Music Reach Density

- A diffusive model

BT5051 CFA exercise

Krushan Bauva (BE17B037)
November 02, 2019
Dept. Of Biotechnology
Indian Institue of Technology, Madras

Contents

1. Motivation

1. 1. What’s the cost of fame?

Pink Floyd’s 1975 album *Wish You Were Here* is an undisputed rock classic. With the cover—two businessmen shaking hands, one of them on fire—it explored themes around

fame, absence, greed, ambition, the music business, and featured a homage to departed founding member Syd Barrett and the demons that plagued him.

Wish You Were Here was the followup album to *Dark Side of the Moon*, a huge success that gave the band international fame. *WYWH* was a more melancholic and reflective record, looking back at their success and what they’d lost along the way. The result was one part lament for their fallen comrade Syd Barrett and one part vitriolic takedown of the industry that they blame for their strife.

Its well known that the music industry will chew up and spit out creative individuals and the band explored that theme in the middle of the album.

“Welcome to the Machine”, one of the songs included in this album, is sung from the music industry professional talking to a creative. Without caring who the artist really is, the label will sell them with the manufactured cliched biography. The ominous, imposing music and mechanical sounds reflect the cold, inhuman aspect of the music biz. This takedown of the industry is also felt in *“Have a Cigar”* where the song is written from the perspective of a music executive who cynically woes the band with money and success and asks the now famous lyric "Oh, by the way, which one's Pink?".

The album has references to the real interactions the band had with the executives who frequently pushed numbers and chart success on them without actually caring about the musical aspirations.

The album packaging reflects these jaded themes too. Look at the front cover which features a business deal being done with one sided literally getting burnt. On the back of the album we see a faceless businessman in a desert reflecting the barren inhuman nature of the industry executives in Pink Floyd’s eyes.

Being a Pink Floyd fan, this album has always had a special place in my heart and it really made me wonder what music business really is like. That’s why I took this up to understand and model how music propagates and receives fame and popularity.

2. Diffusion – background

The word diffusion derives from the Latin word, *diffundere*, which means "to spread out." Diffusion is net movement of anything (e.g., atom, ions, molecules) from a region of higher concentration to a region of lower concentration. Diffusion is driven by a gradient in concentration. For example: if you spray perfume at one end of a room eventually the gas particles will be all over the room.

The concept of diffusion is widely used in many fields, including physics (particle diffusion), chemistry, biology, sociology, economics, and finance (diffusion of people, ideas and of price values). The central idea of diffusion, however, is common to all of these: an object (e.g. atom, idea, etc.) that undergoes diffusion spreads out from a point or location at which there is a higher concentration of that object.

2. 1. Fick's first law of diffusion:

Fick's first law relates the diffusive flux to the concentration under the assumption of steady state. It postulates that the flux goes from regions of high concentration to regions of low concentration, with a magnitude that is proportional to the concentration gradient (spatial derivative), or in simplistic terms the concept that a solute will move from a region of high concentration to a region of low concentration across a concentration gradient. In one (spatial) dimension, the law can be written in various forms, where the most common form is in a molar basis:

$$J = -D \frac{d\varphi}{dx}$$

where:

- J is the diffusion flux, of which the dimension is amount of substance per unit area per unit time. J measures the amount of substance that will flow through a unit area during a unit time interval.
- D is the diffusion coefficient or diffusivity. It's dimension is area per unit time.
- φ (for ideal mixtures) is the concentration, of which the dimension is amount of substance per unit volume.
- x is position, the dimension of which is length.

2. 2. Fick’s second law of diffusion:

Fick's second law predicts how diffusion causes the concentration to change with respect to time. It is a partial differential equation which in one dimension reads:

$$\frac{\partial \varphi}{\partial t} = D \frac{\partial^2 \varphi}{\partial x^2}$$

where:

- φ is the concentration; $\varphi = \varphi(x,t)$ is a function that depends on location x and time t.
- t is time.
- D is the diffusion coefficient.
- x is the position.

3. Material Balance – background

A material balance, also called mass balance, is an application of conservation of mass to the analysis of physical systems. By accounting for material entering and leaving a system, mass flows can be identified which might have been unknown, or difficult to measure without this technique.
Mathematically, the mass balance for a system is as follows:

$$Input - Output + Generation - Consumption = Accumulation$$

4. Introduction

A music popularity index is a recorded music ranking, classified by popularity, which can be measured in many different ways. Nowadays, they are very common in musical websites, since they offer useful statistics suitable for many applications, such as musical recommendations. There are hundreds of methods to measure the popularity of an artist/song, and most websites have their own measurement system.
A music popularity index generally takes into account both these parameters:

- Reach: The number of unique people reached by the artist/song, found in social networks, blogs, home, etc. It takes into account the unique listeners, watchers of videos, registered fans, fans in the mailing list, etc. It could be described as the total potential fan base, and captures an element of effort on the part of the artist and their fans to spread their content and reach more people.
- Influence: Measures the reaction to those exposed to the artists' content. It takes into account the number of songs or videos played per listener, play-through percentage on songs, the total engagement time, etc. It could be described as the merit of the artist's content.

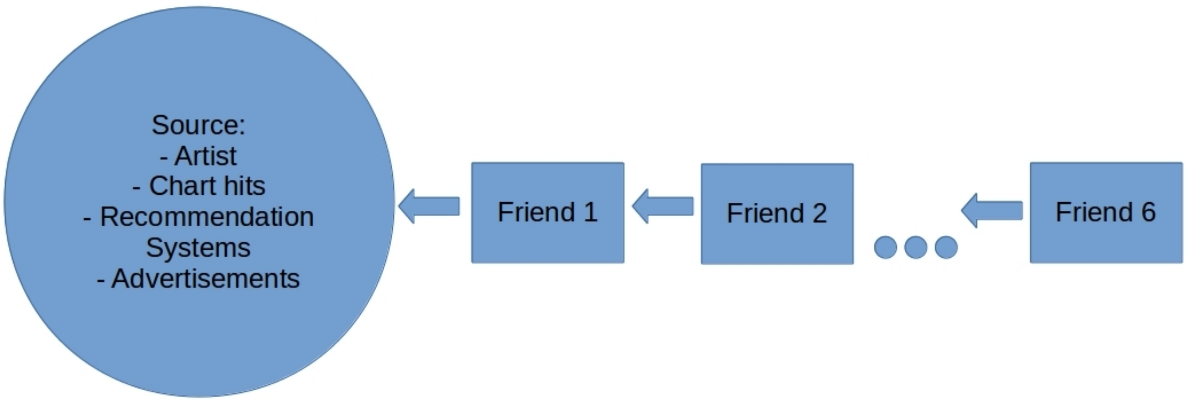
Here, in this analysis of music popularity, we will just be considering the reach of a song which is the number of unique people who heard this song and not the influence.
I will be using music popularity, music reach and music influence interchangably throughout the analysis but it is to be taken as a measure of the number of unique people who heard the song.

5. How do you come across new music?

The various modes or platforms which expose you to new musical content can be categorised under the following:

- Artists that you listen to and whom you follow.
- Chart hits
- Recommendation systems
- Advertisements
- Friends and aquaintances.

If we try to map this, it looks somewhat like this:



Where one would fall somewhere on this map for a particular song. You could be the source and create some content or you could either be the friend 1 who directly gets exposed to new content from the source or you could be the friend 6 who gets to know about the song from friend 5 who in turn gets to know from friend 4 and so on. This serves as an accurate representation of this particular social network.
Hence we can say that this new song starts propagating from the source to a social region where the people are ignorant of this new song. This movement is analogous to the movement of solute molecules through diffusion. In molecular diffusion, the concentration gradient is the driving force. In this case, the music reach density gradient is the driving force.
Remember this mapping is for those people who get exposed to the song or to whom the song reaches. This map does not include people who do not get exposed to that particular song.

5. 1. Why does the network end at friend 6?

Six degrees of separation is the idea that all people are six, or fewer, social connections away from each other. Often called as 6 Handshakes rule. As a result, a chain of “a friend of a friend” statements can be made to connect any two people in a maximum of six steps. It was originally set out by Frigyes Karinthy in 1929 and popularized in eponymous 1990 play written by John Guare. It is sometimes generalized to the average social distance being logarithmic in the size of the population.

This theory is sometimes also refered to as a shrinking world or a small world.

6. Basic Terminology

6. 1. Social distance (r)

It is the measure of closeness of people in a social setting. If two people are aquainted with each other, the social distance between them is 1. So, according to the previous discussion, the maximum social distance between two people will be 6.

6. 2. Music Reach Density (M(r,t))

M(r,t) dentoes the density of users to whom the content was able to reach at a social distance of r at any time t. This density is the ratio of the number of users who heard that particular song at social distance r and time t over the total number of users at the same social distance.

6. 3. Music Reach Flux (H(r))

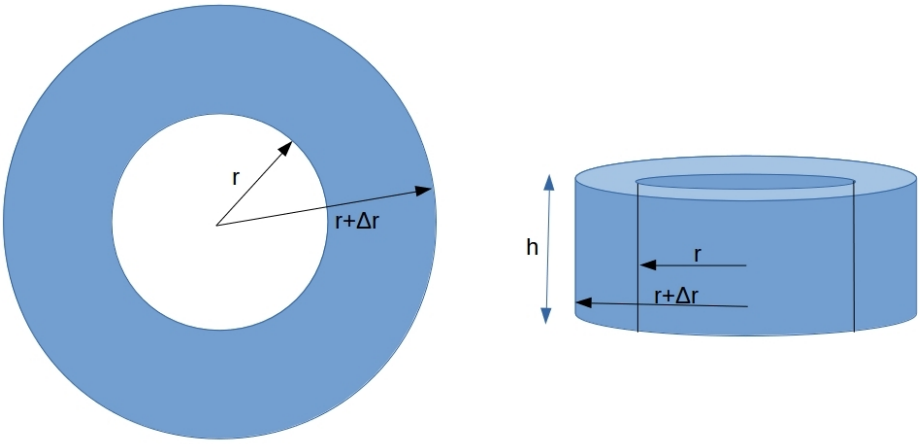
It denotes the music reach density flux at a social distance r.

6. 4. Diffusivity (D)

This term is analogous to the diffusivity in molecular diffusion. Here, it indicates the capability of the music to be propagated in a social setting by diffusion.
This term is strongly dependent on the structure of the song (the music theory behind it). In a study, Norberto Grzywacz, a Professor of Neuroscience and Physics says, “The most popular songs tend to include relative rare chords, that is, they typically have high harmonic surprise. These songs also tend to have choruses with relatively low harmonic surprise preceded by sections with many rare chords.”
Harmonic surprise can be described as where the music deviates from the listeners expectations. Scientists have predicted that these changes in structure coud elicit a pleasurable reward response in the brain. In other words, harmonic surprises can increase the likelihood a song will be a hit.

7. Analysis

Considering this social setting to be a two dimensional network with music flowing radially outwards and applying shell balance for music flow.



$$Input - Output + Generation - Consumption = Accumulation$$

$$Input = H_r(2\pi r h)$$
$$Output = H_{r+\Delta r}(2\pi(r+\Delta r)h)$$
$$Accumulation = \frac{\partial M(r,t)}{\partial t}(2\pi r \Delta r h)$$

Substituting the above terms in the material balance equation

$$H_r(2\pi r h) - H_{r+\Delta r}(2\pi(r+\Delta r)h) = \frac{\partial M(r,t)}{\partial t}(2\pi r \Delta r h)$$

$$\frac{H_r(2\pi r h) - H_{r+\Delta r}(2\pi(r+\Delta r)h)}{(2\pi r \Delta r h)} = \frac{\partial M(r,t)}{\partial t}$$

$$\frac{1}{r}\frac{H_r(r) - H_{r+\Delta r}(r+\Delta r)}{(\Delta r)} = \frac{\partial M(r,t)}{\partial t}$$

Now apply lim $_{\Delta r \to 0}$

$$\frac{1}{r}\lim_{\Delta r\to 0}\frac{H_r(r)-H_{r+\Delta r}(r+\Delta r)}{(\Delta r)}=\frac{\partial M(r,t)}{\partial t}$$

$$\frac{-1}{r}\frac{\partial (rH)}{\partial r}=\frac{\partial M(r,t)}{\partial t}$$

Analogous to Fick's first law,

$$H=-D\frac{\partial M(r,t)}{\partial r}$$

$$\frac{1}{r}\frac{\partial (rD\frac{\partial M(r,t)}{\partial r})}{\partial r}=\frac{\partial M(r,t)}{\partial t}$$

D (diffusivity) is constant, therefore

$$\frac{D}{r}\frac{\partial (r\frac{\partial M(r,t)}{\partial r})}{\partial r}=\frac{\partial M(r,t)}{\partial t}$$

Now the initial and boundary conditions are

IC: At t = 0, M(r,t) = 0

BC 1: At r = 0, M(r,t) = finite, i.e. $\frac{\partial M(r,t)}{\partial r}=0$

BC 2: At r = 7, M(r,t) = 0 (Because of six degrees of separation)

If we assume that M(r,t) is separable as (for a simplistic model)

$$M(r,t)=f(r).g(t)$$

then

$$\frac{\partial M}{\partial t}=f\frac{dg}{dt} \quad \text{and} \quad \frac{\partial M}{\partial t}=g\frac{df}{dr}$$

Therefore

$$f\frac{dg}{dt}=\frac{D}{r}\frac{d\left(r.g\frac{df}{dr}\right)}{dr}$$

$$f\frac{dg}{dt}=\frac{g.D}{r}\frac{d\left(r\frac{df}{dr}\right)}{dr}$$

$$\frac{1}{g}\frac{dg}{dt}=\frac{1}{f}\frac{D}{r}\frac{d\left(r\frac{df}{dr}\right)}{dr}$$

Since the LHS is a function of t alone and the RHS is a function of r alone, for the above equation to hold at all times, each side must be equal to a constant, say k_1^2

$$\frac{1}{g}\frac{dg}{dt}=k^2$$

$$\int\frac{dg}{g}=\int k_1^2 t$$

This implies

$$g=C_1 e^{k_1^2 t}$$

From the other side of the equation

$$\frac{1}{f}\frac{D}{r}\frac{d\left(r\frac{df}{dr}\right)}{dr}=k_1^2$$

$$\frac{D}{r}\frac{d\left(r\frac{df}{dr}\right)}{dr}-k_1^2 f=0$$

$$\frac{1}{r}\frac{d\left(r\frac{df}{dr}\right)}{dr}-\frac{k_1^2}{D}f=0$$

Taking $\frac{k_1^{\ 2}}{D}=k^2$

$$\frac{1}{r}\frac{d(r\frac{df}{dr})}{dr}-k^2 f=0$$

BC 1: At r = 0, f = finite, i.e. $\frac{df}{fr}=0$

BC 2: At r = 7, f = 0 (Because of six degrees of separation)

To solve these equations, knowledge of Bessel functions and their relationships is required. The solution is of the form

$$f=C_2 J_0(kr)+C_3 Y_0(kr)$$

where J_0 is a Bessel function of the first kind

$$J_0(kr)=\sum_{m=0}^{\infty}\frac{(-1)^m(\frac{kr}{2})^{2m}}{(m!)^2}$$

Y_0 is a Weber's Bessel function of the second kind

$$Y_0(kr)=\frac{2}{\pi}[\bar{Y}_0(kr)-(\ln 2-\Gamma)J_0(kr)]$$

where $\Gamma=\lim_{n\to\infty}(1+\frac{1}{2}+\frac{1}{3}+...+\frac{1}{n}-\ln(n))=0.57721...$ (Euler's constant)

and $\bar{Y}_0$ is a Neumann's Bessel function of the second kind.

$$\bar{Y}_0(kr)=J_0(kr)\int\frac{dr}{r[J_0(kr)]^2}$$

C_3 = 0 (from BC 1; otherwise the term would not be finite since $Y_0(0) = -\infty$)

$C_2 J_0(7k) = 0$ (from BC 2).

Now, C_2 cannot be zero since that would result in a trivial solution, (f = 0). Therefore,

$$J_0(7k)=0$$

This happens multiple times when 7k = 2.4048... (= k_1), 5.52009 (= k_2), 8.6537... (= k_3), and so on.

Thus, there are infinite solutions

$$f_n = C_{2n} J_0(k_n r) \qquad \text{for } n = 1, 2, 3, \ldots \infty$$

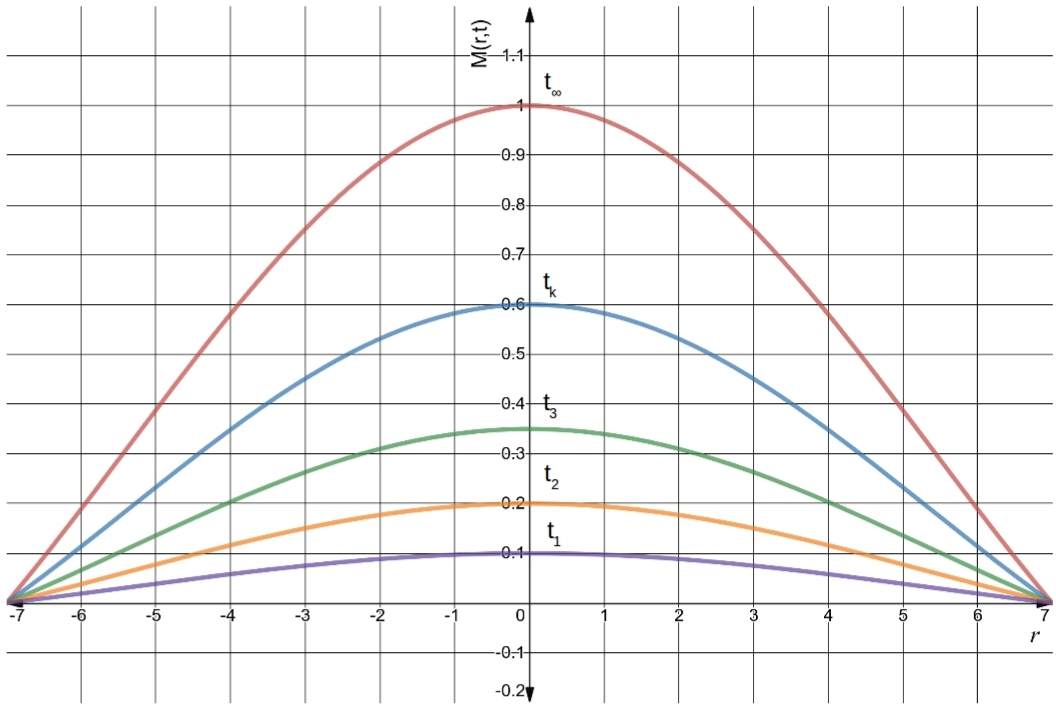
This implies that

$$M(r,t) = f(r).g(t)$$

$$M(r,t) = C_1 C_{2n} J_0(k_n r).\mathrm{e}^{(k_n^2 t)} \qquad \text{for } n = 1, 2, 3, \ldots \infty$$

Thus we found the variation of music reach density (M(r,t)) as a function of social distance (r) and time (t).

A representation plot of M(r,t) verses r for different values of t is given in the figure below.



8. Future Work

8. 1. Music Charts

A music chart, also called a record chart, is a ranking of recorded music according to certain criteria during a given period of time. Many different criteria are used in worldwide charts, often in combination. These include record sales, the amount of radio airplay, the number of downloads, and the amount of streaming activity.
There are a lot of record charts compiled by a lot of companies on various criterias, but we can work on the widely used and recognized Billboard Charts.

Since what we calculated is music reach density, we can find music reach as a function of time by integrating it over the entire area, i.e.

$$Music\ Reach = \int_{r=0}^{r=7} M(r,t)\, 2\pi r\, dr$$

We can then find unknown parameters by fitting the model with real time data from these music charts.

We could also relax the assumption that M(r,t) is separable as f(r) and g(t) and come up with a better model for the same situation.

Apart from this diffusive model, a lot of people are also using data science skills to model and predict music popularity. Links in the reference section.

9. Acknowledgement

I would like to thank Prof. G. K. Suraishkumar for giving me the opportunity to perform this analysis as a part of his course, BT5051 – Transport Phenomena in Biological Systems offered in the independence semester.
All the images in the report are my own work, expect the album cover of Pink Floyd's Wish You Were Here.

10. References

1. Continuum Analysis of Biological Systems – Conserved Quantities, Fluxes and Forces by Prof. G. K. Suraishkumar.
2. Understanding Pink Floyd's Classic Album 'Wish You Were Here' https://www.rockarchive.com/news/2017/understanding-pink-floyd-wish-you-were-here
3. A youtube video on Understanding Pink Floyd's 'Wish You Were Here' by youtuber 'Polyphonic' https://www.youtube.com/watch?v=kJC-cFKP770
4. Diffusion https://en.wikipedia.org/wiki/Diffusion
5. Fick's laws of Diffusion https://en.wikipedia.org/wiki/Fick%27s_laws_of_diffusion
6. Mass Balance https://en.wikipedia.org/wiki/Mass_balance
7. Music Popularity Index https://en.wikipedia.org/wiki/Music_popularity_index
8. Six Degrees of Separation https://en.wikipedia.org/wiki/Six_degrees_of_separation
9. A song's structure can be linked to its popularity https://phys.org/news/2017-08-song-linked-popularity.html
10. Pop songs: does size matter? https://www.independent.co.uk/arts-entertainment/music/features/pop-songs-does-size-matter-2036779.html
11. Music Industry https://en.wikipedia.org/wiki/Music_industry
12. Bessel Function of the First Kind http://mathworld.wolfram.com/BesselFunctionoftheFirstKind.html
13. Plot of bessel funtions with different scaling https://www.desmos.com/calculator/cuvyv45ghs
14. Music Charts https://en.wikipedia.org/wiki/Record_chart
15. Billboard Charts https://en.wikipedia.org/wiki/Billboard_charts
16. Billboard Hot 100 Dominance https://plot.ly/~Sjs05k/185.embed
17. Ultimate Music Database http://www.umdmusic.com
18. Billboard Hot 100 Analytics https://towardsdatascience.com/billboard-hot-100-analytics-using-data-to-understand-the-shift-in-popular-music-in-the-last-60-ac3919d39b49
19. Analyzing the Billboard Hot 100 https://mikekling.com/analyzing-the-billboard-hot-100/
20. Understanding Song Popularity with Automated Machine Learning https://blog.datarobot.com/understanding-song-popularity-with-automated-machine-learning