Interferencia entre Símbolos en Comunicaciones Digitales
I.S.I. (Inter Symbol Interference)

Introducción Pragmática

Las comunicaciones digitales, tal como la palabra dice, tratan sobre transmitir de un lado a otro, (*presumiblemente lejos*), un "mensaje o información" que consta de secuencias de datos que representan información. Básicamente la información, siempre que sea de carácter finita, se puede representar mediante un alfabeto finito o conjunto de símbolos acotado; en consecuencia éstos se pueden enumerar y codificar digitalmente.

Hasta acá todo parece entendible y fácil, pero como los números son una cosa abstracta, y para pasar "cosas" de un lado a otro hay que enviar necesariamente algo medible del otro lado, por ejemplo: vibración en forma de energía mecánica (*sonido*), fluctuación eléctrica, variación magnética, ondas electromagnéticas (radio, luz: óptica), etc.

Esta información deberá, por cierto, ser codificada de algún modo; por ejemplo la más antigua y fácil de entender es el telégrafo/telex, si llega corriente se hace un agujero en la cinta de papel (*1:hoyo*), sino nada (*0:blanco*), y luego se agrupaban con algún mecanismo de a varias unidades, agregando elementos para distinguir cuando comienza cada símbolo para poder siempre hallar el próximo dígito (*codificación*).

Estos mensajes se podrían desguasar en símbolos mínimos o unidades básicas simples, como vimos recién: corriente/no-corriente, del telégrafo. Estos símbolos combinados de alguna manera, lo cual llamaremos con la palabra 'codificación' formarán *letras*, que combinadas entre sí formarán *palabras*, que jutas forman *oraciones*, *ideas, texto, conocimiento* y todo lo demás.

Inicialmente la codificación digital se realizó mediante unos y ceros, siendo una unidad fácil de concebir, llamada ordinariamente binaria (*bi: dos*) y la aplicación más simple consistió en dispositivos como una batería de un lado, el cable como elemento barato y largo, que salvaba las distancias, terminando en un electroimán del otro lado que perpetraba su tarea perforando papel, (*telégrafo*). Conforme se avanzó con la técnica, aparecieron otras maneras de codificar la electricidad en los cables y luego sin cables mediante ondas de radio, hasta lograr modular la luz dese una bombita, pasando por un diodo emisor de luz (*LED*) hasta algo más sofisticado aún como un Láser / Máser[1], siendo ésta una manera "moderna" de reinventar el arcaico espejito o la linterna para avisar algo a lo lejos, con aquel escasamente recordado "Código Morse".

De lo que se trata es de enviar un mensaje en un determinado tiempo, y éste está compuesto por unidades mínimas que llamaremos "*símbolos*". Como estos *símbolos* deben ser enviados a lo largo del tiempo, lo más cómodo e intuitivo tal vez sea enviarlos dividiendo el tiempo disponible para esta tarea, en unidades iguales para evitar que se apiñen y mandar ese mensaje así, codificado en símbolos, secuencialmente cada fracción de tiempo, pues dividirlo en tiempos diferentes, implicaría todo un mecanismo

[1] MASER: Microwave Amplification with Stimulated Emission of Radiation: Laser de micro ondas.

mucho más complicado y de lo que trata la ingeniería es no complicarlo si no hay una ganancia sustancial o necesidad mandatoria.

Imaginemos el telégrafo, ¿cuán rápido puedo mandar un mensaje? si el papel del rollo, se mueve a, digamos para ejemplificar, 1 cm por segundo y el agujero del punzón fuera de ø 1.0 mm. Evidentemente lo lógico sería que no puedo hacer más de 10 agujeros en un cm si cada agujero mide un poco menos de 1/10 de cm, pues de otro modo se cortaría el papel a lo largo de los agujeros que se tocarían, y no podría diferenciar un nuevo espacio o agujero; bueno esto es básicamente una clase de "*interferencia entre símbolos*": el próximo agujero, no debe estar pisando el agujero anterior. En este caso una limitación lógica, así como está resultaría en 10 bits por segundo, y ésta sería lo que se llama la "capacidad del canal": 10 baudios.

Si bien este ejemplo no es preciso, presenta la idea; uno podría pensar que acelerar el papel a más velocidad permitiría más cantidad de agujeros por segundo, eso es cierto, pero enseguida nos tropezaríamos con el tiempo que tarda en bajar el punzón del electroimán actuador y luego un resorte lo volvería a su lugar, eso ya depende de la cantidad de corriente, el electroimán y su mecanismo, el largo del cable del telégrafo, la tensión de la batería, el peso del punzón, tipo y fuerza del resorte, etc. Lo importante es entender que donde hay mecanismos físicos, siempre habrá limitaciones de velocidad que se traducirán en una capacidad del canal y nada se puede escalar indefinidamente.

Particularmente para este tipo de casos Claude Shannon en 1948 planteó la teoría de comunicaciones, analizando la cantidad de información en función del ruido entendible como la inherente probabilidad de aparición de errores en ese canal, usando razonamientos análogos, para luego generalizarlos; bautizando la "capacidad de un canal" en símbolos por segundo, en donde el "símbolo" más básico es el bit o unidad lógica si/no, tal como se mencionó un poco antes en este escrito, mediante alguna magnitud física generable, transmisible y finalmente medible del otro lado.

La Señal

A esta magnitud, que pretende contener alguna clase de información, la llamaremos "señal" y nos abstraeremos desde aquí en los vericuetos de como convertir esa señal eléctrica o numérica de y hacia cosas físicas tangibles, para nosotros ya será una señal de cierta magnitud y la tenemos, y que varía en una dimensión, por ejemplo el tiempo, solo eso, el resto de los detalles, se irán descubriendo, conforme aparezca la necesidad.

Las limitaciones se plantearán por simplicidad, mediante análisis armónicos como el Fourier, que viene al rescate desde hace bastante. Si la señal es continua, se usará la versión analítica o continua, y si la señal es discreta se usarán las versiones discretas, a esta altura no es discriminatorio del problema para entender la I.S.I.

El Canal

Para poder analizar todo el proceso de comunicación, debemos de poder modelizar el canal, como alguna transformación predecible o no que ocurrirá sobre la "señal", la cual transcurre en el tiempo, y para nosotros el tiempo es desde siempre, hasta siempre., al menos teóricamente.

La Causalidad

Uno de los conceptos medio complicados de pescar de entrada es la causalidad y sus variantes; vamos a tratar de entender que es un canal, cuyo proceso sea causal o no.

Algo 'causal' es algo que ocurre y cuyo efecto precisamente ocurre *posterior* a toda acción considerada como causante.

Si aquí se considera el modelo de canal: la causa de toda alteración sería la señal, en otras palabras, un canal no debiera de poder cambiar

Si un canal es *causal*, quiere decir que su comportamiento está perfectamente decidido a partir del momento que comienza la señal, no antes y no podrá "adivinar" nada del futuro, en su propio tiempo [2]

Los sistemas físicos, de tiempo real, siempre son causales, en ningún momento una salida de algo, dependerá de una entrada que aún no ocurrió. Sin embargo en numerosas estafas detectadas en el mundo en los sistemas de apuestas, usadas en países remotos que televisaban carreras de caballos, perros, autos, peleas, etc.; la señal se 'demoraba' un poco, y nadie podía darse cuenta pero "la casa" tenía la ventaja respecto a las apuestas, esto era un sistema 'no causal' en su totalidad, pero 'causal' para quienes estaban dentro del recinto donde recibían la señal demorada. Valga el ejemplo.

En cambio si se procesa señales ya grabadas y se tomas el total de la señal, es posible usar sistemas no-causales, para obtener cosas pues en determinado momento de decisión en el proceso, sí se tiene la información de más adelante, pues ya fue grabada con anterioridad. En estos casos los filtros obtenibles y las detecciones que se logran son a veces mucho mejores que en tiempo real, que es no-causal.

Es importante entender que el concepto de 'causal' solo se puede aplicar cuando hay una sola dimensión de variación de los datos, no cuando hay más de una como por ejemplo en una imagen digital que es un conjunto grande e datos ordenados topológicamente, no es causal entre sí para dimensiones internas como arriba y abajo, pues no hay dirección de variación única, pero si, el conjunto completo es causal entre cuadros (donde la dimensión de variación es el tiempo). En cambio en este tipo de imágenes, hay posibilidad de definir transformadas de Fourier bidimensionales y otras que reconocen caras, gestos, emociones, cuentan elementos, y muchas que hacen trucos increíbles, de esos que se ven en las películas y publicidades. Todo esto es mayormente filtros digitales y procesamiento de imágenes, y hasta muchas veces, los efectos más espectaculares son realizados por filtros no lineales, pero la teoría de todo eso arranca con el tratamiento de señales lineales, simple y básico.

Invariancia en el Tiempo

Resulta que entonces de aparece cruzado otro concepto que es el de "*que algo cambie en el tiempo*" y no me refiero a la señal, sino a los parámetros del canal. Imaginemos un caso un tanto disparatado, que ejemplificará lo que hay que entender: un telex, en donde

---

[2] Hay excepciones que explicaremos al final pues existen y de hecho se usan sistemas/filtros localmente no-causales, los cuales se usan en predictores y filtros muy perfectos, pero hay que entenderlo bien: no *adivinan*, sino predicen o resuelven algo, tan solo un poco más tarde, teniendo cuenta cosas de hace 'un tiempo', jamás mayor a la tardanza en la predicción.

los cables tienen 500 km, y hay viento variable que hace "oscilar" los cables, si bien el cable se estirará y contraerá muy poco (elásticamente) este cambio a lo largo de cientos de kilómetros se sumará y podría alterar el tiempo que tarda la señal en llegar, y si estamos mandando muy 'apretados en el tiempo' los pulsos hacia el solenoide, tal vez para cuando llegue un determinado ***bit***, el anterior se retrasó más porque justo se estiró el cable teniendo que recorrer más distancia, y como la velocidad no es infinita, de pronto llega uno más atrasado y aparece un dilema: ¿es el primero o el segundo? - no los podré diferenciar!

Esto pasa mucho más seguido, cuando se transmiten ondas de algún tipo por medios en movimiento, como ondas de ultrasonido en el agua o radiofrecuencia por la atmósfera, la cual se halla dilatándose y contrayéndose por acción de los cambios de presión producidos por condiciones meteorológicas variables (tormentas, rayos, tifones, vientos, remolinos), cuya acción cambia los parámetros de transmisión, para cuando esperamos un símbolo, éste viene atrasado en forma variable, y no predecible, y nos confundimos.

Esto es, precisamente otra clase de "***interferencia entre símbolos***" (ISI), causada por la no-invariancia en el tiempo de un canal. Llevando esto a la tecnología actual (años 2014+) esto nos ocurre todo el tiempo, hay movimientos sísmicos tanto lentos como rápidos en todos los enlaces de microondas; los enlaces satelitales tienen oscilaciones orbitales y deriva; los terremotos afectan mucho (*modulan en el tiempo*) casi todos los enlaces de comunicaciones con ruido caótico no ergódico. Y lo más frecuente es que nuestro clásico celular o un enlace de radio, lidia con esto todo el tiempo, cuando nos movemos ya sea caminando, en auto o una reflexión de una onda de radio pega en la antena sumando reflexiones que se mueven: vehículos, o cualquier máquina. Es un problema diario en señales y sistemas de comunicaciones. Los conocidos sistemas MIMO[3] usan exhaustivamente procesamiento de señal para sacar ventaja de reflexiones y configuración de canales del tipo variantes en el tiempo y espacio.

Dispersividad

Que un medio sea dispersivo es un fenómeno que tiene que ver con el cambio que un medio introduce en un paquete de información que pasa por él, y en especial se dice que un medio es dispersivo cuando algún parámetro es función de la frecuencia, y en ciertos casos, si el parámetro es el retardo de tiempo, éste no es inversamente proporcional a la frecuencia. Esto último precisamente lo vamos a desarrollar.

Cuando se hace pasar por el medio una señal, por la teoría de descomposición ortogonal de Fourier, que recita que "*toda señal finita o infinita, e inclusive discreta, puede descomponerse en forma unívoca en una suma de señales sinusoidales*", todas con determinados parámetros únicos (fase y amplitud). Si la representación de esta señal en el campo de las frecuencias y fases es finita o no, no es importante en este momento y solo depende de la presentación de la señal misma y como se considera el cálculo.

Sin embargo lo importante se puede entender con tan solo analizar una única onda sinusoidal ideal, de duración infinita con cierta amplitud. Supongamos que tiene un período T_1, ese tiempo es la demora durante la cual el versor giratorio imaginario que genera la onda, recorre 2π radianes. La fase de cualquier onda demorada una cantidad: T_{d1} en el tiempo , será entonces:

[3] MIMO "*multiple-input and multiple-output*", sistema usado para mejorar el alcance y efectividad de transmisión en canales en la banda WiFi (2.4GGHz) mediante varias antenas y juego de fases activo.

$$2\pi \times (T_{d1} / T_1).$$

Esto nos dice que si un "sistema" que demora T_{d1} a una señal de período T_1, se podría representar como una "*demora de fase*" $-2\pi \times (T_{d1} / T_1)$. Nótese que el signo negativo es porque "retrasa" (*sino estaría adivinando la onda antes de que ésta exista*)

Hasta acá el concepto suena sencillo y no presenta dificultades, veamos ahora si a ese mismo medio lo probamos con otra frecuencia diferente, cuyo período es T_2 y al medirlo a ésta frecuencia, se demora un valor igual T_{d2}, usando el mismo razonamiento, la fase resultante será entonces:

$$- 2\pi \times (T_{d2} / T_2)$$

Supongamos inicialmente que $T_{d2} = T_{d1}$ o sea, es un medio en donde todas las señales de diferentes frecuencias tardan el mismo tiempo $T_{d2} = T_{d1}$ en pasar de una punta a la otra. Surge una pregunta inmediata:

¿cambiará la forma de la onda luego de pasar por el medio?

La respuesta no es tan inmediata, pero fácil de deducir, supongamos que no hubo cambios en la amplitud de cada componente de esas dos frecuencias, llamémoslas

$$f_1 = 1/T_{d1}$$

y

$$f_2 = 1/T_{d2}$$

Si Fourier dice que la descomposición en componentes es única, y solo hubo un cambio de fase en cada una, por cierto diferente en ambas, porque sus frecuencias son diferentes pero el tiempo de demora es igual para cada frecuencia, debido a la demora por el viaje de la señal en ese medio, entonces las fases de cada componente serán, respectivamente:

$$\lambda_2 = - 2\pi \times (T_{d2} / T_2)$$

$$\lambda_1 = - 2\pi \times (T_{d1} / T_1)$$

Si ahora:

$$T_{d2} = T_{d1}$$

Reemplazando, dividiendo ambas ecuaciones y simplificando, obtenemos:

$$\lambda_2 / \lambda_1 = T_1 / T_2 = f_2 / f_1$$

Interpretando esta fórmula, vemos que la relación de fases es inversamente proporcional a los períodos y directamente proporcional a las frecuencias, en otras palabras si una onda pasa por un medio cuya fase es cero o si es diferente de cero, es directamente proporcional a la frecuencia, ésta onda se atrasará un tiempo igual a la derivada de esa fase con la frecuencia, esto es llamado ***retardo de grupo***. y es muy usado en telecomunicaciones, y dentro de un rango restringido de frecuencias, es deseable que un sistema o canal tenga un *retardo de grupo* constante, para que de este modo garantice que ningún componente que pueda contener un símbolo, se demore más que otro, contribuyendo a "confundir" algún símbolo, no importando el método de detección.

Particularmente veremos luego los ***diagramas de ojos*** muy usados para determinar visualmente la calidad de un canal ante determinada modulación, en este caso *modulación* es sinónimo de "*cómo demonios se logra la codificación de los símbolos para poder transmitirlos*" y hay mucho para estudiar de eso en telecomunicaciones.

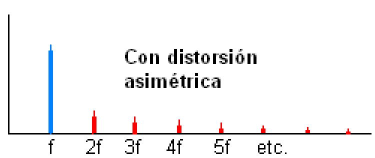
Si usamos estos conceptos y los introducimos en la notación compleja de la transformada de Fourier, el demorar una señal en fase, es equivalente a multiplicar con un exponencial complejo con el ángulo en el exponente, y como es una constante (aunque compleja) sale fuera de la integral y se interpreta como una demora de tiempo, siendo una de las propiedades de la "transformada de Fourier" que una demora en el tiempo, corresponde a una fase lineal en las frecuencias, y dadas las propiedades de simetría de estas transformadas, lo opuesto se cumple, pero ya no es interpretable fácilmente con ondas senoidales.

La no-Linealidad en la Naturaleza

Supongamos que tenemos un medio como el agua salada, que transmite señales de presión aplicadas con un transductor ultrasónico piezoeléctrico, para luego procesar los rebotes o ecos y determinar distancia al fondo de una embarcación, y la presencia de bancos de peces, etc.

Pero resulta que el piezoeléctrico es alineal, debido a que cuando se deforma por el campo eléctrico aplicado, la deformación es proporcional solo al principio, luego se 'achancha', generando tal vez una deformación simétrica de la onda, compatible con una armónica impar, que si se analiza en detalle, es asimilable a una no-linealidad, igual pasaría si aparece una saturación de la electrónica asimétrica agregando componentes pares.

Mostramos en el gráfico próximo, a modo de ejemplo, cómo serían los componentes armónicos, obtenidos por un análisis de Fourier, de una onda senoidal pura, de una frecuencia 'f' luego de una señal con dos tipos de distorsión:



Supongamos que ahora "la estrategia" es mandar un paquete de ondas, de cierta frecuencia y duración, tal vez coincidente con la resonancia del sistema para mejor ganancia, y luego se esperan los 'ecos'. Luego midiendo el tiempo transcurrido y las amplitudes relativas, podría determinarse si hay fondo, peces o elementos variables.

Hasta aquí la no-linealidad parece que no afecta en nada, es como si no importase; pero veremos que si afecta y mucho. Si hacemos el análisis armónico de una distorsión no-lineal, si ésta es simétrica, generará exclusivamente armónicas impares y si es asimétrica, agregará también las pares. Es decir "aparecerán frecuencias que nunca estuvieron en el paquete inicial", y esto por supuesto que modificará el comportamiento del eco, si solo usamos la amplitud, tal vez no importe, pero si queremos procesar "bancos de peces" y el agua por cierto, es dispersiva, es decir las frecuencias más altas (*armónicos*) sufrirán un retraso diferente que las fundamentales, y hasta podría generar 'ecos' de menor valor, señalizando bancos de peces inexistentes, a partir del único rebote del fondo: sorpresa!

Aquí, precisamente en este ejemplo no muy riguroso se puede apreciar que la no-linealidad trae numerosos problemas; bueno, siendo sinceros, confieso que no existe una cosa llamada linealidad más que en la teoría, todo los sistemas son definitivamente no-lineales. El asunto es cuánto y en que rangos lo es, y saber esto implica justamente todo el trabajo de ingeniería: determinar los rangos de linealidad y cuánto es esta, para poder usar cálculo y todas estas teorías, cuando las cosas no son lineales es muy complicado predecir, muy a pesar que hay sistemas que basan todo su principio de funcionamiento en no-linealidad, como un detector de radio, un multiplicador, un simple diodo o transistor es alineal o no-lineal desde su fórmula más básica, sin embargo los rangos en los que se los usa, son tales que se los realimenta y conforma de modo que sean lo más lineales posibles, o viceversa, según la aplicación o necesidad.

Linealidad

La definición estricta de linealidad es esta (*de los libros*): sea ***y*** la salida de un sistema y ***x*** la entrada siendo '***t*** ' la variable independiente, pudiendo ser tanto el tiempo en modo continuo o un índice entero indicando el orden del elemento de la secuencia, si es una señal discreta, en esos casos se suele usar '***n***' por comodidad, pero es indistinto:

$$x1(t) \rightarrow y1(t)$$

y

$$x2(t) \rightarrow y2(t)$$

entonces

$$a \cdot x1(t) + b \cdot x2(t) \rightarrow a \cdot y2(t) + b \cdot y2(t)$$

para todo escalar ***a*** y ***b***, finito y diferente de cero.

En otras palabras, si un sistema responde de una determinada manera a dos señales, diferentes la respuesta a la combinación lineal de esas señales será igual a idéntica combinación lineal de las salidas individuales. Dicho de otro modo, indica que el sumar algo a la entrada solo afecta la salida por sí mismo, y no depende de ningún otro elemento, la salida de toda combinación será la suma (*por eso lo de "lineal"*) de la salida a todo estímulo, como si fuese independiente aplicado.

Ejemplos: Al principio de esta sección, mencionamos agua '*salada*', y luego esta propiedad no ha sido usada ni mencionada hasta aquí; y ahora vamos a demostrar que el agua *salada* puede producir efectos de dispersividad, provocando ISI adicionalmente.

La velocidad de transmisión de una onda sonora en un medio depende de dos cosas, la elasticidad del mismo, y su densidad, los tiempos o demoras, serán pues inversamente proporcionales a su velocidad de transmisión, la cual también podría variar con la frecuencia, pero éste es un fenómeno mucho menor para este caso del agua. El agua salada es mas densa que el agua dulce, debido que la sal agrega el peso del sodio y del cloro, diferentes al oxígeno y agua. En cualquier lugar donde hay agua salada, hay corrientes de agua con diferentes grados de salinidad, y éstas suelen estar constantemente en movimiento por diferentes motivos que no ahondaremos, pero el resultado medible y observable es que se producen láminas con diferente densidad que serán traspasadas por esta señal 'ultrasónica' que busca hallar la distancia al fondo y la presencia de bancos de peces. Cuando pasa de un medio mas denso a otro se produce un fenómeno de refracción, pues al cambiar la velocidad de una onda en su medio se cumple una ley llamada ley de Snell-Descartes, estudiada en óptica y física básicas. Esto produce 'ecos' indeseables, además de los cambios de 'velocidad'. Ignoraremos los 'ecos' pues este es un tema mucho mas complejo y nos enfocaremos en los cambios de velocidad. Si una onda traspasa un medio que tiene diferentes 'láminas' de cierto espesor como ser 'w_1'+'w_2'+...+'w_N' luego cada una tendrá diferente 'densidad' y en consecuencia el tiempo que tardará en traspasar esa lámina 'w_N' será proporcional a su espesor y la densidad salina. El tiempo total será entonces la suma total de los tiempos parciales, omitiendo la formulación, se entiende que éste será variable conforme varíen los espesores de esas 'capas/láminas' de agua con diferente salobridad. Si éstas capas están en movimiento, por los flujos del agua y/o el desplazamiento relativo de la embarcación (*que podría estar en movimiento*), aparecerá un tiempo variable en la trayectoria total, dependiente de cosas que no son solamente la presencia de peces o el rebote del fondo.

Si solo estuviésemos transmitiendo un pulso 'paquete' de ondas, para ver cuándo rebota en algo, cuánto tarda 'el paquete-rebote' en llegar. Éste cambio variable de tiempo de demora del rebote, introducirá una indeterminación en la medición, que no nos permitirá acertar la distancia. Si bien esto puede parecer de poca importancia: nótese que a veces muy sutiles cambios no considerados en el modelo en especial en la geometría y/o parámetros intrínsecos de un medio de transmisión, puede ocasionar '*serias*' modificaciones en nuestras mediciones. Esto es porque bastantes veces la expresión que halla la magnitud medida, si la representamos claramente en forma de una ecuación, puede tener un valor fijo/constante 'restado' a esa magnitud a veces muy cercano al valor que creemos invariable, y una muy pequeña variación de algo que se considera erróneamente "constante" (*como en este caso la velocidad del sonido en el agua salada*), genera una enorme variación de los resultados, para peor si ésta diferencia estuviese presente multiplicando en un denominador.

Si en cambio lo que estamos haciendo es enviando un mensaje codificado a un submarino en las profundidades del océano, estas fluctuaciones de tiempo de transmisión variables actúan similar al ejemplo del telégrafo con cables que se estiran rítmicamente con el viento o caóticamente ante un movimiento sísmico; introducirán demoras 'variables' que cambiarán los momentos de llegada de cada símbolo codificado, y como este momento es algo que debo poder predecir en forma precisa en el receptor, pues puedo perderme o malinterpretar algún símbolo que llegue fuera de tiempo, ya sea atrasado o adelantado; esto es precisamente ISI.

En resumen, este tipo de ISI de los ejemplos, estarán producidas por una no-invariancia en el tiempo del medio de transmisión, producidas, en este caso, por cosas como dispersión y tiempos de tránsito variables en el tiempo.

Linealidad e Invariancia

La linealidad e invariancia en el tiempo, no se cumple casi nunca a rajatabla, en la gran mayoría de todos los sistemas reales, en especial los que llevan semiconductores y otros elementos activos del mundo de la electrónica, inclusive no se cumple tampoco siempre con transformadores y circuitos magnéticos, para cualquier rango del valor de esas señales.

Todo esto de la invariancia y linealidad, solo se cumplen en la teoría, pero justamente el conocer los rangos en que estas cosas se cumplen, pudiendo aprovechar toda la artillería que nos proveen los cálculos, asumiendo que todo es lineal como en un cuento de hadas, es precisamente de lo que trata casi toda la ingeniería que trata con señales analógicas y digitales, siendo la *alinealidad* en particular, uno de los temas más recurrentes, difíciles y molestos, disfrazados y medidos en diferentes ámbitos con términos tales como: distorsión armónica, intermodulación, caracterizaciones de límites como rango dinámico y saturación, entre muchos otros temas relacionados y situaciones que aparecerán en diversas materias avanzadas y aplicaciones de la electrónica.

De hecho en muchas ciencias la *alinealidad* es casi "la reina" de las leyes que gobierna las reacciones químicas, físicas, y muchas otras; lamentablemente en ciertos rubros y aplicaciones los rangos de linealidad son escasos y a veces casi inexistentes; por lo que se debe lidiar con otras técnicas de cálculo, inclusive se apela frecuentemente a linearización artificial, para poder usar el poder de procesamiento digital y cálculo 'pesado' en imágenes de satélites, mediciones, en medicina: electromiografía, electrocardiografía, electroencefalografía, ecodoppler, tomografía axial computada, resonancia magnética nuclear, y tantos otros apartados de aplicación.

Contrariamente todos los mecanismos que 'aprenden' y realizan operaciones biológicas, o simulaciones de procesos biológicos y cognitivos, como redes neuronales artificiales, y otros sistemas usados en un área muy vasta que se engloba dentro de lo que se llama "*inteligencia artificial*" y "*aprendizaje automático*"; precisamente allí ***la alinealidad*** es la "reina" más codiciada, contrariamente a lo que sucede en sistemas lineales; se ha demostrado que introduciendo alinealidades numéricamente adrede (*algunas llamadas kernels*) se logran sistemas de detección, modulación y demodulación, corrección de errores, reconocimiento de patrones con propiedades emergentes que rozan lo asombroso y de Sci-Fi; sin embargo casi toda la matemática subyacente es lineal, y descansa sobre pilares de muchas operaciones lineales complejas como transformadas de todo tipo, con algunos agregados que '*hacen el truco*', especialmente en los algoritmos; pero esto sin dudas, es agua de otro costal, solo no quería dejar de comentarlo aquí.

Invariancia en el Tiempo, en la Realidad

La invariancia en el tiempo, en el caso de semiconductores y circuitos electrónicos complejos, donde éstos abundan, tanto como otros componentes discretos; se produce básicamente en los parámetros de cada componente, variando desde con imprecisiones sobre el lote de fabricación hasta con la temperatura y por último con el envejecimiento. Por lo general las variaciones con temperatura y envejecimiento son tan lentas que su efecto temporal es imperceptible, sin embargo el efecto de estas sobre los parámetros finales de un circuito complejo, a veces son de temer y se deben de calcular y predecir mediante métodos bastante complejos y de corte netamente estadística como el método de Montecarlo, para asegurar al menos estadísticamente que, pese a las variaciones dadas por tolerancias, cambios de temperatura y envejecimiento; el comportamiento deseado no se aparte de las tolerancias indicadas en las especificaciones.

En el caso de sistemas que procesan 'señales' y son puramente digitales como procesadores digitales de señales, éstos pudiesen parecer inicialmente insensibles a estos efectos, puesto que todas las operaciones son 'ideales', las sumas no fallan y todo lo demás es teórico. Sin embargo no hay engolosinarse con esto, hay un tema casi siniestro y difícil de darse cuenta en la implementación de sistemas digitales que involucran cálculos que es la estabilidad numérica y el 'ruido' introducido por las operaciones numéricas con precisión no infinitas. Es decir, tal vez una operación de cálculo de un filtro implique sumar un número de veces el resultado de la cuenta anterior, mas la anterior de la anterior y esto se almacene luego para ser usado en la próxima cuenta, esto se da en particular en los filtros digitales de configuración llamada tipo IIR (*Infinite Impulse Response*) y si el truncado de la operación debiese de converger en cero, tal vez la resta o suma no sean perfectas y dan un resultado de +/- 1 LSB (*Least Significant Bit*) por más punto flotante que se realice, y esto es equivalente a introducir un 'ruido', pero las características de este ruido, no son gaussianas, y todo lo contrario, son dependientes de los datos, produciendo comportamientos muy difíciles de predecir, y en consecuencia, hay que prestar suma atención a estas cosas.

Otro motivo de errores en telecomunicaciones, se dan cuando el 'reloj' de un sistema de procesamiento de señal, es derivado mediante sistemas tipo "PLL[4]" y obtenidos desde la señal misma, o algo especial, obtenido a partir de la señal (*se llama reconstrucción de reloj y/o sincronización local*). Cuando ésta viene cambiada por modificaciones en la velocidad de trasmisión, éste reloj aparece desfasado de los propios datos, pues estas reconstrucciones se supone que son muy lentas, en cambio los datos cambiaron muy rápidamente su momento de llegada. En estos casos aparece una ISI importante y difícil de determinar, cuya solución pasa por re-sincronizar más rápidamente el momento de detección de los símbolos y/o usar métodos adaptativos de más corto entrenamiento, aumentando la vulnerabilidad a ruidos.

Como diría una frase muy americana que aplica bien a esto: "*there is nothing like free lunch*" indicando que toda acción correctiva tiene consecuencias de uno y otro modo, y no hay soluciones mágicas, solo soluciones de compromiso.

---

[4] PLL (acrónimo del inglés *Phase Locked Loop*) Es un tipo de circuito digital realimentado "*Loop*", enganchado "*Locked*" con la fase "*Phase*" a una frecuencia de entrada, logrando multiplicarla por números racionales [P/Q].

Corrección de ISI

Cuando se transmite señal o se recibe información desde un receptor, se puede modelizar el canal como un sistema LTI por más que no sea, puesto que debido al teorema de linealidad e invariancia en el tiempo, por más que un sistema no sea LT se lo puede representar como un sistema LTI al cual se la "suma" una perturbación.

Precisamente esta perturbación es la parte no lineal y no estacionaria en el tiempo, sin embargo las influencias "lineales" e "invariantes" en el tiempo, suelen ser lo suficientemente lentas o pequeñas para poder ser compensadas exitosamente, pues sino no tendíamos ni celulares, ni comunicaciones de radio ni de Internet hoy día.

Modelización de un sistema como LTI

Lo que se va a explicar es cómo realizar una compensación de interferencia entre símbolos, asumiendo que el sistema es LTI. Para esto debemos de medir y/o estimar el máximo de variación de fase y/o de corrimiento en el tiempo que puede aparecer en el canal, y también si aparecen alinealidades, acotar su influencia.

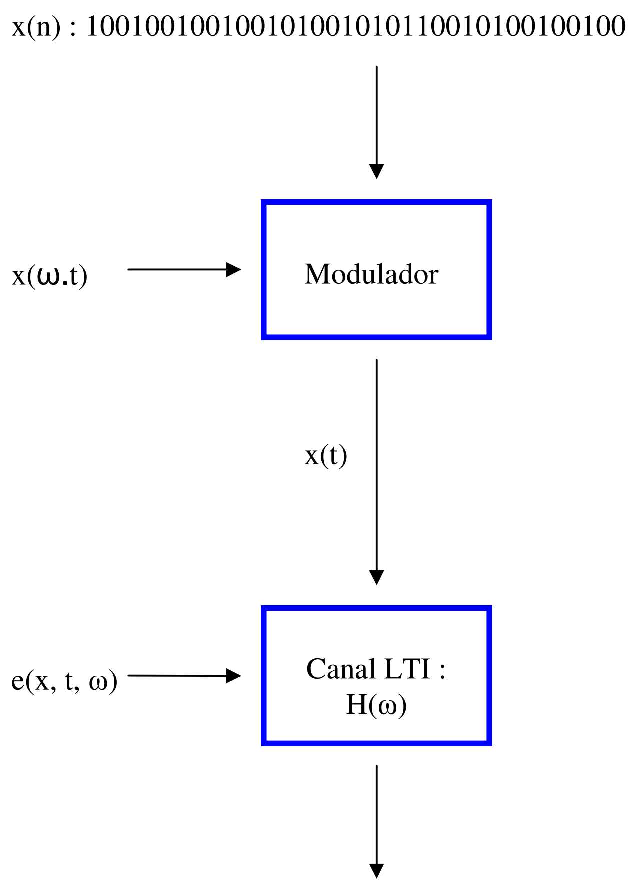
Lo mejor es empezar con un ejemplo de un sistema de transmisión/recepción simple basado en pulsos 1/0, no es importante el medio de transmisión ni el método pero sí que la señal llega como un valor **x(t)**, la cual es analizada en determinados instantes **x(n)** para determinar el valor del símbolo, en ese momento.

El diagrama sería muy simple y es un transmisor, el cual utiliza una señal periódica de frecuencia fija, que "prende" el canal a intervalos regulares, conforme haya un "1" o lo deja apagado cuando hay un "0" en la secuencia de símbolos a transmitir.

Analicemos su comportamiento, supongamos que debemos enviar un mensaje x(n) compuesto por la secuencia de símbolos 10010010010..

El sistema de transmisión, al que llamaremos Modulador, lo único que hará es convertir esos símbolos en algo, simbolizado con x(ω.t), en este caso lo que simplemente hará es colocar una señal de un determinado valor analógico, digamos 1.0 si hay un símbolo "1" y cero 0.0 si hay un símbolo "0", todo muy simple por ahora.

El diagrama en bloques será



y(t) = x(t) * H(ω) + e(x, t, ω)

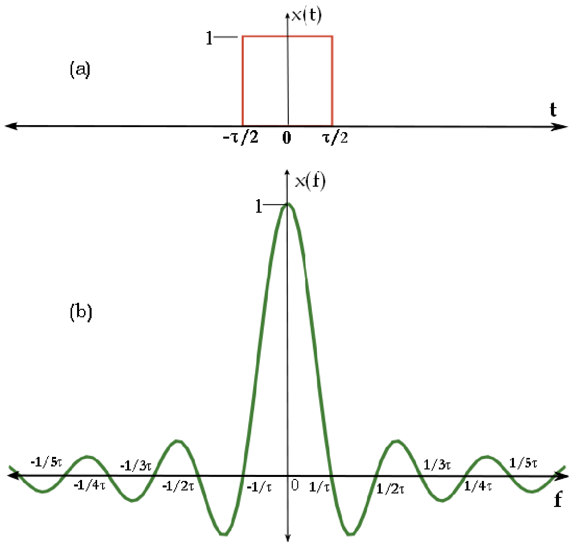
La señal de error "**e**" la asumimos como un a señal adicionada a la salida del sistema LTI. En realidad no importa si se la suma antes o después, puesto que de ser sumada antes, e(x, t, ω) pasaría a ser una señal e'(x, t, ω) = H(ω) * e(x, t, ω). La cual para el análisis del modelo no cambia nada pues es considerada ruido y no nos importa su origen sino su magnitud total a la salida, respecto a la señal: x(t) * H(ω).

El canal H(ω) se puede representar como un sistema basado en parámetros fijos y como es LTI se puede representar como la suma de cada una de sus respuestas a cada momento de la señal x(t), en todo tiempo posterior al comienzo, por ser "causal", en otras palabras cumple con el mandamiento "no adivinarás". Por otro lado al ser fijos los parámetros del canal H(ω) también cumplirán con el segundo de los mandamientos: "no cambiarás con el tiempo".

Y por último también se supondrá que H(ω) es lineal, el mandamiento sería: "si te mando algo en dos partes sumadas de cualquier modo, responderás siempre fielmente como la suma de ambas como si te las mandara por separado". Toda alteración del canal "real" respecto al "ideal" lo metemos en e(x, t, ω) y santo remedio.

Ahora vamos a explicar cómo se produce una ISI y de qué modo cómo se puede, mediante un filtro F(ω), tratar de corregirla, tanto como lo permita el valor de: e(x, t, ω).

Primero para entender que pasa cuando se modula con una onda "cuadrada" analizaremos el pulso producido por un "1", dada su transformada simple, se obtiene la función ***sinc***(*f*) gráficamente será:



Esto muestra el espectro contínuo de un "pulso" simple en el tiempo. Retomaremos esto al final luego de la introducción de algunos conceptos útiles a la hora de analizar un canal de comunicaciones 'ruidoso' y propenso a errores, como lo son casi todos en mayor o menor medida.

Bit Error Rate

Antes de entrar de lleno en el análisis de un canal simple y sus gráficas de simulación, vamos a introducir un concepto muy usado en comunicaciones digitales que se llama tasa de errores de un bit, pero por lo general se la ve con estas siglas: ***BER***, acrónimo de las siglas del inglés "*Bit Error Rate*" y significa la probabilidad de cometer un error de un bit, y como el "*dato*" se mide en bits, es la probabilidad de error del canal, o tasa de error.

Si consideramos una secuencia X de K elementos (*número de bits o símbolos*), y llamamos Y a la secuencia de salida del canal, luego de decodificado, su fórmula sería:

$$P_{BER} = \frac{1}{K}\sum_{i=1}^{K} P_r\,(X_i \neq Y_i)$$

En otras palabras, es la suma bits fallados, dividido por el total de bits transmitidos.

Esta tasa de errores, se busca en realidad que sea muy baja en canales digitales convencionales en especial los de almacenamiento digital, como la escritura/lectura de un cabezal de un disco magnético como un floppy-disk o un disco rígido de computadora. En este caso el canal de entrada es la escritura de datos en el disco en un momento dado, y la salida es la lectura en cualquier otro momento posterior. Sin embargo a pesar de ser menor a 10^{-12}, se utiliza un mecanismo de corrección y firma algorítmica llamado CRC (*Cyclic Redundancy Code*) para asegurar matemáticamente una detección de error mucho mayor precisión y lograr 10^{-15} como BER de todo el sistema, incluyendo el algoritmo. Recordemos que todo canal de este tipo es causal, sin embargo hay 'trucos' que permiten diseñar algoritmos anticausales de corrección, para bajar el BER.

Las variables que atan esta medición del BER con el canal de comunicación son varias como el ruido del canal (*medido como relación de potencias entre señal y ruido*), su máxima distorsión por alinealidad (*generalmente medida como distorsión armónica total*), y algunas más específicas dependiendo del medio de comunicación y las perturbaciones que surjan, por ejemplo en el caso de trasmisiones atmosféricas, la máxima fluctuación de densidad por efectos atmosféricos como las nubes y los vientos.

Lo que hay que entender bien es que en estas comunicaciones, siempre querremos que el BER sea lo menor posible, y para lograr esto se elijen los diversos mecanismos de modulación y corrección. Para cada mecanismo, dado un canal, habrá una curva "ideal" de BER vs. relación señal-ruido, otra de BER vs. distorsión armónica, otra de BER vs. nivel de señal ya que a veces el ruido y la saturación de canal se combinan malamente.

Capacidad de un Canal

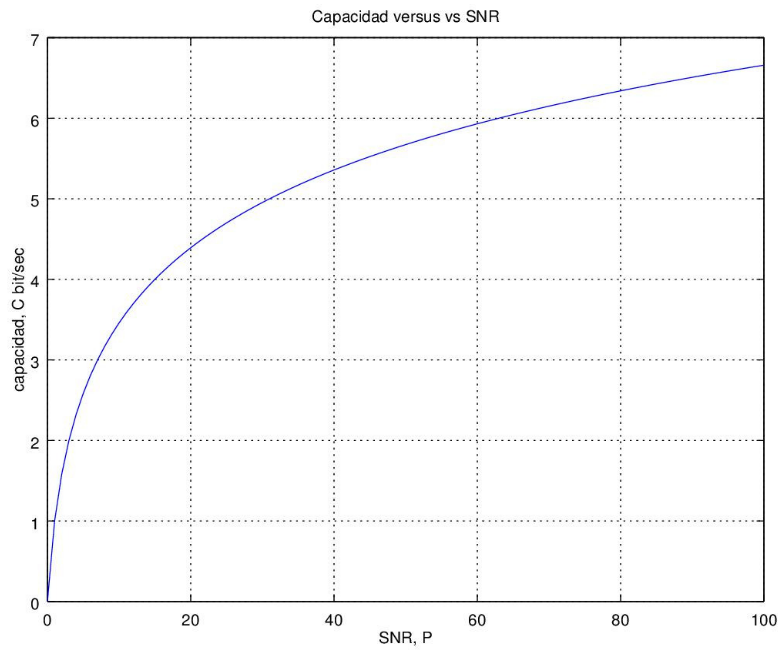
Por lo general, la capacidad de un canal no sobrepasa el límite teórico deducido por ***Shannon***[5] en su famosa teoría de comunicaciones, presentada en 1948, la cual se puede expresar mediante el teorema/ecuación de *Shannon-Hartley*[6].

$$C = B \cdot \log_2 \left(1 + \frac{S}{N} \right)$$

Donde C es la capacidad del canal en bits/segundo, B es el ancho de banda del canal en Hz, S es la potencia de la señal en *Watts* y N es la potencia del ruido del canal en Watts. La relación S/N es conocida como SNR (*Signal Noise Ratio*) o relación señal a ruido y se expresa comúnmente en dB que es una relación logarítimica de potencias.

Si graficamos esta fórmula, vemos que subiendo la SNR solo se logra un incremento logarítmico de la capacidad del canal lo cual no es muy alentador y resulta costoso pues lograr esas SNR no es posible ni práctico, a partir de sistemas reales y físicos. Veamos el siguiente gráfico para entenderlo:

[5] Sin embargo es posible transmitir un poco más rápido que el límite teórico de Nyquist-Shannon, esto se conoce como FTN (*Faster Than Nyquist*) y obedece a ciertas peculiaridades de la manera de cancelar la ISI mediante algoritmos que hacen compleja la detección y aumentan la sensibilidad al ruido.
[6] C. E. Shannon, "A Mathematical Theory of Communication", Bell Syst. Techn. J., Vol. 27, pp.379-423, 623-656, July, October, 1948



Lograr superar estos valores de SNR físicamente es muy complicado, y como se ve, no hay demasiada ganancia de capacidad de canal por arriba de 100 veces, ya que la curva es logarítmica y se aplana cada vez más. De hecho es interesante explorar la zona más empinada, antes de 10 veces de SNR, donde con pequeños aumentos de SNR se logra duplicar y casi cuadruplicar la capacidad

Analizando esta ecuación, parece saltar a la vista que cuanto más bajo sea el ruido mucho más alta será la capacidad del canal, sin embargo esta limitación es poco práctica pues exige mecanismos muy sofisticados de codificación que son sensibles a ruido y a la dispersión del canal, lo cual complica mucho las cosas y aumenta el BER al final. Sin embargo, como se menciona al principio no existe nada gratis, "*no free lunch*".

Modulación y Demodulación

Sacar el jugo a esta ecuación y vencerla con mecanismos de modulación es de lo que trata una buena parte de todo el desafío en la telecomunicación digital hoy, y dado que el poder de cálculo crece según las leyes empíricas de *Moore*[7], aún hoy, en 2014 hay chance cierta de mejorar bastante la eficiencia en casi todos los canales existentes. Basta con ver como la telefonía celular ha multiplicado su velocidad en términos de eficiencia de canal, casi sin aumentar demasiado el uso del ancho de banda asignado, escalando 1G, 2G, 3G y 4G. Y ya vendrán otros 'G' más veloces. El limitante es, por lo general el costo de proceso de cálculo digital en el CPU del terminal móvil ya que la fuente energía es la batería (*su duración disminuye cuanto más cálculo se hace.*.) y el otro límite es la relación señal a ruido de los transistores de los receptores y la potencia máxima admisible para no producir daño a nosotros, los humanos por las radiaciones no ionizantes del terminal, así no nos fríe el cerebro o nos despierta un potencial cáncer.

[7] Moore, CEO de IBM enunció esta curiosa ley empírica, observando el crecimiento del poder de cálculo se duplicaba cada 18 meses, y se viene cumpliendo también bajando el consumo y tamaño de los dispositivos. Hoy un celular barato tiene cientos de veces el poder de un mainframe de hace una década.

Eficiencia Espectral

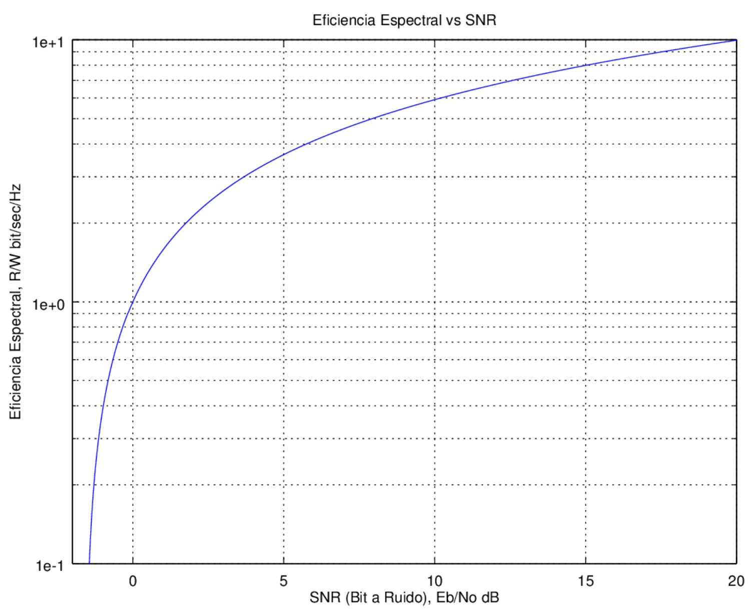
Si hacemos un análisis más interesante, de la ecuación y definimos ***S*** como el promedio de potencia de la señal, llamamos E_b **= S/C** (*Joules por Bit*) la energía promedio por bit transmitido, y ***1/C*** es la duración de un bit, si definimos la densidad de potencia unilateral $N_0/2$ (*Watts/Hz*) podemos jugar un poco y escribir esto:

$$\frac{C}{B} = \log_2\left(1 + \frac{E_b}{N_0}\frac{C}{B}\right)$$

y si introducimos un término nuevo y llamamos **η = C/B** como la ***eficiencia espectral*** en bits/segundo/Hz, re-escribimos la relación, quedando:

$$\frac{E_b}{N_0} = \frac{2^{\mathrm{n}} - 1}{n}$$

Si graficamos la ***eficiencia espectral*** en función de la relación señal a ruido, hallaremos las zonas posibles de existir por el teorema. Si lo graficamos se obtiene:



Hay un lugar al que la curva llega y representa la asíntota de E_b/N_0 cuando el ancho de banda ***B*** tiende a infinito y es ***ln(2) = -1.59dB***, y se denomina límite de eficiencia de potencia de Shannon, Se observa que la zona bajo la curva es la "viable" y por encima de esta curva no es posible enviar información en forma confiable. Los sistemas serán tanto más complejos y complicados cuanto más a la derecha y cerca de la curva azul desean estar. La tecnología hoy empuja, sin dudas, en ambas direcciones sin parar.

Eficiencia de Codificación

Como se ve, hay aspectos más avanzados que, por ejemplo permiten definir la eficiencia mediante un parámetro llamado "*eficiencia de codificación*", que relaciona el número de bits efectivos transmitidos, respecto del total de bits traficados.

Para ilustrar el concepto, lo explicaremos así: como sabemos que el canal puede fallar y deseo bajar la BER total y a su vez saber si falla. Por ejemplo transmito cada bit dos veces, luego los comparo al llegar, si llegan iguales, estoy "*el doble de seguro*" que no hubo error, salvo que ambos hubiesen venido cambiados, en cuyo caso la probabilidad o BER será mucho menor, aproximadamente valdrá BER^2 suponiendo que los errores son independientes entre sí, por teoría de probabilidades; en cambio si ambos bits son diferentes, sabré que hubo un error con la misma certeza BER^2. Para cada esquema de codificación y cada tipo de canal, habrá una eficiencia de codificación y una curva de BER en función de los parámetros de comunicación. La tarea del ingeniero de comunicaciones, es hallar los lugares de menor BER con la mayor eficiencia espectral posible y que estos sean factibles de ser realizados, con las condiciones del canal y el costo de la tecnología de los dispositivos.

Corrección y Detección de Errores

Este tema posee numerosas vertientes desde un simple bit de paridad, el cual mide cuantos unos y ceros se enviaron antes, y si es par pone un uno, caso contrario manda un cero (*igual es una convención, y puede ser al revés*). Esto permite determinar si hubo un error de un bit, pero no permite determinar errores de dos bits, pues un error significa una inversión de un bit y si dos bits se invierten, la paridad no se altera, pasando indetectado el error, sin embargo la probabilidad de no detectar un error, al igual que en análisis anterior mejora con un factor igual a BER^2.

Esto se estudia modelizando los mensajes como un espacio vectorial discreto y medio especial, en el cual se define una clase de 'distancia' entre símbolos, que serían como puntos de ese espacio. Para esto se suele usar un concepto de distancia especial llamada distancia de *Hamming*, que es similar a la distancia de *Manhattan*[8], en espacios euclídeos, resultando ser la suma de las diferencias absolutas, o número de bits cambiados. Siendo ésta, (*la de Manhattan*) una de las distancias generalizadas de *Minkowsky*[9] de grado 1, mientras que la distancia euclídea es de grado dos Raíz cuadrada de la suma de los cuadrados de las diferencias.

Todo esto compone el mundo del estudio de las codificaciones y los mecanismos de corrección de errores; entre ellos hay unos entes matemáticos llamados campos de *Galios*, basados en permutaciones y ellos son generalizaciones interesantes generando códigos de codificación con capacidad de corrección de errores basado en redundancias

[8] La distancia de Manhattan se llama también distancia de 'ciudad' pues mide cuanta distancia hay entre puntos de una ciudad si solo se puede viajar por las calles, en otras palabras: está prohibido ir en diagonal. De allí se llama así por la gran manzana, la isla de Manhattan de los EUA.

[9] La distancia generalizada o de *Minkowsky*, está definida como la raíz n-ésma de la suma algebraica del módulo de las diferencias de coordenadas elevado a la potencia 'n', donde 'n' es el "factor" de la distancia. Existen a su vez un sinnúmero de otras distancias para todo uso, siendo las más conocidas la euclídea (n=2), la de Manhattan (n=1) quedando muchas más, usadas en diversas aplicaciones como las de *Pearson*, *Jaccard*, *Mahalanobis* o *Chebicheff*, *Sokal* y *Michener*, *Bhattacharyya* o *Cavalli-Sforza*, entre muchas otras.

espacialmente formuladas, donde uno de los más conocidos es el código de Reed-Solomon (RS), usado desde los 90 en los CD y mecanismos redundantes de escritura y lectura de datos digitales en medios físicos, magnéticos, ópticos, y hasta en las transmisiones satelitales al/desde el espacio profundo y muchas otras aplicaciones más.

Hay una generalización de la teoría de RS con campos finitos y se llama códigos BCH, acrónimo creado en honor a los matemáticos que lo crearon en 1960: *Raj Bose* y *Ray-Chaudhuri*, aunque según *Wikipedia* también fue inventado independientemente en 1959 por *Alexis Hocquenghem*. Lo cierto es que esos códigos BCH o como se llamen, y quienquiera que los haya inventado, resultan muy útiles a la hora de empaquetar datos binarios en secuencias o '*símbolos*' y permiten crear maneras de codificar y decodificar paquetes de bits muy compactamente y con gran robustez, logrando a la vez una fuerte inmunidad de errores, contribuyendo a la eficiencia de canal y bajando el BER al mismo tiempo.

Lo cierto es que estos algoritmos como el BCH21 (*el número 21 indica el total de bits del 'paquete' o sea su longitud 'n'*) estos códigos llevan bastante cálculo para su de decodificación, que es del orden del cubo de la longitud del paquete en bits $\mathbf{O}(n^3)$, pues se basan en la resolución de polinomios enteros; sin embargo éste cálculo puede ser óptimamente implementado mediante un hardware programable tipo EPLD, o con el uso de microcontroladores dedicados tipo DSP (*Digital Signal Processor*) donde ciertos modelos poseen sets de instrucciones muy específicos, útiles para este tipo de cálculo.

Existen numerosos métodos de codificación; vale mencionar el de LDPC (*Low Density Parity Check Codes*) o "códigos de paridad de baja densidad" de R. Gallager, planteado ya en 1963 pero reflotado recién en los 1990; luego existen los Turbo-Codes, cuya patente expiró en 2013. Actualmente varias combinaciones de estos métodos son usado para mejorar la eficiencia de transmisión como el sistema MIMO para WiFi y WiMAX.

Ecualización Automática y Corrección de Errores

Uno de los más conocidos esquemas de corrección de ISI automática es la utilizada por los modems telefónicos, y los de banda ancha como los de tipo ADSL que aún se usan para brindar internet por pares de cobre. Estos canales de cobre (*conductivos*), ya sea telefónicos o coaxiales, poseen la característica de ser bastante LTI, pero a pesar de esto son muy dispersivos y también un poco alineales, poseyendo de paso alguna distorsión introducida por los amplificadores troncales (*en caso de TV e Internet por cable coaxial)*. Estas redes compensan cada tantos metros la pérdida en dB de la señal (*por la conocida conflagración entre los Sres. Ohm y Joule*).

Estas alinealidades generan armónicos que caen fuera de la banda útil contribuyendo solamente con un poco de intermodulación 'visual' con bandas y basura para los pobres usuarios que ven televisión, pues usan el mismo coaxial y amplificadores de banda ancha; justamente por eso los técnicos de las empresas colocan esos misteriosos "filtros" a la entrada de los televisores y también a la salida de los modems de cable, ellos evitan la ISI entre canales adyacentes en el cable, por intermodulación evitando la 'saturación' de los amplificadores troncales y los sintonizadores digitales de los modems.

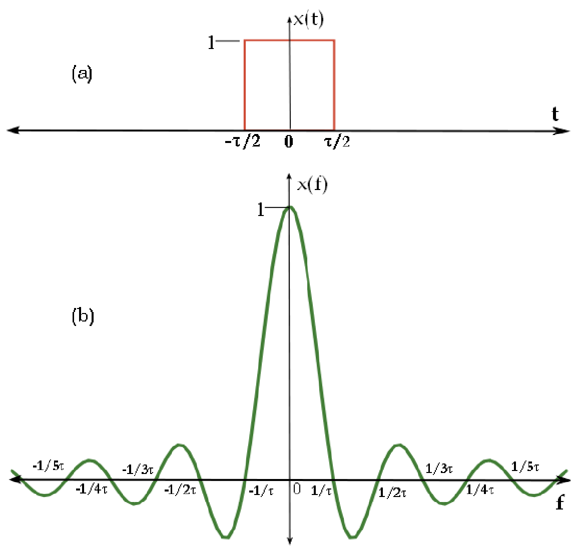
Por ejemplo analicemos un módem telefónico clásico, de los cuales hoy (2014) entran en desuso en nuestro país, el ancho de banda "garantizado" para líneas telefónicas es de aproximadamente 2400 Hz, en consecuencia para lograr esa tasa, (*2400 b/s*) se podría usar un simple pulso gaussiano, pero resulta que logran 56 kbps mediante la utilizació0n inteligente de técnicas digitales para el canal, corrigiendo errores a la vez de compensar el mismo, que por cierto es muy dispersivo, pero suficientemente lineal para lograrlo con razonable éxito; logrando en caso ideal, el equivalente a varias veces (>20) la capacidad del canal en Hz, y esto es verdadera ingeniería de comunicación, imposible sin las técnicas de procesamiento digital de avanzada. Ellas son a la vez sugeridas en normas de la CCITT, las cuales regulan los modos de modulación, los protocolos y recomiendan métodos para los sus sistemas de compensación automática, detección y corrección de errores, salvas de sincronización, etc.

Algo similar ocurre con los modems ASDL y ADSL-2, cuyo esquema es mucho más complejo usando FDM (*división de frecuencias*), y luego en cada canal, usan QAM y un sinnúmero de filtros digitales, implementados mayormente mediante DSP dedicados.

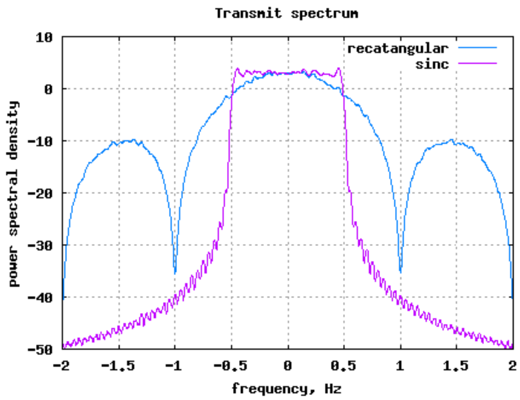
ISI en un canal simple

Volviendo al tema de la ISI vamos a mostrar como se ve un sistema como el inicial presentado, con un canal dispersivo de primer orden (*un solo polo*) cuando la relación de ancho de banda respecto al uso de Shannon varía.

Si consideramos un canal de comunicación con ancho de banda limitado y usando el criterio de *Nyquist*, queremos enviar un pulso, para hallar este pulso que responde a este criterios, se puede simplemente realizar la transformada inversa del propio canal (*de ancho de banda limitado*), resultando en la conocida función ***sinc***(t) siendo exactamente inverso al de la figura siguiente, intercambiando tiempo por frecuencia:



El problema es que este "pulso" no podría existir realmente pues debe arrancar desde ***-infinito*** y terminar en ***+infinito***, cosa obviamente inviable. Para arreglar esto, se puede recortar por ejemplo la longitud de este pulso, en ***+/- 2T***, creándose, obviamente un ancho de banda ilimitado y no ideal, dado por la convolución de este escalón que limita el tiempo (una ***sinc***()) con el escalón ideal en la frecuencia, la figura siguiente muestra la densidad de energía en ***dB***, respecto al ancho de banda, primero de un pulso temporal cuadrado (*en celeste*), y luego de un pulso limitado en ***+2T*** y ***-2T*** tipo ***sinc*** (*en violeta*)



Esto muestra gráficamente que el "ruido" introducido al canal por fuera del ancho de banda permanece bajo, y como todo sistema LTI, si "sustraemos" esa información el resultado es igual que si introdujésemos un ruido inverso (*restamos la señal fuera del ancho de banda del canal real*). Luego cuando tratemos de decodificar la misma, esta "resta" o falta de información, se manifestará como un "ruido" indeseado "sumado", con la única salvedad que no es gaussiano, sino es totalmente dependiente de la señal y su contenido en símbolos y en consecuencia es muy proclive a producir interferencia entre símbolos (ISI) si supera cierto nivel de energía y/o existe ruido agregado.

Diagrama de Ojos

Para visualizar este fenómeno de interferencia, se utiliza un esquema que se llama "*diagrama de ojos*", y se obtiene superponiendo en una misma imagen los diferentes trazos de la señal timada entre los instantes de muestreo T. Como la señal tendrá dos estados posibles +1 y -1 (*correspondiendo a cero y uno transmitidos*) se observaría que ella debiera de pasar en los instantes de muestreo, solo por los valores +1 y -1, sin embargo debido al ruido, pasa por lugares un poco más arriba y abajo, dependiendo de los datos anteriores, y si alimentamos esto con datos aleatorios, se recorrerá todas las posibles señales generando un diagrama que parece un ojo, y cuanto más cerrado esté, mayor será la indeterminación de la señal y la propensión a errores.

<u>NOTA</u>: En la práctica, cuando no hay a mano analizadores digitales, esto se logra ver mediante un simple osciloscopio, en el cual se sincroniza el disparo con el momento de muestreo y el tiempo se ajusta entre T y 2T; y al ser recurrente, el fósforo oficia de memoria, observando diagramas "de ojos" como los que se mostrarán a continuación.

En esos diagramas, precisamente la apertura de ojos, es una medida de la calidad del canal, indicando baja cantidad de ISI y/o ruido. En estos diagramas se puede diferenciar el ruido de la ISI por desadaptación, viendo que el cierre de ojos por ISI suele tener trazos finos recurrentes, mientras que el cierre por ruido, engrosa los trazos, dado que el ruido suele ser gaussiano y al ser totalmente aleatorio produce ese efecto, de engrosar aditivamente cada uno de los caminos producidos por las diversas fases.

Vamos a hacer un modelo del canal y analizaremos que es la ISI en este modelo y ver como disminuirla y llevarla al mínimo y/o anularla.

Supongamos que la señal *s(t)* sería la salida se representa con la siguiente fórmula, donde a_n representa la señal a transmitir, T es el período, g(t) es la función de filtro Supongamos que no hay filtro de forma de onda y que g(t) es δ(t)

$$s(t) = \sum_{n=-\infty}^{\infty} a_n g(t - nT)$$

Si queremos recuperar la señal a_n (*los datos o símbolos*) muestreamos simplemente en intervalos regulares iguales a *kT* (*el período de la señal*).

$$s(kT) = \sum_{n=-\infty}^{\infty} a_n g(kT - nT)$$

Si el sistema es LTI, podemos separar esta sumatoria en dos partes, una cuando $k = n$ y la otra no:

$$s(kT) = a_k g(0) + \sum_{n=-\infty}^{\infty} a_n g(kT - nT)$$

Lo que se puede deducir simplemente es que *s(kT)* es la suma de los datos $\boldsymbol{a_k}$ que nos interesan multiplicado por una función *g(T)* sumado a un segundo término que depende la misma función *g(t)* y de todos los demás datos (*diferente del momento k*) y que *debiese de ser igual a cero* para no tener interferencias desde otros símbolos.

En otras palabras el segundo término expresa la ISI de este modelo y para asegurar que no haya interferencia ISI, bastaría con poder hacer nulo este término y la única manera es que los valores de la función g(t) sean iguales a cero en los momentos kT = (-T, T, 2T, 3T... nT) mientras que la función g(t) debe ser diferente de cero en T = 0 que es el momento del muestreo, donde se recupera el valor $\boldsymbol{a_k}$ simplemente comparándolo con cero, (*en ausencia de ruido*).

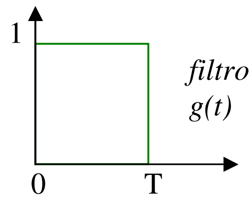
Filtros de Conformación de Modulación

El filtro más simple para transmitir una señal binaria que es representada como una secuencia de deltas multiplicando el dato (secuencia x_n), es el filtro 'hold' de orden cero, o sea un filtro que sea "0" desde siempre hasta que llegue el dato, luego sea "1" desde el momento preciso de llegada del dato (*hold*), hasta el próximo muestreo T, y luego de esto valga cero "0" para siempre, escrito en fórmulas con condiciones sería:

g(t) = 1, para $0 \le t < T$

= 0, para todos los demás valores de t

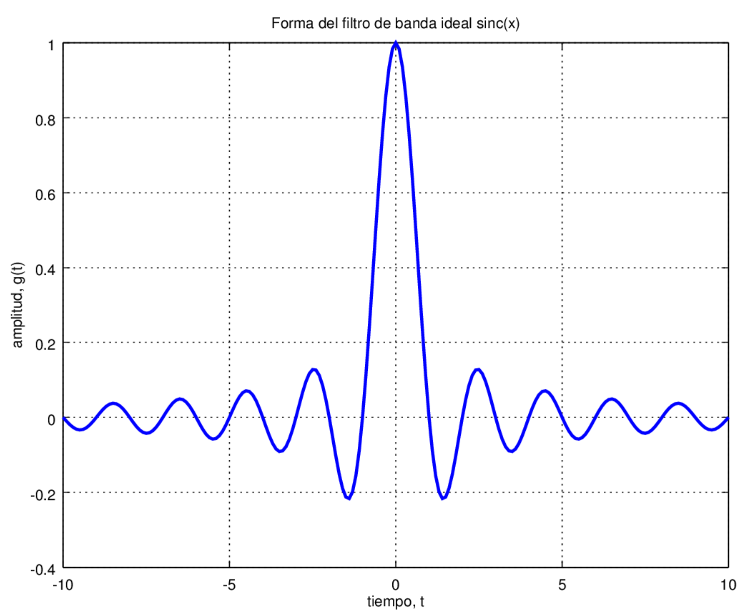
Este filtro *g(t)* es representado simplemente así:



Sin embargo la señal resultante de una señal aplicando este filtro es que poseerá un ancho de banda infinito, representado por la ***sinc()***, no siendo apto para canales con ancho de banda limitado, por lo que podríamos pensar en utilizar un filtro que genere una señal que posea un ancho de banda limitado.

Como sabemos que la transformada y la antitransformada son simétricas *(en forma, salvo un coeficiente)* un canal de ancho de banda "perfecto" es equivalente a un pulso cuadrado en las frecuencias, y equivaldría a un filtro ***g(t)*** dado por la función ***sinc(nT)***. Este es el llamado "canal de *Nyquist*" o filtro de conformación ideal.

Se muestra la gráfica de ***sinc(x)***, con ***x*** en radianes, con la conversión correspondiente se puede expresar en tiempo T de las muestras de los datos. Nótese que este filtro en particular pasa por cero en valores /- 'n x pi' y si el muestreo de la señal se realiza cada T o sea '2 pi' las "colas" de los símbolos anteriores, valdrán cero en el momento del muestreo, eliminando la ISI. Tener en cuenta que esta condición es ideal, y presupone que el canal no tiene dispersión ni retardo variable o fase lineal; por lo general esto no se da y las condiciones de ISI aparecen irremediablemente.



Sin embargo, a pesar de lo amigable e ideal de la situación tampoco es posible crear una señal de estas características, pues se debiera de poder crear un filtro con respuesta infinita y no causal, *(para n<t)*. A pesar de esa imposibilidad, si se limita este filtro a algunas muestras (períodos) a la izquierda (4-6 T), debido a que el módulo de la función

sinc(), baja muy rápidamente *(proporcional a 1/t)* fuera del valor ***t = 0*** por lo que se comete relativamente poco error al transmitir así, pero generando espectro vestigial *(fuera de la banda)* de bajo nivel.

Este espectro vestigial, si pasara por el canal sería bárbaro, -*pero esto no puede pasar*- , ya que un canal 'real' siempre será de un ancho de banda limitado y en consecuencia la ausencia de esta señal (*fuera de la banda*), se verá como un error "agregado" *(en realidad restado, para ser exacto)* esto es debido a que los errores por lo general se miden como la suma de cuadrados de la distancia entre la señal 'real' y la que llega, por eso da lo mismo considerarlo.

Truncado del filtro de conformación de Onda

Algo a tener en cuenta en sistemas reales, es que tanta teoría de señales 'pura' a veces contamina el pensar claramente en el sistema real, por ejemplo si se construye un modulador como el propuesta, 'cortando' literalmente en +/- 5 muestras (*ciclos*) la señal, y se envía esta a un canal real, en realidad se estarían enviando vestigios de señal fuera de canal; y hay que tener cuidado pues a veces los canales luego se multiplexan en frecuencia y se agregan a canales laterales muy cercanos, entonces el exceso de señal (*fuera de banda*) dado por el corte de la señal que es equivalente a la convolución en la frecuencia con *sinc(5/T)* y esto agrega señal fuera de banda, si esta señal espúrea es "sumada" a los demás canales, generará ISI entre canales (*agregando ruido no-gaussiano*) lo cual es un problema grave dado que le estadística de los datos que actúan como ruido, no es necesariamente gaussiana, pudiendo comprometer los sistemas de corrección de ruido, controles de ganancia, etc. de los demás canales (*adyacentes*).

Otro de los problemas frecuentes es cuando esta señal no se filtra expresamente, por más que parezca 'poca' la energía fuera de banda, los amplificadores y filtros suelen no reaccionar linealmente fuera de la frecuencia pasante, generando distorsiones por alinealidades y batidos, que pueden aterrizar en la misma banda pasante o en otra señal de un canal adyacente produciendo una ISI muy dificultosa de eliminar y detectar. Como regla general es conveniente filtrar digitalmente lo más posible, y luego colocar filtros analógicos "reales" para prevenir el solapado o "*aliasing*" fuera de banda.

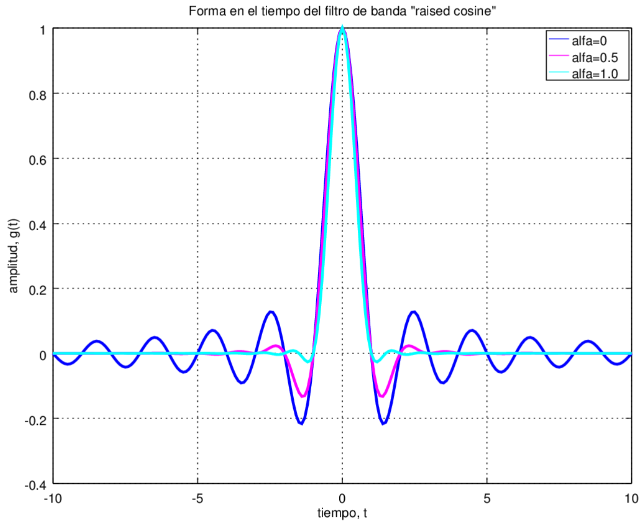
Esto es válido tanto en generación de señal para modulación, como en recepción de señal para su demodulación, donde también se colocan filtros que previenen el *aliasing*.

Por lo general los estándares de comunicaciones ya especifican el tipo de función a aplicar para la generación de la energía en banda y la clase de modulación.

Las soluciones reales pasan por esquemas de modulación bastante probados como la "*raised cosine*" que aplican una ventana adicional (*tipo coseno, no cuadrada)* cuadrada la cual permite un uso más eficiente del canal, sin tanto filtrado adicional posterior. La fórmula de esta ventana es la siguiente:

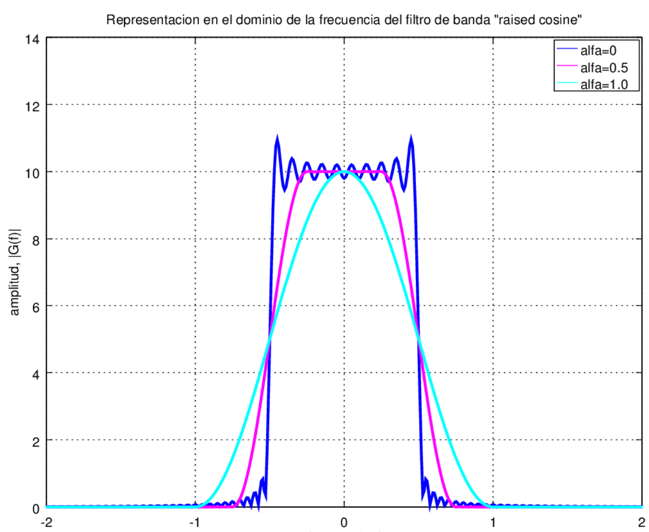
$$g(t) = \left(\frac{sen(\pi t / T)}{\pi t / T} \right) \left(\frac{\cos(\alpha \pi t / T)}{1 - (2\alpha \pi t / T)^2} \right)$$

Y cuya gráfica, para tres valores del parámetro α es:



En esta ecuación, el parámetro α es el parámetro llamado de "exceso de banda" y se varía entre 0 y 1. Con $\alpha = 0$ el filtro se reduce al clásico canal de *Nyquist* sin exceso de banda fuera de +/-1/2T. Cuando $\alpha = 1$ el filtro presenta exceso de uso de banda y no ocupa frecuencias fuera de +/-1/T

Para analizar como este filtro se comporta llenando el espectro, se muestra la grafica[10] de este impulso, en el dominio de la frecuencia.

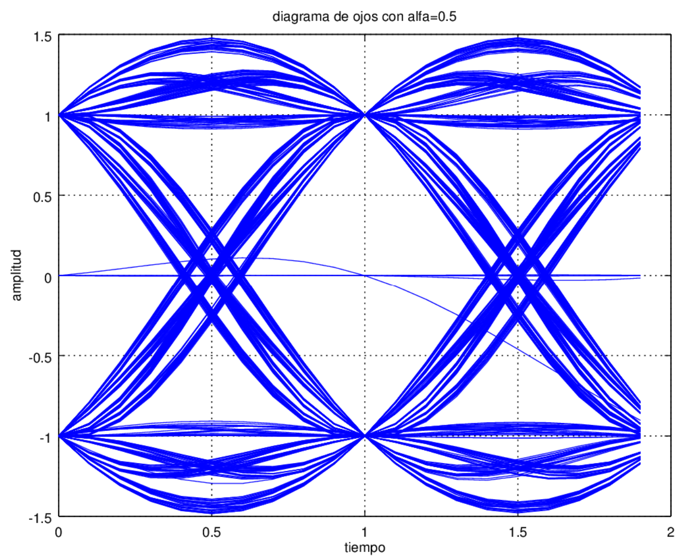


[10] Esto se generó con un código simple de simulación en MathLab.

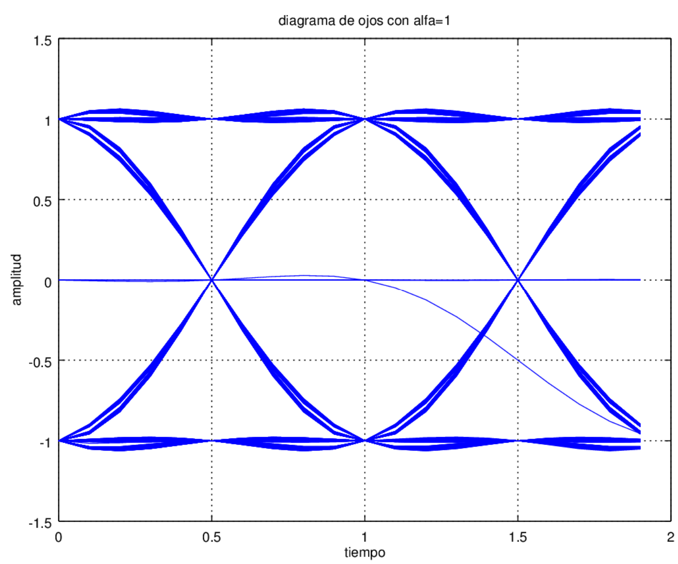
Diagrama de Ojos Ideal

Si a este canal se le ingresa una secuencia aleatoria de datos, y luego se toma la salida sincronizada cada 2T y se superponen en una sola imagen, se obtiene el diagrama de ojos, donde se puede ver la ISI como el fenómeno de cierre de estos ojos, se verá que cuanto más mal se aprovecha este canal, mas cerrados estarán.

Se muestra a continuación, el diagrama de ojos para alfa con valor 0.5



Se muestra a continuación, el diagrama de ojos para alfa con valor 1



Nótese en este dibujo que el ojo está mucho más abierto, sin embargo se está desaprovechando el canal pues solo se usa la un poco más de la mitad del espectro (*en términos de energía*). Lo cual nos lleva a considerar la eficiencia de uso de la banda pasante de un canal como un factor importante.

Uso Eficiente de un canal

Una cosa a tener en cuenta es que cuando se transmite potencia en el espectro cercano a la zona de limitación de banda de un canal se suele tener que lidiar con alinealidades y saturaciones, debido a que suelen ponerse elementos pasivos para limitar transmisión de energía esta zona, pero no siempre están colocadas manteniendo impedancias sino a veces hay efectos de resonancia que se agregan con polos y ceros cercanos a la banda de corte, y eso a veces absorbe energía alinealizando las etapas de salida por cambios bruscos en la impedancia equivalente, tal es el caso de la telefonía en la cual se ponen "*peeking coils*" para compensar la pérdida en banda por longitud de cable y forman circuitos auto resonantes con las capacidades equivalentes de salida de los amplificadores "*drivers*" generando problemas espurios muy difíciles de diagnosticar.

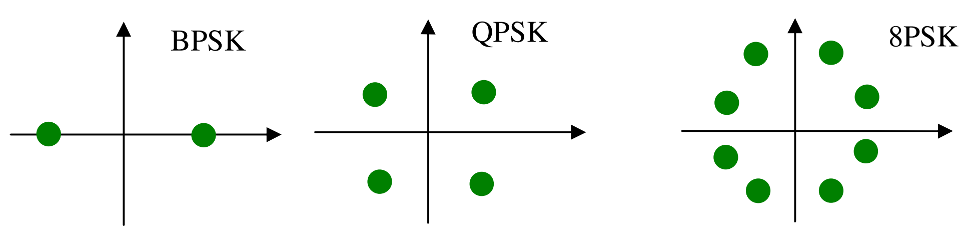
Una medida preventiva de calidad, con sistemas físicos *(cableados, fibra óptica, red eléctrica, etc.)* es tratar de balancear la energía con la forma del canal pasante. En otras palabras el filtro óptimo es el cual tiene igual forma de paso de banda que la banda pasante del canal, si bien esto parece ilógico, tiene su lógica por lo antedicho.

También conviene balancear la energía enviando la misma en el centro del canal pasante, sin olvidar de no abusar de los extremos, y si es posible obviarlos, pues éstos extremos de banda son especialmente dispersivos y producirán sin dudas ISI por fase y tal vez por efectos de saturación e intermodulación. Esto no se aplica siempre sino en canales que está cerca de banda base como la telefonía, donde la banda pasante "*segura*" es de 300-2400 Hz mientras que los canales pueden en ciertos casos llegar a 50Hz y 5000Hz, pero es mejor no usar esas zonas del espectro pasante.

Tipos de Modulación

Los sistemas de transmisión clásicos utilizan esquemas de modulación de fase y de amplitud, y sus nombres suelen denotar el tipo de modulación, los acrónimos usados son PSK (*Phase Shift Keying*) y BPSK (Binary PSK), AM (*Amplitude Modulation*) también hay QAM (*Quadrature Amplitude Modulation*) y nPSK (n-value *Phase Shift Keying*), QPSK (*Quadrature PSK*).

Los diagramas de fase/amplitud de los datos (llamados constelaciones) serán:



Existen un conjunto muy amplio de sistemas de modulación aparte de los mencionados, uno de los más interesantes son el conjunto de los llamados OFDM, que recita (*Ortogonal Frequency Distribution Modulation*) y se basan en señales ortogonales, es decir se modula de algún modo, usando funciones ortogonales entre sí. Hay un tutorial muy interesante de *Ian Poole* en inglés aquí[11]

Por ejemplo, si se usa una función senoidal de cierta frecuencia f1, ella será ortogonal a cualquier otra señal de otra frecuencia diferente f2 ≠ f1 .

Igualmente hay que tener cuidado pues en especial si estas señales son moduladas, aparecerán inevitablemente un espectro resultante de la modulación y dependiente del método, pero si se eligen ingeniosamente las frecuencias, en el caso de senoidales, como que sean múltiplos de un número, pero no entre sí, las señales resultantes de todas las modulaciones, caerán en lugares del espectro únicos, disminuyendo la posible ISI.

Funciones Ortogonales

Existen infinitas secuencias de funciones de este tipo, cuyo producto interno es nulo (*expresado mediante una integral en el caso continuo, o sumatoria, en el caso discreto*).

Las más interesantes son aquellas que tienen un espectro limitado y de algún modo están limitadas en el tiempo. Obviamente las senoidales son el antiejemplo, pues no tienen limitación en el tiempo y tienen un espectro puntual, pero son las 'estrellas' de estas transformaciones tan útiles como las de *Fouriere*.

Hay unas funciones llamadas "ondita" (*del inglés wavelets*) que conforman un conjunto interesante de funciones ortogonales, mediante las cuales se logran cosas piolas en especial para procesamiento de imágenes y reconocimiento de patrones en voz. El espectro de ellas si bien es infinito posee bastante concentración en una zona delimitada.

Otras son las secuencias pseudo-aleatorias generadas mediante juegos de números primos, usadas en criptografía y transmisión codificada de espectro distribuido como los usados en los códigos CDMA (*se verá un poco más adelante un ejemplo*). Estas secuencias son de espectro amplio, plano y muy denso.

Transmisión y Recepción con Funciones Ortogonales

Para todas las funciones ortogonales, el concepto general de su utilidad en las transmisión y recepción de datos usando estos métodos, es la siguiente:

> **Modulación y Generación de Onda**
> La señal se genera usando un conjunto de funciones ortogonales ponderadas por los datos mismo de diversos modos, a veces se pre-procesan los datos mismos numéricamente para que la secuencia tenga valor medio nulo, por ejemplo, o cierto tipo de paridad, agregando datos innecesarios para validar integridad, como el CRC (*Cyclic Redundancy Code*) hasta llegar a codificaciones robustas auto-reparables como los TurboCodes, entre muchos otros.

[11] OFDM http://www.radio-electronics.com/info/rf-technology-design/ofdm/ofdm-basics-tutorial.php

Limitación de Banda
Se suele limitar mediante componentes discretos las señales fuera de banda, para evitar problemas como intermodulación, alisaing, distorsión, e interferencia. A veces esta función se hace con un filtro digital que 'une' señales y luego compone los segmentos de frecuencia, pero igual van filtros discretos pues en la conversión digital-analógica siempre aparecen espectros indeseados a eliminar y cuanto más lejos de la fundamental o la banda pasante, mejor (*usando oversampling*).

Filtrado de Entrada
Se usa por lo general una etapa analógica que acondiciona la señal filtrando datos para evitar el aliasing (*solapado*) del muestreo, y ajuste el rango dinámico con un mecanismo de control automático de ganancia, para evitar el molesto ruido de cuantización (por digitalización con pocos bits), maximizando el rango dinámico. Estos filtros hoy se tratan de hacer lo más simples posibles ya que el trabajo máximo se empuja hacia lo digital.

Demodulación
Aquí es en donde se pone de manifiesto las propiedades de estas funciones, pero como los canales son dispersivos y no siempre se sabe su comportamiento antes de usarlos, se suelen poner filtros acondicionadores de señal para compensar el canal, respecto a la dispersión, minimizando la ISI. Estos filtros suelen ser entrenados con una "salva" y/o usando mecanismos de convergencia propios de la modulació, por ejemplo cuando se modula mediante un PNG (*Pseudo-Noise-Generator*). Este método es un polinomio digital implementado con un registro de desplazamiento realimentado en donde el dato se agrega con un xor. Estos filtros/moduladores y demoduladores son autoconvergentes con su propia su secuencia y los demoduladores se constituyen como filtros ARMA (*Auto Regressive Moving Average*), nombre muy sugestivo de paso. Como digo hay numerosos mecanismos por el cual se corrige la dispersión del canal, para luego hacer uso de las propiedades de ortogonalidad de las funciones de modulación. Luego se procede a reconstruir esas funciones y realizar el producto interno definido para ellas para reconstruir luego el dato que había sido codificado.

Conformado/Recuperación de Mensaje
Dependiendo del mecanismo de modulación, suele realizarse un proceso digital, como proceso inverso; luego se determina en cada canal en paralelo la secuencia de símbolos conjuntos y en algunos casos se realiza el proceso digital inverso al que se hizo previo a la modulación, obteniendo desde una medición de errores hasta la reconstrucción de la señal por paridad correctiva, usando distancias y códigos de Hamming, entre otros métodos.
NOTA: Esta etapa puede contener mecanismos de detección y corrección de errores de n-bits, a su vez hay etapas más macro que podrían detectar errores de mayor orden, de secuencia entre otros. Recordar que mencionada eficiencia de codificación es el factor con el que se juega aquí.

En estos sistemas, lo interesante es que la demodulación si bien es el proceso inverso a la modulación, la contribución de las distorsiones y dispersión del calan y el ruido mismo, hacen que el mecanismo de demodulación pueda intrínsecamente eliminar el ruido (*hasta cierto nivel*) y ser inmune a dispersión de cierto grado.

Importante es notar que el uso de funciones ortogonales permite emplear el producto interno como mecanismo principal de demodulación, pues éste da solamente un valor diferente de cero, cuando la señal es la propia y da cero para todas las demás, en especial las producidas por otros símbolos anteriores, reforzando la inmunidad a este tipo de interferencias que son precisamente las ISI de este trabajo.

Sistemas con Frecuencias Ortogonales

Una variante de OFDM es FDM (*Frequency Division Modulation*) que devino en un estándar de telefonía llamado FDMA (*Frequency Division Multiple Access*) que luego dio lugar al desarrollo de CDMA, con mejor aprovechamiento simultáneo del canal.

En la telefonía celular la ISI constituye una verdadera pesadilla, suele ser producida pro múltiples motivos como el rebote multi-path, el movimiento del terminal o del elemento de rebote, pero no tanto la dispersividad del medio, pero sí las interferencias de otros celulares, dado que hay simultaneidad de transmisiones en los medios (aire) y una antena puede estar conectada simultáneamente con cientos de terminales usando la misma radiobase y tener además interferencia de las demás antenas. En este caso habrá cientos de datos interfiriendo si no hay una sincronía especial (temporal y espacial) y se toman en cuenta una enorme cantidad de problemas ya sea los propios de un solo canal como los rebotes múltiples en edificios y construcciones, sumados a los cambios constantes (non-time invariant) si nos estamos moviendo ya sea caminando o en auto con el terminal móvil, etc. etc.

Para plantear una solución ingeniosa, se inventó y patentó un mecanismo que opino es 'genial' utilizando códigos de modulación binarios que son ortogonales entre sí. Estas son secuencias de tipo pseudo-aleatorias (PN), generadas adrede y especificadas en la norma que presentan la particularidad de que, aparte de ser ortogonales y usar el canal en forma eficiente, su interferencia asemeja un ruido gaussiano. Justamente en este caso el decodificador solamente se degrada un poco cuando aparece sumada señal de los demás canales, tanto más degradada cuanto más canales se usan simultáneamente. Justamente aquí es en donde aparece duramente el problema de ISI, a la cual esta teoría es muy sensible, creando problemas pues a la menos alinealidad esta supuesta "gaussiana" se tornaba demasiado causal, generando fallas de código y de sincronismo.

El mecanismo exacto que usa el sistema CDMA (*Code Division Multiple Access*) se llama técnicamente DSSS (*Direct Sequence Spread Spectrum*) y sus detalles exceden este escrito; pero lo interesante es que se usa una ingeniosa manera de llenar el espectro con algo muy 'parecido' al ruido blanco, que luego se puede decodificar solamente teniendo el código exacto de cada secuencia; y todo esto no se pudiera haber logrado de no ser por procesamiento exhaustivo de señal y una teoría matemática-estadística muy robusta.

Un sistema que usó masivamente el CDMA es la telefonía móvil, cuyo nombre estándar de ITU es IS-95, (*pues se creó ese año, en 1995*) que por cierto era muy bueno y en teoría podía hacer un uso muy eficiente de los canales; pero la contra era que la carga de canales (*el número de canales usados en simultáneo*) influía en el ruido y la calidad de todos los canales, siendo muy castigados los que estaban con menor señal.

El otro problema (*casi parece secundario*) es que los amplificadores de RF debían de ser casi (*obsenamente*) lineales (*demasiado*) dado que cualquier alinealidad generaba intermodulación que caía dentro de banda como un ruido y el sistema es nativamente de banda ancha pues todos los canales usan todas las bandas dentro de los canales asignados. También era sensible a ISI por rebotes, dado que los códigos eran todo iguales (aleatorios) podían interferir entre sí en especial con la sincronización del chip (segmento de datos mínimo).

Todo esto tornó costoso y complejo el mantenimiento y constante calibración de las radiobases; pero también complicaba el costo del procesamiento digital dentro de los terminales (celulares) tornándolos más costosos, consumiendo más batería y siendo más sensibles a la calibración de sus etapas de RF que debían de ser super-lineales.

Este y otros problemas comerciales como las patentes de ***Qualcomm***, determinaron que se deje de usar masivamente el sistema CDMA en el mundo entero, salvo contados países que aún los usan como China y algunos de América. Actualmente se usa en sectores rurales para reemplazar telefonía fija con semi-móvil dado que hay mucho material obsoleto muy barato y disponible, lo caro es el mantenimiento.

Otro tema comercial negativo es sin dudas la patente sobre la tecnología de modulación CDMA que obliga a pagar bastantes royalty (derechos), y en esa arena se debaten tanto la utilidad como la posibilidad de implementar en términos económicos, de demasiadas cosas en comunicaciones, software e industria en general.

Si embargo existen muchos métodos derivados que siguen en uso con este tipo de modulación de espectro distribuido y llegaron con presión de la llegada del 3G, como lo es el WDCMA cuyo estándar se llama *IMT-2000 direct spread.*

Además todo esto evolucionó entre 2G, 2.5G y 3G con CDMA2000, se usó esta tecnología entre el 2000 y el 2006 en América y Asia, (Argentina Incluido, la empresa ***Movicom***, luego de adquirida por Telefónica se fusionó y transformó en ***Movistar***)

A pesar de todo el despliegue, no tuvo demasiado éxito, siendo superado por tecnologías como GSM-TDMA 3G y últimamente LTE (4G) de costo decreciente con capacidad de canal e inmunidad a problemas creciente.

Modulación en Cuadratura

Básicamente lo que se explota es que, por ejemplo una señal seno y coseno de igual frecuencia son ortogonales por definición, esto permite la modulación QAM ya que seno y coseno solo se diferencian en $\pi/2$ (90º / grados). La demostración de ortogonalidad es simple pues integral del producto de ambas funciones entre $+\infty$ y $-\infty$ es cero.

La explicación en el espectro es porque la transformada de señales reales, tiene espectro conjugado simétrico, y en el caso de coseno se está usando el doble de energía, en consecuencia cuando se modula en fase +π y - π, equivale a usar independientemente ambos sectores del espectro por esa propiedad de simetría, pudiendo codificar dos bits con una misma frecuencia.

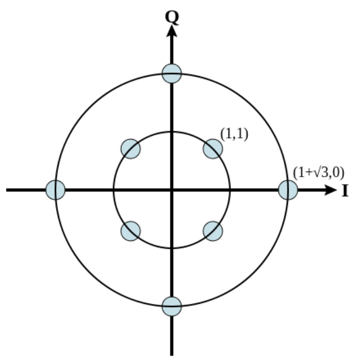
Modulación QAM y QPSK

Debido a que el número de estados que codifican un símbolo suelen ser la multiplicación del número de estados de cada señal ortogonal (seno y coseno) de la modulación, y éstos suelen arreglarse en forma binaria para permitir una modulación de palabras binarias completas, los números de "estados" de las señales serán 2 (el ordinario) 4 para 2 estados en cada señal, 16 para 4 estados y 64 para 8 estados.

En consecuencia los modos de modulación compleja por cuadratura serán 4QAM, 16QAM, QAM-16, 64QAM, 256QAM, QAM-64, etc.

El número delante o a veces atrás de la sigla como en QAM-64, indica el número de estados y no siempre coincide con el número de "ojos" siendo por lo general un cuadrado perfecto binario.

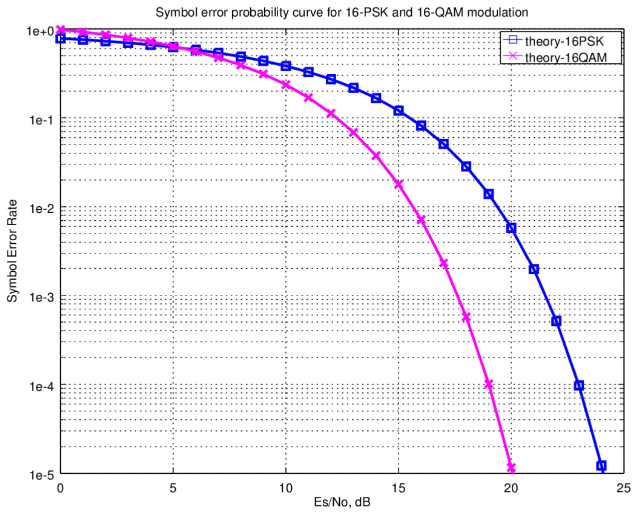
Los errores y las zonas se miden mediante otros diagramas que se llaman constelaciones, y tienen esta forma (*gentileza Wikipedia*) para 8QAM:



Interesante es de notar, que a pesar que a pesar de que la frecuencia y la fase pueden ser considerados una la derivada de la otra, el comportamiento de los sistemas de QAM frente a los QPSK es mejor y es un poco más inmune al ruido, haciendo uso más eficiente del canal *(Shannon)*. La diferencia no es abismal, pero sí son algunos ***dB*** y la demostración de eso existe y se puede hallar en la referencia[12]

En el gráfico siguiente, obtenido con una simple simulación con MatLab/Octava de las fórmulas provenientes de esta teoría, se muestran el resultado de la BER de ambos mecanismos de modulación QPSK-16 y QAM-16 ante el ruido en un canal gaussiano clásico, sin filtrado:

[12] Ver demostración aquí (link válido al 10/2014): http://www.dsplog.com/2008/03/29/comparing-16psk-vs-16qam-for-symbol-error-rate/



Hay toda una rama de la matemática y la ingeniería que se pasa en cómo asignar códigos binarios a cada uno de estos estados de fase, para minimizar las probabilidades de errores debido a transiciones, y dado que el tema es complejo, simplemente lo comento, invitando a quien se interese en profundizar: el tema es realmente apasionante.

© ***Andrés Hohendahl***