = 3

A1 = x(0) - W_8^0x(4) = 1 - 1*2 = -1

A2 = x(2) + W_8^0x(6) = 1 + 1*2 = 3

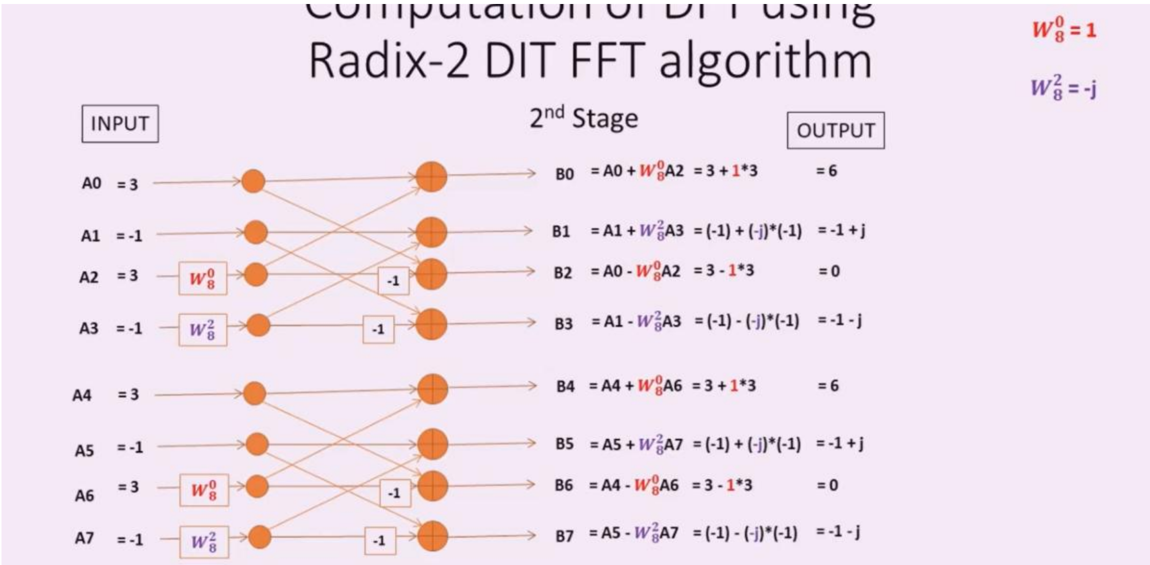
A3 = x(2) - W_8^0x(6) = 1 - 1*2 = -1

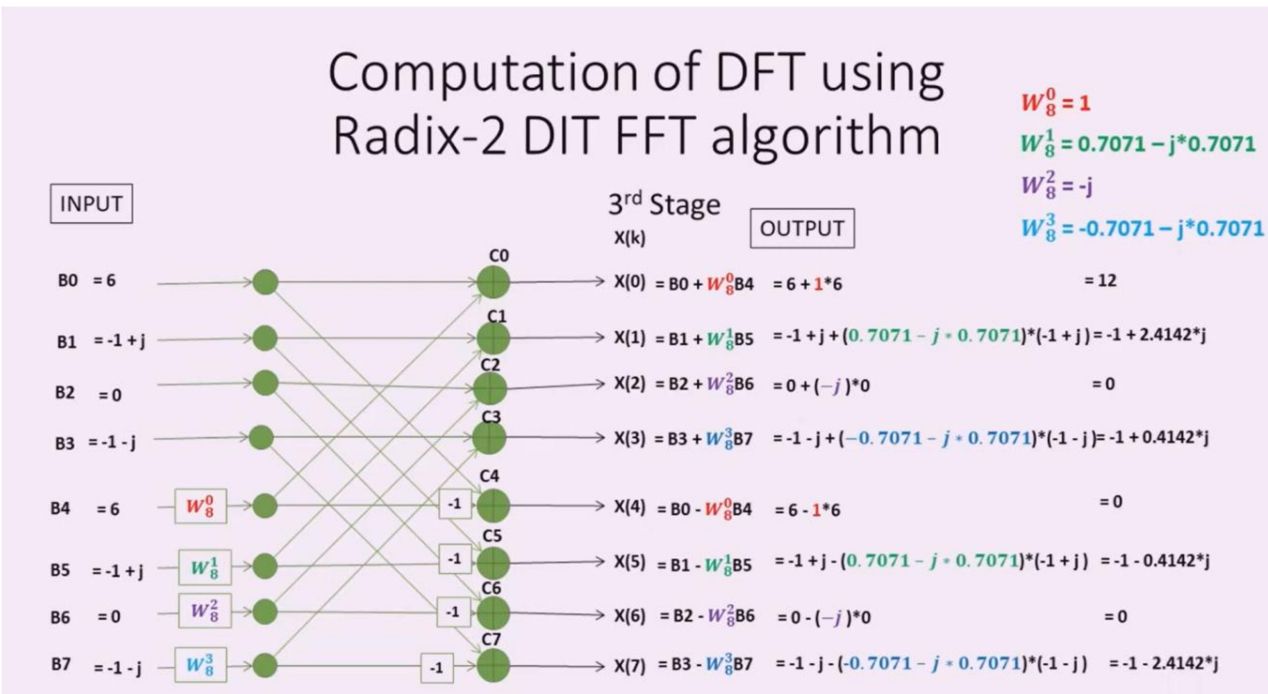
A4 = x(1) + W_8^0x(5) = 1 + 1*2 = 3

A5 = x(1) - W_8^0x(5) = 1 - 1*2 = -1

A6 = X(3) + W_8^0x(7) = 1 + 1*2 = 3

A7 = x(3) - W_N^0x(7) = 1 - 1*2 = -1





B. Moving average filter

The moving average is the most common filter in DSP, mainly because it is the easiest digital filter to understand and use. In spite of its simplicity, the moving average filter is optimal for a common task: reducing random noise while retaining a sharp step response. This makes it the premier filter for time domain encoded signals. However, the moving average is the worst filter for frequency domain encoded signals, with little ability to separate one band of frequencies from another.

There are many versions of this moving average filter such as moving average filter include the Gaussian, Blackman, and multiple pass moving average. These have slightly better performance in the frequency domain, at the expense of increased computation time. However, in this project, for its simple, I will only use the moving average filter and multiple-pass moving average filters.

As the name implies, the moving average filter operates by averaging a number of points from the input signal to produce each point in the output signal. In equation form, this is written:

$$y[i] = \frac{1}{M} \sum_{j=0}^{M-1} x[i+j]$$

Where

- x [ ] is the input signal,
- y [ ] is the output signal, and
- M is the number of points in the average.

For example, in a 5 point moving average filter, point 80 in the output signal is given by:

$$y[80] = \frac{x[80] + x[81] + x[82] + x[83] + x[84]}{5}$$

$$y[80] = \frac{x[78] + x[79] + x[80] + x[81] + x[82]}{5}$$

the moving average filter is a convolution of the input signal with a rectangular pulse having an area of one.

Extra additional:

Frequency Response

Figure 15-2 shows the frequency response of the moving average filter. It is mathematically described by the Fourier transform of the rectangular pulse, as discussed in Chapter 11:

EQUATION 15-2
Frequency response of an M point moving average filter. The frequency, f, runs between 0 and 0.5. For $f = 0$, use: $H[f] = 1$

$$H[f] = \frac{\sin(\pi f M)}{M \sin(\pi f)}$$

Multiple-pass moving average filters involve passing the input signal through a moving average filter two or more times. Figure 15-3a shows the overall filter kernel resulting from one, two and four passes. Two passes are equivalent to using a triangular filter kernel (a rectangular filter kernel convolved with itself). After four or more passes, the equivalent filter kernel looks like a Gaussian (recall the Central Limit Theorem). As shown in (b), multiple passes produce an "s" shaped step response, as compared to the straight line of the single pass. The frequency responses in (c) and (d) are given by Eq. 15-2 multiplied by itself for each pass. That is, each time domain convolution results in a multiplication of the frequency spectra.

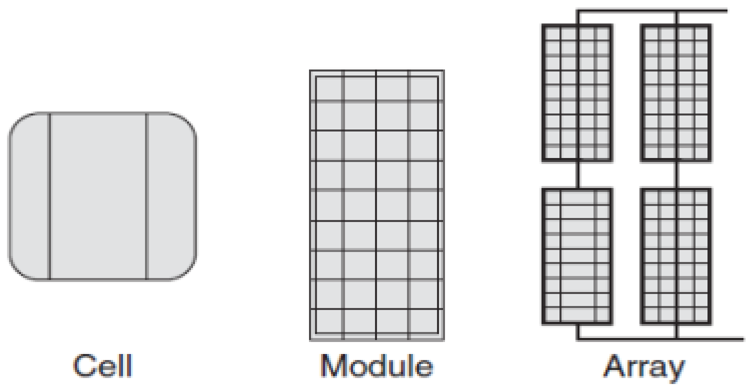
C. Low-pass filter

About the low pass filter, I will use the digital low pass filter, cause my samples is sampled with 1000 Hz (why cause I am measuring the dc value which don't have frequency and if it have its will also for less than 200 Hz due to noise and AC voltage as 54Hz.

D. PV system basic knowledge

1. PV arrays

In this section, I will briefly presence about the PV arrays or PV system.



So the fundamental component of a PV systems is PV cell, which can produce about 0.5V to 0.6V. The working principle of the cell is basic and easy to understand. It is p-n junction, when a photon – or light coming to the pn junction it will excite the electron in the junction and then the electron release. The e- is moved through the circuit and then create a e- flow in the external circuit.

And to increase the power of the system, 1 cell is not enough so it come to connect those cell series and parallel which create a PV panel.

Based on the voltage and current used and required of the manufacture, sometimes, all the cell is connected parallel or series or both parallel and series. But the parallel faced more significant problem then series due to the voltage loss of partial shading: due to the parallel, it limits the total voltage of parallel which is not affect much in the series.

The last level is PV arrays, which is very visible and the final product, the same for the PV field. Modules can be wired in parallel to enhance current or in series to boost voltage. To increase power, modules in arrays are connected in both series and parallel. Typically, modules are first set up in a series string to increase the voltage to the maximum level that safety concerns will allow before paralleling those strings to increase power. By using this technique, $II2RR$ power loss in linked cables is reduced.

The PV system is usually installed with many protections and controlled component as circuit breaker, diode and Inverter – note that the inverter usually compiles with an MPPT – maximum power point tracker.

2. Kind of fault in PV array and detection technique

a. Open circuit

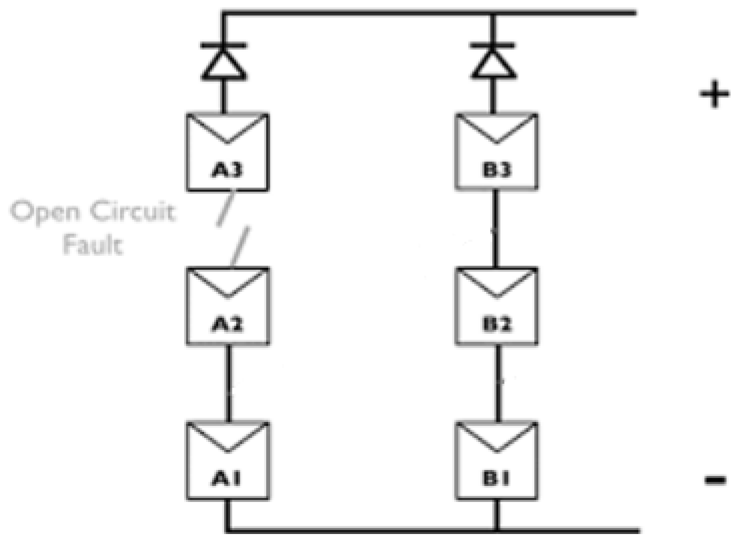
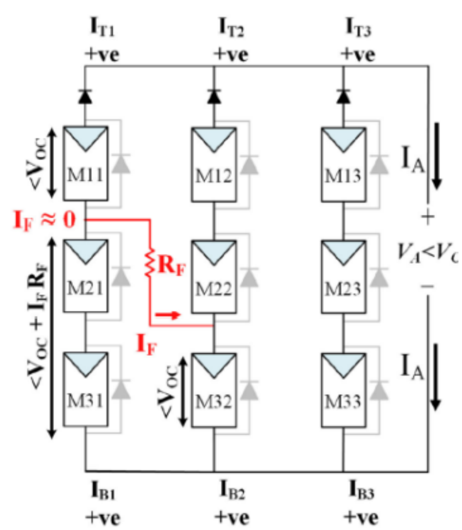


Figure 2.24 Open circuit fault in the PV system [2].

An open circuit fault is an accidental disconnection in a DC current carrying conductor as shown in Figure 2.24. Open circuits are often caused by technical faults when the connections between the panels are not strong. In addition, environmental factors such as wind, storms, and animals also contribute to the open circuit fault. When the open circuit fault occurs, the opening voltage of these cases hardly changes, while the current is short. circuit and peak power decreases linearly with an increase in the number of disconnected wires.

b. Short-circuit



Pole-to-pole faults can occur inside PV arrays and are likely related to large fault currents or DC arcs. Pole-to-pole fault is defined as a random short circuit between two points in an array of different potentials, as shown in Figure 2.23. According to the results, the more panels in series with short circuit failure, the lower the system performance.

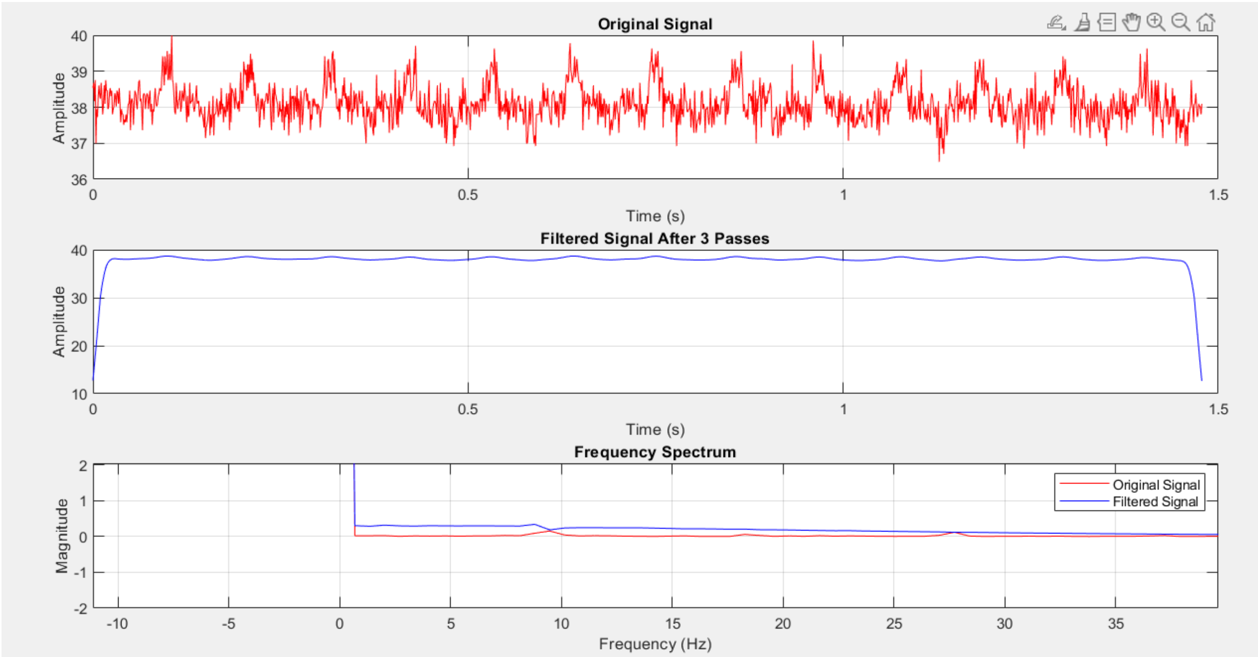
c. Partial shading

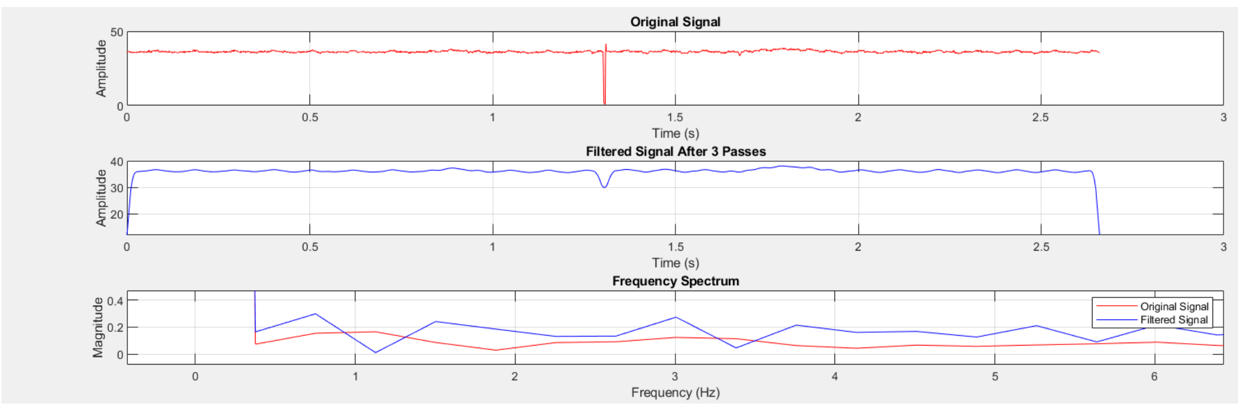
Partial shading is one of the common faults that occur during the operation of the solar cell system, as shown in Figure 2.21. The cause of this fault, in Asian countries and Vietnam, is often due to dirt clinging to and covering part of the solar panel. In addition, fallen trees or shaded trees are also one of the causes of shadow faults. The partial shading causes the performance of the entire system to be significantly reduced, and it is also the cause of hot spots on the module and shortens the life of the module.

II. MATLAB code

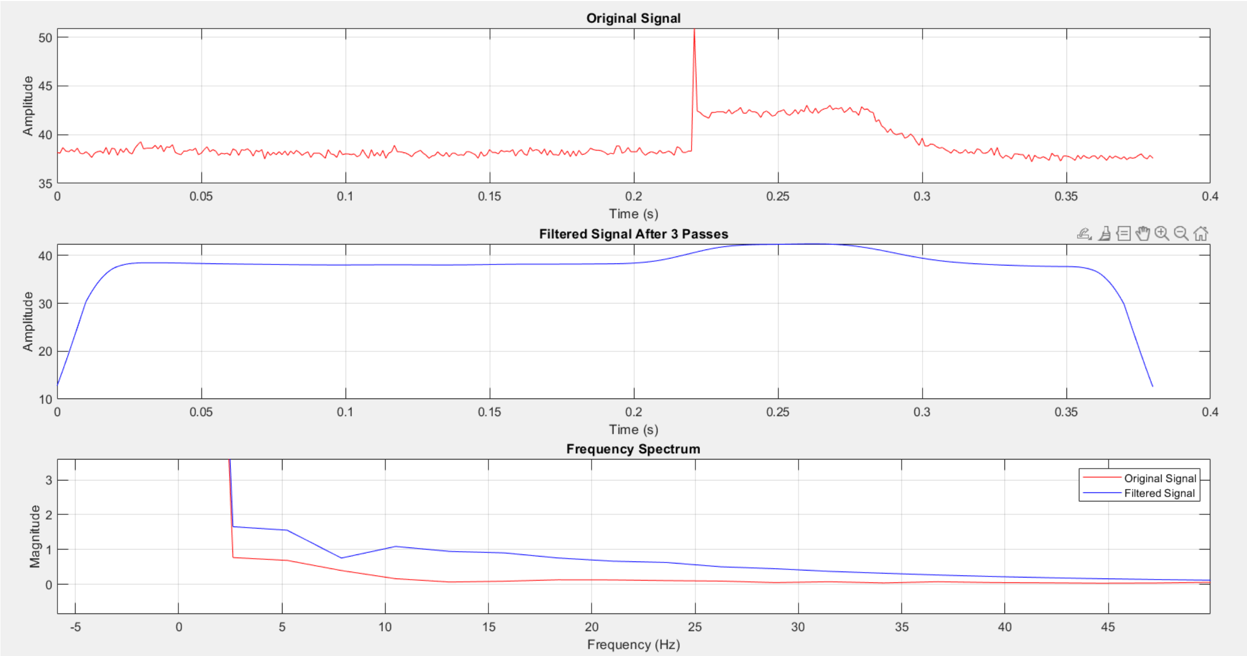
A. FFT

In normal state:

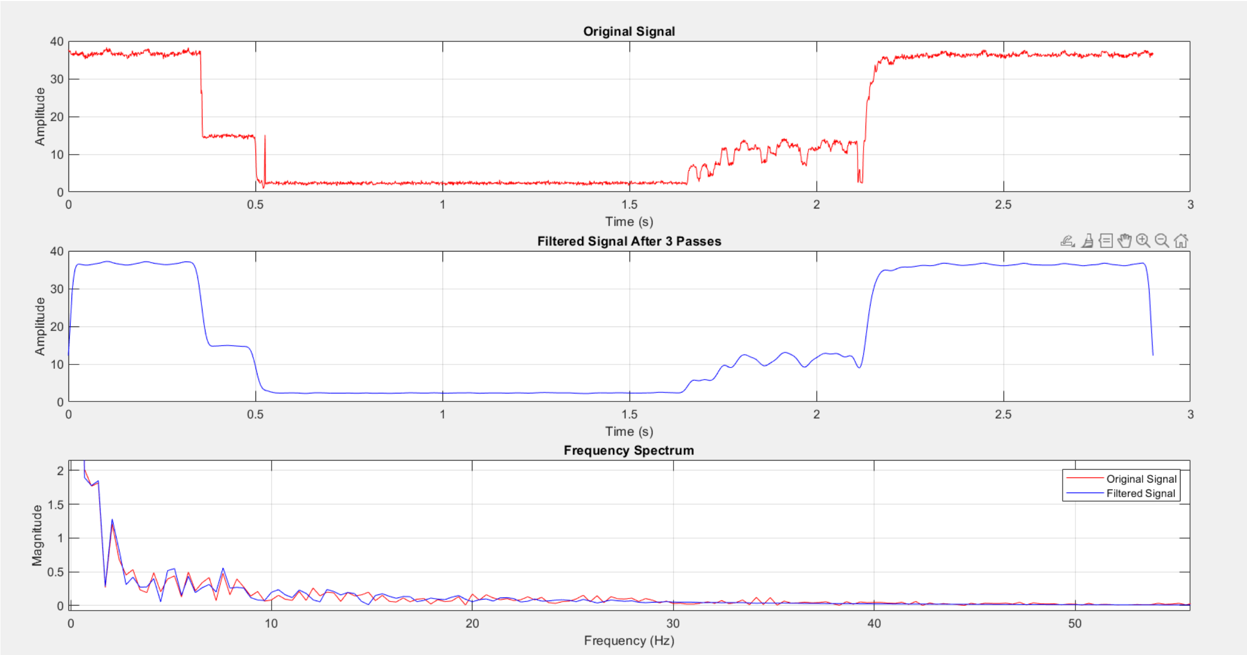




In short-circuit state,



In shading state,



B. Fault detection technique

1. Threshold

It is the most common technique to apply to the fault system technique. In the raw signal, we can apply the upper and lower threshold for the voltage signal, but after applied to some data, it is also required to use the FFT to overcome the peak of the inverter. It is used the MPPT tracker so optimize the power deliver to the system, the peak is the result of that process. To over come this, a stage where the signal is transferred into frequency domain using FFT technique and then we do analysis.

If the frequency domain has 2Hz component which is larger than 1 in amplitude, we can say that there is fault in our system. If not, we can consider it as the peak created from the MPPT.

2. Statistic method

Using mean value and variation.

1. **Training Phase** (Establish Normal Behavior):
   - Collect data under normal operating conditions.
   - Compute the mean and standard deviation for voltage (V) or current (I).

$$\mu_V = \frac{1}{N}\sum_{i=1}^{N} V_i, \quad \sigma_V = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(V_i - \mu_V)^2}$$

2. **Real-Time Monitoring**:
   - Measure real-time data (V_{real}).
   - Check if V_{real} satisfies:

$$|\mu_V - V_{\text{real}}| > k \cdot \sigma_V$$

   - If this condition is true, a fault is flagged.

III. Conclusion

In this project, I have introduced you the basic idea and visualization of DFT and FFT, then some basic knowledge about the PV system and behavior of the PV system. Also, I implement this into finding the fault of the PV and some PV fault behavior.

IV. Reference

https://www.youtube.com/watch?v=spUNpyF58BY&t=635s

https://www.sciencedirect.com/science/article/pii/S0030402624001967

My bachelor thesis